



# **Adoption of ICT4D frameworks to support screening for depression in Nigerian Universities**

By

**Blessing Onuwa Ojeme**

Thesis presented for the degree of Doctor of Philosophy (PhD) in the Department of Computer Science, Faculty of Science, University of Cape Town, Rondebosch, South Africa

Supervised by:

Professor Thomas Meyer

Department of Computer Science, University of Cape Town, South Africa

Associate Professor Audrey Mbogho

Department of Computer Science, Pwani University, Kenya

July 2018

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

## Plagiarism Declaration

I, Blessing Ojeme, declare that,

- a) The content of this dissertation is my original research, except where otherwise stated.
- b) This dissertation has not, and will not, be submitted for any degree or examination at any other University
- c) This dissertation does not contain other persons' data, pictures, graphs or other information unless specifically acknowledged as being sourced from other persons.
- d) This dissertation does not contain other persons' writing unless specifically acknowledged as being sourced from other researchers.
- e) This dissertation does not contain text, graphics, or tables, copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the references section

Signed by candidate

Blessing Ojeme  
PhD Candidate

## List of Publications

**Paper A (Springer Publication).** Blessing Ojeme and Audrey Mbogho (2016). Predictive Strength of Bayesian Networks for Diagnosis of Depressive Disorders. *Proceedings of the 8th KES International Conference on Intelligent Decision Technologies (KES-IDT 2016)-Part II*, Tenerife, Spain, June 15-17, Vol. 56, Pp. 373-382. Available at: [http://link.springer.com/chapter/10.1007/978-3-319-39630-9\\_31](http://link.springer.com/chapter/10.1007/978-3-319-39630-9_31)

**Paper B (Elsevier Publication).** Blessing Ojeme and Audrey Mbogho (2016). Selecting Learning Algorithms for Simultaneous Identification of Depression and Comorbid Disorders. *Proceedings of the 20th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES-2016)*. York, UK, September 5-7, Vol. 96, Pp. 1294-1303. Available at: <http://www.sciencedirect.com/science/article/pii/S1877050916319846>

**Paper C (IEEE Xplore Publication).** Blessing Ojeme, Audrey Mbogho and Thomas Meyer (2016). Probabilistic Expert Systems for Reasoning in Clinical Depressive Disorders. *Proceedings of the The 15th IEEE International Conference on Machine Learning and Applications (IEEE ICMLA'16)*, Anaheim, California, USA, December 18-20., PP. 599-604. Available at: <https://ieeexplore.ieee.org/document/7838209/>

## **Acknowledgements**

I am deeply grateful to my supervisors, Professor Thomas Meyer and Associate Professor Audrey Mbogho, who generously rendered help and encouragement during the process of this study. Whatever I have accomplished in pursuing this undertaking is due to their inspiring teaching, guidance, suggestions and thoughtful advice. You have instilled in me a love of curiosity and learning that would guide me for the rest of my life

Thank you to the head of the computer science department, Associate Professor Hussein Suleman, and other academic staff, for their valuable mentorship as well as encouraging me to find my own research path.

I am grateful to the Council of Scientific and Industrial Research (CSIR) and the Department of Science and Technology (DST), South Africa, for investing in me by providing funding for my research and study through the Centre for Artificial Intelligence Research (CAIR), South Africa.

I am forever grateful to the Accountant-General of Delta State, Nigeria, Mr Cyril Chima Agbele, for being financially and morally supportive during the PhD program. He is one of a kind.

The work presented in this thesis has also been significantly shaped by inspiration and knowledge I've received through discussions with many academic staff of the department – Professor Edwin Blake, Dr Melissa Densmore, Associate Professor Deshen Moodley, Dr Brian DeRenzi, Associate Professor Michelle Kuttel, and Professor James Gain. I have also benefitted immensely from my interaction with mental health professionals: Dr Kingsley Akhigbe, Dr Ehigiator Adayonfo, and academic staff of the University of Benin, Nigeria. I am grateful to you all. I also owe thanks to my many friends and colleagues at UCT (Selvas Mwanza, Amreesh Phokeer, Thomas Reitmaier, Jorgina Paihama Pereira, Fiona Ssozi, Joan Byamugisha, Maletsabisa Tsabi Molapo, Lighon Phiri, Aderonke Sakpere, Chika Banjo, Yamiko Msosa, Mvurya Mgala, Bhavana Harrilal, and many others ) who have helped me with my research and my social life.

I am grateful to have had the opportunity to work with many people outside academia who have taught me many valuable lessons about research, machine learning, human computer interaction and the “real world.” I have benefited immensely from them.

My greatest debt is, as always, to my lovely wife, Chidinma, and kids, David, Deborah and Hannah, who persistently rendered me the much-needed companionship and support on this part of the road of life.

## Abstract

Health is fundamental to development and access to healthcare is a major health and development issue particularly in developing countries where preventable diseases and premature deaths still inflict a high toll. In Nigeria, for instance, under-financing, inefficient allocation of limited medical resources has led to quantitative and qualitative deficiencies in depression identification, and to growing gaps in facility and equipment upkeep. The focus of the present study is Nigerian University students who are at higher risk of clinical depression than other populations. Besides high crime rate, acute unemployment, terrorism, extreme poverty and serial outbreak of diseases, which are everyday life situations that trigger depression for a large proportion of Nigerian population, Nigerian University students are faced with additional problems of poor living and academic conditions. These include constant problems of accommodation and overcrowded lecture halls caused by increasing population of students, recurrent disruptions of academic calendar, heavy cigarette smoking and high level of alcohol consumption. Effective prevention of medical condition and access to healthcare resources are important factors that affect peoples' welfare and quality of life. Regular assessment for depression has been suggested as the first important steps to its early detection and prevention. Investigations revealed that, besides the peculiar shortage in mental health professionals in Nigeria, the near absence of modern diagnostic facilities has made the management of this potentially detrimental problem impossible. Given this national health problem, and that it would take some time before resources, especially human, can be mustered, calls have been made by several bodies that other viable means that take cognisance of the difficulties of assessing mental healthcare be sought. This study is an attempt at exploring opportunities to increase flexibility in depression prevention and detection processes. The study investigated the effectiveness of developing computer-based methodologies, derived from machine learning and human computer interaction techniques for guiding depression identification process in Nigerian universities. Probabilistic Bayesian networks was used to construct models from real depression datasets that included 1798 data instances, collected from the mental health unit of University of Benin Teaching Hospital (UBTH) and primary care centre in Nigeria. The models achieved high performance on standard metrics, including: 94.3% accuracy, 94.4% precision, 0.943 F-Measure, 0.150 RSME, 0.923 R and 92.2% ROC. The findings from the information gain and mutual information show high correlation between "depression" and "alcohol or other drug consumption"; "depression" and family support and availability of accommodation", but low correlation between "depression" and "cigarette

smoking”. The results also show high correlation between “depression” and a synergistic combination of “impaired function and alcohol and other drug consumption”. Following the User-Centered design approach, a desktop-based screening tool was developed for use by University academic staff, as a first step, for regular screening of staff and students for depression, and where necessary, schedule appointment with the appropriate mental health authority for further diagnosis. Though the interesting results from the heuristic evaluations illuminate the challenges involved, it demonstrates the significance and relevance of end-user factors in the process of designing computer-aided screening intervention, especially with respect to acceptance of the system for use in non-clinical environment. The findings presented in this doctoral study provide compelling evidence of the huge potential that the collaboration of machine learning and usability techniques has for complementing available resources in the management of depression among University population in Nigeria. It is hoped that, given the persistent challenges of depression, the findings will be part of the ongoing global research to encourage the adoption of ICT4D frameworks for the prevention of more serious cases by empowering other population for an early first-line depression screening.

## Table of Content

### Chapter One

#### Introduction

1.1	Background information on depressive disorders	1
1.2	ICT4D tools and healthcare in Nigeria	3
1.3	Proof-of Concept	5
1.3.1	Data collection	6
1.3.2	Data analysis and results	7
1.3.2.1	Characteristics of the participants	8
1.3.2.2	Social demographics of clinicians	8
1.3.2.3	Knowledge and use of ICT tools in work place	9
1.3.2.4	Attitudes towards the adoption of ICT tools	9
1.4	Justification of choice of study specific areas	11
1.5	Research goal and objectives	12
1.6	Research contributions	13
1.7	Research questions	13
1.8	Thesis organisation	14

### Chapter Two

#### Understanding the ICT4D frameworks and depression screening relation

2.1	Introduction	15
2.2	Frameworks used in ICT research	16
2.2.1	Psychological theories	16
2.2.1.1	Theories of reasoned action	16
2.2.1.2	Technology acceptance model	17
2.2.1.3	Theory of Adoption and diffusion of innovation	18
2.2.3	Attributes of innovation	19
2.2.3.1	Perceived impact	20
2.2.3.2	Perceived ease of use	21

2.2.3.3	Perceived trial utility	21
2.2.3.4	Perceived result observability	21
2.3	Frameworks for developmental studies	22
2.3.1	The capability approach	22
2.3.2	Sustainable livelihoods approach	23
2.3.3	Millennium development goals	24
2.4	Frameworks guiding the study and justification	25
2.5	Depression and depressive symptoms	26
2.5.1	Screening for depression and depressive symptoms	28
2.5.1.1	Hospital Anxiety and Depression Scale	29
2.5.1.2	Beck Depression Inventory	29
2.5.1.3	The Patient Health Questionnaire – 9	29
2.6	Linking ICT4D, machine learning, HCI and depression screening	30
2.7	Bayesian network approaches in medical diagnosis	32
2.7.1	Bayesian networks classifiers	35
2.7.2	Conditional independence	36
2.7.3	Bayesian networks modelling process	37
2.8	Techniques in human computer interaction	39
2.8.1	Usability and user experience	40
2.8.2	Design process and principles	41
2.8.2.1	User interfaces	42
2.8.3	Usability evaluation methods	42

## Chapter Three

### Review of related literature

3.1	Introduction	45
3.2	ICTs and healthcare in Nigeria	45
3.3	ICT4D for socio-economic development in Nigeria	47
3.3.1	ICT Policy for development in Nigeria	47
3.4	Acceptance of ICT applications in the health sector and theories of	

adoption and diffusion of innovations	48
3.5 Machine learning application in healthcare	50
3.6 Usability of medical applications	54
3.7 Effectiveness of usability evaluation methods in medical applications	54
3.8 Machine learning and HCI collaboration	57
3.9 Summary of the review, gaps, limitations and opportunities	58
3.9.1 What this study does differently	58
3.9.2 Limitations	58
3.9.3 Opportunities	59
 Chapter Four	
Study methodology	
4.1 Introduction	61
4.2 Research design	61
4.2.1 Justifying the mixed research method chosen for the study	63
4.3 Existing depression diagnosis procedure in Nigeria	63
4.4 Design of a new machine learning-HCI screening system	65
4.5 Data collection and description of dataset	66
4.5.1 Data collection	66
4.5.1.1 Abstraction form	68
4.5.1.2 Use of data collectors	68
4.5.1.3 Data validation	70
4.5.1.4 Ethics	71
4.5.2 Data selection	73
4.5.3 Data pre-processing	74
4.5.4 Data transformation	74
4.5.5 Information gain	76
4.6 Problems encountered during data collection	80
4.7 Applying machine learning methods in the experimental set up	80
4.7.1 Naïve Bayes	81

4.7.2 Neural networks	82
4.7.3 Tree-based boosting	85
4.7.4 Decision Tree (C4.5)	86
4.7.5 k-nearest neighbours	87
4.7.6 kernel methods	88
4.8 Challenges of machine learning application in medicine	90

## Chapter Five

### Implementation, presentation of results and discussions

5.1 Introduction	92
5.2 Software tools used in the study	92
5.2.1 Waikato environment for knowledge analysis (Weka)	92
5.2.2 BayesiaLab	93
5.3 Presentation of experimental results	93
5.4 Evaluation metrics for knowledge-based models	94
5.4.1 Cross-validation	97
5.5 Results and discussion of knowledge-based models	98
5.5.1 Relationship between symptoms and depression	98
5.5.2 Knowledge-based models	103
5.6 Human Interfaces to Bayesian networks-Based expert systems	108
5.6.1 Design process	109
5.7 User needs, problem description and System requirements	110
5.7.1 Functional Requirements	111
5.7.2 Non-functional Requirements	112
5.7.3 Developing alternative design: low fidelity prototype	113
5.8 Description of system architecture	116
5.9 Audience	117
5.9.1 Clinicians	118
5.9.2 Medical students	118
5.10 Alerts in decision support application	118

5.11	Application interface: High-fidelity prototype	119
5.11.1	Evaluation of the first prototype	129
5.11.2	Heuristic evaluation	129
5.11.3	Heuristic Evaluation Method	130
5.11.4	Evaluation procedure	132
5.11.5	Recruitment of participants	134
5.11.6	Data collection method	136
5.12	Ethical considerations	140
5.13	Results and analysis of first prototype evaluation	141
5.14	Application interface: second prototype	142
5.14.1	Modifications on the main interface	142
5.15	Evaluation of the second prototype	147
5.15.1	Practical implication of the second evaluation	149
5.15.2	Strengths and weaknesses of heuristics evaluation method	150

## Chapter Six

### Summary, achievements, conclusions and future work

6.1	Summary	151
6.2	Strengths and limitations of study	151
6.2.1	Strengths	151
6.2.2	Study limitations	153
6.3	Key achievements of the study and contributions	154
6.3.1	Summary of research journey	156
6.4	Contribution to the field of ICT4D	157
6.4.1	Novelty	158
6.5	Conclusions	159
6.6	Challenges faced in the study	159
6.7	Future work	160
	References	164
	Appendices	198

## List of Tables

Table 1.1	Mental health professionals & mental health facilities per 100,000 population in Nigeria	3
Table 1.2	Pre-study participants	6
Table 2.1	Conceptual model of depression and depressive symptoms	27
Table 4.1	Abstraction form	68
Table 4.2	Original dataset and attribute values	69
Table 4.3	Descriptive statistics of age bracket and depression grades from dataset	72
Table 4.4	Selected attribute set	73
Table 4.5	Probabilistic and non-probabilistic algorithms	81
Table 5.1	10-Fold cross validation	98
Table 5.2	Mutual information and contribution of symptoms to depression	100
Table 5.3	Synergistic combination of depression symptoms towards depression	102
Table 5.4	Results from attributes in dataset	103
Table 5.5	Template for confusion matrix	104
Table 5.6	Summary of model results in terms of accuracy	105
Table 5.7	Summary of model results in terms of Precision	105
Table 5.8	Summary of model results in terms of F-Measure	106
Table 5.9	Summary of model results in terms of RSME	106
Table 5.10	Summary of model results in terms of Pearson correlation coefficient (R)	107
Table 5.11	Summary of model results in terms of Receiver operating characteristic	107
Table 5.12	Summary of model results in terms of Running time	107
Table 5.13	Jacob Nielsen's usability heuristics	131
Table 5.14	Usability scale based on usability evaluation	133
Table 5.15	Panel of evaluators	135
Table 5.16	Usability problem description and location in the first prototype	137
Table 5.17	Number and severity of heuristic violations	139
Table 5.18	Summary of feedback from evaluators on second prototype	148

## List of Figures

Figure 2.1	Conceptual model and propositions	20
Figure 2.2	Simple Bayesian network with three random variables	36
Figure 2.3	Relationships in a directed acyclic graph	37
Figure 2.4	Schematic representation of a conditional probability table	38
Figure 2.5	Usability evaluation methods	43
Figure 4.1	Traditional diagnostic process	65
Figure 4.2	Framework of the study	65
Figure 4.3	Research method overview	66
Figure 4.4	Flow of information for data collected from patient to researcher	71
Figure 4.5	Dataset descriptive statistics	72
Figure 4.6	Architecture of probabilistic-HCI diagnostic model	81
Figure 4.7	MLP with 2 hidden layers	83
Figure 4.8	MLP with 2 hidden layers for the study	84
Figure 4.9	Tree-based Boosting	86
Figure 5.1	Bayesian networks model showing depression prediction performance	104
Figure 5.2	Modified user-centered design approach for the study	110
Figure 5.3	Description of users taking medical history	113
Figure 5.4	Hand-drawn sketches of design alternative 1	114
Figure 5.5	Hand-drawn sketches of design alternative 2	115
Figure 5.6	Hand-drawn sketches of design alternative 3	116
Figure 5.7	Application interface login page	120
Figure 5.8	Four menus of the main application interface	121
Figure 5.9	Contents of the (A)file, (B)modules, and (C) utilities	122
Figure 5.10	Consent page	123

Figure 5.11 Staff and student details page	124
Figure 5.12 Screening result	128
Figure 5.13 Graphical representation of panel of evaluators	136
Figure 5.14 Count of usability problems and their severity	141
Figure 5.15 Graphical illustration of heuristic violation	142
Figure 5.16 Four menus and shortcuts of the main application interface	143
Figure 5.17 Modified diagnostic and file menus	143
Figure 5.18 Shortcut of modules on the taskbar	143
Figure 5.19 Request for assessment	144
Figure 5.20 Report of students already assessed	145
Figure 5.21A Screening questions	146
Figure 5.21B Screening questions	146
Figure 5.21 C Screening questions	147
Figure 5.22 Screening result	147

# Chapter One

## Introduction

### 1.1 Background information on depressive disorders

It is well known that clinical depression is a commonly encountered emotional distress in various types of clinical populations and the general public. Depression affects society as a whole, regardless of sex, age, colour, religion, culture, race or status. It is estimated that about 350 million people, of all ages, suffer from depression worldwide, out of which about 800,000 people die every year of suicide, a condition caused by undetected and treated depression (WHO, 2012). Depression is characterised by sad mood, loss of pleasure or interest, withdrawal from social activities, feelings of guilt, worthlessness, insomnia, hypersomnia, disturbed appetite, loss of energy, inability to concentrate, and in severe cases suicidal ideation (American Psychiatric Association, 2013), and its diagnosis is dependent on the presence of a number of these symptoms consistently for at least two weeks, causing impairment in the daily activities and/or noticeable problems in relationships with others (Abas, Ali, Nakimuli-Mpundu, & Chabanda, 2014). According to the Diagnostic and statistical manual of mental disorders, American Psychiatric Association (2013), at least five of the nine criteria must be met to be diagnosed with depression.

The world health organisation, WHO (2012) notes that though depression is a global public health problem, it has much higher risks in developing countries due to peculiar shortage of mental health professionals and the near absence of modern technology use in its management. In Nigeria, for instance, studies suggest that depression is also highly neglected (Abas et al., 2014) and accounts for over 30% of all outpatient attendance; over 20% of all admissions in most hospitals in the country (Gureje, Uwakwe, Oladeji, Makanjuola, & Esan, 2010); and that over 35% of all deaths in specialised psychiatric hospitals in the country is caused by depression. WHO (2012) reveals that 9 out of every 10 persons with mental disorders do not receive any services, and that the high negligence of depression in Nigeria is increasingly threatening public health;

Despite being a socially advantaged population, students of tertiary institutions have been shown to be at higher risks of depression than other population (Eisenberg, Gollust, Golberstein, & Hefner, 2007; A. Ibrahim, Kelly, Adams, & Glazebrook, 2013; Othieno, Okoth, Peltzer, Pengpid, & Malla, 2014). In Nigeria, Karl Peltzer et al (2013), Adewuya et al (2006)

and Afolabi et al (2008) have identified higher rate of depression among students of tertiary institutions than other population in the country, mainly due to problems with inadequate family support, accommodation problems, poor academic performance, heavy cigarette smoking and high level of alcohol. Whereas there are publicly available options for its detection and treatments, this disorder is often under-diagnosed and under-treated (Abas et al., 2014). The economic and social burden of depression in Nigeria and many other developing countries is enormous (WHO, 2012), but the work of Gureje et al (2015) notes that this burden can be lowered with early and regular screening for depression using modern computing tools.

Preventive medicine and early diagnosis have become popular solutions and major practice for doctor, globally; they are helpful in understanding and managing medical conditions (CDC, 2009). The works of Daimi and Banitaan (2014) and Obadeji et al (2015) suggest that regular screening, as the first step to early detection of depression could generate a significant positive effect on public health. As has been suggested by many researchers (Boylan, 2007; Wilson-Shaw, Pistrang, & Herlihy, 2012), an intervention that will complement the available resources is needed. Given the acute shortage of mental health professionals and facilities in Nigeria (Table 1.1), and that it would take some time before resources especially human can be mustered, it is only reasonable that other viable means of assessing mental health care, such as building automated screening tools and training a number of potential users, be sought. This ICT4D intervention has several potential benefits, the most important of which is a reduction in consultation costs, and early and regular assessment of depression, which is a pre-condition for effective treatments. As an intervention study, the techniques developed in the study were categorised into two: the first is predictive technique, for modelling based on machine learning; and the second is well-designed, easy-to-use and intuitive screening interfaces integrated into the predictive models for use by non-clinicians in a non-clinical environment.

Table 1.1 Mental health professionals & mental health facilities per 100,000 population in Nigeria (Jack-Ide, Uys, & Middleton, 2012; World Health Organization, 2011b).

	Resources	Per 100,000		Resources	Per 100,000
1	Psychiatrists	0.15	7	Other mental health workers	8.03
2	Psychologists	0.07	8	Beds in mental hospitals	2.53
3	Medical doctors	0.49	9	Psychiatric beds in general hospitals	0.20
4	Nurses	2.41	10	Mental health outpatient facilities	0.03
5	Social workers	0.12	11	%health budget spent on mental health	0.4
6	Occupational therapists	0.05	12	%GDP spent on health	5.0

The goal of the first category (predictive technique) is to seek novel non-trivial knowledge in the form of patterns to help explain the contributions and/or relationships between the features of depression and the disease itself. To do this, an overall methodology that integrates understanding and use of exploratory data analysis tools for predictive modelling is needed. The strategy is based on testing different machine learning classifiers and then select the best performer based on chosen metrics. The goal of the second category (screening interfaces) is to increase access to the first-line depression assessment tests using techniques in human computer interaction (HCI). Building a framework that enables the collaboration between machine learning technologies and humans by harnessing their relative strengths to accomplish what neither can do alone will go a long way to assisting clinicians in the process of examination and treatment of depression

## 1.2 ICT4D tools and healthcare in Nigeria

Information and communications technology for development (ICT4D) is a relatively new multidisciplinary research field that focuses on using information and communications technologies (ICTs) to achieve developmental objectives and uplift the living standards of both the rural and the urban poor (Toyama & Dias, 2008). Among these development objectives are health, income growth and education. According to Toyama & Dias (2008), ICT4D research and practice consists of two major components: ICT and development. While ICT describes digital technologies with certain properties, features and functionalities, that help to reduce poverty and enhance sustainable development, development describes socio-economic and

human development (Thapa & Hatakka, 2017). In the context of the present study, ICT4D is used to describe a wide range of computer technology innovative applications, with the central goal of facilitating effective and efficient operations in screening for depression in the resource-constrained areas in Nigeria, through the use of ICTs. In this definition, the study argues that any search for development objectives through the use of ICT should be channelled in line with existing resources and locally available technologies. ICT4D technologies include computers (hardware, software and procedures), the internet, telephones, radio and television (Toyama & Dias, 2008). The present study focuses on combining computer technologies, and humans together to solve real-world problem in a timely manner.

Increasingly, ICT4D-aided interventions have continued to give invaluable supports to enhance the efficiency and effectiveness of the healthcare industry. Health ICT4D intervention areas include health monitoring devices, such as pagers, remote operation equipment and electronic patient records, heart rate sensors and alarms, and networked monitoring stations (Wiseman, Cox, & Brumby, 2013). Other health information and communication devices include pagers, networked monitoring stations, remote operation equipment and electronic patient records. These resources are needed to support health professionals and other stakeholders for efficient and effective management of the healthcare sectors (Haux, Howe, Marschollek, Plischke, & Wolf, 2008).

In the domain of mental health, the use of ICT4D-aided screening can be traced back to the works of Spitzer and Endicott (SPitzer & Endicott, 1969; R L Spitzer & Endicott, 1968) in which they asked a question ‘can the computer assist clinicians in psychiatric diagnosis’? An answer to this question was given by Schmid et al (Schmid, Bronisch, & Von Zerssen, 1982) in *DIAGNO*, a diagnostic program which allowed users to select either ‘yes/no’ to questions from the administered psychiatric questionnaire. *DIAGNO* later developed into different versions, one of which is called *DiaSika* (Schmid et al., 1982), used in France and Germany. Today, one of the most popular applications developed for use by both clinicians and non-clinicians is *MoodGYM*, a computerised cognitive behavioural therapy programme for depression, anxiety and general psychological distress (Twomey & O’Reilly, 2016), developed by the centre for mental health research at the Australian national University.

While studies in the past have shown beneficial effects that ICT4D-aided interventions can have on health delivery outcome in the developed world, it has attracted little or no significant attention in most developing countries, including Nigeria (Harris, 2016). For instance, compelling evidence show severely limited use of ICT4D-aided tools in mental healthcare in

Nigeria, even till this day (Adeleke, 2015; Idowu, Cornford, & Bastin, 2008; Jimoh, Pate, Lin, & Schulman, 2012). Depression screening, which is the focus of this study, is still based on manual inspection of individual sufferers in hospital domains by clinicians and rarely aided by computerized systems.

With the aim of supporting the existing infrastructure to alleviate the problem of mental healthcare delivery in Nigeria, it is important that all possible areas of solutions be investigated so as to select the one that works for a particular society. The needed solution should be channelled towards rapid expansion of ICT4D-aided healthcare delivery including health information technologies, especially screening, diagnostic and treatment plans. Owing to the limited impact the available services and past interventions have had on mental health management, the study has, as its primary goal, to explore the interplay of two complementing ICT4D developmental technologies, machine learning and human computer interaction (HCI) with a view to increasing flexibility in screening for depression (Kleine & Unwin, 2009). This would also be in response to calls by the WHO (2012) and other stakeholders (Abas et al., 2014) for more comprehensive mental health diagnostic approaches that address the difficulties in developing countries. Machine learning and HCI techniques are both ICT technologies that have the common goal of enhancing the effectiveness of a system and making it easier for humans to use to solve real-life problems (Holzinger, 2013). While machine learning systems are good at computation and analysis of data at the lowest level, HCI, a multidisciplinary research field, investigates issues related to the design and implementation of the interface that enables interactions between humans and computers so that a person's needs are satisfied in the most effective way.

### **1.3 Proof-of-concept**

In the build-up to this study, the researcher had conducted a survey with two study population: the first was with healthcare professionals (psychiatrists, child/adolescent psychiatrists, psychologists, nurses, social workers, psychotherapists and doctors) who worked at the University of Benin Teaching hospital (UBTH) at the time of the study; the second was with staff (academic and non-academic) and students (undergraduates, masters and PhD) of the University of Benin, Nigeria. The surveys were conducted with the following objectives in mind:

1. assess the feasibility of the study

2. assess the availability of relevant information from the two data sources and seek the permission of the clinicians to collect data in their facilities
3. estimate the level of risks and prevalence of depression among University students
4. interview clinicians to identify the difficulties they face with depression screening and isolate those that can be solved with ICT technologies
5. identify a cost-effective workable local solution that utilises existing infrastructure and technologies to address some of the identified difficulties with respect to the objectives of the study.
6. consider the elements of the study design: the research questions, the timeframe required to collect the data, the study population, inclusion and exclusion criteria, and the accuracy of the collected data.

### 1.3.1 Data collection

Data was collected through self-administered questionnaire (appendices C and D), in paper format, collecting information about background, resources, adoption of ICT tools in medical and mental health services.

Table 2.2 Pre-study participants

	<b>Participant</b>	<b>Total No</b>
1	Psychiatrists	3
2	Child/Adolescent Psychiatrist	2
3	Psychologists	2
4	Nurses	6
5	Social workers	4
6	Psychotherapists	4
7	Doctors	8
	<b>Total</b>	<b>29</b>

	<b>Participant</b>	<b>Total No</b>
1	Undergrad Students	22
2	Masters students	17
3	PhD students	12
4	Academic staff	12
5	Non-academic	10
	<b>Total</b>	<b>73</b>

As shown in Table 1.2, a total of 102 (29 medical and mental health professionals and 73 University staff and students) participated in the surveys. The sample size of each group was determined by participants' availability for the survey. The researcher conducted a three-hour training for each group on study objectives, data collection procedures, contents of the questionnaire, data confidentiality, participant rights and data quality issues with the participants prior to the actual data collection. Data collection and data quality were monitored consistently by the researcher.

**Clinicians:** The researcher administered a questionnaire (Appendix C) to each clinician to be completed. This was divided into a cover letter (Appendix A) and three sections. The cover page provided brief information about the researcher, purpose of the study and instructions on how to fill up the questionnaire. It also provided information about maintaining anonymity and confidentiality, and that participation was voluntary. Lastly, it clearly stated that data would be used only for research purposes. The first section of the questionnaire was intended to collect the socio-demographic details of participants including age, sex, professional category, and years of experience on the job. The second section inquires about the knowledge and use of ICT tools in medical practice while the third section talks about the common type of mental disorders, their age bracket and their attitudes towards the adoption of ICT tools as a workable solution to the peculiar difficulties.

**University staff and students:** The researcher administered a questionnaire to each participant to be completed (Appendix D), which was, again, divided into three sections following a cover letter. The first section was about socio-demographic details including age, sex, staff or student category, the number of years spent in the University. The second section inquires about the level of support from parents, family size, housing problems, academic performance, health risk behaviour, such as levels of alcohol consumption and cigarette smoking, and the outcomes. The third section talks about the use of ICT tools and their attitudes towards using same to solve their health problems. Accommodation problem was assessed by asking the students whether they are officially accommodated in their rooms or sharing with a fellow student (parents home, with a relative, rented room). Academic problems were assessed by asking the students to rate their performance (poor, good, excellent). Smoking and alcohol consumption levels were assessed by asking the students to fill the quantity consumed (none, low, medium, high).

### **1.3.2 Data analysis and results**

At the completion of data collection, the researcher tabulated the data manually in Microsoft Excel sheet, converted it to a comma separated value (CSV) and was analysed using a software tool, BaysiaLab 6 (Conrady & Jouffe, 2015). The relation between the study population and the relevant variables were described using descriptive statistics.

Ethical approvals for the surveys were obtained from the Faculty of Science Research Ethics Committee of the University of Cape Town (Appendix E) and the UBTH Ethics and Research

Committee (Appendix F). Informed verbal consent was obtained from the head administrator of each unit. Written consent (Appendix B) was also obtained from each study participant following a clear explanation of the purpose, data collection procedures and data confidentiality issues of the study.

### **1.3.2.1 Characteristics of the participants**

Though not part of the objectives of the present study, the researcher felt that identifying and presenting the characteristics of the study participants are necessary for the following reasons:

- I. Help the readers to understand the background of the participants
- II. Provide a picture of the suitability of the participants for the study
- III. Provide information on how well the participants represented the characteristics of the systems to be developed and hence provided appropriate contextual information for the presentation of the findings of the study.
- IV. Shed some light on the availability, use or non-use of ICT4D tools by the participants

This is presented under three categories:

1. Socio-demographics of participants
2. Knowledge and use of ICT tools in work place
3. Attitude towards the adoption if ICT tools

### **1.3.2.2 Social demographics of participants**

. Out of a total of 102 (29 clinicians and 73 University staff and students) questionnaires administered in the surveys, 89 (87.3%) were completed, returned and analysed as part of the motivation for the study. This consisted of 26 (89.7%) from clinicians and 63 (86.3%) from University staff and students.

Among the clinicians, 16 (61.5%) were male while the remaining 10 (38.5%) were female. Almost half of the clinicians, 12 (46.2%), were Nurses, 6 (23.1%) were Doctors, 2 (7.7%) were Social workers, and 3 (11.5%) were each of Psychiatrists and Psychologists. 7 (26.9%) of the clinicians had less than 5 years work experience, 12 (46.2%) had between 6 and 10 work experience, 5 (19.2) had between 11 and 15 work experience while only 2 (7.7%) had more than 16 years work experience.

Among the staff and students, 43 (68.3%) was students while the remaining 20 (31.7%) was staff. The students consist of 26 (41.3%) undergraduates, 11 (17.5%) masters students and 6 (9.5%) PhD. The staff are 10 (15.9%) each of teaching and non-teaching. 35 (55.6%) of the

respondents were between 18 and 28 years old, 16 (25.4%) between 29 and 39 years old, and 12 (19.0%) were above 40 years. It was also found that more than half of the respondent staff, 12 (60%), were also post graduate students in the institution but would prefer to be categorized as staff in the survey.

### **1.3.2.3 Knowledge and use of ICT tools in the work place**

It was interesting to note that all 26 (100%) respondent clinicians had ICT awareness, were computer literate (can do basic tasks with office applications and use internet services), and believed that ICT had important roles in medical practice. However, 20 (76.9%) of the clinicians believe that ICT tools were mainly useful in the area of MIS while 6 (23.1) believed that ICT tools can be used to aid decision making in mental health. The respondents commented that they were less likely to adopt ICT in the discharge of their services for the following reasons: 1) physical signs have lesser role to play in mental health diagnostic decision; 2) electronic health records were relatively new and poorly integrated in clinical practice in the domain of mental health and that 3) electronic interfaces of the few screening tools they had seen were either poorly designed or do not fit into clinicians' daily routine; 4) ICT tools serve more as MIS that create a huge overhead on their clinical work, and that beyond the conventionally offered electronic records, offer little to enhance patient-clinician communication, clinician-clinician communication and diagnostic decisions.

Expectedly, all 63 (100%) staff and students were computer literate, had unhindered access to desktop computers and laptops provided for them by the University management. 23 (36.5%) of the staff and students were satisfied with their jobs while 40 (63.5%) were dissatisfied. 19 (30.2%) stated poor salary as reason for dissatisfaction, 10 (15.9%) poor learning environment, 14 (22.2%) facilitated-related problems and 20 (31.7%) management problem. 10 (15.9%) of the staff and students confirmed that they completed their given tasks with desktop computers between 2 to 4 hours while 53 (84.1%) completed a given task in more than 4 hours. Interestingly, 46 (73.0%) have heard of the practical application of machine learning and HCI techniques for solving real-world problems, including medical diagnosis. 31 (49.2%) have been involved in machine learning and HCI projects in the past while 25 (39.7%) felt it was successful.

### 1.3.2.4 Attitudes towards the adoption of ICT tools

Majority of the clinicians, 18 (69.2%) were satisfied with their current job. Poor training opportunity 2 (7.7%), facility-related problem 5 (19.2%) and management problem 1 (3.8%) were the reasons given for job dissatisfaction. All 26 respondents confirmed the availability of desktop computers and laptops for every clinician provided by hospital management. Interestingly, all 26 (100%) confirmed the important roles of ICT tools in their medical practice, but while 21 (80.8%) said ICT tools would improve evidence-based medical practice, 5 (19.2%) felt otherwise. 20 (76.9%) admitted that adoption rate of ICT tools in medical practice was still very low in Nigeria. 15 (57.7%) of the clinicians confirmed their acceptance of a workable local solution that won't violate the ethics in medical practice, while 17 (65.4%) showed interest in the use of machine learning and HCI could provide.

For staff and students, 10 (15.9%) lived in their parents' home, 5 (7.9%) lived with a relative while 48 (76.2%) lived in rented home. In terms of emotional bond with family members 29 (46.0%) had no support from family, 24 (38.1%) had low support, 7 (11.1%) had medium support while 3 (4.8%) had high support. Family background was measured using the rating of family background by income distribution and inequality in Nigeria (Bakare, 2012). In the study, a family is rated as follows: wealthy if it's within the highest 25%; well-off if it's within the 50% to 75% or poor if it's within the lowest 25% in terms of wealth). In this regard, more than half of the staff and students, 42 (66.7%) came from poor family background, 14 (22.2%) was well-off, and 7 (11.1%) was wealthy. On health risk behaviour, 28 (44.4%) had high level of cigarette smoking, 25 (39.7%) medium, 6 (9.5%) low and only 4 (6.3%) did not smoke at all. 31 (49.2%) had alcoholic consumption, 21 (33.3%) medium, 3 (4.8%) low and 9 (14.3%) none. For academic performance, 9 (14.3%) had excellent performance, 11 (17.5%) had good performance, 21 (33.3%) had poor performance. Academic performance was not applicable to 22 (35.0%) of the respondents. 35 (55.6%) of the respondents confirmed that they were often depressed by various challenges. 28 (44.4%) said they felt depressed yearly, 31 (49.2%) monthly while 4 (6.3%) was weekly. 45 (71.4%) often felt the need to see a clinician to examine their health status while all 63 (100%) was dissatisfied with their visit to the hospitals because of the high costs, long waiting times to see a clinician and long waiting times to retrieve their medical records. 19 (73.1%) of the clinicians stated that depression was the most common type of mental disorder while another 15 (57.7%) confirmed that the most commonly depressed population being treated in their facility are those between 21 and 40 years. 54 (85.7%) said they would prefer a workable alternative solution to complement the efforts of the limited

medical resources. 49 (77.8%) felt that a combination of human experts and ICT could reduce costs and long waiting times at the hospitals presently experienced.

Given the above analysis and the inability of the currently available medical resources and practice to yield the desired results, in terms of early management of depression among the teaming population, the need for complementary and affordable ICT4D-aided frameworks that support the existing system for screening of depression in Nigeria universities is paramount.

#### **1.4 Justification of choice of study specific areas**

Firstly, the driving forces which make Nigeria a critical context for this study are as follows: 1) insufficient studies and data on the risk factors and prevalence of depression in Nigeria (Adewuya et al., 2006; Peltzer et al., 2013); 2) everyday life presenting additional situations such as high crime rate, unemployment, terrorism, extreme poverty and serial outbreak of diseases, that trigger depression for many (Abdulmalik et al., 2013; Ahmed Kayode, 2014; Gureje et al., 2015; Razzouk et al., 2010); 3) more studies have shown the peculiar shortage of mental health professionals and equipment to attend to the teaming population of depression sufferers (Table 1.1), poor mental health data management, and weak evidence-based decision-making practices in Nigeria (Gureje et al., 2015; Salihu, 2015); 4) many healthcare professionals across the world are adopting and introducing ICT tools to reduce costs, serve patients and extend into new areas (Lluch, 2011).

Secondly, the initial motivating factor for which Nigerian University students stand as an interesting case for this study was the results of the pre-study surveys, which conform with those found in the literature (Adewuya et al., 2006; Eisenberg et al., 2007; A. K. Ibrahim, Kelly, Adams, & Glazebrook, 2013; A. Ibrahim et al., 2013; Osemene & Lamikanra, 2012; Peltzer et al., 2013). In summary, the results suggest that Nigerian University students are at higher risk of depression than other populations for the following reasons: 1) most of the students undergo role transition to universities to live with other students with less adult supervision and poor living and academic conditions; 2) constant problems of accommodation, inadequate and overcrowded lecture halls caused by increasing population of students, unavailable study materials, recurrent disruptions of academic calendar and lower socioeconomic status, 3) social variables such as pressure from parents and society to pass examinations, modern lifestyle, large family size; 4) additive behaviour such as heavy cigarette smoking and high level of alcohol consumption. The works of Fisher et al (Fisher, 2011) and Mehdi and Maryam (Mehdi & Maryam, 2014) have also shown that many resource-constrained countries lack data or

resources for data collection on depression to provide local evidence about the type and extent of the problems. This data is needed to guide prevention policy, develop protocols for improved practice and adapt local resources for use. These factors, in addition to the regular symptoms of depression (American Psychiatric Association, 2013) have presented enormous medical challenges to University communities and available medical services in Nigeria (Semrau et al., 2015).

Lastly, since the study is situated in the domain of ICT4D, which advocates for using ICTs to alleviate socio-economic problems facing disadvantaged population, any pursuit for development objectives through ICT should be tailored towards already existing resources and locally available technologies to help in problem-solving. From literature and the pre-study survey, a gap in the adoption of ICT4D framework for supporting the methodologies for depression screening in Nigerian universities was identified. Additionally, the choice of the use of computer-based technologies is facilitated by the availability of desktop computers and laptops in all departments of the University of Benin, and the ease with which the computer-based depression screening fit into the activities of the University academic staff and students.

## **1.5 Research goal and objectives**

A search of the literature at the time of writing this thesis did not reveal any comprehensive integration of ICT4D frameworks into the daily routine practice of healthcare professionals with reference to the peculiar difficulties to managing depression screening in Nigeria. As its main goal, the study investigates the extent to which ICT4D frameworks can support the methodologies for screening for depression among the teaming population in Nigerian University community. This was done in two phases: the first was to build machine learning models for depression, and the second was to build HCI systems that compute and communicate machine learning results in ways that are compatible with the decision-making process of healthcare professionals. To achieve this goal, the following specific objectives were necessary:

1. To identify depression screening challenges faced by clinicians and then isolate those that that can be solved using ICT4D applications
2. To identify the extent to which ICT4D applications are used for routine disease screening in selected healthcare institution in Nigeria
3. Investigate the strengths of various classification techniques for the purpose of developing new models.

4. Address an important issue related to the most important predictors and relationships between the predictors
5. Integrate the knowledge-based framework into a graphical user application, which can assist University academic staff in the regular assessment of staff and students for depression
6. Investigate the effectiveness of supporting University academic staff in assessing their staff and students for depression using ICT4D application framework.

## **1.6 Research contributions**

The study contributions are summarized in twofold: with respect to predictive modelling, it is the development, evaluation and implementation of probabilistic machine learning algorithms to clinical datasets for predictive modelling. With respect to real-life applications, it is the integration of computationally-managed diagnostic models into an easy-to-use screening tool for use in non-clinical settings and the interpretation of them. They are as follows:

1. Empirical evidence of the workability of ICT4D framework in building a depression screening tool.
2. Empirical evidence of which methods and techniques for revealing complex interdependencies and relations between (1) symptoms of depression (attributes), and (2) attributes and targets.
3. Extensive empirical evaluation of the above-stated methods using real-life clinical datasets from University population and second putting the method to work in a real-world audit to explore its practical value in non-clinical settings.
4. Illustration that existing heuristics in HCI can be used to evaluate smart systems in real and practical environment.
5. Empirical evidence of the effects of using subject matter expertise and experienced usability evaluators in evaluating a medical application.
6. Empirical evidence of the knowledge and attitude of clinicians towards the use of ICT tools in work place

## **1.7 Research questions**

To address the above stated research objectives, the following research questions were formulated and answered by the study.

1. Are there ICT4D tools that can help address some of the disease screening challenges faced by mental health professionals in Nigeria?
2. What is the combined framework that integrates ICT4D tools and novice users together?
3. How can the linkage between machine learning and HCI tools be strengthened to support the framework for screening for depression in Nigeria in the context of ICT4D project?
4. What machine learning techniques would be appropriate to screen for depression among University staff and students?
5. Is there an association between the symptoms of depression and depression that are of statistical relevance?

## 1.8 Thesis organisation

To achieve the objectives of this research study, contents of the remaining chapters of the study are organised in the following structure:

Chapter two describes the elements of the present study, ICT4D, machine learning classification process, HCI, and depression screening. It also discusses the link between machine learning and HCI to create depression screening tools for heterogenous non-technical user population.

Chapter three reviews previous studies related to the lack of impact of past ICT4D interventions in the public health systems in Nigeria before looking at the potential of ICT4D tools in screening for diseases. It also explores relevant literature in machine learning and human computer interaction approaches to mental health decision support systems in resource-limited areas.

Chapter four describes the research methods used to achieve the goal and objectives of the study; the pre-existing methodological basis of depression classification and the need for the work described in this study. Specifically, it looks at the drawbacks of the existing detection techniques and procedure, and the motivation for a complementary technique and procedure for managing depression, highlighting the places of difficulties. Also described are the research's philosophy, design, data collection procedure and details of the different algorithms used at different phases of the development process.

In chapter five, details of the tools and programs used to facilitate the implementation process are discussed. With a view to addressing a part of the study research questions, the analysis of the results from various machine learning-based decision support models are made

in this chapter. A comparative analysis of their performances is given. The chapter concludes by presenting an empirical evaluation of the knowledge base models and the developed user-friendly graphical interfaces.

Finally, in chapter six, the summaries, contribution and achievements of the work done in the study are presented. The limitations of the study are discussed, and conclusions are drawn. The chapter closes by reviewing and discussing a variety of plans for possible future directions for problems to be solved using improved versions of the methods described in the study.

## **Chapter Two**

### **Understanding the ICT4D frameworks and depression screening relation**

#### **2.1 Introduction**

The field of ICT4D focuses on how ICT technologies, such as computers (hardware, software and procedures), the internet, telephones, radio and television, can be implemented for social and economic developmental purpose in resource-constrained areas of both the developed and developing countries (Barjis, Kolfshoten, & Maritz, 2013; Toyama & Dias, 2008). These technologies are useful in delivering information to where it is required, in a timely manner. ICT4D was intended to jump-start the process of improving quality of life, social quality, empowerment, and health development (Ganju, Pavlou, & Banker, 2016; Thomas, Li, & Oliveira, 2017). As an interdisciplinary research field, the conceptual approaches to the field of ICT4D are from diverse academic and industrial domains, such, ICT, development, business, economics, sociology, communications and psychology (Avgerou, 2017; Kleine, 2010). Whichever is adopted is dependent on the purpose of the study. With regards to the purpose of this study, the description of ICT4D has two major components, ICT and development (Avgerou, 2017; Toyama & Dias, 2008), is adopted. Focus will be on how ICT applications can be implemented for developmental purpose in areas of limited medical resources. Although both ICT and development have been independently well-defined, the knowledge gap in the link between ICT interventions and development, in the context of developing countries, is still an ongoing area of inquiry (Thomas et al., 2017). Bridging this research gap requires an understanding of both ICT and development drawn from their frameworks. As Ann et al (2015) notes, frameworks are useful tools for analysing the interactions and interrelationships between the diverse elements involved in the adoption and use of health technologies. The study therefore argues that the application of these frameworks are necessary considerations for designing intervention programs to enhance ICT acceptance in the healthcare industry in Nigeria. The next subsections examine several of these models through their frameworks and a few were chosen for the study based on their popularity in the healthcare industry. It must be stated, however, that these descriptions are not exhaustive but contextual and scoped for the presented study.

## **2.2 Frameworks used in ICT research**

ICT describes information management technologies which facilitate, by electronic means, the creation, storage, and dissemination of information. These technologies, according to Kleine and Unwin (2009), include the ‘old’ ICTs of radio, television and telephone, and the ‘new’ ICTs of computers and network (hardware and software), wireless technologies, the Internet, and the services associated with these technologies. The study focuses on computers and their applications for intervention in medical resource-constrained areas in Nigeria. As a multidisciplinary research field, ICT is not studied in isolation; it is studied within the wider frameworks upon which it is situated (Sein, Hatakka, Thapa, & Sæbø, 2016). The approaches to ICT research are therefore drawn from multiple fields, derived from both community and organizational settings. These are described in the following subsections:

### **2.2.1 Psychological theories**

Psychological theories, such as theories of reasoned action (TRA) and planned behaviour, the technology acceptance model (TAM), and adoption and diffusion of innovation theory (DOI), are useful for understanding the motivations for ICT adoption and use in workplace (Kim & Crowston, 2011). These theories, developed in the field of social psychology to aid the understanding of a variety of human behaviours, are sometimes called the theory of intention-behaviour (Gagnon, 2015) since they focus on people’s intention to engage in a certain behaviour (ICT adoption and use) and can assist in explaining the motivations and impediments to adoption and variations in usage patterns between contexts and individuals (Otieno, Liyala, Odongo, & Abeka, 2016).

#### **2.2.1.1 Theories of reasoned action**

The theories of reasoned action (TRA), a well-known psychological theory formulated by Ajzen and Fishbein (Ajzen & Fishbein, 1980) postulates that the realization of a given behaviour is predicted by the individual intention to perform this behaviour (Gagnon, 2015), which is in turn formed by two antecedents: attitude towards the behaviour and subjective norm. TRA has been used to explain the nature of intentions and the limits of the application of ICT adoption and use research as a fundamental theoretical framework. TRA has also been in a number of other research fields in combination with other theories and models as a foundation to such studies (Otieno et al., 2016). TRA models have been successfully used in

the health-related fields and medical innovation (Khasawneh & Ibrahim, 2008) to explain different behaviours of healthcare professionals towards ICT adoption and use. For instance, Beadnell et al (2008) used the TRA to predict intentions to use condoms with casual partners, as well as to steady-partner safer sex behaviours: mutual monogamy and condom use. Valtonen et (2015) in an experiment using the framework of the theory of planned behaviour, described how experiences of learning with ICT in pedagogically meaningful ways can affect pre-service teachers' intentions to use ICT for teaching and learning. The results showed no differences in pre-service teachers' attitudes and behavioural intentions towards the use of ICT for teaching and learning.

### **2.2.1.2 Technology acceptance model**

Technology acceptance model (TAM) originated from the TRA and was first developed by Davis (1989) to provide detailed explanation of the factors for understanding the decision process of ICT. The factors in TAM, which constitute the most important predictors of people's attitudes toward using ICT are, perceived usefulness and perceived ease-of-use (Gücin & Berk, 2015). Perceived usefulness describes the extent to which an individual believes using the ICT will enhance his performance while perceived ease-of-use describes individual believes the given ICT will reduce the intensity of their work. In the healthcare domain, Gücin and Berk (2015) note that while perceived usefulness of technological innovations may be the most distinctive factor for healthcare professionals, ease of use is of big importance for patients. Perceived ease of use is affected by personal norms and perceived control beliefs. The effects of technology acceptance and use by healthcare professionals and patients have been studied in the social and behavioural literature. For instance, Kane (2014) notes that the mobile technologies guarantee many benefits including disease screening and treatment access, self-assessment and disease management, for the patients. Similarly, the work of Khan and Woosley (2011) has shown that several countries have adopted ICT in their healthcare because of the benefits of ICT usage by healthcare professionals, including, reduction in screening and treatment errors, reduction in treatment period and cost, quick transfer of medical records. Fontenot (2014) further notes that technology usage in healthcare provides costs savings for patient, healthcare professionals and government, and facilitates the transformation of a patient to a healthier, cheaper citizen.

### **2.2.1.3 Theory of Adoption and diffusion of innovation**

The conceptual framework in much of this study is Rogers's adoption and diffusion of innovations (DOI) (Rogers, 2015). The theory is based on the fact that understanding the factors that shape innovation diffusion would enable decision makers to promote acceptance and increase technology use by generating effective policy initiatives at the system level, and establishing more decisional systems at the organization level (Ann et al., 2015). Innovation is an idea, new product, a program, or a technology that is new to the adopting unit. In the healthcare domain where the DOI has received much attention, adoption describes the discrete decision to accept or reject a health technology, while diffusion describes the process by which a health technology is communicated through certain channels among the members of a social system over time (Ann et al., 2015; Gagnon, 2015). As ICTs grow in popularity, understanding adoption and use of them is very critical in terms of design, development and deployment of new ICT. The major objective of the diffusion of innovation (DOI) theory is to understand the adoption of innovation in terms of four elements of diffusion including innovation, time, communication channels, and social systems (Rogers, 2015). DOI states that people's technology adoption behaviour is determined by their perceptions of technology characteristics, such as relative compatibility, relative advantage, trialability, observability, and complexity of the innovation (Rogers, 2015).

The key concepts of Rogers' DOI identified that the process of introducing a new technology and method into the workplace and moving through to its sustained use, was composed of five stages occurring over time. These are knowledge, persuasion, decision, implementation, and confirmation (Rogers, 2015). An individual learns about an innovation, then forms an attitude toward the innovation before reaching a decision on whether to adopt or reject the innovation. This is followed by putting the innovation into practice, and finally, either agree and adopted the innovation as best practice or reject the innovation, at which point the innovation is discontinued (Rogers, 2015). These innovations are presented in the form of technologies, such as new screening procedures, surgery procedures, treatments, or new medical drugs. Successful innovations become widespread among the clinicians community, following a diffusion process pattern (Duch, Rallo, & Guardiola, 2014). Technology, according to Duch et al (2014), can, like the innovation concept, be more than a physical tool or artefact. It can be the skills and knowledge needed to operate a tool, it can also be the knowledge used

to solve problems or the pedagogical tools to extend the knowledge of others pattern (Duch et al., 2014).

In the healthcare industry, several researchers have used Rogers' DOI as its theoretical framework, either separately or in combination with other models and theories to explain adoption and use of ICT. In Zhang et al (2015), the DOI was used to provide valuable insight about the feasibility of introducing consumer electronic health services in a primary healthcare setting. The study described the factors influencing the low rate of adoption of the electronic health invention to include the incompatibility of the new service with the patients' preference for oral communication, ineffective communication of the availability of the electronic appointment service to the patients, a perceived lack of value of the new online service for the majority of patients, and some functional limitations of the service itself. The study further notes that lack of access to a computer or the Internet at home, low level of computer literacy, and the low level of socio-economic status of the study population as factors causing the low rate of adoption of the new online service. On the other hand, the electronic appointment service was perceived to be an advantage for full-time working patients who could only make an appointment to see a physician after business hours. Duboise et al (2013) also described validity, reliability and maturity of the science; communication of the science; economic drivers; patient and physicians ability to apply the findings; and incorporation into guidelines., as five factors affecting technology diffusion. Confidence in the device, developed by good communication of trusted data on safety and efficacy, promotes diffusion, as does the ability to apply the device in the local clinical setting. Economic aspects may slow or hasten diffusion depending on the insurer willingness to pay and hospital funding policies. The study also concluded that incorporation into clinical practice guidelines can greatly influence the behaviour of medical professionals and accelerate diffusion.

### **2.2.3 Attributes of innovation**

The attributes of an innovation describe the characteristics of the innovation that affects the rate at which it is adopted. Although the theory of adoption and diffusion of innovation proposes many theories to explain how a new concept or technology diffuses in a given population, the present study shall consider only a few attributes of an innovation and the characteristics of the associated adopter categories as they relate to the adoption rate of a particular product. Rogers summarizes previous innovation research studies and identified five perceived characteristics that make an innovation more or less desirable to adopt. The rate of

adoption for an innovation is faster if members of a system perceive it to be better than what the innovation it supersedes (perceived relative advantage); is consistent with the existing values, previous experiences and needs of potential adopters (perceived compatibility); is not relatively difficult to understand or use (perceived complexity); can be experimented to be used without a significant commitment in time, effort or expense (perceived trialability); and possesses benefits that visible to the user and others (perceived observability (Rogers, 2015).

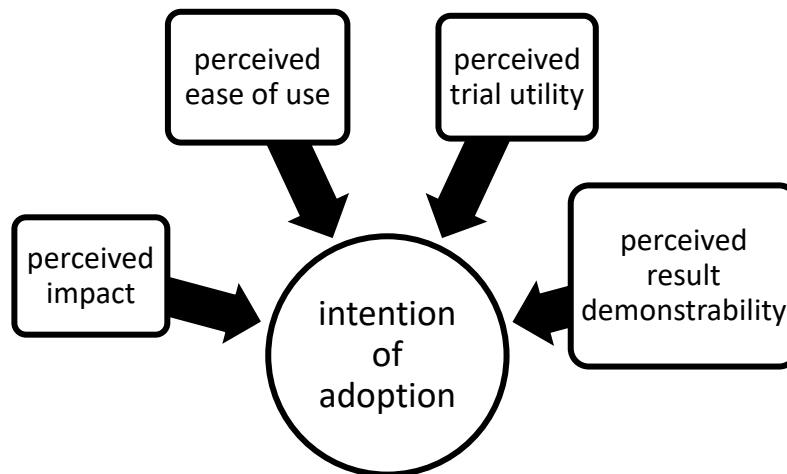


Figure 2.1: Conceptual model and propositions

For this study, the adapted conceptual framework and the reasoning for its concept composition is shown in Figure 2.1. The innovation is the development of screening tools that empower a heterogenous user population to help prevent adverse effect of health events such as depressive disorders that may lead to more serious and complicated health issues. The channels that will be discussed in the study are computer (software, hardware, and procedures) technologies such as machine learning and human computer interactions. .A description of the framework is briefly described:

### 2.2.3.1 Perceived impact

Relative advantage relies on the technology it replaces. This means that the depression screening tool to be developed becomes problematic if majority of the end-users regard the tool as more complex than the system it replaces. However, since the study already established that the screening tool is not replacing the existing traditional method, which is still in use, but complementing it, it may be less relevant to measure advantages relative to existing technology. Instead, the perceived impact of the innovation will explore what impact the use of the innovation has on all users work processes, with no overt reference to the existing system. This

means that it will be entirely up to the users to describe what they feel the concept of perceived impact entails.

### **2.2.3.2 Perceived ease of use**

The depression screening system is without doubt a complex innovation due to the underlying machine learning architecture and all its corresponding software. However, the end-users of the tool are likely never to be exposed to this complexity, and are only interested in the actual use of the application. Notwithstanding that exploring the relationship between these attributes is not part of the objective of the present study, it is important to note that ease of use attribute has a significant effect on attitudes toward usage, which is another variable in the TAM model. It is safe to assume that perceived ease of use also will have a strong effect in the intention of adopting the depression tool in this study.

### **2.2.3.3 Perceived trial utility**

The third attribute of the conceptual model, perceived trial utility, is derived from the original framework of Rogers and specifically the trialability attribute (Rogers, 2003). Since the study focuses on the perceived importance of a prototype depression tool, the perceived trial utility attribute will be operationalized by exploring how important this testing period is in order for the researcher to form positive intentions toward adoption of this innovation. Rogers noted that the trialability of an innovation is positively associated with its adoption. There is no reason to believe that the perceived utility of this prototype application will not have a similar effect on intentions of adopting the depression tool.

### **2.2.3.4 Perceived result observability**

Observability reflects how observable the use of the innovation is to others that are exposed to the users of the innovation. It will be interesting and purposeful to examine how the users' perception of the result demonstrability affect their intentions of adopting the depression screening tool. Perceived result demonstrability is therefore an important attribute in this study, as it may reveal how visible the results from usage is to the users during the testing. Because of this, perceived impact and perceived result demonstrability is likely to be somehow related in terms of their effect on intention of adopting the depression screening tool.

## 2.3 Frameworks for developmental studies

Development has many perspectives and several interesting theories have been developed to describe it, but the concern of the present study is development within the context of economic and social development (Kleine & Unwin, 2009) through efficient mechanisms enabled by ICTs, with emphasis in the medical resource-constrained area of Nigeria. A few of these are described in sections 2.3.1.

### 2.3.1 The capability approach

The Sen's capability approach (Sen, 2007) offers a way of thinking about development as an enlargement of capabilities. The capability theory is grounded in the fields of development and economics, and was pioneered and advocated by Amartya Sen in conjunction with Martha Nussbaum out of the need to measure progress in development, and the dissatisfaction with existing tools for monitoring and evaluating development (Sen, 1997, 2000, 2001, 2007). The core idea in the capability theory is that what people are effectively capable of doing (capabilities), should be the focus of wellbeing evaluations and government policy, as opposed to what they actually do (functionings) (Nussbaum, 2005, 2011; Sen, 2007). What this means in practice, is much more attention to matters directing policy efforts towards health, education and sustainability, economics, and promoting human well-being (Sen, 2007). The list of capabilities that are essential for a good life, as given by Nussbaum (Nussbaum, 2005) include normal life span; bodily integrity; bodily health; imagination and thought; senses, emotions; practical reason; affiliation; other species; play; and control over one's environment. Knowing the importance of ICT, Sen (2007) extended it to include capabilities, such as computer literacy. Examples of functionings are such diverse things as working, resting, being literate, being healthy, being part of a community, being able to travel, and being confident (Robeyns, 2005).

In the ICT4D literature, two categories of studies that used the capability frameworks have been identified: the studies that operationalized (or put into practice) the frameworks (Lorgelly, Lorimer, Fenwick, Briggs, & Anand, 2015; Simon et al., 2013; Yap & Yu, 2016) and those that applied the frameworks (Anand & Dolan, 2005; J. Coast, Smith, & Lorgelly, 2008; Dang, 2014; Hill, Douptcheva, Lerch, & Sauvain-dugerdil, n.d.; Karimi, Brazier, & Basarir, 2016; Law & Widdows, 2008; Mitchell, Roberts, Barton, & Coast, 2017; Y. Wang, Kung, & Byrd, 2018). Most of the studies that in the operationalisations category designed a framework using Amartya Sen's work (Lorgelly et al., 2015; Simon et al., 2013; Yap & Yu, 2016). What is common among the operationalisations is that they look at the actual outcomes, the extension

of freedoms, but they categorize the outcomes differently and focus on different concepts within the framework

### **2.3.2 Sustainable livelihoods approach**

The Sustainable livelihoods approach (SLA) is one of the most widely used livelihoods frameworks in development practice and has been adapted by different development agencies such as the British Department for International Development (DFID) DFID (2008). A livelihood, according to DFID, consists of capabilities, assets and activities needed for a means of livelihood. A means of living is sustainable when it can cope with and recover from stresses and shocks and maintain or enhance its capabilities and assets whether now and in the future, while not undermining the natural resource base (DFID, 2008). Recognising that livelihood insecurity is a constant reality for many poor people, and that insecurity is a core dimension of most poverty, the SLA has as its core the elimination of poverty in poorer countries through tangible and intangible assets (physical, natural, financial, human, and social) of firms, communities and individuals (Mazibuko, 2013). It does this by putting people at the centre of development, thereby increasing the effectiveness of development assistance through the objectives, scope and priorities for development. The SLA explores livelihood resources and strategies that enable or constrain the achievement of sustainable livelihoods for different groups and institutional processes (Serrat, 2017). Although the application of the SLA is flexible and adaptable to specific local settings and to objectives defined in participatory manner, it inspires many core principles.

Among the critiques of the SLA are its too much focus emphasis on the micro level (community health, for example), inadequate attention to the macro-level (the state or the international society, for instance), and its focus on self-help, assumptions that those living in poverty can always make 'rational' choices and difficulties in defining and measuring the capital types and sustainability (Petersen & Pedersen, 2010)

Another drawback of the usability of the SLA framework is that there are too many issues to address, and this makes the framework too broad and superficial to help design and analyse anything in-depth (Petersen & Pedersen, 2010). In this regard, Barnes et al (2017) called for disaggregation of the unit of analysis with emphasis on individuals through analysis of dimensions such as gender, age, wealth, and the distribution of control over resources.

In the area of ICT, the SLA has been applied for identifying knowledge gaps required for providing a way for the rural poor to escape deprivation and empower them to diversify the contributions to their livelihood through ICT (Grinfeld, 2007).

### 2.3.3 Millennium development goals

Recognising that health is central to development and poverty reduction, the millennium development goals (MDG), launched in year 2000, has three of its eight goals directly related to health, as noted by WHO (2005). These are: Reduce infant mortality (goal 4); Improve maternal health (goal 5); and Combat HIV and AIDS, malaria, and other diseases (goal 6). Several ICT4D studies have been carried out because of the MDGs. For instance, *infoDev* (global development financing programme, hosted by the World Bank conducted evaluations on ICT4D projects (Batchelor & Norrish, 2005) focussing answering questions related to the relationships between poverty and ICTs.

Studies from several development institutions, such as the United Nations Development Programme, UNDP (2015) world bank (World Bank, 2013), the DFID and the Organization for Economic Cooperation and Development OECD (2003) have examined the relationship between ICTs and efforts to reduce poverty and achieve the other MDGs, and have come to the firm conclusion that effectively incorporating ICTs into development programs are useful tools in efforts to reach the MDGs. The UNDP (2015) note that ICTs themselves will not eradicate poverty but will be a useful tool for creating an enabling environment in pursuit of the MDGs. Similarly, the world bank argued that in spite of the various issues associated with deploying ICT projects, there is a growing evidence that the use of ICTs can be a critical and required component for addressing some facets of poverty in the following ways:

- 1) provide more efficient means of production
- 2) bring past unattainable markets within the reach of the poor
- 3) improve the delivery of government services, and
- 4) facilitate management and transfer of knowledge.

It is evidently clear from the studies cited above that the development of ICTs in the health sectors of Nigeria and many other developing countries, can support and drive the socio-economic development (Haluzá & Jungwirth, 2015) of these countries and in particular assist them to achieve the various MDG targets in the health services (Chetley et al., 2006; Panir, 2011)

## 2.4 Frameworks guiding the study and justification

First the study acknowledges the primary goal of ICT4D, which is promoting the socio-economic development objectives of disadvantaged communities in both urban and rural areas through the direct or indirect application of ICT innovations. According to major development institutions such as the UN, World Bank, DFID, and OECD, these development objectives may include healthcare delivery, education, income growth, etc through the application of ICTs. Given that the intention of the present study is not to argue against or for any definition or theory, the study streamlines the definition of development in the context of its objectives by borrowing from the various definitions from these international agencies. In this study, development refers to both social and economic development at micro-level. The micro-level, as used in this study, means providing depression screening interventions through ICT innovative applications to Nigerian universities, which normally contribute to making them live in poor health. Evidence from several studies in Nigeria (Afolabi et al., 2008; Baiyewu, Yusuf, & Ogundele, 2015) and other developing countries (E. Coast, Leone, Hirose, & Jones, 2012; Das, Do, Friedman, McKenzie, & Scott, 2007), point to the negative effects of poor health costs on welfare. The inability of the severely depressed, for instance, to work, thereby straining family resources, strengthens the relationship between poverty and depression. Costs associated with depressive disorders are recognised as a major cause of household poverty, particularly undetected depression that entail direct expenses on treatment and opportunity costs through lost incomes (Baiyewu et al., 2015).

In this regard, the study referred to ICT4D as the use of technologies that facilitate, through electronic means, the creation, management and dissemination of information to enhance the socio-economic wellbeing of its users. Put simply, in the context of the study, ICT4D refers to using machine learning and HCI technological innovative applications to facilitate effective and efficient depression screening for University staff and students by University academic staff (the users). The study argues, in this definition, that the innovation's impact to members of the University academic staff and students can be described in both social and economic development context and are expected to enhance the living standards of the University community.

Economic development in this study, is described as a way of enhancing some aspects of the living standards of the University population by reducing the cost associated with depression screening and improving their mental health. Social development, on the other hand, addresses those needs that may not necessarily be described in economic terms but can

positively affect the living standards. For example, in the pre-study survey, all 63 (100%) University staff and students interviewed expressed dissatisfaction with their visit to hospitals because of the high costs, long waiting times to see a clinician and long waiting times to retrieve their medical records. This corroborates the findings of several studies in medical literature of high costs and long waiting times experienced by patients in hospitals (Ansell, Crispo, Simard, & Bjerre, 2017; Globerman, 2013; Oche & Adamu, 2013). The living standards (costs of consultation and long waiting times) of University staff and students may be raised if ICTs-based depression screening tools are created and used in universities. The attempt to alleviate this and many other health challenges faced by members of the University community is what the study refers to as social development. It includes development as a process of using ICTs to expand the freedoms that people enjoy (Sen, 2007).

## **2.5 Depression and depressive symptoms**

Depression, a common health problem, ranking third after cardiac and respiratory diseases, has been found to be a major predictor of disability, mortality and poor quality of life (A. Ibrahim et al., 2013; Morrison, Shin, Tarnopolsky, & Taylor, 2014). An estimated 350 million people of all ages are suffering from depression, worldwide with varying degrees of prevalence rates, primarily due to different assessment methods, study designs, and heterogeneity of the populations under study (Lim et al., 2018). In some studies different factors such as age (Schaakxs et al., 2017), gender (Luppa et al., 2012) and populations under study (A. Ibrahim et al., 2013) have been associated with varying degrees of prevalence rates of depression. For instance, the work of Ibrahim et al (2013) suggest an evidence that University students are at higher risk of depression, despite being a socially advantaged population.

The International Classification of Disease (ICD-10) WHO (1992), and the Diagnostic and Statistical Manual of Mental Disorders 5th edition (DSM-5) American Psychiatric Association (2013), which identified depression as being characterized by the presence of cognitive, emotional, somatic, and behavioural symptoms, are the commonly used classification systems for diagnostic purposes in mental health. These classification systems help psychiatrists to measure and quantify mental illness though resulting in reliable diagnosis.

Depression is characterized by the presence of cognitive, emotional, somatic, and behavioural symptoms American Psychiatric Association (2013; WHO (1992). Depression can also be categorised as either a diagnosis or a dimension (symptoms) (Stein, 2012)). Depression represents a diagnosis of a depressive disorder from the categorical perspective while it

represents the presence of depressive symptoms of different grades from the dimensional perspective.

Three examples of diagnoses, as defined by the fulfilment of a set of criteria in the DSM-5 and ICD-10 classification system, are the major depressive disorder (MDD), Persistent depressive disorder (called dysthymic disorder in the ICD-10 classification system) and Adjustment disorder. According to the DSM-5, persistent depressive disorder is a new diagnosis, which includes both chronic major depressive disorder and the previous dysthymic disorder. The main reason for this change was no evidence for meaningful differences between these two conditions. This is represented in Table 2.1.

Table 2.1 Conceptual model of depression and depressive symptoms

<b>DSM-5</b>	<b>ICD-10</b>
Major depressive disorder Presence of 5 or more symptoms	Major depressive disorder Presence of 5 or more symptoms
Persistent depressive disorder 3 or 4 symptoms	dysthymic disorder 3 or 4 symptoms
Adjustment disorder 2-4 depressive symptoms	Adjustment disorder 2-4 depressive symptoms

Depression is characterised by sad mood, loss of pleasure or interest, withdrawal from social activities, feelings of guilt, worthlessness, insomnia, hypersomnia, disturbed appetite, loss of energy, inability to concentrate, and in severe cases suicidal ideation

Of these symptoms, sad mood and loss of interest or pleasure are two main criteria, and the remaining symptoms are additional ones (American Psychiatric Association, 2013). An MDD is established if one of the main criteria and, at least four other criteria are consistently present during a two-week period (American Psychiatric Association, 2013). In addition, the symptoms shall be followed by a decline in functional social activities. DSM-5 offers explicit symptomatic criteria allowing for a reliable diagnosis of depression (American Psychiatric Association, 2013). A diagnosis of dysthymic disorder is established if the patient experiences three or more of the depressive symptoms, including sad mood and, at least, two additional symptoms, present for at least two years. Two to four of the depressive symptoms, including sad, are required to be consistently present for two weeks in order to establish an adjustment disorder. This should be followed by significant impairment in occupational, social, or other vital functioning areas (American Psychiatric Association, 2013).

### **2.5.1 Screening for depression and depressive symptoms**

In mental health, the standard method for detecting depression is an interview conducted according to the DSM-5 and ICD-10 classification systems (American Psychiatric Association, 2013; WHO (1992). The interview is usually administered by a clinician or trained mental health professional. Structured and semi-structured diagnostic interviews are the commonly used approaches for establishing the presence of depression. The Structured Clinical Interview for DSM Disorders (SCID) (First, 2015) is considered as the gold standard in mental health for interviewing and applying the DSM-IV criteria for diagnosis. The SCID became an efficient, easy-to-use instrument which helped medical professionals to make standardized, reliable, and accurate diagnoses (First, 2015). Schedules for Clinical Assessment in Neuropsychiatry (SCAN) (Wing et al., 1990) is another commonly used semi-structured interview created by the world health organisation (WHO) to detect and measure mental disorders that may occur in adult life. Though not created directly using any of ICD-10 and DSM-IV, the SCAN could be used for both systems. The WHO Composite International Diagnostic Interview (CIDI) (World Health Organization, 1990) is a fully structured interview used to diagnose mental illnesses. It was generated in collaboration with WHO and the United States. Alcohol, Drug Abuse, and Mental Health Administration (ADAMHA). It takes approximately one to two hours to administer a full structured or semi-structured interview. This may be challenging in clinical practice as many patients are frail and require personnel resources. It is however, possible to select some parts from the interview that are related to specific diagnosis of interest, which takes shorter time.

### **2.5.2 Screening tools for depression and depressive symptoms**

In this section, some of the most commonly used tools for screening and monitoring depression disorders are described. These include: Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983), Beck Depression Inventory (BDI) (Beck, 1961), and the Patient Health Questionnaire – 9 (PHQ-9). However, it is noteworthy of mention that screening tools are not diagnostic. If a patient has been assessed by a certain questionnaire and identified to experience depressive symptoms, the patient would require further assessment with a clinical interview to fully establish a potential diagnosis.

### **2.5.2.1 Hospital Anxiety and Depression Scale**

The hospital anxiety and depression scale (HADS) is a 14-item questionnaire developed by Zigmond & Snaith (1983) for screening for depressive symptoms and anxiety in a medically ill population, based on patients' self-report (Zigmond & Snaith, 1983). The HADS has seven items on each subscale; depression and anxiety, respectively, and it's rated from 0 to 3, with 0 representing 'no problem, although the rating can be reversed in some items. The feedbacks on this scale are dependent on the on frequency of symptoms in the past week. According to Snaith and Zigmond (1986), the questionnaire does not have all the symptoms of somatic and cognitive conditions of depression in accordance with DSM criteria. Four of the items are about anhedonia (one of the major diagnostic criteria for MDD in ICD and DSM classification systems). The cut-off score of 7 or 8 from the total depression sub-scale score of 21 is used to indicate 'possible' depression, while 10-11 indicative of for 'probable' depression.

### **2.5.2.2 Beck Depression Inventory**

The Beck depression inventory (BDI) is a twenty one-item patient self-report measure of the severity of depressive symptoms. The questionnaire assesses both the cognitive and somatic symptoms of depression over the last two weeks. The scores are interpreted as follow: 0-3 means 'low depression'; 4-6 represents 'mild depression'; 7-9 indicates 'moderate depression' while 10-21 indicates 'severe depression' (O'Connor, Rossom, Henninger, Groom, & Burda, 2016) It has high sensitivity and specificity and it's quite reliable in assessing depression severity symptoms (O'Connor et al., 2016). The shortcomings of BDI are its high item difficulty (requires the patient to be able to read and understand the questions) and poor discriminant validity against anxiety.

### **2.5.2.3 The Patient Health Questionnaire - 9**

The Patient Health Questionnaire - 9 (PHQ-9) is a 9-item streamlined patient self-report instrument, which is a version of the Primary Care Evaluation of Mental Disorders (PRIME MD) instrument (Robert L. Spitzer, Kroenke, & Williams, 1999). This tool was first designed to assess and detect specific psychiatric disorders in primary care using the DSM diagnostic criteria, which consists of two parts: a one-page questionnaire and a 12-page clinical evaluation guide. The questionnaire has nine (9) symptoms of depression according to the DSM-5 diagnostic criteria American Psychiatric Association (2013). Each question requires a user to

rate the frequency of a depressive symptom experienced over the last two weeks. These items include: 1) loss of pleasure or anhedonia, 2) sad or depressed mood, 3) insomnia or hypersomnia, 4) fatigue or loss of energy, 5) appetite disturbances, 6) guilt or worthlessness, 7) diminished ability to think or concentrate, 8) psychomotor agitation or retardation, and 9) suicidal thoughts. Scores for each item range from 0 (“not at all”) to 3 “nearly every day” with a total score ranging from 0 to 27 (Gelaye et al., 2013)

## **2.6 Linking ICT4D, machine learning, HCI and depression screening**

In this section, a description of framing the study as an ICT4D research is given. This is followed by a description of the link between the four components of the study, namely: ICT4D, machine learning, HCI, and depression screening.

As stated in section 1.4, one of the reason for using Nigerian University as a case study for depression screening is that it could be studied within the ICT4D context. The staff and students, who are the key members of the University, are categorised as disadvantaged population in the sense that despite being at high risk of depression due to problems with job dissatisfaction/poor academic performance, inadequate family support, accommodation problems, heavy cigarette smoking and high level of alcohol or other drug consumption (Adewuya et al., 2006; Peltzer et al., 2013), the resources to cater for them are limited, hence they can be said to be a potential target group for ICT4D interventions.

Given the vision of ICT4D, which is essentially rooted in the idea of universal access to information as a necessity for development, Hu et al (Hu, Perer, & Wang, 2015) note two major features in data-driven healthcare, 1) medical professionals generate voluminous amounts of data as an outcome of their daily activities and interactions with the millions of patients, which need to be analysed, and 2) the complexity of these data being collected in paper and stored in fragmented records in different departments, make it almost impossible for medical professionals to discover hidden patterns in them and analyse using manual methods. These features call for a new powerful tool to convert structured (e.g. diagnosis codes, medications), unstructured healthcare data (clinical notes), and image data (X-rays) into useful knowledge, which is expected to be greatly aided by artificial intelligence (AI) techniques in a very short time frame. Machine learning, an important subfield of AI research, provides powerful techniques in the context of real-world healthcare data and uses cases to build up data-driven healthcare analytics framework with greatly enhanced accuracies at much lower costs (Hu et al., 2015). Machine learning application is described in detail in section 2.6

The central goal of the study is to develop an ICT4D framework for collaborations between, HCI technology, humans and machine learning models in the domain of depression, taking advantage of their relative strength to achieve what each cannot achieve alone. Part of the specific objectives is to develop and enable machine learning models to become more accessible, effective, efficient tools, so that users from different backgrounds can use them with satisfaction. One approach is to combine machine learning methods and HCI techniques (Holzinger et al., 2013; Shneiderman, 2002). Machine learning-HCI collaboration has been described as the process of identifying novel valid, and potentially useful data patterns, with the goal to understand these patterns. Shneiderman (2002) identified two ways of combining HCI techniques and machine learning algorithms. One way is to provide support tools for users with both components so that users can then explore data with direct manipulation user interfaces. Another way is to use machine learning as a first pass and then examine the results visually. Examples included air-traffic control and video games (Holzinger et al., 2013).

Machine learning models are good at computation and analysis on data, humans are better at extracting knowledge from their experience, and transferring the knowledge obtained to different domains. The field of HCI research, on the other hand, is concerned with the broader human context of computing systems, with the primary goal of developing technology and practices that improve the usability of computing systems. Thus, the design requirements of the study, implementation and evaluation of the system, made use of usability standards in international standard for organisation (ISO) 9241 (ISO, 2017), a popular technique in HCI, which includes the efficiency, effectiveness, and satisfaction with which the user interacts with a system.

As noted by Grudin (2009) and Heuer (2013), the skills of machine learning systems and HCI are complementary. Although investigation on how machine learning models can be made more accessible and usable to people working in different domains is an ongoing research, Heuer (2013) noted that the effectiveness of machine learning algorithms is dependent on three factor: quality of machine learning algorithms, quality of data, and user interface paradigms that machine learning experts employ to do their work. Grudin (2009) further note that the collaboration of machine learning and HCI methods could become very relevant, in the sense that machine learning community could increase their usage of HCI methods by giving users the right amount of support and assistance and enable developers to create better user experiences

The challenge, however, is how to create a framework that enables humans who may or may not have deep knowledge about the internal workings of machine learning systems or usability technique, to use machine learning models for decision making in safety-critical domain, in ways that earn the trust of professionals. The study started by developing interpretable machine learning models for data computation, and then created a graphical user interface (GUI) design that support transparent interaction between novice users and machine learning algorithms.

A number of machine learning algorithms have been proposed for addressing challenges in medical diagnosis, many of which have been reviewed by Kavakiotis (2017). Some of the more notable ones, discussed in chapter four, are Bayesian networks, decision trees, K-nearest neighbours ( $k$ -NN), Artificial neural networks, and Support vector machines.

Part of the focus of this study is using machine learning algorithms as pipeline for pattern recognition in clinical depression data. Developing machine learning models given patients' data and expert knowledge for making predictions in clinical depression is of great practical importance (Patel, Khalaf, & Aizenstein, 2016a). Machine learning tasks can be broadly divided into five main classes: Association rule, Classification, Regression, Clustering and Optimization tasks (Murphy, 2012). A complete discussion of all the divisions is outside the scope of this study but classification task, which is the focus of the study, is briefly described.

Classification techniques are used to assign features to the pre-determined class targets. The class role is played by a selected feature in the dataset. In statistics, the feature is called the dependent variable. While classifying the features, the classifier creates a classification model, which can then be adapted to new data. For instance, the breast cancer diagnostic model developed by a classification technique may continue to be applied to the decision-making support system for the purpose of the diagnosis of a patient whose data were not used to create the prediction model. Classification is a two-step process consisting of training and testing steps. During the training step, the algorithm analyses the data meant for learning and creates a classification model. During testing, the model is checked for accuracy by using a new set of data. Solving classification tasks includes the selection of appropriate algorithms. A few commonly used classification methods include Bayesian networks, artificial neural networks, decision trees, and many others (Murphy, 2012). Bayesian networks, which the study concentrates on, is described in section 2.6

## 2.7 Bayesian network approaches in medical diagnosis

The mechanics of using probability theory in medical expert system decision-making usually rests on an early theory by Thomas Bayes (1763), to the fields of probability and statistics. This is the origin of Bayesian networks. Bayes' theory is a useful mathematical device that combines evidence from data with prior information. The Bayesian solution to the prediction problem consists of specifying a family of probability distributions (a model), and assigning a prior probability to each distribution in the family. Healthcare areas in which Bayesian networks are used to support problem-solving include diagnostic reasoning, prognostic reasoning, treatment selection, and discovering functional interactions (Armero et al., 2016; Lucas, Gaag, & Abu-Hanna, 2004)

The study concentrates on medical diagnosis, which is a logical process by which the clinician follows a chain of events from cause to effect and, thereby, reaches a decision. This has stimulated the interest of a wide variety of researchers for a long time. Bayes theory is derived from the product rule of probability, and relates the conditional and marginal probabilities of events X and Y, provided that  $P(Y) \neq 0$ .

Bayes' theorem is stated mathematically:

$$P(X/Y) = \frac{P(Y/X)}{P(Y)} \times P(X) \approx \text{posterior} = \frac{\text{likelihood}}{\text{evidence}} \times \text{prior} \quad 2.1$$

$P(X)$  = prior probability (or “unconditional” or “marginal” or pre-test probability) of event X occurring. It is “prior” in the sense that it does not take into account any information about Y; however, the event Y need not occur after event X. In other words, the prior  $P(X)$  describes the information we have about the variable before seeing any data.

$P(X/Y)$  = posterior probability (or conditional or post-test probability) of event X occurring given that event Y has occurred. It is called “posterior” because it is derived from or depends upon the specified value of Y. Clinicians determine the *a priori* probability of a disease, and then incorporate laboratory and imaging findings to calculate the *a posteriori* probability (Ghosh & Valtorta, 1999).

$P(Y/X)$  is the likelihood or conditional probability of Y occurring given that event X has occurred.

$P(Y)$  is the evidence or marginal likelihood or model likelihood of event Y occurring, and acts as a normalizing constant to guarantee the posterior probabilities sum to one.

$\frac{P(Y/X)}{P(Y)}$  is the Bayes' factor or likelihood ratio. The term likelihood is used for the probability that a model generates observed data.

Introducing Bayesian networks into the present study, which is depression diagnostic framework, allow the decision support system to deal with probabilistic uncertainty in a mathematically sound way.

As stated in equation 2.1, Bayes' theorem applied to diagnosis may be written:

$$P(D/S) = \frac{P(D)P(S/D)}{\sum_j P(D_m)P(S/D_m)} \quad 2.2$$

This expresses the conditional probability of disease D, given symptom complex S as a function of the unconditional probability of disease D and the conditional probability of symptom complex S, given disease D.

Equation 2.2 was extended for reasoning under uncertainty in medical diagnosis (Shojaei Estabragh et al., 2013):

$$\frac{P(D)P(S/D)}{P(D)P(S/D) + P(S/\neg D) * P(\neg D)} \quad 2.3$$

$P(\neg D)$  = prior probability of disease being false

$P(S/\neg D)$  = probability of finding symptom S even when disease D is false

To overcome the difficulty of diagnosis where  $P(S)$  is not known, the following formalism is used:

$$P(D/S) = \frac{P(\neg D)P(S/\neg D)}{P(S)} \quad 2.4$$

Dividing equation (2.3) by equation (2.4) gives:

$$\frac{P(D)P(S/D)}{P(\neg D)P(S/\neg D)} \quad 2.5$$

Expressing the conditional probability in equation (2.2) for multiple diseases  $D_1, D_2, \dots, D_m$  and multiple symptoms  $S_1, S_2, \dots, S_n$  looks like this:

$$P(D_i)/S_1, S_2, \dots, S_n = \frac{P(S_1/D_i)P(S_2/D_i) * \dots * P(S_n/D_i)P(D_i)}{\sum_{k=1}^m P(S_1/D_k)P(S_2/D_k) * \dots * P(S_n/D_k)P(D_k)} \quad 2.6$$

Bayesian networks is not only a computer reasoning method, but a common clinician problem-solving tool in everyday life; that is reasoning based on past knowledge and experience (Seixas, Zadrozny, Laks, Conci, & Muchaluat Saade, 2014). As a machine learning technique, it is inspired by a model on how humans solve problems in their daily lives. For instance, a clinician solves a medical diagnosis problem by applying past knowledge and experience on a new case (Chattopadhyay, 2017). Bayesian networks is applied in medical domains such as diagnosis, treatment, classification, and knowledge acquisition, and also fits into complex daily problem-solving because it can combine with other machine learning techniques (Horný, 2014). The motivations for applying Bayesian networks in medical diagnosis, according to Koller and Friedman (2009), Horny (2014) and Heckerman (1997), are:

1. method works in ways similar to normal human reasoning process, and acceptable to clinicians in detecting diseases. That is, solves a new problem by heavily relying and applying past knowledge and experience. This is the prime reason for preferring to use Bayesian networks for this study.
2. because of its numerical capabilities, it is a convenient environment to represent and process incompleteness of knowledge and uncertainty of information, which are essential parts of medical domains
3. cognitively adequate knowledge extracted from patients' records
4. well-developed probability theory formalism (with which it assigns numerical degree of belief between 0 and 1
5. ability to integrate different types of knowledge
6. the knowledge in the domain grows with time so, it is important that the system can learn new knowledge. Bayesian networks can learn by adding new cases into the case base
7. the cases in the knowledge base can be used for treatment plans for patients and for training purposes of the less experienced clinicians.

### **2.7.1 Bayesian networks classifiers**

Bayesian networks are specific types of graphical models represented as a directed acyclic graph (DAG), which are capable of displaying dependent and independent relationships among random variables, clearly and intuitively through the established theory of probability (Koller & Friedman, 2009). They consist of nodes,  $N = X_1 \dots X_n$ , which represent discrete or continuous random variables (representing object and events in the real world) and a set of

directed arcs  $A = X_1 \dots X_n$ , between pairs of nodes, which represent influential relationships between the nodes (Conrady & Jouffe, 2015). The arcs (edge) in a directed acyclic graph point from one node, called parent node, to the other node, called child node. However, the absence of an edge linking between two variables is an indication of independence between them given that the values of their parents are known. For a discrete variable, the strength of the relationship between the variables is represented by conditional probability distribution. Each node is associated with either a prior (or unconditional) probability, or a conditional probability table (CPT) which quantifies the influence of its parents. The conditional dependencies in the graph are estimated using well-known computational methods (Korb & Nicholson, 2004). This probabilistic framework for both prior and future events makes Bayesian networks suitable for estimating uncertainty that is inevitable with prediction (Lu, 2005). Bayesian networks models can be used for probabilistic prediction tasks such as classification and regression. Many interesting and difficult real-world problems such as medical diagnosis, pattern recognition, forecasting, cluster analysis, association discovery or anomaly detection (Conrady & Jouffe, 2015), which fall into the supervised or unsupervised problem can be described as classification tasks, with each requiring the construction of an algorithm or classifier that assigns a target class label to instances described by a set of attribute names and values (Su, Andrew, Karagas, & Borsuk, 2013).

## 2.7.2 Conditional independence

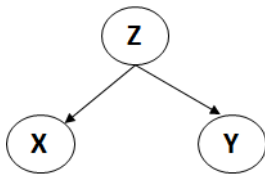


Figure 2.2: Simple Bayesian network with three random variables

Given three sets of random variables  $X$ ,  $Y$ , and  $Z$ , the variable  $X$  is said to be conditionally independent of  $Y$  given  $Z$ , as illustrated in Figure 2.2, if and only if the probability distribution governing  $X$  is independent of the value of  $Y$  given  $Z$ ;

$$P(X/Y, Z) = P(X/Z)$$

2.3

In general, a Bayesian networks is of the form: if the nodes of a network are  $X_i$  and the set of parent nodes of  $X$  is  $pa(x_i)$ , then the probability of any particular set of values for all the nodes in the network is given by

$$P(x_1, \dots, x_D) = \prod_{i=1}^D P(x_i | pa(x_i)) \quad 2.4$$

where  $pa(x_i)$  represents the parental variables of variable  $x_i$ . Written as a directed acyclic graph, with an arrow pointing from a parent variable to child. A Bayesian networks is a directed acyclic graph with the  $i^{\text{th}}$  vertex in the graph corresponding to the factor  $P(x_i | pa(x_i))$

For nodes with no parents,  $P(x_i | pa(x_i))$  should be read as just  $P(x_i)$ , the prior probability

### 2.7.3 Bayesian networks modelling process

Bayesian networks framework for inferencing and decision making is quite attractive and has a long history of real-world applications including medical diagnosis, computer troubleshooting, speech recognition, and traffic control, to name a few (Kamm & Tretjakov, 2009). Bayesian networks modeling process involves determining the nodes and their values, the network structure, and finally adding their numerical probabilities (Pollino & Henderson, 2010). A Bayesian networks is designed as a directed acyclic graph, with nodes  $X = x_1, \dots, x_n$ , from the domain and directed arcs connecting pairs of nodes  $x_i \rightarrow x_j$ , representing direct probabilistic dependencies between nodes. If the nodes are discrete, the strength of the relationship between nodes are quantified by a conditional probability table for each node (Barber, 2011). The only constraint on a directed acyclic graph is that it must have no directed cycles (that is, no node in the graph can be revisited simply by following the direction of the arcs). This is illustrated in Figure 2.3

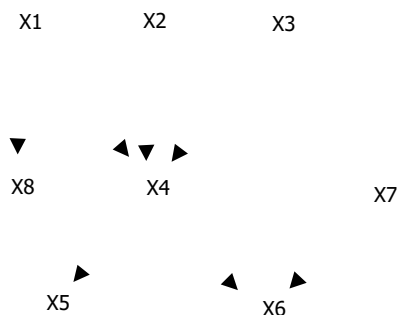


Figure 2.3 Relationships in a directed acyclic graph

In a directed acyclic graph, a direct arc from node  $X$  to node  $Y$  shows that node  $X$  has a direct influence on node  $Y$ , written as  $X \rightarrow Y$ . In this case,  $X$  is referred to as a parent of  $Y$ , with  $Y$

being referred to as a child of  $X$ . In the directed acyclic graph example (Figure 2.6), the parents of node  $X_4$ , written as  $\text{pa}(X_4) = (X_1, X_2, X_3)$  while the children of  $X_4$ , written as  $\text{ch}(X_4) = (X_5, X_6)$ . The family of a node is itself and its parents. In this case, the family of  $X_4$ , written as  $\text{fa}(X_4) = (X_1, X_2, X_3, X_4)$ . The Markov blanket (MB) of a node is itself, its parents, children and the parents of its children. In this case, the Markov blanket of  $X_4 = (X_1, X_2, X_3, X_4, X_5, X_6, X_7)$ . If there is a directed chain of nodes, one node is an ancestor of another if it appears earlier in the chain, whereas a node is a descendant of another node if it comes later in the chain. In this case,  $X_1, X_2, X_3$  are ancestors to  $X_5, X_6$  and  $X_5, X_6$  are descendants of  $X_1, X_2, X_3$ . Any node without parents is called a root node, while any node without children is called a leaf node. In our example,  $X_1, X_2, X_3, X_7$  are root nodes, while  $X_5, X_6, X_8$  are leaf nodes. Any other node (non-leaf and non-root) is called an intermediate node. In the light of the cause-effect relationship of the Bayesian networks structure, root nodes represent original causes, while leaf nodes represent final effects (Korb & Nicholson, 2004).

The network directed acyclic graph and conditional probability table provides a convenient means of representing assumptions of conditional independence (CI). Each node  $x \in X$  in the graph denotes a random variable  $X$  (which is a feature in a dataset), and has a value corresponding to the probability of the random variable,  $P(X)$  (Ethem, 2010). The nodes and the arcs define the structure of the network, and the conditional probabilities are the parameters given the structure. Two kinds of information that can be modelled by Bayesian networks are topology of the graphical structure (or qualitative aspect) and the parameters representing probabilistic information (or quantitative aspect, conditional probabilities tables).

In a Bayesian network for discrete domains, each variable has a conditional probability table CPT, which represents the relationships between the nodes. Figure 2.7 illustrates the use of CPT. For each variable, all possible combinations of its parents are specified and the probabilities of the child taking the value of each of these combinations are then specified in the CPT.

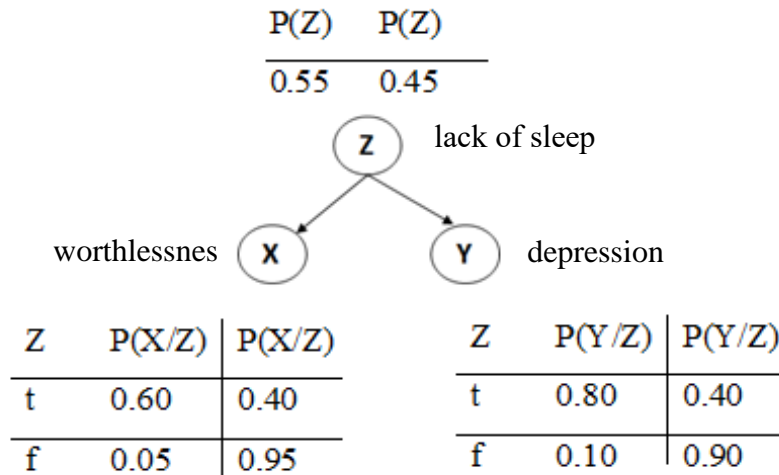


Figure 2.4 Schematic representation of a conditional probability table

In Figure 2.4, let Z = worthlessness, X = lack of sleep, and Y = depression

The arcs between the nodes indicate the causal dependencies between the nodes, and the CPT represents strength of influence of the nodes on each other as follows:

The prior probabilities of worthlessness Z = 0.55 is the probability that Z will be true (t) and 0.45 the probability that it will be false (f). The two probabilities sum up one because the total probability across all the possible states that a variable can take is one, given that X can only hold two possible states, T or F. The CPT of the next node, sleeplessness, X, shows that the probability of X being true given that its parent variable Z is known to be true is 0.60. Similarly, given that the Z is known to be true, the probability of X being false is 0.40. The CPT of depression Y can also be represented in a similar manner.

A Bayesian network represents the joint probability distribution (JPD) of all the nodes in it and this JPD can factorized into a product of the node's local conditional independent distributions.

For example, the JPD of all the variables in Figure 2.7 can be represented as follows:

$$P(Z, X, Y) = P(Z)P(X/Z)P(Y/Z)$$

One constraint to specifying the arcs in Bayesian networks is that directed cycles are not allowed i.e. a variable cannot be its own ancestor or descendant. For example, if two variables,  $A_i$  and  $A_k$   $A_i$  have a causal dependency where  $A_i$  (parent node) causes  $A_k$  (child node), this is represented as  $A_i \rightarrow A_k$ . This means that a situation where  $A_i \rightarrow A_k \rightarrow A_n \dots \dots A_i$  is not allowed. A Bayesian network can therefore be defined as a Graph G of the following form:

$$G = (V, E); \text{ if } (x, y) \in E \text{ then } (y, x) \notin E$$

That is, Bayesian networks allows only directed edges, but not directed circles ( $X \rightarrow X$ )

Having described a few of the several benefits of using Bayesian networks to implement a depression risk assessment model, it can be seen as a potentially viable option in this study.

## **2.8 Techniques in human computer interaction**

Human computer interaction (HCI), a huge and significant subfield of ICT4D (Ho, Smyth, Kam, & Dearden, 2009), involves the design approach that has the goal to obtain usability and also to ensure good user experience. HCI is the study and theory of the interaction between humans and complex technology, and how to design, create, and evaluate technologies to facilitate those interactions. It is concerned with how current input and output technologies affect interaction, and the situations in which these technologies and techniques might be put to best use (Toyama, 2010). In other words, HCI studies and models complex, real-world, human activities and situations. Being a multidisciplinary field covering diverse fields such as computer science, psychology, sociology, ergonomics and industrial design psychology, physics, biology (Hewett, Baecker, Card, Carey, Gasen, Mantei, Perlman, 2009; Limerick, Coyle, & Moore, 2014), HCI researchers often learn from these different disciplines in order to attain the goal. The central objective of HCI research is to make systems more usable, more useful, and to provide users with experiences suitable for their specific background knowledge and objectives (Endsley, 2016). The effectiveness of HCI, however, is dependent on a number of important parameters such as usability and comfortability, which would encourage the end-users to get the services safely and more reliably (Franklin & Sridaran, 2012)

The focus of HCI research is mainly on the use of technology in workplaces where computers have been the most prominent (Bardzell & Bardzell, 2015). As the use of computers in society increases, so does the variety of people who interact with the technology, and the contexts of use. It is therefore important to investigate HCI issues for these variety of users and new environments. One area in which computer usage is becoming prominent is the hospital (Appari, Eric Johnson, & Anthony, 2013; Paz & Pow-Sang, 2015; Teufel, Kazley, Andrews, Ebeling, & Basco, 2013). It is therefore necessary to expand the HCI research agenda to include the use of computers in screening for diseases. It is also important to investigate how non-expert users as a unique group of users and to understand how their interactions differ from those of experts.

Developer of collaborative human computer systems are faced with the huge task of developing software for several users (at design time) while making it work as if it were designed for each individual user (only known at use time). To complete a task with the

computer system before now, humans had to adapt themselves to the way that the technology operated. Beside requiring much longer training time, this was a very tedious approach to learn how to use computing systems. Over time, researchers have turned the situation around to where the computing systems can be adapted to the processes human beings use in completing a task. The design of the user interface with which humans interact with computers is an important area of HCI (Wright, Blythe, & McCarthy, 2006).

### 2.8.1 Usability and user experience

In the last two decades, there has been an increased emphasis on designing medical devices to improve the precision in disease screening processes (Acharya, Thimbleby, & Oladimeji, 2010; Franklin & Sridaran, 2012). According to Franklin and Sridaran (2012) well designed medical devices are necessary to ensure the health and safety of device users as well as provide safe and effective clinical care for patients. A good quality and usability is necessary to achieve the desired success in the medical devices to (Franklin & Sridaran, 2012). The design of the user interface with which humans interact with computers is an important area of HCI (Wright et al., 2006). The concept of usability was the pioneering work of Gould and Lewis (1985) though has been formulated and evaluated by several researchers over the years. As an important technique in HCI, usability is described and measured by Nielsen (Nielsen, 1993) in terms of five characteristics: *learnability* (system is easy to learn in order to help users perform tasks the first time they interact with the interface), *efficiency* (system is efficient to use so that users can perform tasks, quickly, once they have learned the system), *memorability* (system is easy to remember so that if users return to the system after a period of not using it, they can use it easily), *errors* (system has low error rate so that users can easily recover from any errors that may occur while interacting with it, and *satisfaction* (system is pleasant to use so that users derive satisfaction while using it). The focus of usability is on the criteria for improving the product's simplicity to learn, ease-of-use and effectiveness, and ways to simplify everyday life by improving interaction between people and the interactive products (Endsley, 2016)

User experience (UX), an important component of usability, focuses on parameters or impressions associated with user's satisfaction, aesthetic appeal, and emotional fulfilment as they use a system or product (ISO, 2017). The concept of user experience has been defined and applied in many research areas. For instance, during the design phase in medical applications, user experience developers are able to identify priorities in developing tools to help patients

manage chronic diseases, administer the correct dose of medication, or communicate more effectively with their healthcare provider (Willis, 2014). During the design process of the interactive user interfaces for depression screening, the concept of usability and user experience were considered.

## **2.8.2 Design process and principles**

Designing for a user experience is a continuous process that does not end when one single function is completed. Instead, it challenges the designers to evolve the product's design alongside its users' evolving behaviours, limitations and needs. The process of design consists of a series of practical and creative activity, which gives rise to a system or product that aims to assist end-users in achieving their objectives. The commonly used UCD approach for medical applications, which the study is based on described the design process in five orderly steps (Endsley, 2016; McCurdie et al., 2012; Vincent & Blandford, 2011). These steps are discussed in detail in chapter five.

The design process in this study resulted in an innovative interactive artefact, which was later evaluated by a combination of end-users and professionals in medical field and usability designs, as suggested by a host of researchers in usability literatures (P. Lilholt, Jensen, & Hejlesen, 2015; Yen & Bakken, 2009). Before the design process began, it was important to gain an in-depth understanding of how the current system works, who the users are and the challenges encountered when handling depression screening. Consequently, as part of the design process, this study recruited evaluators with strong domain expertise including mental health professionals, experts in the area of heuristic evaluation and other end-users in the evaluation of the designed prototypes. According to Yen and bakken (2009), while HCI experts are good at detecting general interface problems, end-users are better are identifying severe interface obstacles to their task performance and workflow. Put more simply, HCI experts identify more of "ease of use" problems while end-users detect more of "usefulness" problems. Mental health professionals were recruited to detect any flaws in the medical domain. In the present the study, all the participants recruited were intended end-users of the system to be designed

### **2.8.2.1 User interfaces**

In the design of the user interfaces of this study, the information technology that is available and close to the users' proximity and understanding was considered. The study is intended for

the PC, laptop and tablet screens. From our pilot study, these systems were provided and maintained by the institution's management. This has the potential of improving the way results from the machine learning models were presented.

### **2.8.3 Usability evaluation methods**

Usability evaluation methods (UEMs) describe those methods composed of series of well-defined activities to assess human interaction with a system or product, mainly to identify issues or areas of improvement in this interaction in order to increase usability (Paz & Pow-Sang, 2015). There are several evaluation methods but one of the most widely used is the User-Centered Design (UCD) (Endsley, 2016), an approach and philosophy for designing and developing usable products and systems that place the user at the centre of the process (Dearden et al., 2010). The UCD evaluation method, an integral part of the UEMs (Figure 2.1), is founded on getting user feedback in each step of the design process. Receiving such feedback involves several usability methods at any step of the design process. Several usability evaluation methods have been developed, mainly to establish, in a systematic way, the level of usability of graphical user interfaces and to identify usability problems. These methods have been classified differently by different studies, this study adopts the classification by Nielsen (Nielsen, 1993), which grouped usability evaluation methods into four general classes: automatic (involving the use of software to evaluate a user interface); empirical (involving real users to interact with a user interface); formal (incorporating the use of models to evaluate a user interface), and informal (where evaluators use rules and their skills, knowledge and experience to evaluate a user interface). Some commonly used methods in the empirical methods are the evaluators testing, user testing and the tool-based testing (Figure 2.5). In the evaluations testing methods are the Heuristics, Pluralistic walkthrough, Cognitive walkthrough, Guideline reviews, Consistency inspections, and the Standards inspection methods. The User testing methods have the user-based testing, Think-Aloud method, Constructive interaction (also known as co-discovery learning), Retrospective testing, Questionnaires and interviews, and the Focus groups methods. The Tools-based methods have the Software tools (automatic usability evaluation) and Software tools (transaction log file).

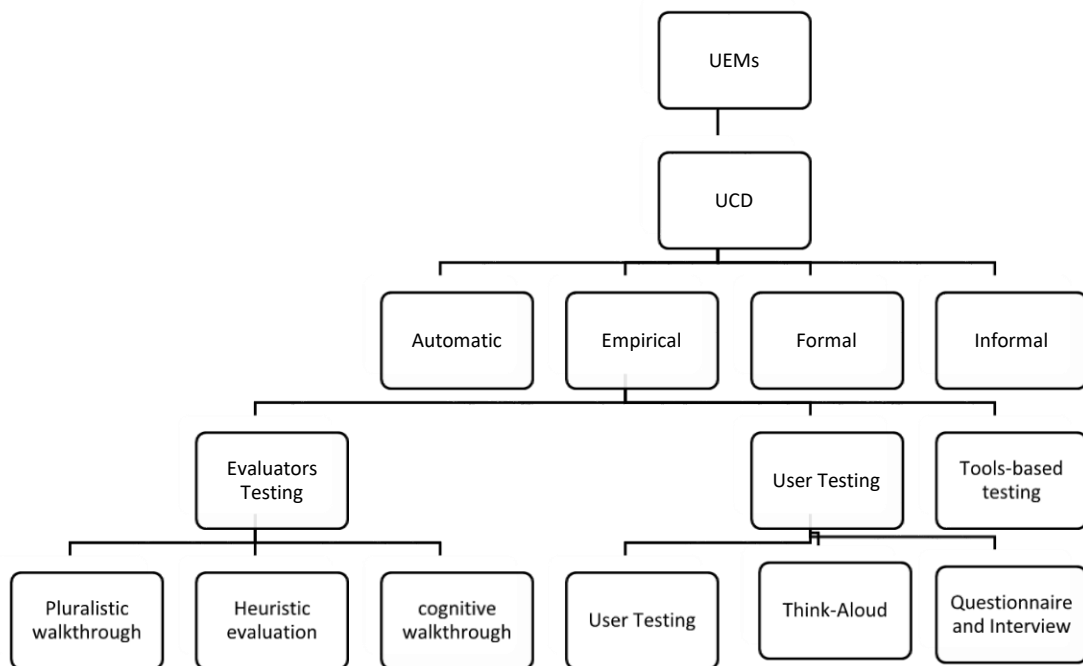


Figure 2.5 Usability evaluation methods

Given that part of the objectives of this study is to investigate which machine learning model would be integrated into a user interface screening tool to assist non-clinicians take assessment tests and get an estimation of the presence of depression for staff and students in a University in Nigeria, the study adopted the empirical approach where the real users of the developed artefacts will interact with it for feedback. Depending on the stage of the development process in which the evaluation takes place and on what design aspects should be evaluated, it is common in the HCI literature where more than one evaluation methods were used to evaluate a system (Bligárd & Osvalder, 2014). Furthermore, the Heuristic evaluation method from Evaluator-based testing was chosen for the study. This is discussed briefly in the section 2.8.4 while a more detailed discussion and its application to the study is given in chapter five.

## Chapter Three

### Review of related literature

#### 3.1 Introduction

This study is informed by a number of prior studies, which have been described by many authors on the possibility of adopting ICT4D frameworks, such as ICT and development frameworks, to support the methodology for depression screening in selected areas in Nigeria and other countries having similar changes with healthcare delivery services.. There are two reasons for this: 1) to examine the available knowledge in the study area in order to comprehend the connection between the problem and the body of knowledge in the subject area; 2) to establish the need for the study and acquaint the researcher with the methodologies and approaches that have been used by other researchers to find answers to similar research questions. The chapter describes studies on ICT4D and ICT4D frameworks and interventions that are relevant to the purpose and objectives of the present study. It also discusses studies that have attempted to solve medical screening challenges in developing countries using similar methods. Lastly, the study discusses findings, gaps and opportunities from the reviewed literature.

#### 3.2 ICTs and healthcare in Nigeria

In Nigeria, the three tiers of government (federal, state and local) are vested with the responsibilities of providing healthcare services to Nigerians (Nigeria Federal Ministry of Health, 2014). The federal government, however, has the responsibility of providing policy, drug regulations, diseases control, vaccinations, and most cases trainings. The federal government also runs the management of teaching, orthopedic and teaching hospitals (Nigeria Federal Ministry of Health, 2014).

One of the most important social services that serves as the basis of a country's infrastructural development is good healthcare delivery (WHO, 2013). Although the use of ICT across the world has risen, there is still a huge gap between the developed and developing worlds, especially in relation to developing countries such as Nigeria (Adeleke, 2015). In the last two decades, many developed countries have made significant progress in improving the health status of their citizenry. The disease burden is increasingly defined by disability instead of premature mortality, as it is in many developing countries and the major causes of death and

disability have changed from communicable diseases in children to non-communicable diseases in adults (WHO, 2013). However, in Nigeria and other Sub-Saharan African countries, the picture is quite different as communicable, maternal, nutritional, and newborns' diseases continue to dominate (WHO, 2013). As shown in section 1.1 (Table 1.1), Nigeria has an acute shortage of medical and mental health professionals (less than 1 per 100, 000 people) (Jack-Ide et al., 2012; WHO, 2013).

Healthcare delivery has gradually shifted from curative medicine to preventive medicine over the last two decades (Guilleminault et al., 2017). This shift has brought about not only an increased need for communication among healthcare providers as well as the publics, but has also necessitated the need for increased communication with the community. With Nigerian population at about 186 million (NPC, 2013; Worldometers, 2016) and still increasing, there is a need to find ways of improving efficiency and the quality of healthcare delivery systems in the country. A study by Adeleke (2015) notes that the growing field of ICT4D definitely has a key role to play in bringing the needed healthcare improvement and efficiency to the resource-limited and marginalised population of Nigeria. While it identified the advent of the use of ICTs in the health sector as the ideal vehicle for the improvement in disease detection and management particularly in Nigeria, the study showed that existing ICTs have had positive impact in the area of access to healthcare facilities and services, in many other African countries, including Kenya, Ghana, Ethiopia, Rwanda, Senegal, Tanzania, Malawi and Uganda (Adeleke, 2015). ICT tools that facilitate communication and the processing and transmission of information and the sharing of knowledge by electronic means, comprises all electronic digital and analogue ICTs such as computers (hardware, software and procedures), the internet Web 2.0 and 3.0, telephones (fixed and mobile), radio, and television, electronic-based media, such as digital text, audio/video recording, social networking and web-based communities (Barjis et al., 2013; Toyama & Dias, 2008)

There are thus increasing evidence of the transformative power of ICTs in the healthcare sector in Nigeria. The WHO (2013) has described electronic health (e-health) as the combined use of health-related ICT tools in the health sector including computers, software, mobile phones, and digital platforms, to enable healthcare professionals support and deliver health services to patients and their relatives (WHO, 2013). The main electronic health applications include, computer-aided screening and diagnostic, decision support tools, electronic health records, telemedicine, Internet-based technologies and services, digital imaging, and different kinds of health monitoring systems (WHO, 2013). Mobile Health (m-Health), another field of

new field of e-Health, has also been described as medical and public health practice supported by mobile devices including, wireless devices, mobile phones, patient monitoring devices, and personal digital assistants (Varshney, 2014). M-health Education is a new set of applications designed for mobile devices for training, testing, supporting, and supervising healthcare workers, and provide health information to individuals (WHO, 2013).

### **3.3 ICT4D for socio-economic development in Nigeria**

Although many theories have been used to explain it, as earlier stated, the link between ICTs and socio-economic development in the developing countries is still an ongoing area of inquiry (Thomas et al., 2017). This shall be described with specific reference to the Nigerian healthcare system in section 3.3.1

#### **3.3.1 ICT Policy for development in Nigeria**

The emergence and adoption of ICTs in fast-tracking development in all sectors of the Nigerian economy, and in the inseparable sociocultural and political spheres of life, is well known in the literature (Ezema, 2016; Gambo & Soriyan, 2017; Olatokun, 2009). Previous studies have noted the strategic difficulty that health institutions and industrial sectors will face if they did not take steps to harness and implement ICTs as tools for leveraging their activities in the emerging global economy (Akande, 2004; Asuke et al., 2016; Cd, Cn, Me, & Ho, 2017; Olatokun, 2009; UNCTAD, 2014).

Some of the several efforts the Nigerian government has made towards the development of ICTs in the country include: the launching of the National Telecommunication Policy in September, 2000, the development and launching of the National IT Policy in 2001, the development of a comprehensible science and technology policy in 2001, , the establishment of the national information technology development agency (NTDA) in 2001, and the launching of the national space research and development satellite systems programmes by the national space research and development agency (NASRDA) in 2001 (Babalobi, 2010). These efforts are complementary to other development initiatives, such as the granting of licenses to mobile telephone network operators (MTN, Airtel and Globacom) in 2001 (Ndukwe, 2011)

Nigeria's information and communications technology for development (ICTT4D) was launched in 2010,. The national information development agency (NTDA), in collaboration with the united nation's economic commission for Africa (UNECA), co-ordinated the development of the National ICT4D strategic action plan. This Action Plan provides

implementation of strategies for a 5-year period for the main sectors of the national economy- education, health, infrastructure, human resources, and development. agriculture, legal/regulations, media/community, among others - as part of an integrated approach to achieving national development within the context of the Federal Government of Nigeria's Seven Point Agenda, the national Economic Empowerment Development Programmes (NEEDS) and various socio-economic development programmes and initiatives, all geared towards the improvement of ICT penetration in all sectors of the nation's economy (FGN, 2012)

While there has been this remarkable investments and growth in ICT usage in many sectors of Nigerian economy, there is a worrying lack of empirical evidence on effects of ICTs upon the underserved areas and the lives of the poor. In the ICT literature, there has also been consensus that ICT usage in Nigeria's tertiary institutions is still very low and shallow (Ayoola, 2015; Babalobi, 2010). Thus, the National ICTs Policy, approved by the Federal government in January 2012, has been developed to support the development of Nigeria's Vision 2020 (FGN, 2012), with the goal of yielding reasonable expectations if compared to similar investments made in the mobile telecommunication ("create a conducive environment for the rapid expansion of ICT networks and services that are accessible to all at reasonable costs, and for the transformation of Nigeria into a digital and knowledge-driven economy" (FGN, 2012).

### **3.4 Acceptance of ICT applications in the health sector and theories of adoption and diffusion of innovations**

As earlier stated, a deeper understanding of the complex processes of the adoption and diffusion of ICTs requires the combination of theories from diverse perspectives, including communications, information systems, political science, sociology and international relations (Wisdom, Chor, Hoagwood, & Horwitz, 2014). In terms of technology and society compatibility, these theories could help explain why a particular technology is more compatible in certain societies than in others (Rogers, 2015). The Diffusion of Innovation theory that was elaborated by Everett Rogers (2015) became one of the most influential modernization theories (Rogers, 2015). It has become the blueprint for development communication. Rogers' intention was to understand the adoption of new behaviours, and that innovations diffuse over time according to an individual's stages. In in the Nigeria healthcare, for instance, economic factors, such as income level, the availability and price structures of ICT products and services, and

bandwidth and other supporting infrastructures have been noted as major factors affecting the effective utilisation, adoption and diffusion of ICT applications (Zayyad & Toycan, 2018). Income can be expected to be positively associated with the demand for modern ICT products. Economic factors influence the means used to access the ICT applications, such as the internet bandwidth. The availability of Bandwidth determines the adoption and diffusion of Internet. In Nigeria and many developing countries, bandwidth is very low, and lower bandwidth results in a longer time being needed to transfer data, and hence a low relative advantage for Internet use (Zayyad & Toycan, 2018).

Another factor of the value system is related to the skills required to use ICT applications. Literacy and computer skills are almost the prerequisites for ICT use (Zayyad & Toycan, 2018). High level of illiteracy, poor infrastructural development in the rural and urban areas, availability and affordability certainly have profound effects on the acceptance and diffusion of the new innovations by both the users and non-users of healthcare services. A large proportion of the population in developing countries, including Nigeria – the site of this study, is illiterate, and a still higher proportion lacks computer skills (Oyegoke, 2013). Other studies have equally reported similar findings (Ahlan & Isma'eel Ahmad, 2015; Asemahagn, 2016). Innovations are adopted by those individuals from the higher socioeconomic strata, rather than the lower ones. As Rogers (Rogers, 2015) had previously stated, the acceptance patterns of an innovation are dependent on several elements, including the characteristics of the innovation itself, the channel of communication, the nature of the social system, and time. ICT adoption and utilisation consists of successive phases that eventually lead to its use or rejection. In addition to the inherent problems as highlighted, deficiencies in the knowledge and skills of patients and health professionals to use ICT solutions present other challenges. This was particularly noted during the pilot study in chapter one. Other challenges relate to the challenges ICT poses for healthcare organisations. Even when implemented, the benefits of ICTs and applications cannot be realised if the intended users are unable to use them. One challenge is to train the intended users in the use of ICT applications so that they can improve on their health or quality of service (Ruxwana, Herselman, & Conradie, 2010).

In summary, the findings of the studies under review and others like them are a pointer to the fact that a number of factors influence the adoption and utilisation of ICTs in the Nigerian healthcare sector, including: the extent to which a people believe that using a particular technology will meet their need or enhance their job performance; the perceived benefits of using the technology,; availability of useful information by using the ICT application; the

extent to which the person believes that using a particular technology will be free of effort, and that the performance benefit of usage are outweighed by the effort involved. Others factors relate to the levels of ICT access, access to supporting communication infrastructure and the ICT-related skills (Ruxwana et al., 2010).

### **3.5 Machine learning application in healthcare**

A number of recent studies have shown that healthcare is one of the most promising and important areas of machine learning application. It is, however, important to mention that significant obstacles remain before machine learning is fully integrated into healthcare delivery. The barrier to integrating new products into healthcare is much higher than other industries since even small mistakes can have life-threatening consequences for patients. The techniques being applied now in research must be made more robust, a clear chain of accountability must be present, and justification for how why and how care decisions are made must be made clear. Nevertheless, some healthcare areas where machine learning applications have been successfully applied described:

- 1) Radiology. Image-based diagnostics like radiology lend themselves to the application of deep learning. There are large amounts of labelled image data to work from and a degree of uniformity that's unmatched in many other vision applications (Gibson et al., 2018; Ravi et al., 2017; Suzuki, 2017).
- 2) Diagnostics. Machine learning techniques have the potential to help clinicians make better diagnostic decisions, manage patient triage and screening programs, and identify high-level population health trends This has been demonstrated in several studies Google Brain (Google, 2017) and IBM Watson (IBM, 2017) are all very active in this area as well, among others.
- 3) Health Monitoring. Machine learning is also driving health diagnostics and monitoring into the hands of consumers. Apple built-in heart rate monitor to collect data on irregular heartbeats and alert patients who may be experiencing atrial fibrillation (Hijazi, Page, Kantarci, & Soyata, 2016; Smarsly, Dragos, & Wiggenbrock, 2016)
- 4) Personalized medicine. Personalized, or precision, medicine seeks to tailor medical interventions to the predicted response of the patient based on their genetics or other factors. Applications include selecting the best medicines for each patient and

developing custom medications that target pathways based on an individual patient's genetics.

- 5) Electronic Health Records. The major EHR vendors—including Allscripts, Athenahealth, Cerner, eClinicalWorks, and Epic—all made announcements at HiMSS about ways that they would be incorporating AI into their products. Allscripts announced a partnership with Microsoft to develop an Azure-powered EHR, while Epic unveiled a partnership with Nuance to integrate their AI-powered virtual assistant into the Epic EHR workflow. Trump Administration advisor Jared Kushner even made an appearance advocating for greater EHR interoperability as a step towards applying AI, machine learning, and big data (Adkins, 2017; Zheng et al., 2017).

Because of its good mathematical formalism, many early medical application systems extensively used Bayes' theorem for their experimental designs for dealing with uncertainty with favourable results (Shortliffe & Lindberg, 1976). The Bayesian framework also has methods of updating probabilities when new data are added. Though its application to diagnosis has been disputed on the grounds that conditions under which the theorem is meaningful are rarely met and that sources of the probabilities often are vague, several applications can readily be found in recent literature that applied the Bayesian networks in medical diagnosis (Armero et al., 2016; Ojeme & Mbogho, 2016a, 2016b; Ojeme, Mbogho, & Meyer, 2016; Sumathi & Poorna, 2015; Usher-Smith, Emery, Hamilton, Griffin, & Walter, 2015).

Sumathi (2015) described a Bayesian networks framework for detecting depression grade and suggested treatments in accordance with patient details provided to the framework. Though not implemented, the framework systematizes domain expert knowledge and observed datasets of patients and helps to map cause-effect relationship between variables. The study framework had great limitation in that it was unable to identify key factors of depression. In a study by Fleiss et al (1972), three expert systems (Bayesian classification, discriminant function classification and logical decision tree classification) which generated psychiatric diagnoses were compared with the diagnoses made by physicians. It was found that the expert systems' diagnoses were as good as those of the clinicians' on a cross-validation sample drawn from the same patient population. In our previous studies, (Ojeme & Mbogho, 2016a), we had made an effort to test the predictive strength of Bayesian networks with real clinical depression datasets and obtained a good accuracy against other machine learning algorithms. Then, in an effort to understand the interactions between depression and co-existing physical disorders, we also

developed various multidimensional machine learning models with the concept of Bayesian networks. Again, the interesting results obtained were published (Ojeme & Mbogho, 2016b). After more dataset was collected, we went beyond merely producing predictions like the previous ones by carrying out another study (Ojeme et al., 2016) revealing more reasoning ability of Bayesian networks in depression after performing dimensionality reduction to remove all redundant features using an unsupervised learning algorithm, the principal component analysis (PCA). We then used graphical knowledge-based system models to precisely quantify the importance of individual symptoms, and identified the most efficient path towards the target nodes. Other studies where Bayesian networks were tested in mental health domain include Curiac et al (Curiac, Vasile, Baniias, Volosencu, & Albu, 2009) who presented a Bayesian network-based analysis of four major psychiatric diseases: schizophrenia (simple and paranoid), mixed dementia (Alzheimer disease included), depressive disorder and maniac depressive psychosis. Chang (2014) also developed a prototype that utilizes ontologies and Bayesian network systems for inferring the possibility of depression.

In Chattopadhyay (2017), an application of neuro-fuzzy system was described in a study which applied Mamdani's technique mapped onto a feed forward back propagation neural network (BPNN) for the differential diagnosis of depression. The study modelled the manual process of clinical diagnosis of depression, obtaining a 95.5% classification on a 302 real-world adult depression cases and 50 controls. Though the results obtained do not directly compare with the model object of the present study (whose aim is the screening of depression by non-clinicians, they do prove the potential of machine learning techniques in medical domain, showing also that no prior knowledge is required to be included in the model in order to construct a good prediction tool. In another similar study, Faust et al (Faust, Ang, Puthankattil, & Joseph, 2014) presented a hybrid model consisting of probabilistic neural network (PNN), support vector machine (SVM), Decision tree (DT), K-nearest neighbor (K-NN), Naïve Bayes classifier (NBC), Gaussian mixture model (GMM), and fuzzy-sugeno classifier (FSC) for the diagnosis of depression based on the electroencephalography (EEG) signal. When tested with patient's data, the results showed an accuracy of 99.5%, with the probabilistic neural network classifier performing better than the other classifiers in discriminating between normal and depression EEG signals. In a similar study, (Hosseini-fard, Moradi, & Rostami, 2013) described a nonlinear analysis of electroencephalography (EEG) signal for discriminating 45 depression patients and 45 normal controls. For discriminating the two groups, k-nearest neighbor, linear discriminant analysis and logistic regression as the classifiers are then used. A classification

accuracy of 83.3% was obtained by correlation dimension and logistic regression classifier among other nonlinear features. The classification accuracy increased to 90% when all nonlinear features were combined and applied to the logistic regression. Again, the different number and type of collected variables prevent us from a comparison with the results of the present study (where data of previously diagnosed cases were collected), however, this application is an evidence that machine learning model-based systems can be efficient tools for medical healthcare clinicians in the diagnosis of depression. Recently, Sato et al (2015) applied a novel machine learning technique for the analysis of functional magnetic resonance imaging (fMRI) to predict individual vulnerability to major depression. The classification algorithm was able to distinguish remitted major depression from control participants with 78.3% accuracy. Several other past studies have also explored predictive models for diagnosis and treatment response of depression. A complete survey of these methods and a comparison between some of the most commonly used algorithms can be found in Patel, Khalaf, & Aizenstein (2016b).

As noted by several studies, Oteniya (2008), Seixas et al.(2014), the main advantage of Bayesian network classifier over other classification methods is the opportunity of considering the prior information about a given problem. However, the main disadvantages of Bayesian classifier are (1) the numerical attributes require discretization in most cases; (2) it is not suitable for large data sets which contain many attributes (time and space issues). Several other previous studies have also explored predictive models for diagnosis and treatment response of depression.

### **3.6 Usability of medical applications**

Usability of an artefact is directly linked with the context of users' task, the environment of use and the artefact. According to Nielsen satisfaction Nielsen (2010), usability is defined in terms of the following five usability attributes: learnability, efficiency, memorability, errors and Usability is therefore not a one---dimensional property of an artefact. Factors such as how well users learn and use an application, how often an error occurs while using the system, how well a user remembers and recognises features based on previous experience with the system, and overall user satisfaction are all associated with usability. These factors are usually measured during usability testing. The usability of a product can be assessed using usability evaluation or usability testing or a combination of these two methods (Usability.gov, 2012). Usability in medical devices is of particular importance because the slightest usability issue can have a

negative impact on patient care (HIMSS, 2009). In some cases, can lead to fatal accidents (Braun, 2005). Usability testing early in the development process can identify potential problems with an artefact.

### **3.7 Effectiveness of usability evaluation methods in medical applications**

Although the different usability evaluation methods (heuristic evaluation, think-aloud, cognitive walkthrough, cognitive talk analysis, and interviews) vary with regard to the type of users (real users or opinion of experts), the number and type of problems identified by them, time constraints, efficiency and the cost of using these methods (Jake-Schoffman et al., 2017; Usability.gov, 2012), their general aim, which is to identify usability problems that prevent users from interacting easily with an interface, are similar (Bhutkar, Konkani, Katre, & Ray, 2013). A number of work has been carried in the HCI literatures concerning several usability evaluation methods for different applications, but they provide little knowledge about usability studies on medical applications. Consequently, in this section, we provide a synthesis of relevant and appropriate usability evaluation methods used and the empirical studies available in medical applications and in related areas. The is to help identify the boundary, weakness, and strength of current research.

The empirical findings of these comparative methodological studies indicated which methods were more effective in detecting usability problems with regard to several criteria: the number of usability problems, the type of usability problems, and the cost of using each method. However, early discovery of usability problems is critical in medical applications because usability problems can induce inefficiency and lead to errors. The intricacy in workflow in clinics and hospitals and the attitude of users can sometimes make the use of computer technology very complex so, the challenge for developers is to make these complex systems usable.

In Hosky et al (2012), the importance of using different usability evaluations for system design process was highlighted in a study to compare four evaluation methods used during the development of outpatient clinical documentation software. These methods include: clinician email response, online survey, observations and interviews. The findings from the study showed that no single method identifies all or most usability problems. Instead, each method proved to be good for evaluations at a different stage of design and characterizes different usability aspect. Email responses elicited from clinicians and surveys report mostly technical, biomedical, terminology and control problems and are most effective when a working

prototype has been completed. Observations of clinical work and interviews inform conceptual and workflow-related problems and are best performed early in the cycle. Appropriate use of these methods consistently during development may significantly improve system usability and contribute to higher adoption rates among clinicians and to improved healthcare quality.

Every device that is used in high-risk applications, such as medical diagnosis and treatment, has to aid the users in preventing errors. For a depression screening, for instance, this means that the tool should provide user guidance and assistance for all tasks of daily work. As Schnall et al (2016) noted, early detection of usability issues is critical in medical applications because usability issues can induce inefficiency and lead to errors. However, medical applications are fraught with errors. Various methods of usability tests exist, each providing diverse insight into usability issues. Usability evaluation and usability testing are two ways of assessing usability.(Usability.gov, 2012). Usually, usability experts perform usability evaluations and conduct usability tests with end-users of a product or system. Usability evaluations and usability tests can be used to identify potential usability issues with a product. Most often usability evaluation reveals issues that can be quantified with a follow-up testing session. In some scenarios, usability specialists may not be qualified to judge if required information is present in the product. In such cases, usability experts and subject matter experts collaborate to conduct usability evaluations (Schnall, Cho, & Liu, 2018).

According to Lacerda et al (2015), usability tests during the design process help to reduce errors in clinical decision support systems intended for use in clinical practice. Through a systematic literature review, the study investigated which usability evaluation methods have been used more frequently in a variety of medical applications. The goal of the study was to provide a preliminary discussion about how usability evaluation methods have been applied in the healthcare for smartphones. The findings from the study show that heuristic usability evaluations of mobile health applications were reported for both patients and healthcare professionals use. In the evaluations reported for patient's use, two were related to the self-management of diseases such as diabetes (Isaković, Sedlar, Volk, & Bešter, 2016) and thrombosis (Martínez-Pérez, De La Torre-Díez, & López-Coronado, 2013) one was developed for the control of the user's caloric intake (Chomutare, Fernandez-Luque, Arsand, & Hartvigsen, 2011), and the other one is a reference guide for everyone to understand blood test reports (Harvard Health publications, 2014).

In Narasimha et al (2016), an investigation on the usability issues associated with geriatric patients using home-based video telemedicine systems was carried out. The four home-based

video telemedicine systems chosen for the study are: (1) Doxy.me, (2) Polycom, (3) Vidyo and (4) VSee. In a between-subjects experimental set-up, 20 participants were assigned randomly to one of the four conditions. The participants were asked to complete a demographic questionnaire and the representative tasks on the telemedicine platform. This was followed by a think-aloud session at the end of which, the participants completed a NASATLX workload survey, an IBM Computer System Usability Questionnaire (IBM-CSUQ), and a post-test subjective questionnaire. Difficulties faced by the participants include downloading application plug-ins, locating icons and the size of the icons.

In the medical literature, only a few studies have been carried out on quality control or regulations to ensure that medical applications are user-friendly, accurate in content, evidence-based, or efficacious. The work of, Boudreaux et al (2014) described seven approaches for evaluating and selecting medical applications. These include (1) Review of the scientific literature, (2) Searching application clearinghouse websites, (3) Searching application stores, (4) Review of application descriptions, user ratings, and reviews, (5) carry out a social media query within professional and, if available, patient networks, (6) Pilot the application, and (7) Elicit feedback from patients. The study concludes that because of the enormous range of quality among medical application, strategies for evaluating them becomes necessary for adoption to occur so that it aligns with core values in healthcare.

In Reolon et al (2016) was presented the results of a systematic review of literature to analyse the existence of specific usability heuristics for m-health applications. The results showed that only two studies reported the usage of heuristics customized to mobile applications, but not further taking into consideration specific characteristics of healthcare applications. The study noted the clear lack of studies on the customization of usability heuristics for m-health applications, which has the potential to contribute to the adoption of lightweight evaluation techniques and, thus, contributing to the improvement of the usability of m-health applications.

### **3.8 Machine learning and HCI collaboration**

The complementary effect of the fusion between the techniques of machine learning and HCI has been well discussed in the literature (Amershi, Cakmak, Knox, Kulesza, & Lau, 2013; Heuer, 2013; Lieberman, 2009; Marco et al., 2016; Moustakis & Herrmann, 1997). As noted by Amershi et al ((2013), a growing research community at the relationship between machine learning and human-computer interaction are making interaction with humans an essential part

of building machine learning systems. These efforts include employing interaction design principles to machine learning classifiers, using human-participant testing to evaluate machine learning systems and inspire new methods, and changing the input and output channels of machine learning systems to better leverage human capabilities. This complementary efforts was also corroborated by the work of Ilbeygi, and Shah-Hosseini (2012), which presented the proposal for the collaboration of HCI and Fuzzy Inference System (FIS) for developing intelligent systems for emotion recognition from facial expression recognition. The study further described new algorithms for facial feature extraction that demonstrated acceptable performance and precision (93.96% for Emotion Recognition of six basic emotions). Heuer (Heuer, 2013) in his work “On the intersection of human-computer interaction (HCI) and artificial intelligence”, highlighted the mutual benefits from a closer collaboration of machine learning and human computer interactions. Similarly, Lieberman (Lieberman, 2009) note that the user-centered approach and testing methods of HCI can benefit machine learning to obtain user interfaces, while machine learning applications can serve HCI’s goals of providing usability and good user experience.

Much earlier, Moustakis and Hermann (1997) presented reports on a survey of 112 professionals and academicians specializing in HCI, who were asked to state level of machine learning use in HCI study. Feedback from the survey were captured in a structured questionnaire. The results of the survey Analysis showed that about one-third of the participants had used machine learning in different HCI tasks. The study concluded by noting a gradual decline in the number of researchers who use machine learning in HCI tasks, and that the main causes are a) misperceptions about machine learning b) lack of awareness of machine learning’s potential, and c) scarcity of concrete case studies demonstrating the application of machine learning in HCI.

### **3.9 Summary of the review, gaps, limitations and opportunities**

#### **3.9.1 What this study does differently**

Previous studies in the domain of ICT intervention in the Nigerian healthcare have examined the use of ICTs by health workers, especially clinicians, so as to update their knowledge and enhance healthcare delivery. This study, on the other hand, investigates the existing ICTs used in medical application, not only from the perspectives of the clinicians but also those of the intended end-users who are benefitting from the services. Access to healthcare services in Nigeria seems to be poor, as the clinician/patient ratio (Table 1.1) is unacceptably low,

especially when compared to that in developed countries. This study will focus on the use of ICTs in disease screening, and its uniqueness is that the focus is not only on the clinicians, but also on the heterogenous users who stand to benefit from their services. At the time of writing this thesis, there was no study in the literature that applied a collaboration of machine learning and HCI techniques for disease screening in Nigeria. The present study outcomes are the possible machine learning and HCI designs for the depression screening in Nigerian universities and the possibility of extension to other populations. Though the study outcomes were interpreted based on existing ICT4D frameworks such as adoption and diffusion of innovations by Rogers (2015), the context of the present was different.

### **3.9.2 Limitations**

It is important to note that the study does not claim to have reviewed all published papers, books and technical reports on the subject topic. Several factors may have caused the non-inclusion of some important published works in the survey. The most important factor was the exclusion of non-English publications in the survey. Whilst it is true that some work on machine learning models for inferencing in medical diagnosis may have been constructed and published in other languages, this was a necessary limitation as there was no consistent method for checking whether non-English studies were relevant for the review.

Additionally, given that this review focused mainly on specific research questions, it is important to note that it could not have covered the entire research in the field of ICTs and mental health. This is because, beside the search keyword index used in the study failing to find some important published works in other areas of ICTs and mental health, it is also unable to find published works in ICTs and mental health in the databases that were inaccessible to the researcher.

Lastly, beside ensuring that only peer-reviewed conference and journal papers, books and technical reports were included in the review, no other form of quality checks were conducted on the review publications.

Given the limitations of the reviewed studies, it is important therefore not to interpret the findings and selection of studies in this review as being representative of the totality of research effort concerning adopting ICT4D framework to support the methodology for depression screening in Nigeria.

### 3.9.3 Opportunities

The literature outlined in the preceding sections and others like them revealed some opportunities and highlighted some important gaps that motivated this study. The opportunities are summarized below:

1) there is need to create a local computing solution that the people understand to complement existing facilities in reducing the risk of depression in Nigeria University community. In any resource-constrained environments, it is important that all possible areas of solutions be investigated so as to select the one that works for a particular society. The needed solution in Nigeria, with shortage in mental health professionals and diagnostic facilities should be channelled towards rapid expansion of ICT-aided healthcare delivery, including health information technologies, especially screening plans. The goal is to create a screening tool by combining machine learning and HCI techniques to increase flexibility in depression prevention and detection processes

2) use of computers for healthcare delivery is underexplored in developing countries, especially with respect to creating diagnostic-assisted tools

3) Though the results from the reviewed papers are not directly comparable with the model object of this present study (whose aim is the development of computing solutions that enhances regular assessment of University students for depression by non-clinicians), they do demonstrate the potential of machine learning model-based systems and HCI as efficient tools for clinicians in the diagnosis of depression

4) though the studies reviewed do not have real primary depression datasets, do not evaluate the models with metrics described in this study, and the interfaces were built for use by only experts clinicians, the present study builds on them, using them as baselines in the experimental analysis. The methods were chosen because, within computational biology, industrial psychology and public health, they have been applied to an increasing variety of problems, including various types of disease detection, and their performance could be evaluated with the datasets used in the present study.

These gaps and opportunities implied that there was a need to provide a workable depression screening solution on personal computers and laptops installed in all departments at the University of Benin, specifically designed for non-clinicians and designed based on users' needs. Therefore, the next logical question was, which machine learning techniques and which HCI methods would support non-clinicians and non-technical users? Once such machine learning techniques were designed, it was deemed important to integrate their effect on easy-

to-use screening interfaces on desktop computers and laptops. Consequently, the following research questions were posed: What is the combined framework that integrates ICT4D tools and novice users together? How can the linkage between machine learning and HCI tools be strengthened to support the framework for screening for depression in Nigeria in the context of ICT4D project? What machine learning techniques would be appropriate to screen for depression among University staff and students? This study was conducted to address these and a few other research questions. The related work provided some opportunities that could be explored when designing machine learning techniques on for depression screening. These are summarized below. 1) When designing machine learning techniques on a desktop computer, consider: a) obtain dataset; 2) select data subset; 3) pre-process data; 4) transform data; 5) extract useful patterns or structure in data with several algorithms; 6) evaluate performance; 7) select the best performing algorithms integrate knowledge on interface. Chapters four and five discuss the design and implementation of machine learning techniques and screening interfaces and indicates how the identified gaps and opportunities were integrated within the design process.

## Chapter Four

### Study methodology

#### 4.1 Introduction

This chapter deals with the methodology that was adopted in conducting the study, describing and justifying (where necessary), the methods and procedures used. These include sections on the research design that was used to carry out the study, the methods of data collection and the data collection instruments. The chapter discusses validity and reliability issues applied to the study dataset, ethical considerations and methods for data analysis. Finally, the chapter discusses problems encountered during data collection and presents an evaluation of the methodology used to conduct the study.

#### 4.2 Research design

Research design, according to Creswell (2014), is a plan of how a researcher intends to carry out a study. It is the general plan for the collection, measurement and analysis of data, with the central goal of answering the research questions and achieving the research objectives. This includes the summary of what the researcher will do, from writing the research questions, objectives and its operational implications, to the analysis of data (Morse, 2016). According to Creswell (2014), a good research design must provide maximum information and offer an opportunity for considering many different aspects of the problem. Chilisa and Kawulich (2015) noted that the nature and the context of the study is a determinant of the research method employed since a good research method for a certain study might be inappropriate for another study. Positivist (called scientific), Interpretive (known as anti-positivist), and the mixed methods are the three major research paradigms that have been identified in the Western tradition of science (Creswell, 2014). In terms of research methods some commonly used ICT4D research methods include questionnaire, surveys, ethnography, in-depth interviews and focus groups (Burrell & Toyama, 2009; Kleine, 2015; Loh, 2015). Questionnaire and survey instruments are quite common in intentional ICT4D projects, where they can include users and other local residents.

The positivism paradigm holds that the scientific method is the only way to establish truth and objective reality. It assumes a quantitative methodology; based on previous observation or previous history and reasoning as a tool of understanding a certain problem or behaviour. It

holds the view that the methods, techniques and procedures used in the natural sciences produce the best framework for investigating the social world. Positivist paradigms, which believes in the use of survey methods and questionnaires for data collection, are concerned with what has caused a particular relationship and what the effects of this relationship are, they also prefer quantitative data which can be transformed into numbers and statistics (Creswell, 2014).

The commonly used ICT4D for data collection in Interpretive or qualitative research method are ethnography, participant observation, grounded theory, documentary methods, field works and the use of an unstructured interview for data collection (Creswell, 2014; Kleine, 2015; Loh, 2015; Ramlo, 2016), concentrate on subjectivist approach to studying social phenomena which have importance to a range of research techniques. Anti-positivism researchers criticize positivists because they believe that statistics and numbers are not useful about human behaviour. Assumptions on the multiplicity of realities also inform the research process. For instance, the research questions may not be identified before the research begins but rather may evolve as the study advances (Mertens, 2008). The research questions are usually open-ended, descriptive and non-directional (Creswell, 2013).

Given the nature of strengths and limitations of both the positivist and Interpretive research approaches highlighted in the several studies (Chilisa & Kawulich, 2015; Creswell, 2014; Ramlo, 2016), the 'mixed method research approach' was considered (McKim, 2017) as a third possible research approach. The philosophy behind the mixed method is that though each research paradigm has its own approaches and methods, the researcher may adopt research methods cutting across research methods to solve the problem or to answer research questions (Chilisa & Kawulich, 2015).

The ontological and epistemological positions of the three research methods in turn influence the choice of the methodology and the data collection methods employed in a research project

#### **4.2.1 Justifying the mixed research method chosen for the study**

This study focuses on adopting ICT4D frameworks to support depression screening methodology. During the study, several types of qualitative and quantitative data were collected. The qualitative data were gathered using self-administered questionnaire and interviews from clinicians, University staff and students. The questionnaire examined such issues as knowledge and use of ICT tools in work place and attitudes towards the adoption of ICT tools in routine medical practice. Quantitative data were recorded using two methods: the

self-administered questionnaire and artefacts. The questionnaire invited participants to describe themselves in terms of age, gender, and number and severity of heuristic violations. These artefacts included user interface design features in chapter five.

Given that the diverse nature of the study research questions requires the use of more than one of the research paradigms and their methods and that the data needed to be explored to generate new knowledge are in multiphase, the mixed research method of inquiry, which uses both qualitative and quantitative research practices, was considered to be the most appropriate approach to conduct the study. A mixed research method design, according to Creswell (2014), includes the combination of various quantitative or qualitative approaches in the same study. It involves the collection, analysis and integration of quantitative and qualitative data in a single or multiphase study (Ramlo, 2016). Although certain methodologies tend to be associated with a given research paradigm, Morse (Morse, 2016) suggests that the scope, objectives, and nature of study are consistent across research methods and across research traditions. Furthermore, the steps applied in the present study are consistent with the works of Morse (Morse, 2016) Ramlo (2016) who noted that the steps of a mixed methods, which are similar to those in traditional research methods include: identifying a research problem and research objective; designing the purpose of the research and research questions(s); choosing a research method; collecting data; analysing data; interpreting/validating data; and communicating results or findings. According to Bergman (2008), some of the advantages that come with employing a mixed method in a study with different data collection approaches include:

- It provides *corroboration*: Combining qualitative and quantitative research corroborates research findings, mutually, thus providing greater validity.
- It can *offset* disadvantages: A study can take advantage of the strengths in qualitative and quantitative research by offsetting any of the disadvantages found in either of the two.
- It is *comprehensive*: it offers the researcher the ability to provide a more thorough account of the field of investigation by using both qualitative and quantitative methods.
- It allows for *instrument development*: clearer and more structured scale items can be devised from a qualitative probe of the inquiry area.

### 4.3 Existing depression diagnosis procedure in Nigeria

In healthcare system, generally, both clinicians and patients rely on an accurate diagnostic process to detect the correct illness and draw out a treatment plan. In Nigeria, with shortage of mental health services (Jack-Ide et al., 2012; World Health Organization, 2011b), General-Practice clinicians are the first-line contact in the detection and treatment of depression at the primary care (Gureje et al., 2015; Obadeji et al., 2015). Here, traditional depression diagnostic practice (see Figure 4.1) typically involves a series of thinking and analyses and clinician-to-patient interview (sometimes through a questionnaire) where judgments are made from the patient's demographic data (name, age, sex, marital status, religion, nationality, educational level and status), appearance and behaviour, subjective self-reported symptoms, source of referral (self-referral, parent or guardian, police or relatives), past psychiatric/depression history, past medical history and current life circumstances (Baasher, Carstairs, Giel, & Hassler, 1975; Chang et al., 2014). The views of relatives or other third parties may be taken into account. A physical examination to check for ill health, the effects of medications or other drugs may also be conducted. This is followed by the evaluation of the relative importance of those symptoms and their categorisation according to their significance. Next is the differential classification where the clinician rules out some other diseases by excluding one disease after another in order to arrive at a definite depression diagnosis.

Though still in use till this day, several studies in the metal health literature (Ahmed & Bhugra, 2007; Ayonrinde, Gureje, & Lawal, 2004; Gureje et al., 2015; Obadeji et al., 2015) have expressed concern over the inability of this manual model address the challenges of managing depression in Nigeria and other developing countries. In particular, the studies highlighted the following limitations:

- a) The model is static and does not address or resolve the most challenging problem of shortage of mental health professionals and limited diagnostic facilities to cate for the ever-increasing number of depression sufferers
- b) The model is slow, time-consuming and often lead to long waiting queues before patients get medical attention.
- c) The model lacks the ability to transform collected data into actionable knowledge. Most often, the data collected end up in the dirty shelves in different hospital departments.

As its central goal, this study attempts to address these challenges by adopting ICT4D frameworks to support the methodology for depression screening in the Nigerian University

community. As earlier mentioned, the specific tools employed are techniques from machine learning and HCI (Figure 4.1).

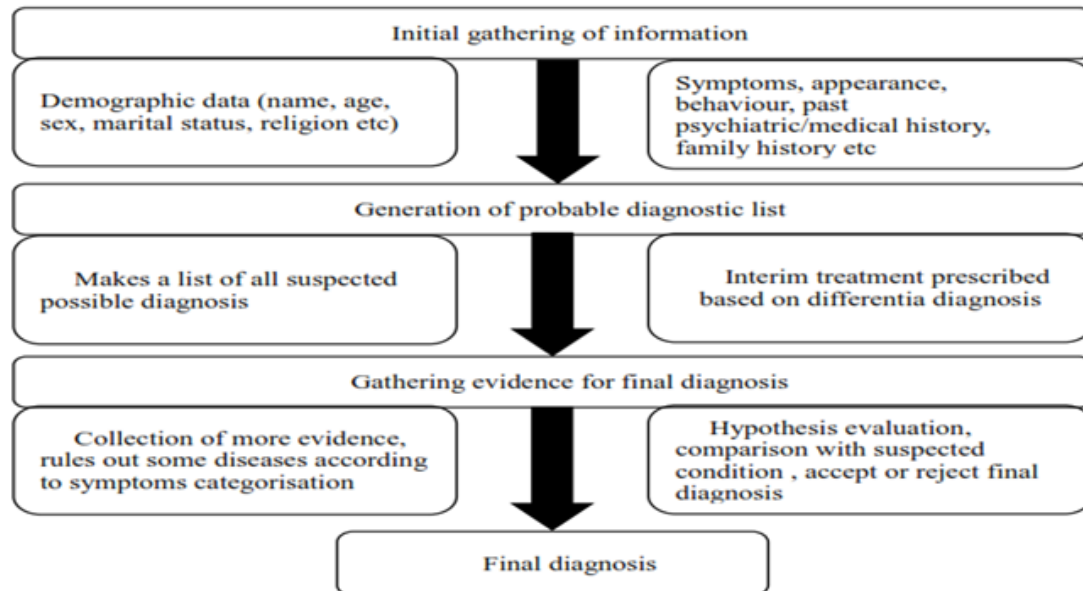


Figure 4.1 Traditional diagnostic process

#### 4.4 Design of a new machine learning-HCI screening system

This section shows the steps undertaken to develop the University students' depression screening application with focus on machine learning with Bayesian networks and HCI

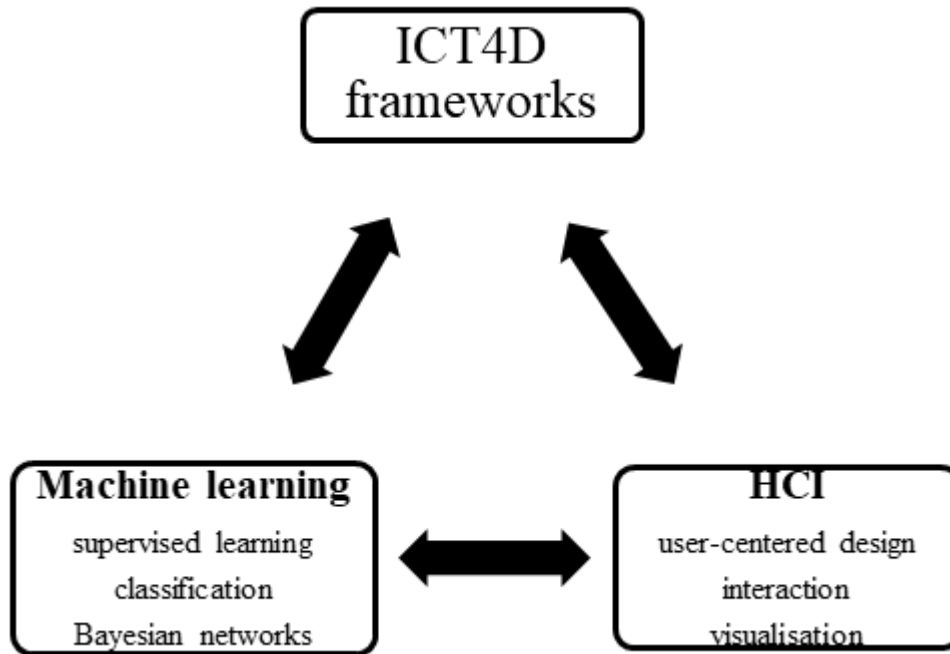


Figure 4.2 Framework of the study

Figure 4.3 shows a more detailed research method overview, from data collection to implementation of the model. Data gathering process, which is first step in the processes is discussed next.

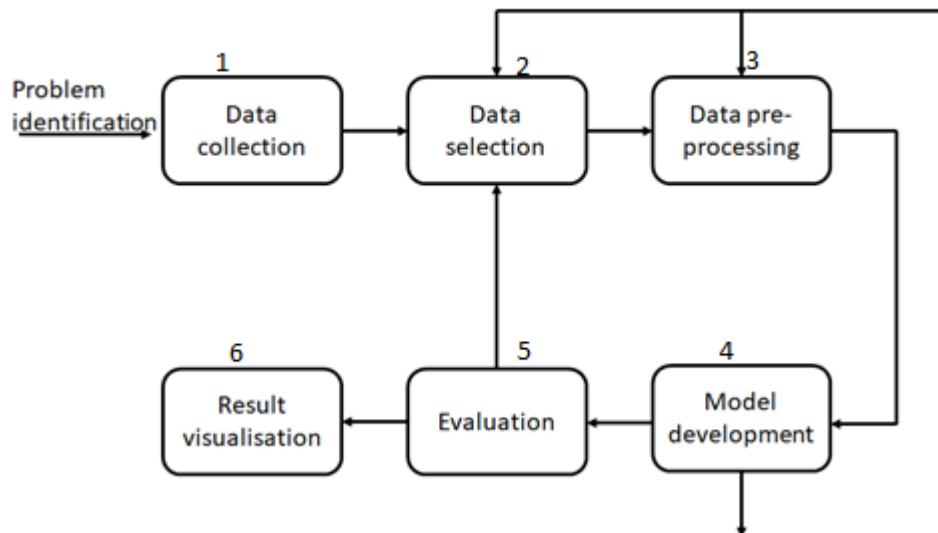


Figure 4.3 Research method overview

## 4.5 Data collection and description of dataset

Besides the data collection methods used in ICT4D research, another factor that influenced the choice of data collection methods for the study was the research questions. For instance, to address the first research question: Are there ICT4D tools that can help address some of the disease screening challenges faced by mental health professionals in Nigeria? This research question was answered through the detailed design process (Section 4.3). The problem identification, requirements gathering, design, evaluation and results outcome of the designed prototypes indicated the ICT4D applications that could be helpful in addressing the challenges of depression screening in Nigeria.

To address the fourth research question: what machine learning techniques would be appropriate to screen for depression among University staff and students? the predictive strength of various machine learning techniques was required. This was described and tested, and the results obtained was published (Ojeme & Mbogho, 2016a).

There are six orderly steps in the research method overview in Figure 4.3. Steps 1-3 are discussed in sections 4.5.1, 4.5.2 and 4.5.3 while steps 4, 5 and 6 are discussed in sections 5.3, 5.4, and 5.5.

### 4.5.1 Data collection

Data plays an important role in an experimental study such as this. To achieve this role however, data must be of high quality. High quality data are defined by the Completeness, accuracy, and consistent of data elements. The quality of any large real-world data set depends on several factors, among which the source of and the process of data collection are often the crucial factors (Cai & Zhu, 2015). It requires careful planning and preparation for fieldwork. In healthcare, for instance, diagnostic errors account for more than 8% of adverse occurrence in medical diagnosis and up to 30% of malpractice claims (Nendaz & Perrier, 2012). Mechanisms of errors may be related to the working environment but cognitive issues are involved in about 75% of the cases, either alone or in association with system failures. Majority of diagnostic errors are not related to deficiency of knowledge but to data collection errors, data integration, and data validation and verification that may lead to premature identification closure (Nendaz & Perrier, 2012). To this end, a substantial amount of attention was given to the data preparation, which typically involves data acquisition methodology, data formatting and cleaning (strings, numbers, delimiters), and data storage and handling, normalization (acronyms, typos), filling in/removing missing values, de-duplication, and merging multiple

dataset/features/sources (Gareth, Witten, Hastie, & Tibshirani, 2013). Data acquisition includes methodological effort such as knowledge elicitation method from clinicians, method of harmonising opinions from the experts in case of disagreements, and incorporating a computer-based decision-making such that would be accepted in their daily diagnostic practice. As discussed in section 1.2.1, prior to embarking on the study, the researcher conducted a pilot data collection survey with two study population (Table 1.2).

Following the methodologies for data collection in ICT4D research, the researcher employed the use of workshops (in the form of seminars), semi-structured interviews and record extraction from previously diagnosed cases for the collection of data. This was necessary to establish good relations with the domain experts (since they will have to set aside considerable time to provide me with information) and avoid the bias of knowledge elicitation from human experts, as highlighted by Oteniya (2008). Since the study did not involve any direct interaction with patients, but only used anonymized records of previously diagnosed patients, it was unlikely that the use of the data for this study would increase the chance of causing any kind of harm to the patients whose records were used to develop the datasets, and to the physicians who administered diagnoses. That is, to maintain confidentiality of patients and physicians, personal details of patients and physicians were removed from the records before data was extracted. Thus, data collection was carried out using the instruments described in section 4.5.1.1.

#### **4.5.1.1 Abstraction form**

Abstraction form is a standard instrument used to systematically collect data from scientific reports. The purpose of the abstraction form was to guide the collection of data from the clinical records of patients who were previously diagnosed with varying degrees of depression. For each item on the abstraction form, the section of the medical record was specified. For instances, both patients-reported symptoms and psychiatrists' observations were sources of reported symptoms. The abstraction is divided into 12 sections, shown in Table 4.1

Table 4.1 Abstraction form

S/N	Items
1	Inclusion and exclusion criteria
2	Patient demographics
3	Medical history
4	Physician care
5	Activities of daily living
6	Reported symptoms
7	Medications
8	Special tests
9	Laboratory tests
11	Admission/discharge status
12	Completeness checks

#### 4.5.1.2 Use of data collectors

A data collection team made up of two (2) psychiatric nurses and of two (2) medical records personnel experienced with medical data were selected by the psychiatrists to help in the record extraction process. They were screened for abstraction skills using test clinical records. After demonstrating adequate skills, four (4) data collectors were invited to participate in the training sessions on the use of abstraction form. The training, which lasted for two days, covered the specifics of accessing study sample medical records, abstracting data from medical records of patients with depression, and monitoring data collection. The data collection was supervised by the psychiatrists. Expectedly, inter-observer discrepancies among the data collectors was almost a threat to the process but this was quickly addressed by the combined team of psychiatrists and psychologists. The researcher worked with the data collection team throughout the data collection period. The whole process was to enhance the quality of research data as their unbiased evaluation of medical diagnosis has a great impact on treatment planning.

It is important to state here that since the data was not recorded in a form which leaves it for research purpose, much time was spent in the extraction and conversion process. This was done to ensure that only records that are relevant for answering our research questions are extracted. The dataset (see Figure 4.4, Tables 4.3, 4.4) consisted of anonymised records of 1798 data instances, (1020 male and 778 female cases from 12 to 92 years old with a mean age of 42.55 and standard deviation of 13.92). It originally had 27 main attributes and one class attribute (depression). The dataset attributes, their codes, values (level of presence/absence) and data types are shown in Table 4.4. The class attributes consist of four grades (No depression (0), mild depression (1), moderate depression (2) and severe depression (3). The datasets were

collected from the University of Benin Teaching Hospital (UBTH) and one primary care centre in Nigeria. All depression cases selected fulfilled clinical criteria for depression as defined in DSM-5 (*Diagnostic and Statistical Manual of Mental Disorders fifth edition*, n.d.) and ICD-10 (World Health Organization, 1992).

Table 4.2 Original dataset and attribute values

S/N	Attributes	Code	Values		Data type
			Absent	Present	
1	Sad mood	SM	0	1	Integer
2	Thought of suicide	SU	0	1	Integer
3	Loss of pleasure	LP	0	1	Integer
4	insomnia	IN	0	1	Integer
5	Hypersomnia	HY	0	1	Integer
6	Loss of appetite	LA	0	1	Integer
7	Psychomotor agitation	PA	0	1	Integer
8	Psychomotor retardation	PR	0	1	Integer
9	Loss of energy	LE	0	1	Integer
10	Feeling of worthlessness	FW	0	1	Integer
11	Lack of thinking	LT	0	1	Integer
12	Indecisiveness	ID	0	1	Integer
13	Recurrent thought of death	TD	0	1	Integer
14	Impaired function	IF	0	1	Integer
15	Weight gain	WG	0	1	Integer
16	Weight loss	WL	0	1	Integer
17	Employment status	ES	0	1	Integer
18	Depression in family	DF	0	1	Integer
19	Stressful life events	SL	0	1	Integer
20	Financial pressure	FP	0	1	Integer
21	Level of family support	LS	None: 0; low: 1; medium: 2, High:3		Integer
22	Availability of Accommodation	AA	None: 0; low: 1; medium: 2, High:3		Integer
23	Alcohol consumption	AC	None: 0; low: 1; medium: 2, High:3		Integer
24	Other drug consumption	DC	None: 0; low: 1; medium: 2, High:3		Integer
25	Job satisfaction	JS	None: 0; low: 1; medium: 2, High:3		Integer
26	Academic performance	AP	None: 0; low: 1; medium: 2, High:3		Integer
27	Cigarette smoking	CS	None: 0; low: 1; medium: 2, High:3		Integer
	<b>Depression diagnosis</b>	<b>DG</b>	<b>None: 0; mild: 1; Moderate: 2, severe: 3</b>		<b>Nominal</b>

### 4.5.1.3 Data validation

Medical records contain information given by the patient to the physician. Most times, this information is either incomplete or not standardized. For example, the patient may incorrectly recall information from his/her earlier medical history, or may report incomplete or inaccurate symptoms. Also, the clinician may take an incomplete history or record information incorrectly. Again, there might be some missing information in older records, without the benefit of subsequent advances in medical knowledge. These errors occurring at the patient and clinician levels are difficult to avoid in practice and neither the patient nor the clinicians are involved in the real data collection process for research purposes. As a result of this potential threat, further validation of the collected data was needed. This was done by a group of eight mental health professionals, made up of three Psychiatrists, one Child/Adolescent Psychiatrist, two Clinical Psychologists, one Clinical Social Worker and one Nurse Psychotherapist. A control group (No depression) of 372 cases were also collected. The two datasets were merged by the researcher and mental health professionals, who also manually matched their attributes to achieve compatibility. In a few cases where the attributes did not match, the researcher and the mental health professionals consulted the team at the primary health care centre for clarifications. A summary of this data collection phase is displayed in Figure 4.4



Figure 4.4 Flow of information for data collected from patient to researcher

### 4.5.1.4 Ethics

Ethical approval for the study was obtained from the Faculty of Science Research Ethics Committee, University of Cape Town on March 10, 2015 and the Ethics and Research Committee of the University of Benin Teaching Hospital on August 17 2015.

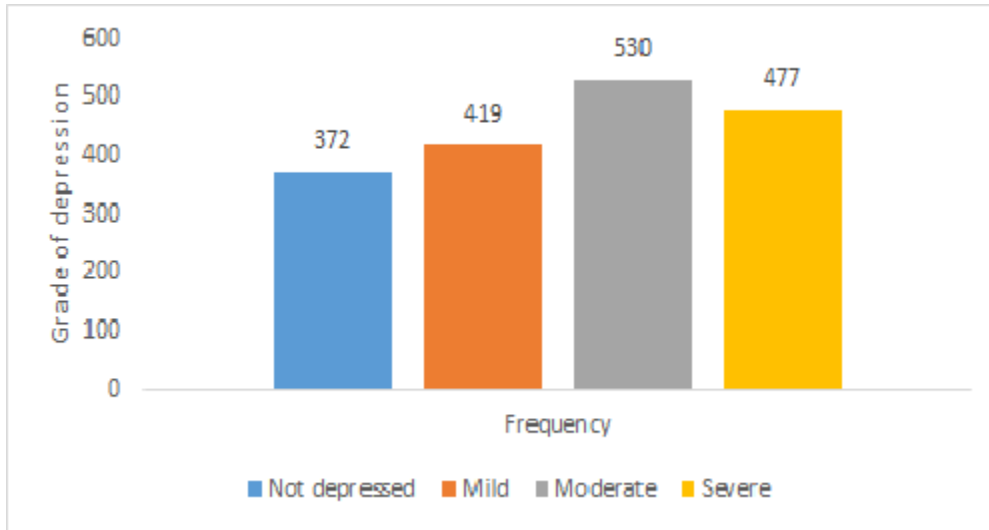


Figure 4.5 Dataset descriptive statistics

The distribution of different depression grades among patients' age-bracket were analysed using descriptive statistics.

Table 4.3 Descriptive statistics of age bracket and depression grades from dataset

Age Interval (years)	Severe depression	Moderate depression	Mild depression	No depression	Total number of instances
12 - 30	283	322	249	158	1012
31 - 60	101	99	132	147	479
61 - 92	93	109	38	67	307
<b>Total</b>	<b>477</b>	<b>530</b>	<b>419</b>	<b>372</b>	<b>1798</b>

**Basic assumptions:** The following three assumptions were made about our study data:

1. In order to be able to optimally find correlations between features and classes, the features in our data are assumed to be conditionally independent given the classes.
2. There are no cases in the dataset for which there are variables for which a no-value is specified.
3. No ordering exists on the cases in the database and all cases are independent of each other given the target

### 4.5.2 Data selection

Selecting data relates to the merging of related features or analysis task from the collected dataset. After careful considerations by the combined team of eight mental health professionals, some features in the dataset were merged. For instance, job satisfaction and academic performance were merged, availability of accommodation and level of family support were merged, and alcohol and other drug consumption were also merged. Consequently, the predictor variables (serial number 1-24 in Table 4.4) and the diagnosis (marked in boldface) were selected.

Table 4.4 Selected attribute set

S/N	Attributes	Code	Values		Data type
			Absent	Present	
1	Sad mood	SM	0	1	Integer
2	Thought of suicide	SU	0	1	Integer
3	Loss of pleasure	LP	0	1	Integer
4	insomnia	IN	0	1	Integer
5	Hypersomnia	HY	0	1	Integer
6	Loss of appetite	LA	0	1	Integer
7	Psychomotor agitation	PA	0	1	Integer
8	Psychomotor retardation	PR	0	1	Integer
9	Loss of energy	LE	0	1	Integer
10	Feeling of worthlessness	FW	0	1	Integer
11	Lack of thinking	LT	0	1	Integer
12	Indecisiveness	ID	0	1	Integer
13	Recurrent thought of death	TD	0	1	Integer
14	Impaired function	IF	0	1	Integer
15	Weight gain	WG	0	1	Integer
16	Weight loss	WL	0	1	Integer
17	Employment status	ES	0	1	Integer
18	Depression in family	DF	0	1	Integer
19	Stressful life events	SL	0	1	Integer
20	Financial pressure	FP	0	1	Integer
21	Level of family support availability of accommodation	LS	None: 0; low: 1; medium: 2, High: 3		Integer
22	Alcohol or drug consumption	AC	None: 0; low: 1; medium: 2, High: 3		Integer
23	Job satisfaction/Academic performance	JS	None: 0; low: 1; medium: 2, High: 3		Integer
24	Cigarette smoking	CS	None: 0; low: 1; medium: 2, High: 3		Integer
	<b>Depression diagnosis</b>	<b>DG</b>	<b>None: 0; mild: 1; Moderate: 2, severe: 3</b>		<b>Nominal</b>

### 4.5.3 Data pre-processing

Data collection phase may produce dataset containing missing values, inaccurate, inconsistent and incomplete data. Inaccurate data containing incorrect attribute values may sometime be due to data entry errors, faulty data collection tools, errors in data transmission, and data collectors may enter these incorrect values just to fill mandatory fields during data collection (Han & Kamber, 2011). Incomplete data can occur for various reasons. For example, some attributes values were not important during data entry and some attributes values were

not always available. Inconsistency occurs when there is a record that is in conflict with other records on the dataset.

For the present study, one of the data pre-processing tasks carried out was the conversion of the dataset, both on a structural level (transposition to row and columns) as well as an instance level (value conversion and attribute normalisation). Since there were no missing values in the study dataset, other data pre-processing tasks carried out were, datatypes conversions, conversion of data to tables, and merging of features. As shown in the resulting dataset from the pre-processing in Table 4.5, the training cases in the dataset were in Boolean format, based on presence (1) or absent (0) of symptoms (for symptoms with serial number 1 – 23), and the degree of presence (symptoms with serial number 21 – 24).. These pre-processing activities were necessary in order to satisfy data quality elements and meet the requirements of the machine learning analysis tools, as suggested by Murphy (2012) and Cai and Zhu (2015)

#### **4.5.4 Data transformation**

Part of the objectives of this study was to compare the performance of supervised learning algorithms based on the selected symptoms of depression. One big challenge in applied machine learning, however, is choosing features that can express the properties of the data relevant for the task at hand. There are four consequences of a dataset with large feature space that resulted from irrelevant features (Bro & Smilde, 2014; Silipo, Adae, Hart, & Berthold, 2014). One is that dealing with large feature spaces can easily become prohibitive due to time and memory constraints. Second, there is a high chance of many features being irrelevant, and thus, introduce extra noise to the dataset. This could degrade prediction's accuracy. Third, the amount of data required to generate a prediction function that expresses the properties of the data typically grows with the dimensionality of the feature space, and therefore, a feature space with fewer dimensions is desirable. Lastly, high data dimensionality makes it difficult for human analysts and machine learning algorithms to visualize and reveal hidden structures and patterns in data, which is not readily evident, for accurate predictions (Maaten & Hinton, 2008). By ignoring features that contribute nothing to the efficacy and efficiency of the learning algorithm, the dimensionality of the dataset and computational complexity are reduced and at the same time and improvement in the classification accuracy.

At heart of the data transformation procedure, dimensionality reduction is usually employed to denoise a research dataset and choose the most predictors. This describes the reduction of a high dimensional dataset  $X = (x_1, \dots, x_m)$  into a low-dimensional data  $Y = (y_1, \dots, y_d)$  by

removing redundant features and outliers from the dataset while preserving as much of the significant structure of the high-dimensional data as possible in the reduced-dimensional representation (Jeong, Ziemkiewicz, Ribarsky, & Chang, 2009). This is similar to the way physicians extract the most significant symptoms of an illness during medical diagnosis (Chattopadhyay, 2017).

Two main methods of dimensionality reduction are feature selection and the more general feature extraction (Sakthivel, Nair, Elangovan, Sugumaran, & Saravanmurugan, 2014). A subset of the existing features is selected in feature selection, while new features are constructed from the existing ones either using class information (supervised) or not (unsupervised). In feature extraction. There are several feature extraction methodologies and techniques that can be used to reduce feature space for more efficient performance of machine learning algorithms. Some of these techniques have been well applied in the single-target classification tasks and are directly applicable to the multi-target domain. They are grouped into two: filter methods, and wrapper methods. The study concentrates on the filter methods for the following reasons:

- a. Filter methods are much faster compared to wrapper methods as they do not involve training the models. This independency of the filter method has been reported as a benefit in several studies (Alhaj, Siraj, Zainal, Elshoush, & Elhaj, 2016; Bannasar, Hicks, & Setchi, 2015) Wrapper methods, on the other hand, are computationally very expensive.
- b. Using the subset of variables from the wrapper methods make the model more prone to overfitting as compared to using subset of variables from the filter methods.
- c. Filter methods are independent of any machine learning algorithms. Instead, they extract features by ranking them on the basis of their scores with the output in various techniques. Commonly used filter methods include information gain (univariate algorithm), Relief-F, principal component analysis (PCA), linear discriminant analysis (LDA), Ranking by Chi-square, correlation, p-value from significance test (Chandrashekar & Sahin, 2014).

However, the intention of this study is not to further select or ignore features that have been carefully selected by a combined team of mental health professionals. As part of the objectives, the study aims to establish the strengths of each feature and the inter-relationships between the features and the class target. Due to its efficiency, simplicity, remarkable success in bioinformatics (L. Wang, Wang, & Chang, 2016), text categorisation (Gao, Xu, Meng, Qi,

& Lin, 2014), and several other applications (Chandrashekar & Sahin, 2014), this study employed the information gain to compute the relative importance of the predictors (symptoms, in this case) and calculate the dependence relationships between the symptoms, and the symptoms and the diagnosis

#### 4.5.5 Information gain

The information gain is an approach that ranks features based on a relevancy score which is based on each individual attribute. The information gain of an attribute says how much information with respect to the classification target the attribute has (Alhaj et al., 2016). That is, the information gain looks at each feature independently, computes its information gain and measures how important and relevant it is to the class label Based on mutual information, with possible interactions between features (Bennasar et al., 2015), the information gain selects only those features that have significant information gain (mutual information with the class variable). Because a common measure for the information is the entropy (that is, the measure of the degree of average uncertainty in the random variable) (Shannon, 1948), we start the computation of information gain by calculating the entropy for the target class. Entropy, as defined by (Shannon, 1948) is:

$$info(D) = - \sum_{i=1}^m p_i \log_2(p_i) \quad 4.1$$

Where:

$info(D)$  represents the information needed to classify a tuple in  $D$ , also known as the entropy of  $D$

$D$  denotes the dataset sample

$p_i$  represents the fraction of  $D$  in respect of target class

$m$  is the number of possible outcomes

The extreme entropy values for  $info(D)_{max}$  are 1 (totally random) and the deterministic value is 0 (perfectly classified).

The next step in computing the information gain is to calculate the expected information required to classify a tuple from  $D$  based on the partitioning of attribute  $A$

The expression is:

$$info_A(D) = \sum_{j=1}^v (|D_j|/|D|) * info(D_j) \quad 4.2$$

where  $D_j$  is the subset of  $D$  containing distinct value of  $A$  and  $v$  is the number of distinct value in  $A$

The information gain measurement can now be computed as the difference between the prior entropy of classes and posterior entropy

$$Gain(A) = info(D) - info_A(D) \quad 4.3$$

According to Sluga and Lotric (2017), mutual information measures how much information is communicated, on average, in one random variable about another. In other words, mutual information reports how much to which the joint probability of the predictor (in this case, symptom), and the target (depression) deviates from what it would be if the predictor was independent of the target (Conrady & Jouffe, 2015; Korb & Nicholson, 2004). As an example, suppose  $X$  denotes the roll of a fair 6-sided die, while  $Y$  denotes whether the roll is even (0 if even, 1 if odd). Evidently, the value of  $Y$  says something about the value of  $X$  and vice versa. That is, these variables share mutual information. For the present study, mutual information helped us to measure the amount of information that one symptom of depression has about another or depression itself, thereby reducing the uncertainty of one symptom to the knowledge of the other. Consequently, the value of mutual information of each symptom, and the overall contribution of each symptom to the target nodes (depression) were computed.. The mutual information between a predictor  $E$  and a target  $C$  is defined as:

$$MI(C, E) = H(C) - H(C/E) \quad 4.4$$

$$\approx \sum_{c \in C} \sum_{e \in E} P(c, e) \log_2 \frac{P(c, e)}{P(c)P(e)} \quad 4.5$$

$$\approx \sum_{e \in E} P(e) \sum_{c \in C} P(c/e) \log_2 \frac{P(c/e)}{P(c)} \quad 4.6$$

Where  $H(C)$  represents entropy of class variable (in training set)

$H(C/E)$  represents entropy of class variable given feature  $C$

This enables the computation of the mutual information between a classification target and any possible predictors (Conrady & Jouffe, 2015). Consequently, the predictor that provides the maximum information gain and, thus, the greatest predictive importance can be found. The concept of mutual information as a feature selection method has been well applied in the literature with remarkable success to determine the relevance between features and target classes (Fang et al., 2015; Miche et al., 2017; Qian & Shu, 2015). In Qian and Shu (Qian & Shu, 2015), an effective mutual information-based feature selection algorithm with forward

greedy strategy was developed for evaluating candidate features in incomplete data, which used the largest mutual information with the target class and also took into consideration the redundancy between selected features. The selection of candidate features was implemented in a dwindling object set. When tested, the experimental results on different real data sets showed that mutual information-based feature selection algorithm was more effective for feature selection than most existing feature selection algorithms

The general algorithm for implementing information gain is as follows:

```

1:  Function IG C/E feature ranking-based entropy
2:  initialisation:
3:  S = 0;
4:  C ? domain of a class label;
5:  E? domain of an attribute values;
6:  For each  $c_i \in C$  do:
7:      calculate  $P(c_i)$  ;
8:       $H_c = S + P(c_i) * \log_2 P(c_i)$  );
9:      S ?  $H_c$ ;
10: End For
11: For each  $e_j \in E$ :
12:     calculate  $P(e_j)$ 
13:     Sum = S +  $P(e_j) * \log_2 P(e_j)$  );
14:     C ? Sum;
15: End For
16: For each  $c_i$  do:
17:     For each  $e_j$  do:
18:         calculate  $P(c[i] | e_j)$ 
19:         M = S +  $P(c_i | e_j) * \log_2 P(c_i | e_j)$  ;
20:         S? M;
21:     End For
22: End For
23:  $H(C/E) = (-1) * \text{Sum} * (-1) * M$ ;
24: IG =  $H_c - H(C/E)$ 
25: return IG
26: End function

```

In the study, the proposed procedure for computing the information gain is as follows:

1. Pre-process depression dataset
2. Calculate the entropy for the target class (depression)
3. Determine the mutual information values between the symptoms and depression in the dataset. This was done by the probability density estimation method to extract the linear and nonlinear relationship between depression symptoms

4. Determine the overall contribution of each symptom to the target class

After processing with the original depression dataset in BayesiaLab (version 6) (Conrady & Jouffe, 2015), the results of the mutual information and overall contributions towards the class variable are determined the results in Table 5.2 of section 5.6

#### **4.6 Problems encountered during data collection**

Like many other ICT4D and machine learning studies, a number of challenges were encountered by the researcher during data collection fieldwork. These are summarised as follows:

- Difficulties in getting mental health professionals together for a series of workshops in order to acquaint them with various ICT-related terms. The researcher had difficulty in explaining and defining terms such as ICT4D, machine learning, human computer interactions, user experience and user interface
- Difficulties in getting staff and student of the University as participants during the pilot study to respond to questions while maintaining a minimum human bias. Participants' difficulty in recalling information was a general problem at all the sessions. Majority of the University staff and students were willing to share information but found some of the questions difficult to understand and confusing, thereby leading to some giving exaggerated view of the usefulness of computers, machine learning and HCI in screening for diseases.
- Difficulties in getting access to de-identified sensitive patient data as the hospital and primary care centre generally do not keep records in usable formats. The researcher experienced problems extracting only the needed information from the dusty, tattered records which were scattered in all departments.
- Because the original dataset is composed of disparate sources, adequate time was spent converting and cleaning the data to get the desired quality for machine learning algorithm processing and analysis. This was resolved during the pre-processing stages of the data.
- Another challenge that confronted the initial depression dataset was the highly unbalanced dataset. The imbalance, 80% in one class and 20% in another class. The researcher however dealt with this challenge by going back to the field to collect more data on the classes that were short

#### 4.7 Applying machine learning methods in the experimental set up

The experimental set up is described in this section. To make the experiments reproducible, details of the study methods and their parameters are first discussed before the implementation platforms. The extensive experimental evaluation measures for the methods, which explain the step-by-step process that was followed for the comparison of the methods' performance are described in chapter five.

The architecture of the Bayesian network model integrated into a user interface design is shown Figure 4.6. However, other commonly used state-of-the-art probabilistic and non-probabilistic algorithms (see Table 4.5), which can be implemented in public domain software are also used to prove the superiority of Bayesian networks. These include Naïve Bayes, artificial neural networks, Tree-base boosting, K-nearest neighbour, and Kernel methods

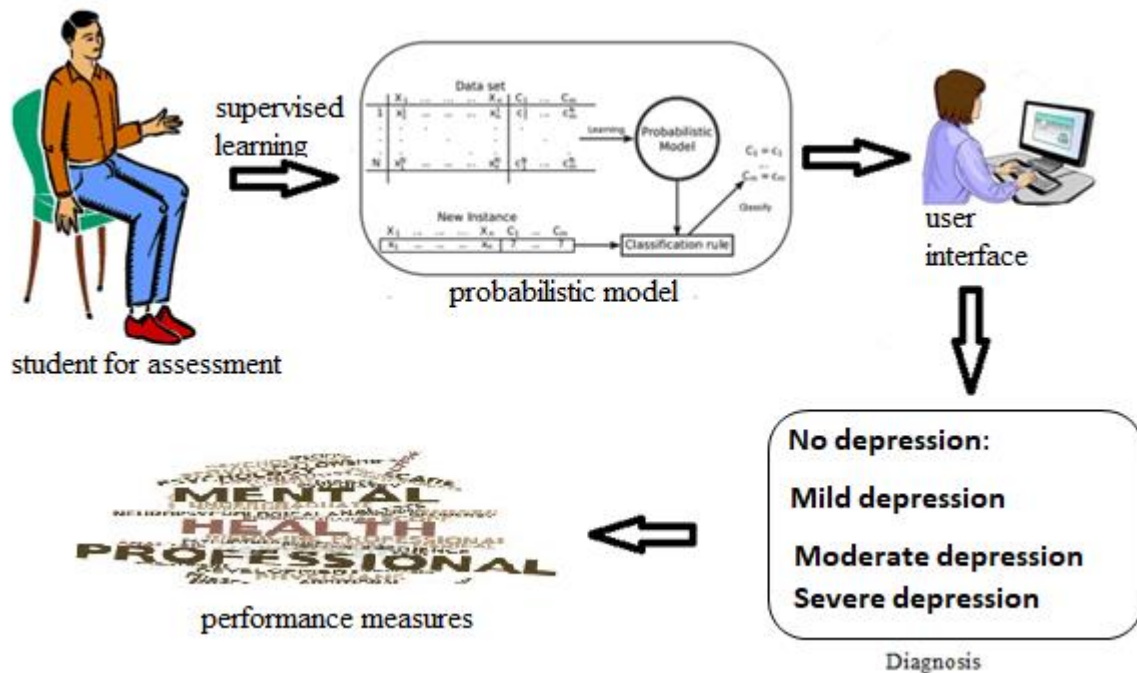


Figure 4.6 Architecture of probabilistic-HCI diagnostic model for the study

Table 4.5 Probabilistic and non-probabilistic algorithms

	Algorithm	
1	Bayesian networks	Probability-based methods
2	Naïve Bayes	
3	Artificial neural network (MLP)	
4	Tree-base boosting (C4.5)	Non-probability-based methods
5	K-nearest neighbour	
6	Kernel methods (support vector machines)	

### 4.7.1 Naïve Bayes

Naïve Bayes algorithms are probabilistic classifiers that are based on Bayes' theorem. Naïve Bayes classifiers are well known for their unrealistic assumption that attributes  $X_i$  are conditionally independent (naïve) given the class  $C_j$  and high performance in single-target classification (Xuesong, Wei, Qinghua, & Victor, 2016). Although this independence assumption is not true in real-life, the assumption can still lead to optimal decisions even if probability estimates are inaccurate due to feature dependence in practice (Zaidi & Cerquides, 2013). One side effect of this model is faster running times which is as a result to lower computational complexity. To classify a given instance, the posterior probability of each possible class value  $C_j$  is computed, and then, the most probable class  $C^*$  is selected. More formally

$$C^* = \arg \max_{C_j} P(C_j) \prod_{i=1}^n P(X_i/C_j) \quad 4.7$$

Although the popularity of the Naïve Bayes stems from the fact that it does not require iterative parameter estimation schemes, and it's relative ease of construction (Wu et al., 2008), a number of studies has noted its inferiority to AdaBoost and SVM in terms of performance (Kotsiantis, 2007; Miyamoto, Hazeyama, & Kadobayashi, 2009). In a study, using a set of 35,000 training and testing datasets (5,000 depressed and 30,000 non-depressed) from the electronic health records of Palo Alto Medical Foundation in California, USA, Huang et al (2014) built on the success of rule-based expert system by demonstrating the predictive strength of Naïve Bayes classifier on depressive disorders. The features were disease and drug treatment terms extracted from clinical notes and patient demographics. When tested against the ICD-9 codes, the results had a sensitivity of 65% and a specificity of 80%

### 4.7.2 Neural networks

The original idea behind the neural network is inspired by the mechanism of patterns recognition in the human brain. Considerable research has been carried out on neural networks since the pioneering work of McCulloch and Pitts (1943), to handle more complex tasks. A commonly used neural network algorithm, the multilayer perceptron (MLP). An MLP is a feedforward neural network (FFNN) with one or more hidden layers. In its general structure, the MLP is made up of three different interconnected layers (Kotsiantis, 2007). These are the

input layers of source neuron, which accept the elements of the input feature vectors, an output layer of computational neurons, which representing the classification result or supply the response of the neural network, and at least one middle or hidden layer of computational neurons, which represent datasets that are not linearly separable and are fully connected to the input and output units. The input signals are propagated in a forward direction on a layer-by-layer basis (J. Tang, Deng, & Guang, 2015). This is shown in Figure 4.7

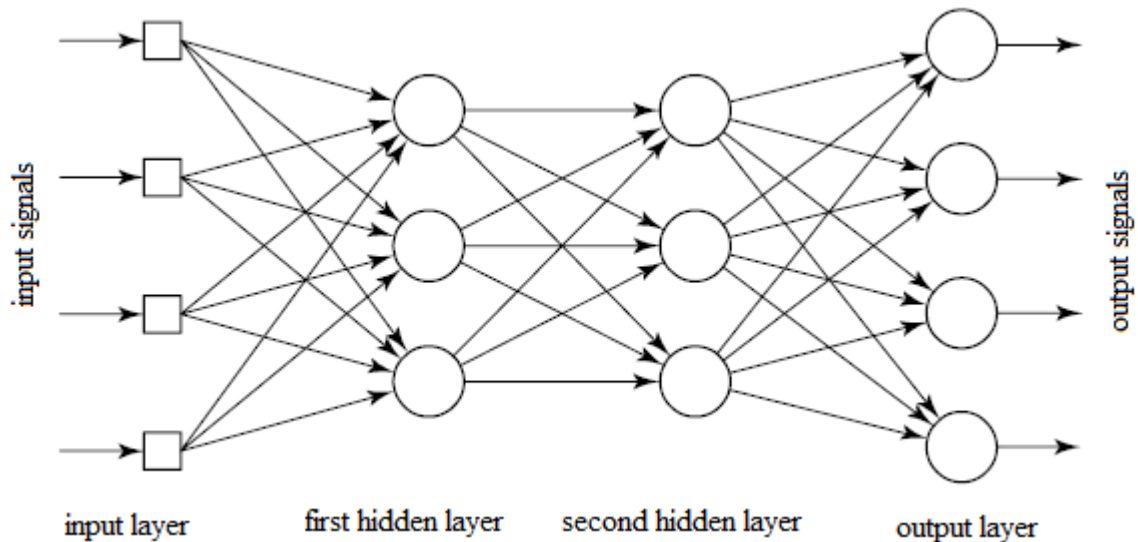


Figure 4.7 MLP with 2 hidden layers

The MLP uses the back-propagation (BP) algorithm and a sigmoid function to teach the ANN how to perform a given task. The BP algorithm has the ability to detect multiple nonlinear correlations from the training examples (Negnevitsky, 2005) and helps to compute how fast the error changes as activity changes by using error derivatives with respect to hidden activities. This is defined as follows:

$$X_j = \sum_{i=1}^n x_i * w_{ij} - \theta_j \quad 4.8$$

$$Y_j = \frac{1}{1 + e^{-x_j}} \quad 4.9$$

where

n is the number of inputs to node j,

$w_{ij}$  is the weight of the connection between each node  $i$  and node  $j$

$\theta_j$  is the threshold value assigned for node  $j$

$x_i$  is the input value for input node  $i$

$Y_j$  is the output value produced by node  $j$

Input signals  $x_1, x_2, \dots, x_n$  are passed through the network from left to right while error signal,  $e_1, e_2, \dots, e_j$  from right to left.

Identity

$$F(X) = 1$$

Binary step

$$F(X) = 1 \quad \text{if } X \geq 0$$

$$F(X) = 0 \quad \text{otherwise}$$

Binary sigmoid

$$F(X) = \frac{1}{1+e^{-\theta X}}$$

Bipolar sigmoid

$$F(X) = \frac{-1+2}{1+e^{-\theta X}}$$

Hyperbolic tangent

$$F(X) = (e^X - e^{-X}) / (e^X + e^{-X})$$

Figure 4.8 shows the architectural design set up for depression diagnosis neural network and its corresponding backpropagation algorithm components. Each input to one neuron represents a symptom of depression and the output represents the diagnosis of depression (No depression, mild depression, moderate depression and severe depression). The weight, a key element of the ANN, represent information being used by the network to solve a problem. Each neuron has an internal state called activation function which helps to compute output signal from input signal. During training, a neuron receives input, determines its strength or weight, calculates the total weighted input, and compares with the threshold value (between 0 and 1). If the total weighted input is greater than or equal to the threshold value, the neuron will produce output, and if the total weighted input is less than the threshold value, no output will be produced.

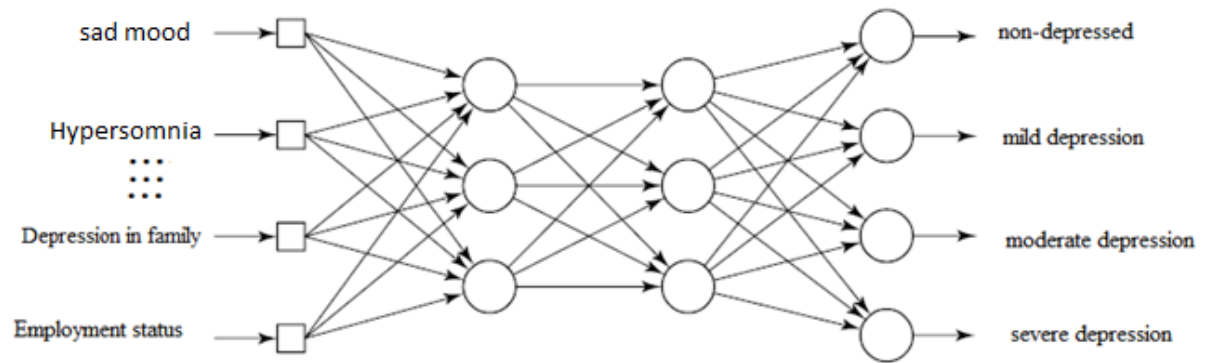


Figure 4.8 MLP with 2 hidden layers for the study

In a study, Suhasini et al (2010) designed a multi-decision support system with Backpropagation (BP) and radial basis function (RBF) neural network. Using 400 patient records divided into four equal part, the model was trained with 3/4<sup>th</sup> of the dataset and tested with the remaining 1/4<sup>th</sup>. Experimental results show that the proposed method achieves an accuracy of 98.75% for identifying the psychiatric problems. Several other studies have applied neural networks in the domain of mental health diagnosis (Bing-Mei, Xiao-Ping, Zhi-Ming, & Xue-Rong, 2011)

### 4.7.3 Tree-based boosting

Boosting is a general method for improving the accuracy of machine learning algorithms. A commonly used boosting algorithm is AdaBoost developed by Freund and Schapire (1997). Adaboost, short for Adaptive Boosting, became a popular classifier due to the modification of combining several classifiers to create a strong classifier, based on a set of attributes. A simple learning algorithm (a weak learner) is called several times, each time selecting the best weak classifier (using a greedy feature selection process). The learner is called weak because it is not expected to retrieve the best classification function to accurately classify the training data. To reduce the computational complexity, the complex classifiers are combined in a cascade structure proposed by Viola & Jones (2004). AdaBoost methods have mainly been used in bioinformatics applications (Yoon & Kim, 2008) and can also work well in domains with missing data values (Zhou & Lai, 2016).

The learning algorithm begins by using training dataset as input (depression and its symptoms, for this study). It then searches for the most suitable test that can be put in the root node. If found, the algorithm then creates a new internal node and calls itself recursively to build a subtree for each subset in the partition induced by the test on the training instances. To

select the most suitable test, the algorithm scores the tests by the reducing the variance (see equ 4.5) induced on the instances. The reduction in variance helps to maximise cluster homogeneity and improves predictive performance. If no acceptable test that significantly reduces variance (as measured by a statistical F-test) is found, then the algorithm creates a leaf and labels it with a representative case, or prototype, of the given instances. The variance of a set of example  $S$  is defined as the average squared distance between each example's class vector  $v_k$  and the set's mean class vector  $\bar{v}$ :

$$var(S) = \frac{\sum_k d(v_k, \bar{v})^2}{|S|} \quad 4.10$$

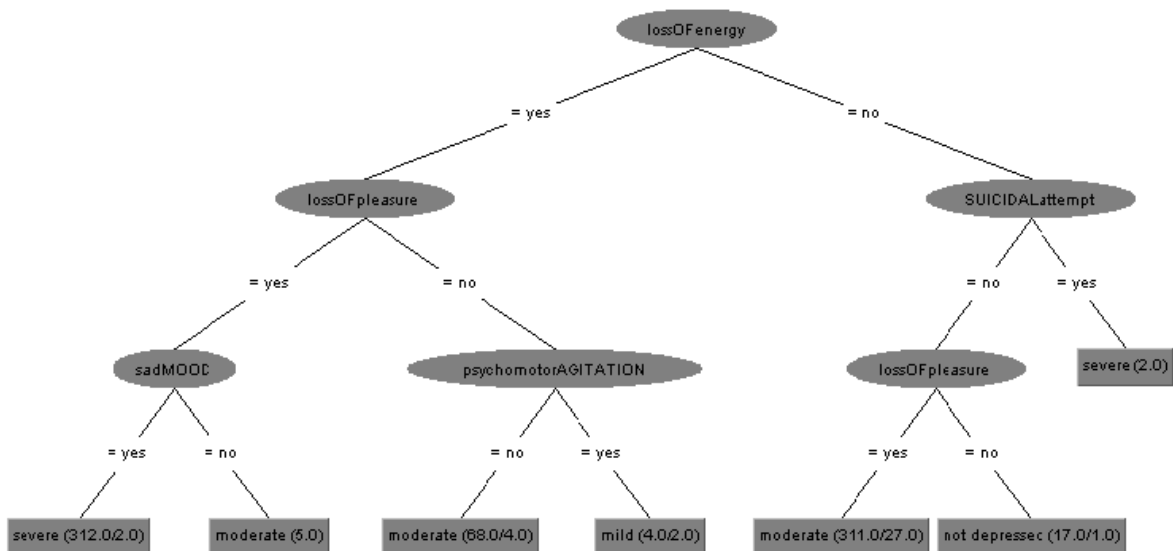


Figure 4.9 Tree-based Boosting

#### Basic algorithm for Tree-based Boosting

- a. Start with full dataset
- b. Find test that partitions examples as well as possible
  - a. Examples with the same class should be put together
  - b. Maximal information gain
- c. For each outcome of test, create child node
- d. Move examples to children according to outcome of test
- e. Repeat procedure for each child that is not 'pure'

#### 4.7.4 Decision Tree (C4.5)

C4.5 is a commonly used decision tree algorithm introduced by Quinlan (1992). It has graphs consisting of nodes and leafs with each node having one or more following nodes or leafs. A leaf contains the classification result. Beginning with the root node, the graph is traversed up to a leaf. At each node, a condition is evaluated to decide about the selected successor. For example, a condition for a node with two successors could be the evaluation if an input exceeds a threshold. The left successor is chosen if the threshold is exceeded, the right successor otherwise. It selects one feature with at least two outcomes, which partitions the set of samples most effectively. In case all samples of a set belong to the same class, or the set is small, the leaf is labelled with the (most frequent) class label (Wu et al., 2008). C4.5 is represented as J48 in many popular machine learning software tools (Hall & Witten, 2010) and performs quite well with numerical and categorical variables. Besides the good combination of low error rate and speed of C4.5 algorithm, its created decision trees are usually easy to understand and interpret by users (Wu et al., 2008). C4.5 performs better if there are only a few highly relevant inputs. The general algorithm for building a decision trees is as follows (Armano & Tamponi, 2018; J R Quinlan, 2006):

- 1: check for base cases
- 2: for each attribute  $A$ , find the normalised information gain from splitting on  $A$
- 3: Let  $a\_best$  be the attribute with the highest normalized information gain
- 4: Create a decision *node* that splits on  $a\_best$
- 5:     Recur on the sub lists obtained by splitting on  $a\_best$ , and add those nodes as children of the node

#### 4.7.5 k-nearest neighbours

K-nearest neighbours (KNN) algorithm, the most commonly used instance-based learning (IBL) algorithm (Muja & Lowe, 2014) is a statistical and efficient lazy learning algorithm that makes a prediction by calculating the output values of the nearest learn data tuples –called the nearest neighbours (Deng, Zhu, Cheng, Zong, & Zhang, 2016). That is, it uses the assumption that samples of a dataset with similar properties will exist in close proximity (Kotsiantis, 2007). These nearest neighbours are determined by means of a suitable distance measure. Where  $k = 1$ , the algorithm simplifies to the simple nearest neighbour (nn) algorithm. The calculation of the output values is done by either voting or by the means of a weighted average. Usually, the voting is used for classification tasks while calculation by means of a weighted average is often

used for regression tasks. Ties must be broken using a suitable tie breaking policy. For example by preferring the output class with the highest a priori probability on the learn dataset. A popular approach is the Nadaraya-Watson estimator (Nadaraya, 1964; Watson, 1964), that can be written for case where  $k = N$  as

$$\bar{y} = \frac{\sum_{i=1}^N F(d_i) y_i}{\sum_{i=1}^N F(d_i)} \quad 4.11$$

$F$  = an arbitrary one-dimensional kernel, which is usually a continuous, bounded and integrable value

$d_i$  = the distance of the  $i$ -th data tuple of the learn dataset to the given input position.

The basic algorithm for K-NN as follows:

Input:  $D = \{(x_1, c_1), \dots, (x_n, c_n)\}$

1: **begin**

2:  $y = (y_1, \dots, y_p)$  new instance to be classified

3: compute  $d(x_i, y)$  for each  $(x_i, c_i)$

4: sort  $d(x_i, y)$  from the lowest to highest,  $i = (1, \dots, n)$

5: select  $k$  points nearest to  $y$ :  $D_x^k$

6: assign to  $y$  the most frequent class  $D_x^k$

7: **end**

where  $D$  represents the dataset,

$k$  represents the number of neighbours,

$p$  is the number of features

$d(x_i, y)$  represents the Euclidean distance

$n$  represents the number of values

The properties of the kNN algorithm, according to (2006) are summarised as follows:

#### 4.7.6 kernel methods

Kernel methods are data analysis techniques that map an existing dataset into a feature space, where existing patterns can be discovered using simple linear relations. For its sound theoretical foundations and geometric interpretations, support vector machines, introduced by Cortes and Vapnik (1995), is the best-known kernel method, though others such as, radial basis function, least squares support vector machines (LS-SVMs) and kernel Fisher discriminant

(KFD) analysis have been proposed for both classification and regression problems in the literature (Engelbrecht & Paquet, 2008). Traditionally, support vector machines are used with a single-output variable to determine the mapping between the input vector  $T = \{(x_1, y_1), \dots, (x_n, y_n)\}$ , with  $x_i \in \mathbb{R}^m$  and  $y_i \in \{\pm 1\}$ , assuming  $n$  examples with  $m$  real attributes, and the single output  $Y$  from a given training data set  $D$ . If a training dataset  $D = \{(x_i, y_i) \mid 1 \leq i \leq m\}$  where  $m$  is size of dataset and  $x_i \in \mathbb{R}^d$  is a  $d$ -dimensional feature vector and  $y_i \subseteq y, y = \{y_1, y_2, \dots, y_i\}$ ,  $x_i$  and  $y_i$  represent the  $i$ th instance and its relevant class, respectively. In the same time, the complement of  $y_i$ , which is  $\bar{y}_i = y \setminus y_i$  is called the irrelevant set of labels.

The goal of SVM is to construct a (set of) hyperplane(s) in a high-dimensional space. A fundamental challenge is how to find a hyperplane that fits well, especially in case where the dimensionality of the feature space is large. SVM solves this problem by taking only a small amount of the training data into consideration (that is, the so-called support vectors), which are used to determine the margin. Another challenge has to do with the finite dimensional space. Usually, this space is not linearly separable. Consequently, a mapping, called the kernel trick (Holloway, Suttie, Dass, & Neubauer, 2011) from the low-dimensional feature space into a higher-dimensional feature space is performed, presuming that the separation is easier in the higher-dimensional space.

In order to apply a support vector machine to a dataset, kernels must be defined for each data type and the kernel matrices must be combined algebraically. When support vector machine is applied in classification tasks, the task is divided among several support vector machines. The processing time is proportional to the number of kernel computations performed. The authors employed modified SVMs, which allows the simultaneous training of a set of support vector machine classifiers by using a single optimization. Then, the discriminatory boundary is defined thus:

$$(w_i - w_l, x) + b_k - b_i = 0 \quad 4.12$$

Where  $w_i, w_l \in \mathbb{R}^d$  and  $b_k, b_i \in \mathbb{R}$  is the weight vectors and bias terms, respectively. Similarly, the classification system's margin on  $(x_i, y_i)$  is defined. It is reasonable that the labels belonging to an instance should be ranked higher than those not belonging to that instance:

$$\min_{(y_k, y_i) y_i \neq \bar{y}_i} \frac{(w_i - w_l, x) + b_k - b_i}{w_i - w_l} \quad 4.13$$

After that, maximize the classification system's margin in the whole training set  $D$  and meanwhile a slack variable is incorporated by considering real-world situation.

Therefore, the SVM defines the margin over hyper planes for relevant–irrelevant label pairs, which explicitly characterizes label correlations of an individual instance. SVM achieves great accuracy, though has a high computational cost, which limits its usability for many applications. SVMs have been employed in the development of several clinical decision support systems to help clinicians to make decisions about diagnosis of disease and interpretation of medical tests. For example, in Yu et al (Yu, Liu, Valdez, Gwinn, & Khoury, 2010), an SVM technique was developed for the classification of persons with and without common diseases. Several other studies have used SVM for diagnosing neurological and psychiatric disorders by utilising a diverse range of neuroimaging technique as in Orru et al (Orrù, Pettersson-Yeo, Marquand, Sartori, & Mechelli, 2012), and Ramirez et al (Ramírez et al., 2013) The SVM is used in automatic computer-aided diagnosis systems for early diagnosis of AD by the means of SPECT imaging (Ramirez et al., 2013).

The proposed procedure for applying the SVM algorithm are as follows:

1. Pre-process the dataset
2. Select SVM model
3. Select a kernel function
4. Tune the parameters
5. Test the model for deployment

#### **4.8 Challenges of machine learning application in medicine**

The rapid advancement in machine learning and computing technology has had positive effects in service delivery in many applications, medicine inclusive. However, despite these benefits, some limitations of employing machine learning methodologies in medicine have been noted by studies in informatics literature (de Bruijne, 2016; Foster, Koprowski, & Skufca, 2014; Hijazi et al., 2016; Holzinger & Jurisica, 2014). These are described briefly:

- 1) The first is the type of data in healthcare databases. Data types in healthcare database are usually heterogeneous in the sense that many patients' tests results are in numeric form, text form, and images. Analysing such mixed data types often poses a big challenge to developers and data scientists and machine learning tools. The different sources of data, including laboratories, medical centres, clinicians, patients and many more, makes data collection and integration processes both laborious and

time-consuming. To overcome this challenge, several studies have recommended that a data warehouse to be put in place where data is pre-processed before data analysis process begins (Berndt, Fisher, Hevner, & Studnicki, 2001; Taniar & Chen, 2011; Wyllie & Davies, 2015)

- 2) The second limitation is the unorganised nature of data in healthcare databases. This include missing attribute values, inconsistent with patient history or family history, and corrupted files. Though the problem of missing attribute values can be solved by constructing or estimating the missing attribute values, complete dataset, without any missing values, makes analysis process and conversion into useful knowledge much more efficient.
- 3) The third limitation is that medical data often contains large number of cases and attributes, which makes it difficult for machine learning algorithms to generate accurate patterns and useful knowledge. This can however, be solved by using feature selection and extraction processes to select the most important features in data.
- 4) The fourth limitation of machine learning application in healthcare is that it requires combined expertise in the fields of machine learning and medical science. Given that it is difficult to have a person with such combined expertise, analysing medical data may require collaboration between experts in both fields. However, a lack of cooperation between the experts in the two domains may produce a negative outcome.

## Chapter Five

### Implementation, presentation of results and discussions

#### 5.1 Introduction

Having established compatibility between the two study datasets and transformed into a format suitable for machine learning algorithms, the next phase is to set up an elaborate experiment that is necessary to empirically confirm the performance of the study methodological design and answer the study research questions. This was done in three phases. The first phase involved running the depression dataset through a four-step mutual information procedure in BayesiaLab 6. This was done to determine the relevance of each symptom relative to the classification target and to determine the strengths of a combination of symptoms with respect to the class target. The second phase involved running different classification algorithms using the depression dataset in Weka 3.8. This afforded the researcher an in-depth knowledge of the performance of the classifiers in a 10-fold cross validation method. The last phase involved integration of the knowledge base models into designed screening interfaces. This was done to enable the researcher test the efficiency and effectiveness of use by both clinicians and non-clinicians in a non-clinical environment. The first two phases were carried out using commercially available predictive analytics software tools, which are described briefly in section 5.1

#### 5.2 Software tools used in the study

A number of Predictive analytics software exist for modelling and inference through algorithms that find patterns in data and enhances experience of working on messy data. These include BayesiaLab (Conrady & Jouffe, 2015), Waikato environment for knowledge analysis (Weka) (Bouckaert et al., 2014), Multi-label extension to Weka (Meka) (Read, Reutemann, Pfahringer, & Holmes, 2016), Analytical, Bayes Net Toolbox, GeNIe (Genie, 1999), Hugin (Hugin, 2016), JavaBayes, MSBNx, (Korb & Nicholson, 2004) and Netica (2010). The framework for the implementation of the models for the study are discussed in sections 5.2.1.

##### 5.2.1 Waikato environment for knowledge analysis (Weka)

Available as an open source tool with a supportive large user community, Weka is an efficient multi-usage machine learning workbench available in two segments: API, and a number of

command line and intuitive graphical user interfaces for the whole machine learning lifecycle. It supports all the important features for machine learning tasks such as, data preparation, visualisation, exploration, inspection, discretisation, numeration, scaling, attribute selection, missing values, outliers, statistics, visualization, balancing, sampling, row selection, building classification, regression, association rules and clustering models (Bouckaert et al., 2014). Weka has many algorithms (Bayesian networks and its variants, Naïve Bayes and its variants, Decision trees (C4.5 represented as J48)), Kernel methods (SVM represented as Sequential Minimal Optimisation (SMO)), Boosting (AdaBoost), Neural networks (MLP), Linear models, Time series, etc), which are provided built-in as well as provided in third party plugins. It is a very versatile environment, which has several advantageous features (including numeric and graphical interface building parts) making it an ideal machine learning development environment (Bouckaert et al., 2014). Weka also has interesting automation feature to save and redo previous tasks on a new dataset and does not have the boundaries of many other machine learning software in terms of algorithm options. Another advantage of Weka is its simplicity, high performance and the ability to load data from all major internet survey programs like survey monkey, survey gizmo, and many others. Weka has the capability to automatically link the students' epidemiological data to the diagnosis section; hence it eliminates entering duplicate data. The database tables were configured to conform to the input requirement set by Weka. For the present study, Weka provided the platform to test the strengths of several classifiers using standard metrics

### **5.2.2 BayesiaLab**

Though has much fewer algorithms than Weka, BayesiaLab has exceptional machine learning capabilities to explore, transform and extract valuable knowledge from raw data. It then generates easily-interpretable reports with every possible chart, for visualisation purpose. BayesiaLab was used in the study for the analysis of data during the proof-of-concept study and the analysis of the statistical relationships between each symptom with respect to depression and synergistic combination of symptoms with respect to depression (described in section 5.5.1)

### **5.3 Presentation of experimental results**

Several test experiments on real clinical dataset were performed to prove the theoretical conjectures made in the study and to evaluate the performance of the proposed Bayesian

network models to discover interdependencies between features and targets. Two evaluations were carried out to ascertain the quality of the proposed integrated systems. The first was the task of evaluating and comparing the knowledge-based models (classifiers) and the second was to evaluate the instantiation of the artifact.

The first experiments had full dataset with 1090 instances, 19 normal features and one target class. The second experiments were to investigate the influence of the PCA-reduced features on the performance of the models. In all experiments, a ten-fold cross validation was performed for the dataset, the running time was computed, and a ranking of the algorithms was provided.

#### 5.4 Evaluation metrics for knowledge-based models

Several evaluation metrics for assessing the quality of a classifier have been described in the literature, but based on five under-listed criteria reported by Mandal (2014) and Martin (2010), the study has considered the following seven metrics, (1) predictive accuracy, (2) precision, (3) F-Measures, (4) Root mean squared error (RMSE), (5) Pearson correlation coefficient (R), (6) Receiver operating characteristics (ROC), and (7) running time. The metrics perform differently under balanced and imbalanced datasets (Saito & Rehmsmeier, 2015).

- A. **Accuracy:** It signifies the level of confidence a classifier has, usually computed as the proportion of correct classification ratio that it is capable of producing.
- B. **Speed:** Though less important than the accuracy, the speed is the classifier's response time from introducing an unseen example to classify, to the instant in which the classifier produces the predicted class.
- C. **Learning speed:** This is the time required by the classifier to obtain the classification rule from a data set.
- D. **Robustness:** Minimum number of examples needed to obtain a precise and reliable classification rule.

##### 1. Prediction accuracy

Though not a preferred metric for an imbalanced dataset, prediction accuracy is one of the most commonly used performance measures for a machine learning classification system. It is common to describe the classification performance of medical diagnosis in four categories: True Positive (TP), False Negative (FN), False Positive (FP) and True Negative (TN). Prediction accuracy is measured on all four values. The True Positive and True Negative represented the number of persons diagnosed by the diagnosis system correctly either as having

or not having depression. The False Positive represented the number of persons diagnosed by the system as depressed whereas they were actually not depressed while False Negative represented the number of persons classified as not depressed whereas they were actually depressed.

Traditional accuracy measures of prediction accuracy of medical diagnostic tests include sensitivity and specificity.

**Sensitivity:** The sensitivity or true positive rate of a learning machine is defined as the proportion of the positive cases that are predicted to be positive. It is defined as:

$$\text{Sensitivity} = \frac{\text{Number of positive correctly classified}}{\text{Number of total positive}} = \frac{TP}{TP + FN} \quad 5.1$$

**specificity:** The specificity or true negative rate is defined as the proportion of the negative cases that are predicted to be negative. It is defined as:

$$\text{specificity} = \frac{\text{Number of negative correctly classified}}{\text{Number of negative correctly classified}} = \frac{TN}{TN + FP} \quad 5.2$$

Therefore, prediction accuracy is defined thus:

$$\text{Prediction accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad 5.3$$

A perfect classifier (that is, 100% prediction accuracy) would have 100% Sensitivity, which means that all TP cases are identified as TP, and 100% Specificity, which means that all TN cases are identified as TN.

## 2. Precision

The Precision is the proportion of the examples which truly have class x among all those which were classified as class x. in other words, Precision is a measure of a classifiers exactness. For instance, a precision of 0.957 achieved by the Bayesian networks classifier in the study is interpreted as 95.7% correct predictions among the positive predictions. A good classifier should be precise, so that its estimates show little or no variation. Variation.

Mathematically, Precision is defined as:

$$\text{Precision} = \frac{TP}{TP + FP} \quad 5.4$$

### 3. F-Measure

The F-Measure conveys a balance between the precision and the recall (Bouckaert et al., 2014), determines the effectiveness of the classifier in classifying TP cases.. in essence, the F-Measure is the harmonic mean of the recall and precision measures

F-Measure is simply defined as:

$$F - \text{Measure} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad 5.5$$

Recall is otherwise called True positive Rate (TPR), which is equivalent to Sensitivity (defined above)

### 4. Root mean squared error (RMSE)

The root-mean-squared error (RMSE) measures the differences between values predicted by a model and the values actually observed. This difference is the average of the error.

Mathematically,

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (x_{obs,i} - x_{model,i})^2}{n}} \quad 5.6$$

where  $x_{obs}$  is observed values and  $x_{model}$  is modelled values at time/place  $i$  and  $n$  is the number of observations. The values of RMSE range from 0 to infinity with 0 indicating a perfect model performance.

### 5. The Pearson correlation coefficient (R)

The Pearson correlation coefficient ( $r$ ) shows the strength and direction of a linear relationship between two variables, X (model output) Y (observed values). It is obtained by dividing the covariance of the two variables by the product of their standard deviations, given a value between -1 and +1. A correlation coefficient of +1 shows a total correlation, 0 is no correlation and -1 is a total negative correlation. Mathematically,

$$R = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad 5.7$$

where  $x_i$  = actual number;  $y_i$  = predicted number;  $\bar{x}$  and  $\bar{y}$  are average numbers for actual and predicted, respectively. As an equivalence of Cohen's Kappa (Feingold, 1992), Pearson correlation coefficient ( $r$ ) is represented in Weka as Cohen's Kappa (Bouckaert et al., 2014)

## 6. Receiver operating characteristics (ROC)

Receiver operating characteristics (ROC) curve is used for visualizing the performance of a classifier. ROC provides the area under the curve (AUC) of the plot of the TPR (y-axis) against the false positive rate (x-axis). An excellent classifier will have ROC area values between 0.9 and 1.0 (90 and 100%) while a poor classifier will have ROC area values between 0.6 and 0.7 (60 and 70%) (Bouckaert et al., 2014). The ROC curve is used for visualizing the performance of a classifier, where the sensitivity (TPR) is plotted against 1-specificity (FPR).

## 7. Runtime

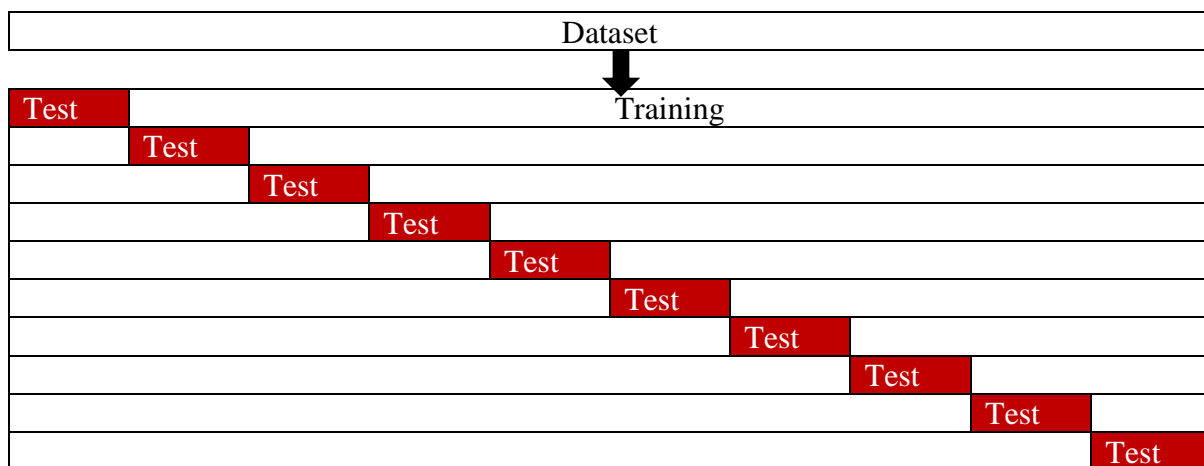
This is the computational costs to first learn a model from the learn data sample and then make predictions for new unknown input vectors. This is usually reported as cumulative training and testing times in seconds.

### 5.4.1 Cross-validation

To obtain a fair accuracy estimation, the dataset used to train the classification algorithms must be independent from the dataset used to test it (Witten, Frank, & Hall, 2011). Weka provided the platform that ensured this independence by performing a stratified 10-fold cross validation (Juan Fernandez del pozo, Pedro, Larranaga, 2013) to split out the depression dataset. Stratification is the process of dividing a study dataset into subgroups before actual sampling takes place (Witten et al., 2011). This strategy is used when sub-data (strata) within the dataset vary greatly and generally ensures a better representation of the study. That is, the original dataset was randomly divided into ten parts, each with approximately the same size. As illustrated in Table 5.1, in each fold, one part was held-out for testing and the learning algorithm was trained on the remaining dataset. The process was repeated ten times so that each part was used as the test data exactly once, where the averaged metric values out of ten runs were reported for the algorithm. The stratification helped preserve the statistical properties of the original dataset, including the joint probability distribution of the class variables and a reduced variance. It also helped to ensure that our limited and unbalanced depression dataset got a good balance between the size and representation of the training and test sets. Each subset

was utilized once as a test set and nine times as part of a training set. After that, its error rate is computed on the holdout set. Therefore, learning procedure (search, evaluate, and classify) is executed a total of ten times on different training sets in order to ensure a valid and robust results. Finally, the ten error estimates are averaged to produce an overall error estimate. This technique was used for all the classification algorithms under study.

Table 5.1 10-Fold cross validation



## 5.5 Results and discussion of knowledge-based models

In this section, the results of the algorithms described earlier in the study are presented in two parts. The first part is the computation of mutual information (relationship) between the symptoms of depression and depression itself, while the second part is the classifiers performance on the dataset

### 5.5.1 Relationship between symptoms and depression

Mutual information measures, in the information-theoretic sense, how much information the presence or absence of a symptom contributes to making the correct classification decision on depression. The study followed the four-step procedure for information gain as stated in section 4.5 and tested with the original depression dataset in BayesiaLab 6 (Conrady & Jouffe, 2015).

As noted by Korb and Nicholson (2004), mutual information is symmetric. This means that the amount of mutual information that a predictor reports on the target is the same as the amount of mutual information a target reports on the predictor. Mathematically,  $MI(C: E) = MI(E: C)$  and is zero if  $C$  and  $E$  statistically independent

Building on this concept, and given that there are two variables: the symptoms and depression, the results of the mutual information and overall contributions towards the class variable (depression), which represents the ranking of the symptoms based on statistical relevance measures, are tabulated in Table 5.2.

Table 5.2 Mutual information and contribution of symptoms to depression

S/N	Parent/Target/depression	Symptom	Mutual information	Overall contribution (%)
1	Depression	Impaired function	0.6309	17.1944
2	Depression	Loss of energy	0.5621	15.318
3	Depression	Worthlessness	0.3650	9.9484
4	Depression	Weight loss	0.3445	9.3879
5	Depression	Lack of thinking	0.3101	8.4521
6	Depression	Alcohol or other drug consumption	0.3098	8.3152
7	Depression	Indecisiveness	0.3069	8.3627
8	Depression	Loss of appetite	0.2022	5.5093
9	Depression	Recurrent thought of death	0.1898	5.1730
10	Depression	Family support and availability of accommodation	0.1774	5.1689
11	Depression	Psychomotor retardation	0.1896	5.1674
12	Depression	Loss of pleasure	0.1618	4.4101
13	Depression	Job satisfaction or academic performance	0.1310	3.8041
14	Depression	Insomnia	0.1179	3.2135
15	Depression	Sad mood	0.0863	2.3532
16	Depression	Suicide attempt	0.0680	1.8526
17	Depression	Weight gain	0.0352	0.9602
18	Depression	Depression in family	0.0277	0.7540
19	Depression	Hypersomnia	0.0214	0.5825
20	Depression	Stressful life events	0.0202	0.5503
21	Depression	Psychomotor agitation	0.0171	0.4654
22	Depression	Financial pressure	0.032	0.0873
23	Depression	Cigarette smoking	0.0152	0.0550
24	Depression	Employment status	0.0097	0.0257

It can be observed from the results in Table 5.2 that ‘impaired function’ had the highest mutual information (0.6309) and highest overall contribution (17.1944%) to the depression.

This was followed by 'loss of energy' with mutual information (0.5621) and a contribution of (15.318%) towards depression. It is surprising to note that the level of alcohol or other drug consumption, with mutual information (0.3098) was the 6<sup>th</sup> highest contributor (8.3152%) to depression, while the 'level of family support and availability of accommodation', with mutual information (0.1774), was the 10<sup>th</sup> highest contributor (5.1689%) to depression. 'Job satisfaction and academic performance', with mutual information (0.1310), was the 13<sup>th</sup> highest contributor (3.8041%) to depression, while cigarette smoking, which had a low mutual information (0.0152) was the second lowest contributor (0.0550%) to depression. 'Employment status' had the lowest mutual information (0.0097) was also the lowest overall contributor (0.0257%) to depression.

For the synergistic combination of symptoms, only the positive contributors were considered, as shown in Table 5.3. It was observed that 'loss of appetite' and 'sad mood' with a mean synergy of 6.099 were the highest contributors to depression. This was followed by the combination of 'sad mood' and 'weight loss' with a mean synergy of 6.7256. 'Alcohol or other drug consumption' and 'impaired function' were the 5<sup>th</sup> highest contributor to depression. The combination that produced the least contribution to depression was from 'suicide attempt' and 'weight gain' and 'recurrent thought of death' with a mean synergy of 0.0003 (serial number 106, Appendix H)

Table 5.3 Synergistic combination of depression symptoms to depression

S/N	Symptom	Symptom	Mean synergy (%)	S/N	Symptom	Symptom	Mean synergy (%)
1	Loss of appetite	Sad mood	6.9968	23	Lack of thinking	Weight loss	0.8029
2	Sad mood	Weight loss	6.7256	24	Depression in family	Impaired function	0.7392
3	Insomnia	Loss of pleasure	4.5318	25	Job satisfaction or academic performance	Sad mood	0.6845
4	Loss of appetite	Loss of pleasure	4.3210	26	Impaired function	Cigarette smoking	0.6705
5	Alcohol or other drug consumption	Impaired function	4.1431	27	Recurrent thought of death	Weight loss	0.6430
6	Insomnia	Loss of energy	4.0504	28	Psychomotor retardation	Weight loss	0.6206
7	Impaired function	Sad mood	3.6677	29	Impaired function	Psychomotor agitation	0.5298
8	Loss of energy	Loss of pleasure	3.2593	30	Depression in family	Indecisiveness	0.5031
9	Insomnia	Loss of appetite	3.1744	31	Depression in family	Lack of thinking	0.4873
10	Loss of appetite	Loss of energy	2.5662	32	Impaired function	Suicide attempt	0.4427
11	Impaired function	Worthlessness	2.1636	33	Depression in family	Worthlessness	0.4084
12	Impaired function	Indecisiveness	1.7308	34	Job satisfaction or academic performance	indecisiveness	0.3392
13	Impaired function	Lack of thinking	1.6686	35	Cigarette smoking	Weight loss	0.3210
14	Impaired function	Recurrent thought of death	1.4990	36	Depression in family	Psychomotor retardation	0.2645
15	Family support and availability of accommodation	Job satisfaction or academic performance	1.3994	37	Depression in family	Weight loss	0.2423
16	Impaired function	Psychomotor retardation	1.3568	38	Suicide attempt	Weight loss	0.2318

The comprehensive reports of the contributions of each of the symptoms and the synergistic combination of symptoms relative to the target class (depression) are shown in Appendices G and H.

### 5.5.2 Knowledge-based models

After the preparation of dataset for a format acceptable by the machine learning classifiers, their performance were tested in Weka with the specified metrics. Table 5.4 shows the results of the proposed Bayesian networks versus the NB, KNN, SVM, C4.5, MLP and AdaBoost,

Table 5.4 Results from attributes in dataset

Algorithm	Evaluation						
	Accuracy (%)	Precision (%)	F-Measure	RSME	R	ROC (%)	Time (s)
<b>Bayesian networks</b>	94.3**	94.4**	0.943**	0.150**	0.923	92.2	10.04*
Naïve Bayes	93.3	93.3	0.933	.0175	0.935**	98.6**	1.24
KNN	92.2	92.9	0.901	0.372*	0.919	95.3	0.55
SVM	92.0	91.7	0.872	0.192	0.811	86.6	9.21
C4.5	90.3	91.5	0.924	0.314	0.912	91.3	0.92
MLP	90.1	89.4	0.856	0.245	0.901	90.2	0.42
AdaBoost	49.1*	47.0*	0.348*	0.368	0.710*	80.3*	0.17**

As indicated in Table 5.4, the results confirm a good performance of the proposed Bayesian networks model (marked in boldface) in terms of the considered metrics, but given that a classifier may have its own inductive bias that could be detrimental, it was found helpful testing out a variety of other supervised learning classifiers. This procedure helps to avoid coercion of consensus by one classifier and to get the performance of other classifiers in the same domain. For visualisation purpose, the best results on each metric is marked in double asterisks (\*\*) while the worst performer is marked in single asterisk (\*).

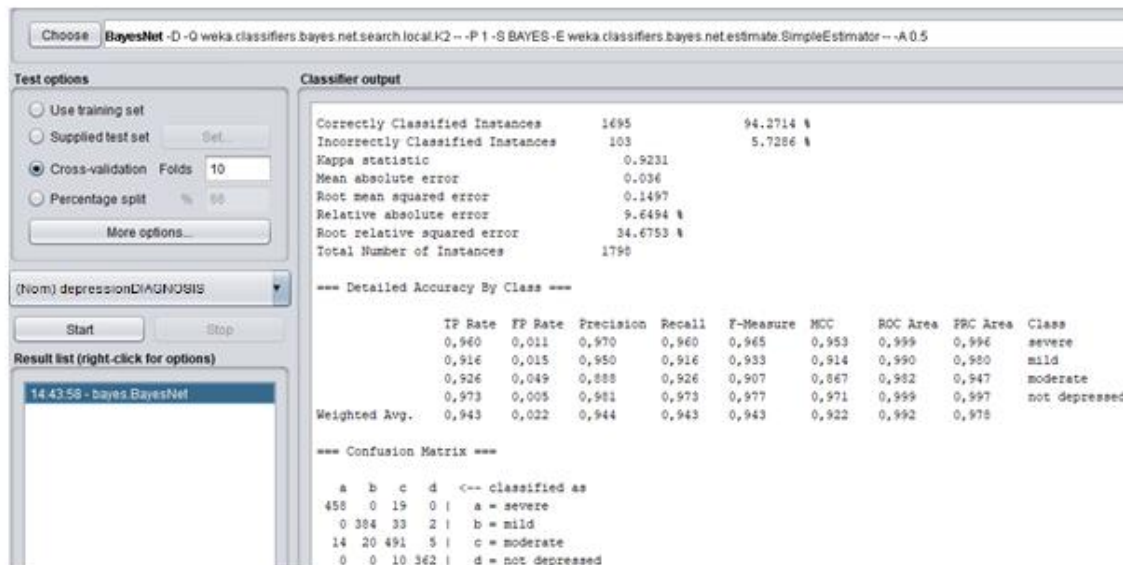


Figure 5.1 Bayesian networks model showing depression prediction performance

It was observed that Bayesian networks showed significant superiority over other classifiers in terms of predictive accuracy, precision, F-Measure, RSME, (see Table 5.2) as a pointer to producing predictions with the lowest deviations from the original dataset. Bayesian network classifier, however, was the slowest in terms of running time (10.04 second). As shown in Figure 5.1, the detailed classification results are calculated and displayed in a matrix format called Confusion Matrix (also called the contingency table) (Bouckaert et al., 2014).

The confusion matrix, a clean and unambiguous way to present the prediction results of a classifier, shows how many instances (patients in this case) have been assigned to each depression (depression). Elements show the number of test examples whose actual class is the row and whose predicted class is the column. For this study, there were four outcomes:

1. True positives (TP) are positive items correctly classified as positive.
2. True negatives (TN) are negative items correctly identified as negatives.
3. False positives (FP) are negative items classified as positive.
4. False negatives (FN) are positives items classified as negative.

Table 5.5 Template for confusion matrix

Confusion Matrix		Classified as:	
		Negative	Positive
Actual class	Negative	TN	FP
	Positive	FN	TP

From the results, there were 477 severely depressed cases, 419 mildly depressed cases, 530 moderately depressed cases and 372 ‘No depression’ cases, as detected by mental health professionals. Out of these figures, it was observed that the Bayesian networks algorithm classified 458 as severe depression, 384 as mild depression, 491 as moderate depression and 362 as ‘No depression’ cases. This represented a total of 1695 (94.3%) correctly classified cases. In the same experiment, it was observed that 19 instances were wrongly classified as severely depressed, 35 instances were wrongly classified as mildly depressed, 39 instances were wrongly classified as moderately depressed while 10 instances were wrongly classified as ‘Not depressed’ cases, making a total of 103 (5.7%) as wrongly classified cases.

Table 5.6 is a quick view of the performance of all classifiers in terms of predictive accuracy.

Table 5.6 Summary of model results in terms of accuracy

<b>Algorithm</b>	<b>Accuracy (%)</b>
<b>Bayesian networks</b>	94.3**
Naïve Bayes	93.3
KNN	92.2
SVM	92.0
C4.5	90.3
MLP	90.1
AdaBoost	49.1*

In terms of the predictive accuracy in Table 5.4, it was observed that Bayesian network classifier, with 94.3%, performed better than the other classification algorithms. This was followed by the Naïve Bayes classifier with 93.3%. It was also observed that the KNN, which comes next didn’t have significant difference with the performances of SVM. Again, the performances of C4.5 and MLP were almost the same with 90.3% and 90.1% respectively. AdaBoost had the poorest performance with 49.1%.

Table 5.7 Summary of model results in terms of Precision

<b>Algorithm</b>	<b>Precision (%)</b>
<b>Bayesian networks</b>	94.4**
Naïve Bayes	93.3
KNN	92.9
SVM	91.7
C4.5	91.5
MLP	89.4
AdaBoost	47.0*

In terms of Precision, Table 5.7 provides another quick view of the performance of all classifiers, with Bayesian networks also outperforming all others. As it was with the predictive accuracy metric, Naïve Bayes came next, followed by KNN. SVM with 91.7% didn't show any significant difference C4.5 (91.5%). MLP had a lower 89.4% while AdaBoost also had the poorest performance with 47.0%

Table 5.8 Summary of model results in terms of F-Measure

Algorithm	F-Measure
<b>Bayesian networks</b>	0.943**
Naïve Bayes	0.933
KNN	0.901
SVM	0.872
C4.5	0.924
MLP	0.856
AdaBoost	0.348*

In terms of F-Measure, Bayesian networks surpassed other classifiers having an identical performance measure (0.943) with the predictive accuracy (94.3%). Naïve Bayes also recorded higher performance. Surprisingly C4.5 produced a higher F-Measure (0.924) than SVM with 0.872. AdaBoost, also had the poorest performance with 0.348.

Table 5.9 Summary of model results in terms of RSME

Algorithm	RSME
<b>Bayesian networks</b>	0.150**
Naïve Bayes	.0175
KNN	0.372*
SVM	0.192
C4.5	0.314
MLP	0.245
AdaBoost	0.378

In terms of RSME, Bayesian networks was better than other classifiers. Naïve Bayes had better performance with full features. Although KNN recorded the poorest performance with this metric, there was no significant difference between its RMSE value and those of AdaBoost and C4.5

Table 5.10 Summary of model results in terms of Pearson correlation coefficient (R)

Algorithm	R
<b>Bayesian networks</b>	0.923
Naïve Bayes	0.935**
KNN	0.919
SVM	0.811
C4.5	0.912
MLP	0.901
AdaBoost	0.710*

In terms of Pearson correlation coefficient (R), Naïve Bayes produced a better performance value (0.935) than Bayesian networks (0.923) while C4.5 and MLP showed superiority over SVM. AdaBoost was also the poorest in performance with a low value of 0.710

Table 5.11 Summary of model results in terms of Receiver operating characteristics (ROC)

Algorithm	ROC (%)
<b>Bayesian networks</b>	92.2
Naïve Bayes	98.6**
KNN	95.3
SVM	86.6
C4.5	91.3
MLP	90.2
AdaBoost	80.3*

With Receiver operating characteristics (ROC), Naïve Bayes also had a far better performance value than Bayesian networks. Surprisingly, again, C4.5 and MLP performed better than SVM while AdaBoost had the lowest performance value.

Table 5.12 Summary of model results in terms of Running time

Algorithm	Time (s)
<b>Bayesian networks</b>	10.04*
Naïve Bayes	1.24
KNN	0.55
SVM	9.21
C4.5	0.92
MLP	0.42
AdaBoost	0.17**

In terms of running time, AdaBoost had the best performance value while Bayesian networks recorded the worst performance, closely followed by SVM.

In conclusion, of the seven (7) metrics used to measure the performance of the classification algorithms, it can be that Bayesian network classifier has demonstrated its superiority over other highly competitive state-of-the-art classifiers in this case, having the best values in four (accuracy, precision, F-Measure, and RMSE) and worst value in one (Running time). This was followed by Naïve Bayes, having the best values in two of the metrics (R and ROC). Although KNN recorded the worst performance value in one of the metrics (RMSE), it followed Naïve Bayes in performance having outperformed SVM, C4.5 and MLP in predictive accuracy, F-Measure, Precision, ROC, R. The results demonstrate a superior performance of probabilistic classifiers (Bayesian networks, Naïve Bayes, and KNN) over non-probability-based classifiers (SVM, C4.5, MLP, and AdaBoost). In all, AdaBoost showed a poor performance except for the running time, where it recorded the highest performance. The results also agree with the results of our earlier study to test the predictive strengths of several supervised classifiers (Ojeme & Mbogho, 2016a), although with smaller dataset.

## **5.6 Human Interfaces to Bayesian networks-Based expert systems**

This section provides an implementation of a knowledge-based system based on the Bayesian diagnostic framework described earlier and to test the feasibility and convenience of use by non-clinicians and non-technical persons in a non-clinical environment. This framework retains some of the most appealing characteristics of existing traditional diagnostic techniques, and improves the efficiency and effectiveness of ICT-aided depression screening. More importantly, it helps in the transfer and exchange knowledge and technology between experts and non-experts in accordance with Rogers Everett theory of Diffusion of Innovations (Rogers, 2015).

Every activity that makes up a study usually has a goal. For this study, the goal is to investigate which machine learning model would be integrated into a user interface screening tool to assist non-clinicians take screening tests and get an estimation of the presence of depression for staff and students in a Nigerian University. To achieve this goal, the study started by finding the best machine learning model. This was achieved in the preceding sections when the predictive strengths of six supervised classifiers were tested using seven metrics. Having done that, the next task was to design a user interface that the machine learning model would be integrated into. Various critical design decisions were involved. First, the software needed to be designed such that it will be intuitive and operates at its most basic and effortless level

for the users to be successful because the more intuitive, the less training required. The objective is to implement the screening component of a working Bayesian network system in a simplified graphical user interface that would assist University academic staff to understand the problem-solving process of clinicians so as to be able to determine what student information is needed to come up with a possible assessment for depression. It will guide and direct the users on what might support or refute the depression diagnosis. After answering the questions, the users can see clearly the results of the classifier.

There are several usability design methods but the User-Centred Design (UCD) (Endsley, 2016), a method and philosophy for designing and developing usable products and systems that place the user at the centre of the process (Dearden et al., 2010), was chosen. The UCD evaluation method is founded on getting user feedback in each step of the development process. Receiving such feedback involves several usability methods at any step of the design process. Several usability evaluation methods have been designed to identify usability problems. These methods have been classified differently by different studies. The work of Nielsen (1993) grouped usability evaluation methods into four general classes: automatic (involving the use of software to evaluate a user interface); empirical (involving real users to interact with a user interface); formal (incorporating the use of models to evaluate a user interface), and informal (where evaluators use rules and their skills, knowledge and experience to evaluate a user interface).

Given that the study seeks to build a usable system following a standard research methodology, the empirical approach, where the real users of the developed artefacts will interact with it for feedback, the empirical usability evaluation method was adopted. As demonstrated in Table 2.1, the most commonly used empirical methods are the Evaluator testing, User testing and Tools-based testing. Again, the Evaluators testing was chosen for the study for the same reason. Among the evaluators testing methods are the Heuristics, Pluralistic walkthrough, Cognitive walkthrough, Guideline reviews, Consistency inspections, and the Standards inspection methods (Endsley, 2016). The concentrates on heuristic evaluation method.

### **5.6.1 Design process**

The UCD approach has been widely used in medical applications (Endsley, 2016; Peischl, Ferk, & Holzinger, 2013; Vincent & Blandford, 2011) and its design process for the present study is described in four orderly methodological phases as suggested by Preece et al (Preece,

Rogers, & Sharp, 2015): user needs (establishing user requirements), developing alternative designs (design alternatives), building prototypes (prototyping) and evaluating designs (evaluating). At the top of these four phases is the task identification step where the goal of the study is defined: developing depression screening system for use by non-clinicians. The first step towards achieving this goal was to test the strength of several supervised classifiers with the aim of choosing the best in commonly used machine learning metrics.

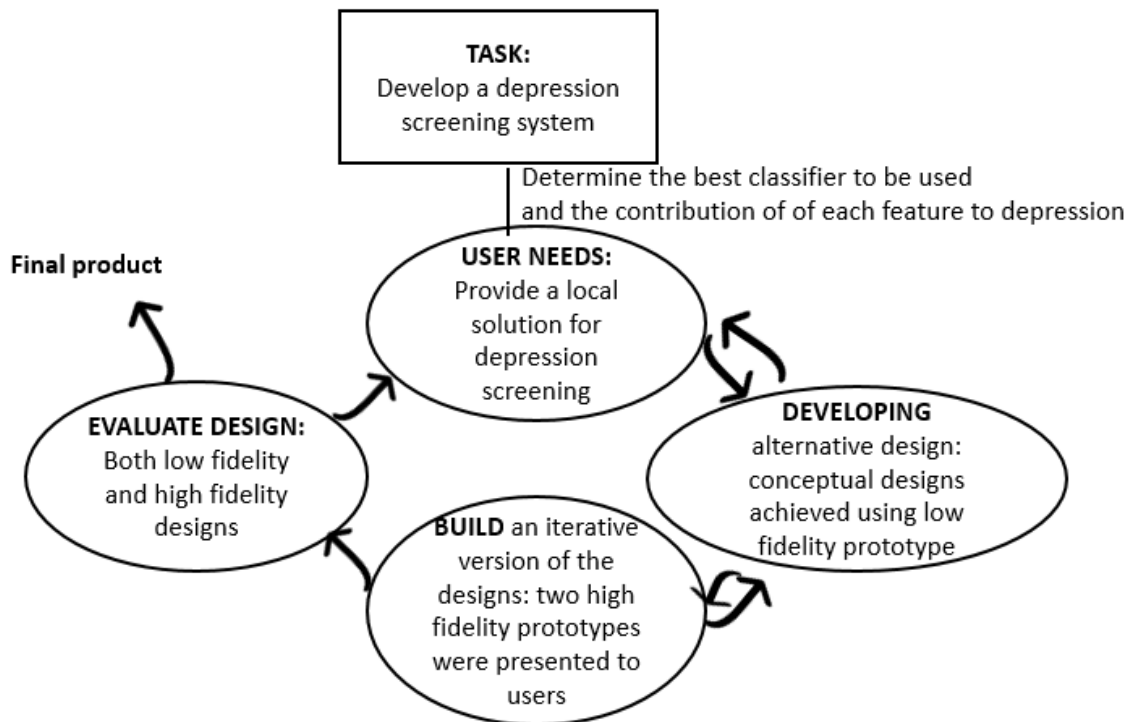


Figure 5.2 Modified user-centered design approach for the study

In section 5.7, the user requirements of the system and problem description are given in detail. In section 5.8, an architecture is designed for guiding the implementation. This is followed by a detailed description of the prototypes from the technical perspective in subsection 5.10. Section 5.10.1 discusses the evaluation of the implemented user-interface, with emphasis on part of the contribution of the study.

## 5.7 User needs, problem description and System requirements

The process started by identifying the task at hand. As well described in section 1.3, this was done during the pilot study conducted by the researcher prior to the commencement of the study. The task was duly identified through a self-administered questionnaire to a team of 29 clinicians (Table 1.2) at the University of Benin Teaching Hospital, and 73 staff (teaching and

non-teaching) and students (Undergraduates, Masters and PhD) of the University of Benin, all in Nigeria. The users of the proposed system, what they want to do with it, and the environment in which it will operate were also discussed. The requirements (inputs and outputs) of the proposed system were agreed upon by the researcher and the users. In accordance with the work of Sharma & Dhir (2016), two types of requirements were used for the study: functional and non-functional requirements. Functional requirements describe those things that the system will be able to achieve while non-functional requirements describe the attributes or constraint that the system must respect. As itemised below, these were identified during the pilot study:

### **5.7.1 Functional Requirements**

The artefact should allow direct interaction with its users.

1. The designed artefact should provide a simple way of viewing depression data for staff and students.
2. The designed system should be intuitive enough and operates at its basic form
3. The system must have a search icon for users to perform free searches, filter search requests following criteria as stated by users
4. The interface must be simple and easy to use.
5. The designed system must have an error prevention mechanism to prevent users from making wrong entries
6. The designed artefact should be able to group and display screened depression data as requested by the users, for instance; age-wise, depression grade-wise or risk-wise.
7. The designed artefact should use icons and terminologies that are easy to understand even for persons without medical, or computer background
8. The designed system must not be expensive to maintain.
9. The designed artefact should provide a possibility to make booking for depression assessment outside the routine period.
10. The designed system must motivate the users to continue using it.
11. The artefact should have different ways of visually presenting the data.
12. The designed artefact should have clear and visible display of only the important functions.
13. The interface should have distinct screens for confusable items
14. The system should have relevant staff and students-specific information
15. The system should have in-built interruptive and non-interruptive alerts

### 5.7.2 Non-functional Requirements

1. The designed artefact should be able to access publicly available yearly reports.
2. The designed artefact should be effective to allow users get instant responses to their requests
3. The designed system should conform to the Nigerian Disability Discrimination Act (NASS, 2012); hence ensures that the system is accessible to disabled users and used for all disabled staff and students as well.
4. The designed artefact should be available for academic staff and students of the University and the clinicians at the mental health unit of UBTH since the Bayesian networks model will be installed in laptops and desktop computers in all the department of the University

#### **Demand specification:**

For this study, the requirements needed to support the existence of the designed artefact are:

- A computer is the recommended device especially when displaying dashboards with a lot of content or for analysis purposes.
  - Other devices such as laptops and tablets may also be used for simple informative tasks but would not be recommended for analytical tasks that require larger viewing surfaces.

#### **Problem domain description**

The first problem addressed by the system is developing a Bayesian network model and then linking it to a user interface that estimates the degree of depression present in a staff or student and alerting the appropriate body for referral to a specialized hospital for a confirmation test and possible hospitalization and treatment. The academic staff of the University of Benin gather medical history from staff and students. At intervals of six months, each head of department in collaboration with the management staff of the University is required to submit a summary of case document to the mental health unit of the hospital for those who need further assessment and treatments.

**Medical History:** All relevant information gathered from staff and students about their illness, past and current drug treatments, significant medical history, social background and family background. Figure 5.3 is an illustration of users taking medical history from staff and students.

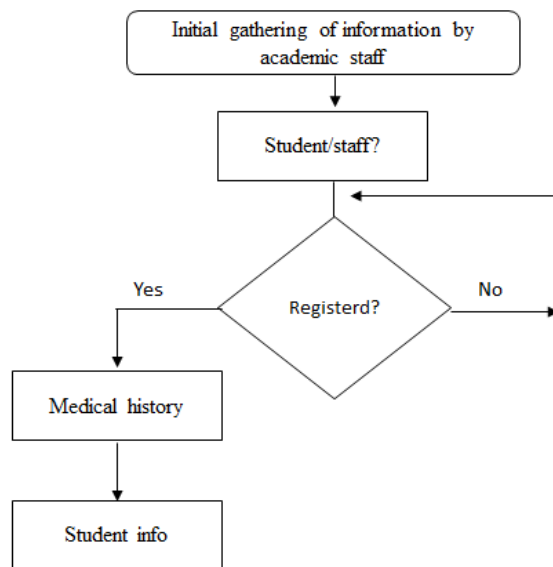


Figure 5.3 Description of users taking medical history

### 5.7.3 Developing alternative design: low fidelity prototype

After setting the requirements and problem descriptions right, the development of design alternatives comes next. Typically, this is generated from the requirements. The starting point of this phase is called conceptual design, followed by the physical design. The conceptual design, according to Preece et al (2015) represents and validates the requirements gathered, and one way of achieving it is through low fidelity prototypes. Just like the requirements phase, it is achieved by the designer working in collaboration with the users. Low-fidelity prototyping (also known as low-tech), according to Sharp et al (2011) represents simple and easy translation of design concepts into testable and tangible artefacts, collecting and analysing the user requirements at the initial stage. They help discover design issues and get them resolved at the early stage. While low-fidelity prototypes do not resemble the final product, they are simple and quick to develop. High-fidelity prototypes, on the other hand, look very much like the final product, conveyed in the same medium as the final product, and take much longer time to develop. As suggested by Preece et al (2015), three different design alternatives were created during the interactive design process. The design alternatives were made focusing on the features of the design in order to avoid wasting much time in this early stage. Figures 5.4, 5.5 and 5.6 present the appearances of the three hand-sketched alternatives.

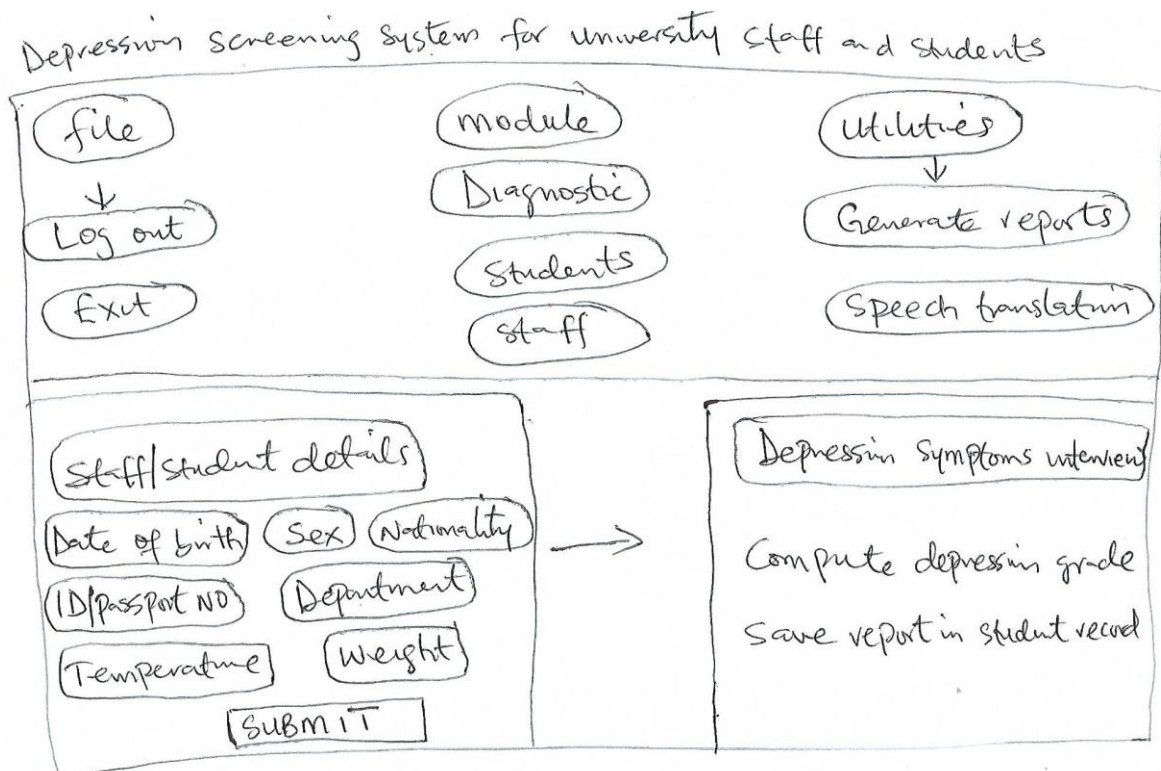


Figure 5.4 Hand-drawn sketches of design alternative 1

As shown in Figure 5.4, the first design alternative is made up of three parts: the top part which divides the interface into two equal halves, contains the list of variables (file, modules and utilities) and their drop-down menus. The lower left part contains the personal details of staff and students of the University who will be using the system. Once the personal details have been entered and the 'submit' button is clicked, the last part of the interface is displayed. This part contains the interview questions for depression symptoms

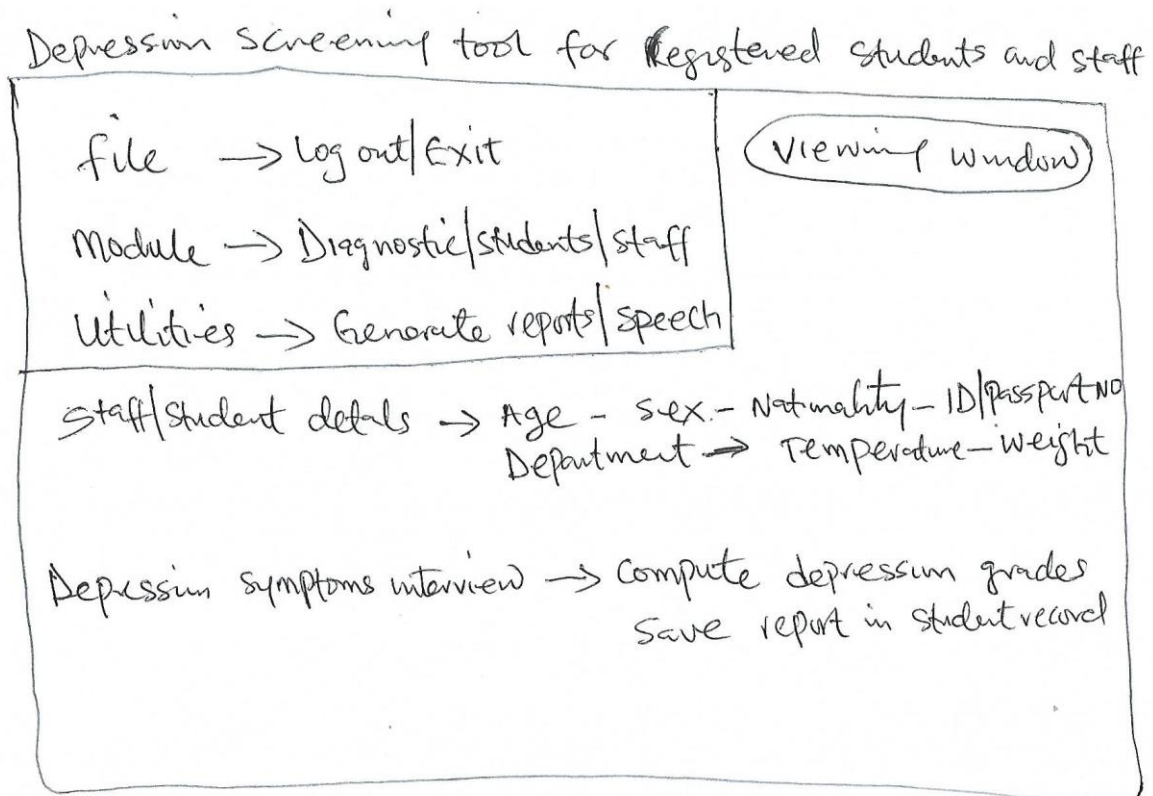


Figure 5.5 Hand-drawn sketches of design alternative 2

The design alternative 2 in Figure 5.5 shows a slightly different arrangement of the features. Though the interface is also divided into parts, the arrangement of the features are different, with the top right part dedicated for viewing any item of interest in the system simply by selecting the variable.

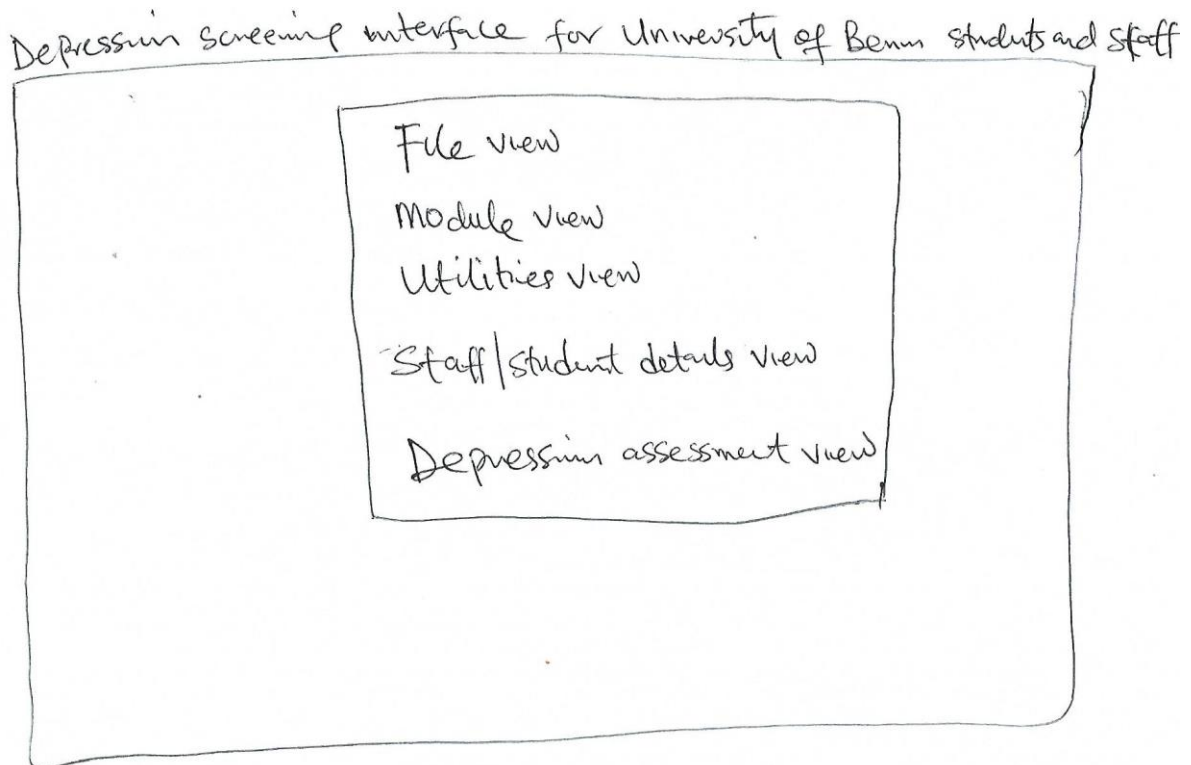


Figure 5.6 Hand-drawn sketches of design alternative 3

Much more minimalist than design alternatives 1 and 2, design alternative 3 shown in Figure 5.6 presents the five different options of viewing the features on horizontally arranged containers across the interface. This design alternative, unlike the first and second design alternatives, limited users from selecting the features before selecting the items in the drop-down menu. Items displayed on the next page, after selecting one of the five features, would depend on whether it's in staff or student mode.

It must be noted that the sketches in the three design alternatives only show the intended buttons and functions and their positioning and labelling; but none of the functions or buttons works at this initial stage of development

## 5.8 Description of system architecture

After analysing the results obtained from the users, consisting of academic staff and students of the University of Benin, some good features from the three design alternatives were combined into one design concept. Among the academic staff were HCI experts, medical professionals (with domain knowledge in mental health) and other end-users. This merging led to the design of two prototypes for the decision support system in the study. The first prototype was the high-fidelity design, and after the users provided feedback, the researcher adjusted the design features into the final prototype.

The development uses the information technology (IT) that is close to the users' skill, experience and understanding, which is an important requirement in ICT4D project. This was expected to improve the process of depression assessment, which is typically a challenging task. The resulting artefacts from the design process presents several user interfaces that allow for depression assessment in the most basic way, send alert to mental health professionals in high risk cases, generate and use database report. This approach is a remedy for the systems in the literature in which the users (mainly clinicians) were presented with screening tools that suffer from poor consideration for user group to human-interface design and clinicians actual work practices (Bee, Brooks, Fraser, & Lovell, 2015; Lawrence-Jones, 2010; Molin, 2016; Wiseman et al., 2013; Zakaria & Ghani, 2013).

Designing easy-to-use interfaces for depression assessment makes it easy to navigate through the database and enables retrieval of any information that would have been hard to manage due to the presentation and amount of information. Graphical interfaces make data more accessible since the academic staff and students can see the exact information they are authorised to see using just a few clicks. Some example cases will therefore be displayed and will demonstrate that it is easier to explore information through simple natural interactions. Additionally, staff, students and other possible users will have easier time to go through their diagnostic reports and recommendations to mental health professionals for further tests and treatment plans.

A simple overview of the first prototype (high fidelity) implementation is presented and discussed in section 5.9 while the second prototype is discussed in section 5.11. They were developed with the programming language, Java. It allows users a model for simulating interaction of depressive disorders with the staff and students, and has a data collection interface. To facilitate the implementation, many existing tools and application programming interfaces were used. Many of them are open-source software, but some are commercial software.

## **5.9 Audience**

The design methodology described in the study aims to add new knowledge to the existing mental health knowledge bases used by mental health interest groups, public health researchers, industrial psychologists, and many others. Though the graphical interfaces were specifically designed to be used by non-clinicians and non-technical University of Benin academic staff as a first step in regularly assessing their staff and students for depression and thereafter

scheduling appointments for those that need help, it could eventually be useful and available for use by the following groups:

### **5.9.1 Clinicians**

The research is expected to establish some models and tools that would be of assistance to clinicians in screening for depression. The objective is to build an expert system that combines the human expertise and the technology intelligence to achieve more accurate disease identification. This system may assist clinicians in decision making and double check clinician's assessment (evidence-based diagnosis).

### **5.9.2 Medical students**

For medical students, the resulting prototype opens up possibilities of learning as it can be used as an electronic learning system by describing tasks in which the data could be used, for instance, connect types of depression, and the study outcomes to do research even at the student level. User interface with good data visualization functionalities will allow a medical student to explore relationships in the data to combine different representation forms and to understand relationships which otherwise could be tedious to find.

## **5.10 Alerts in decision support application**

Interruptive and non-interruptive alert systems are important parts of decision support application (Pevnick, Li, Grein, Bell, & Silka, 2013) (McCoy et al., 2012), and has thus been given a consideration in the design of the present system. Interruptive alerts are those that disturb the workflow and demand a response or an action by the clinician before proceeding. It could be a pop-up with the warning on which the clinicians actively have to do something with in order to proceed. These interruptive alerts for the more serious warnings and their purpose are to force the clinician to recognize a possible hazardous situation. Non-interruptive alerts are for the less serious warnings and they do not disturb the workflow of the clinician due to them appearing as info-buttons or links on which you can choose to open or not. It is preferable today to have interfaces that eliminate the need for interruptive alerts by guiding the user instead, since too many interruptions can cause irritation amongst the users (Payne et al., 2015). In the study, Payne et al (2015) had focused on the user interface design of both interruptive and non-interruptive alerts, conducting a series of meetings with a group of experts

consisting of people with a background in either medicine, informatics or computer interface design to look at the design of drug-drug interaction (DDI) alerts. The result of the study was a list of components that a DDI alert should include:

1. Drugs involved
2. Signal word indicating the level of seriousness
3. Clinical consequences
4. Mechanism of the interaction
5. Contextual information/modifying factors
6. Recommended action(s)
7. Evidence

Horsky et al. (Horsky et al., 2012) have, aside from the list of general user interface design recommendations, provided specific suggestions for alerts and reminders. These recommendations were compiled in the same targeted review. Some of the recommendations they provide specifically for alerts are:

1. Interruptive alerts should be reserved for the two or three highest levels of severity (severe and moderate depression, in this case)
2. Content of the alert should be limited to 1-2 lines with a justification separated by white space.
3. Alert prioritization
4. Meaningful colour coding
5. Revise trigger rules

An extensive literature review on this is found on Horsky et al (2012)

### **5.11 Application interface: High-fidelity prototype**

The merged selected features in the three design alternatives were re-designed and implemented in the PCs and, laptops and tablets in each department at the University of Benin and the mental health unit of the University of Benin Teaching Hospital (UBTH). This was done in two prototypes: the high-fidelity (first) prototype and the final (second) prototype).

The high-fidelity prototype, a simple, easy-to-use graphical application interface, was developed with the widget toolkit, *Java swing*. This was needed to facilitate interactions between the academic staff and students of the University. Given that all University staff (teaching and non-teaching) and students have predefined profiles with the University, the first screen that appears on clicking the application on the desktop is the login page (Figure 5.7),

and access rights to the software are given based on roles. Access rights are given to both staff and students. Once the “ok” button is clicked, the system searches through the permanent database of the University staff or student list, comparing the username and password entered with the ones already on the database. Upon successful login authentication, the system takes the user to the main interface page (Figure 5.7). If, however, the username (staff or student number) or password fields were empty, or they were unregistered, or were incorrectly entered, a relevant error message was displayed.



Figure 5.7 Application interface login page

The main interface consists of three menus (Figure 5.8): the *file*, *modules* and *utilities*. The *file* menu has the *logout* and *Exit* buttons. The *modules* menu consists of three modules: the *diagnostic module*, *student module*, and *staff module*. The *Utilities* menu consists of the *Reports* and *Speech Setting* (Figure 5.9. )

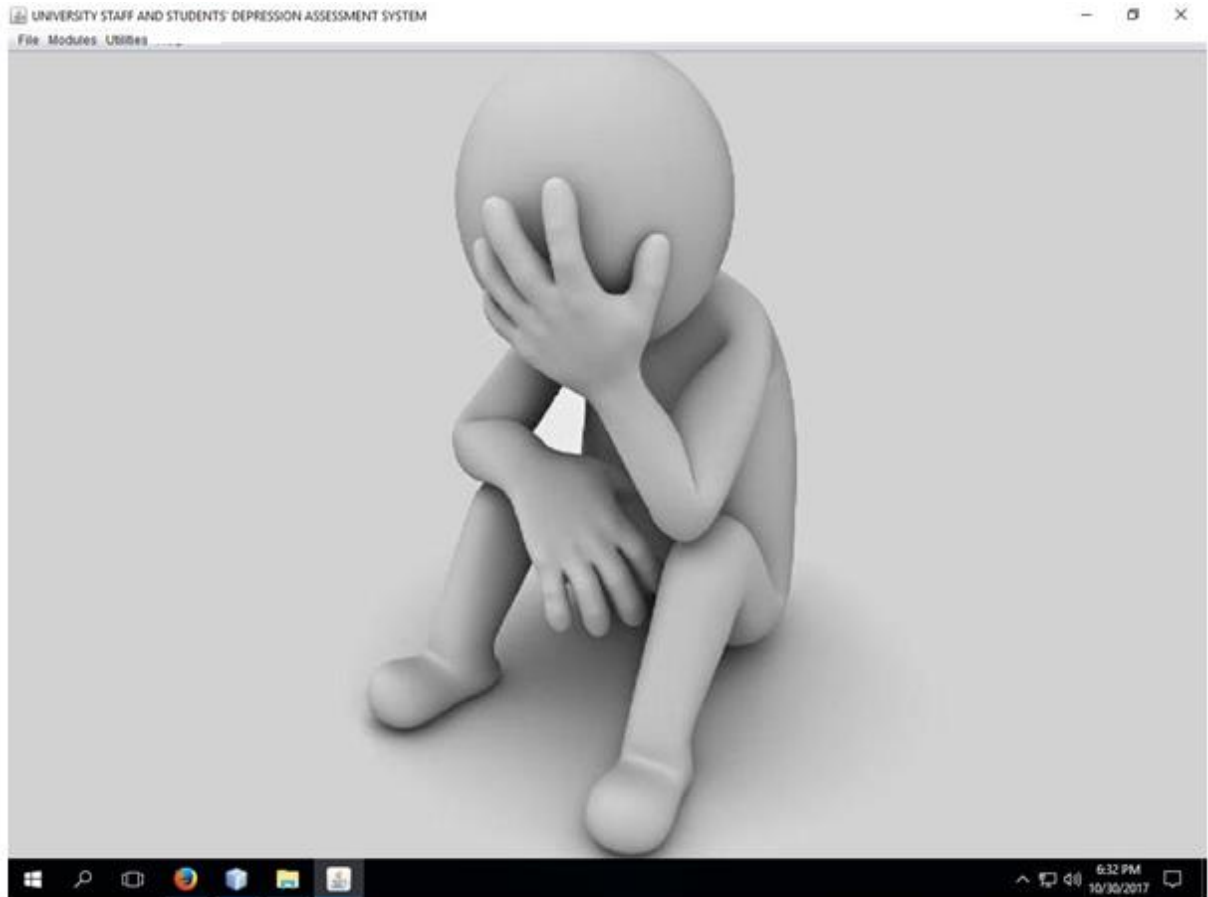


Figure 5.8 Four menus of the main application interface

Each of the modules starts with a request for a username and password before gaining access to the module content. A click of the diagnostic module, for instance, prompts the user for another authentication.

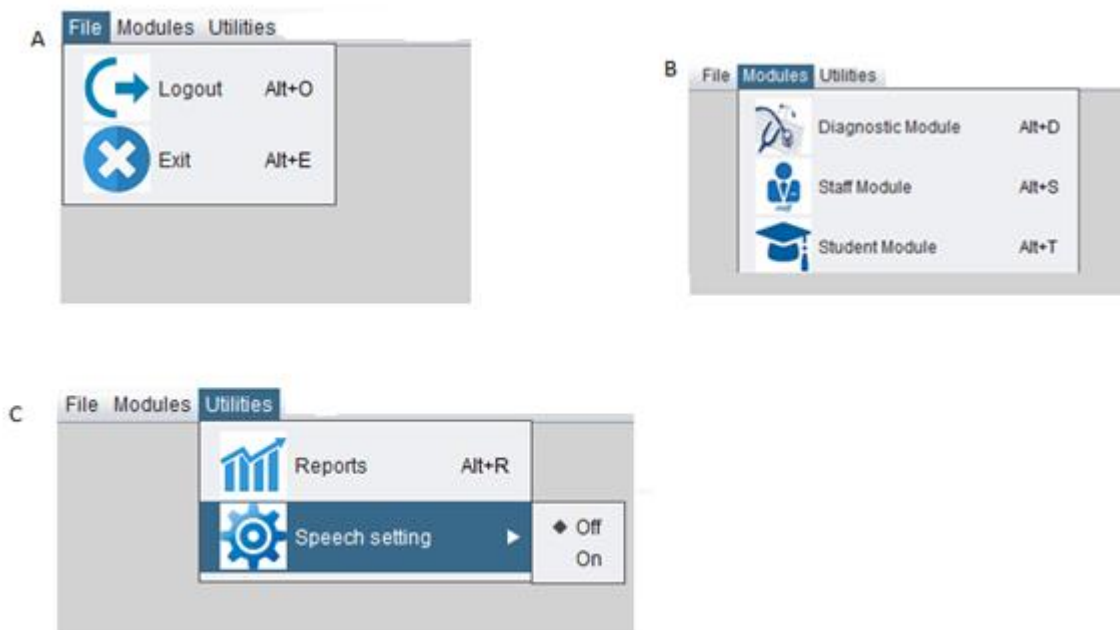
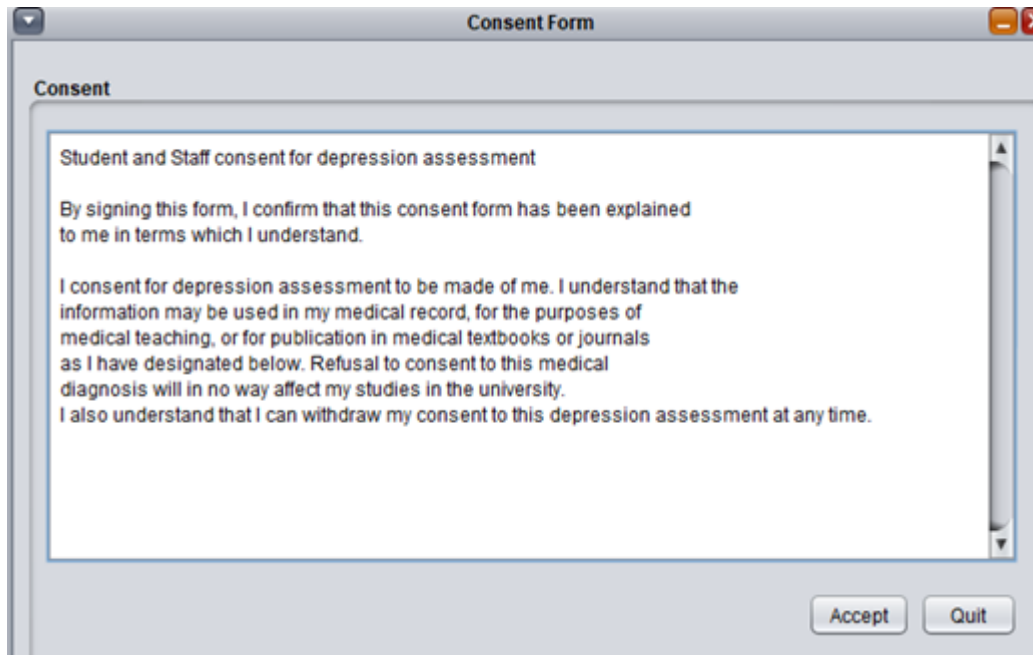


Figure 5.9 Contents of the (A)file, (B)modules, and (C) utilities

The *diagnostic module*, which is where the actual screening data are entered and processed, is only accessible to academic staff. It starts with a user login page, prompting the academic staff to enter his/her name, staff number and password. This was developed with the *Java authentication and authorisation service* (JAAS) framework, a pluggable Java package for providing information security service to authenticate and enforce access controls on users. The username and password are the same as the predefined official staff number with which the University communicates with staff. This login page serves two purposes: 1) to ensure that only authorised users are allowed into the diagnostic module and 2) to keep track of the number of users and also to uniquely identify each user for the purpose of computer log. The users' data were stored in a secure central server of the University of Benin.

Upon successful login to the diagnostic module, the academic staff explains the purpose of the system to the staff or student for assessment before continuing. The system then prompts to go the next screen, the *Consent page* (Figure 5.10). The user's login state was retained unless they logged out.



Consent Form

Consent

Student and Staff consent for depression assessment

By signing this form, I confirm that this consent form has been explained to me in terms which I understand.

I consent for depression assessment to be made of me. I understand that the information may be used in my medical record, for the purposes of medical teaching, or for publication in medical textbooks or journals as I have designated below. Refusal to consent to this medical diagnosis will in no way affect my studies in the university.

I also understand that I can withdraw my consent to this depression assessment at any time.

Accept Quit

Figure 5.10 Consent page

In the consent page (See details in Appendix I), the purpose of the system is explained to the staff or student in very clear terms. If he/she does not agree to the assessment, the system gives a “**THANK YOU**” message. At this time, the user halts the process by clicking on the *quit* button and takes the next person for assessment. But if he/she agrees to the assessment, the user clicks *the Accept* button and the system goes to the staff and students details page (Figure 5.11), where some basic data is collected.

The image shows a web form titled "Student/Staff Details". The form has the following fields and values:

- ID: Ifolfo001
- Name: (empty)
- Surname: (empty)
- Date of Birth: 17 Nov 2017
- Sex:  Male  Female
- ID/Passport Number: A123425675
- Department: Computer Science
- Year of Study: (empty)
- Nationality: Nigeria
- Marital Status: Seperated
- Student Weight (kg): 60
- Student Temperature (°C): 35
- Date Collected?: (empty)

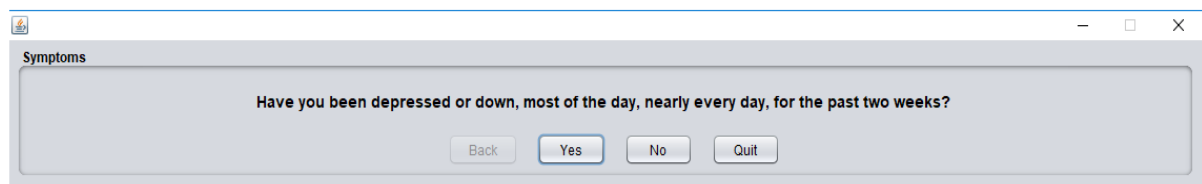
A "Submit" button is located at the bottom of the form.

Figure 5.11 Staff and student details page

Because of the anticipated heavy usage of the system and the possibility of having two or more students with the same name, the data is indexed on name, student number and ID/passport number. Given that the University already has basic details of every staff and student, the user only takes readings of his/her weight, as the other details are captured from the University database upon entering his/her number. If a staff/student says he/she has been assessed previously, the system goes through the permanent database and displays details of already assessed persons. If there is no one by those details in the list, it is then forced to treat this person as a new staff/student for assessment. If, however, the details do exist in the list, the system prompts the user to confirm with the student. He does this by presenting his/her details in the form "Blessing Ojeme with student number ojmbles001 and passport number A04607041 has been found. He is 27 years old, weighs 70 kilograms, and the last time assessed was on 15<sup>th</sup> May 2017". This helps to prevent the user from entering multiple information for the same student.

The 'submit' button takes you to the next page where the symptoms of staff students are collected. In order to ensure fulfilment of all criteria for depression as defined in DSM-5

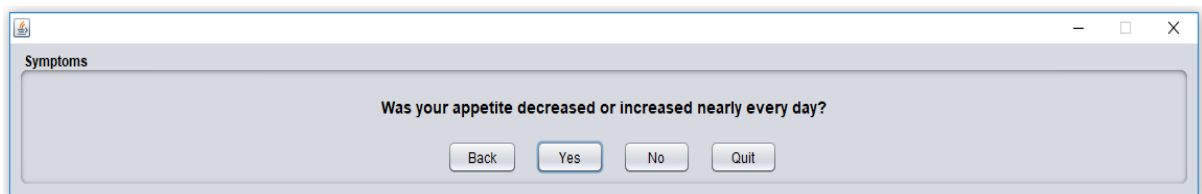
(*Diagnostic and Statistical Manual of Mental Disorders fifth edition*, n.d.) and ICD-10 (WHO, 1992), interview questions and answer categories were derived from the study dataset, using the same datatypes, and the degree of presence and absence of symptoms. This also agrees with several studies of depression in the medical literature (Allgaier, Pietsch, Frühe, Sigl-Glückner, & Schulte-Körne, 2012; Arrieta et al., 2017; Eisenberg et al., 2007; A. K. Ibrahim et al., 2013; Semrau et al., 2015).



Symptoms

Have you been depressed or down, most of the day, nearly every day, for the past two weeks?

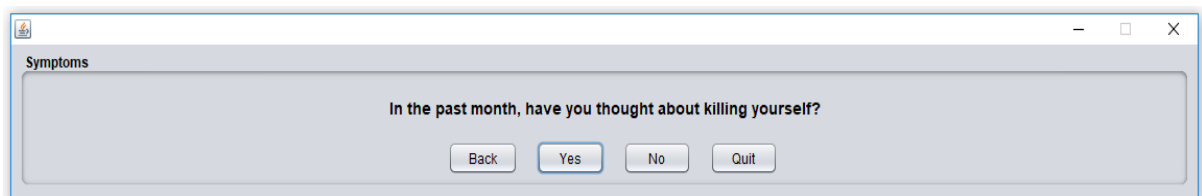
Back Yes No Quit



Symptoms

Was your appetite decreased or increased nearly every day?

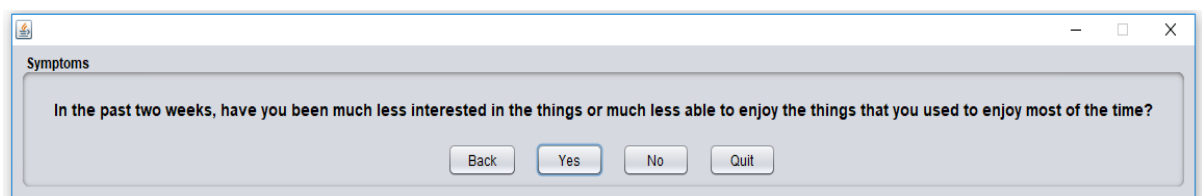
Back Yes No Quit



Symptoms

In the past month, have you thought about killing yourself?

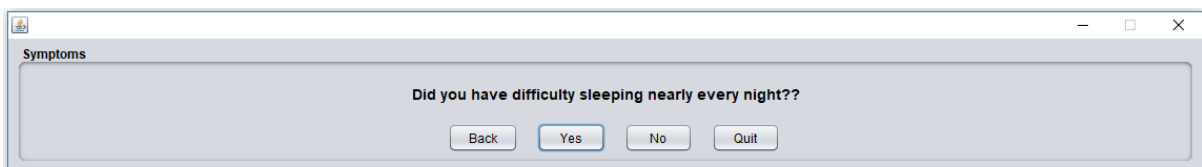
Back Yes No Quit



Symptoms

In the past two weeks, have you been much less interested in the things or much less able to enjoy the things that you used to enjoy most of the time?

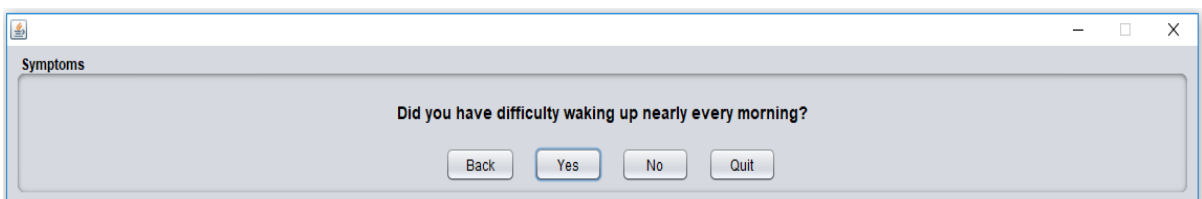
Back Yes No Quit



Symptoms

Did you have difficulty sleeping nearly every night??

Back Yes No Quit



Symptoms

Did you have difficulty waking up nearly every morning?

Back Yes No Quit

Symptoms

Did you have difficulty sitting still or restless almost every day?

Symptoms

Did you move or talk more slowly than normal?

Symptoms

Did you feel tired or without energy almost every day?

Symptoms

Did you feel worthless or guilty almost every day?

Symptoms

Did you have difficulty concentrating, or be easily distracted almost every day?

Symptoms

Did you have difficulty making decisions almost every day?

Symptoms

In the past month, did you repeatedly think that you would be better off dead or wish you were dead?

Symptoms

Did you have difficulty maintaining relationship with friends or perform work efficiently for the past two weeks?

Symptoms

Did your weight increase without trying intentionally by about 5% or 8lb/3.5kg, for a 160lb/70kg person in a month?

Symptoms

Did your weight decrease without trying intentionally by about 5% or 8lb/3.5kg, for a 160lb/70kg person in a month?

Symptoms

Did these symptoms cause significant problems at home, at work, socially, at school, or in some other important ways?

Symptoms

Have you been under serious financial pressure for the past two weeks?

Symptoms

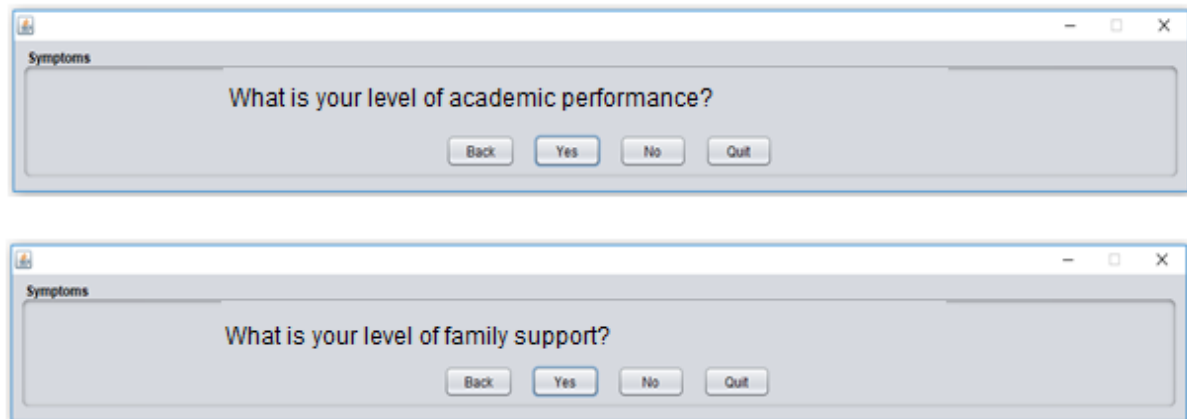
Does any of your family members suffer from depression for the past three months?

Symptoms

What is your level of cigarette smoking?

Symptoms

What is your level of alcohol?



After each student has answered all the relevant questions, a notification of depression status is displayed. The results will display the grade of depression. The user interface is linked to a Bayesian networks engine that controls which questions are asked and what interpretations are given at the end of the screening exercise.

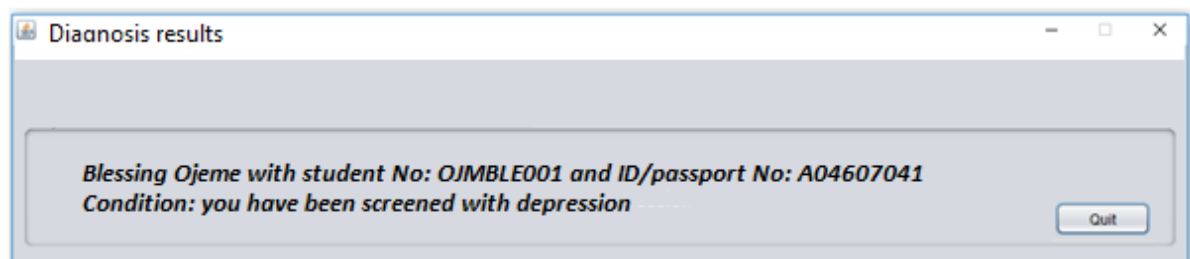


Figure 5.12 Screening result

The *staff module* contains diagnosis and referral records of all staff of the University. Access right to this module is given to all staff. These records are indexed on staff number and password so that a staff can only have access to his/her records. The records in this module are write-protected

Similarly, *student module* contains diagnosis and referral records of all students of the University. Access right to this module is given to all students. These records are indexed on each student number and password so that a student can only have to access his/her records. The records in this module are also write-protected.

In the *utilities menu* are Reports and Speech Setting modules for generating reports and activating explanatory sounds. Also in this menu are specified the modalities for alerting clinicians at the hospital for further depression assessment (for moderate and severe cases),

modifying records of staff and students, adding and deleting users. Access right to this module is given to the management staff, consisting of heads of departments/units, Deans of faculties, the Vice Chancellor, his deputy, and registrar of the University. This module also serves as the *storage* of the application in the sense that after the successful authentication and authorisation in each module, a log of the profile of the staff or student, login and logout date and time, and a summary of activities during the period is loaded and stored in this module. Each of these records can be opened by clicking on a related menu. This opens a screen with a list of all the saved programs. On clicking the required program, the user is asked if they want to load the program or to delete it. When the user clicks on the required program it is loaded back to the main interface.

### **5.11.1 Evaluation of the first prototype**

Detailed evaluation of medical diagnostic tools, an important pre-requisite to their routine use in the clinical settings, helps to demonstrate that the efforts (1) lead to needed change in the right direction, (2) contribute to positive results in different parts of the system, and (3) require more efforts to bring a process back into satisfactory ranges (Varkey, Reller, & Resar, 2007). This is based on the confidence that good performance replicates good-quality practice, and that comparing performance among users and providers will encourage better performance. To this end, the developed Bayesian network model in the study was implemented in an easy-to-use software system for the assessment of staff and students for depression. In an attempt to optimise user-friendly interface and utility in a busy University setting, the screening tools sought to use few parameters which are consistent with a reasonably high ease of use, ease of learning, user satisfaction, error prevention and efficiency of the interface (Oteniya, Cowie, & Coles, 2005; Peng, Ramaiah, & Foo, 2004). As mentioned earlier, only prototypes of a screening tool (not a full-fledged patient/clinically-tested system) were developed, running on personal computers, laptops and tablets, at the University of Benin and the UBTH. It is hoped that this work will lay the foundations necessary for the professional development and deployment of such a full-fledged system, and also serve as a reference model for other government and non-government agencies. Thus, there was no total evaluation or clinical testing of the prototype with real patients. Evaluation was done through heuristic and intended end-user performance, where the designed screens were put to work in an audit to explore its practical value. The methods employed in evaluating the screening tool prototype in this study are discussed in section 5.11.2

### **5.11.2 Heuristic evaluation**

. In this approach, several evaluators assess the user interface to confirm if it conforms to a set of usability principles ('heuristics'). Nielsen (1993) identified a set of ten usability heuristics which were: visibility of system status, match between the system and the real world, user control and freedom, consistency and standards, error prevention, recognition rather than recall, flexibility and efficiency of use, aesthetic and minimal design, helping users to recognise, diagnose and recover from errors, and help and documentation. Despite the criticisms of these Nielsen's ten heuristics as being too general, the study is motivated by the successes it has recorded in medical diagnosis (P. H. Lilholt, Hæsum, & Hejlesen, 2015; Z. Tang, Johnson, Tindall, & Zhang, 2006).

### **5.11.3 Heuristic Evaluation Method**

Heuristic evaluation is a usability evaluation method developed by Nielsen and Molich (Nielsen & Molich, 1990). The heuristics evaluation method is based on a set of heuristics that guide evaluators through the interface to uncover potential problems (Yuan, Finley, Long, Mills, & Johnson, 2013). The guidelines offer a range of heuristics available, however these were grouped into ten (10) Jacob Nielsen's Heuristics (Table 5.13), to facilitate the evaluation process. These heuristics were based on earlier heuristics and guidelines for evaluating medical applications together with guidelines developed to evaluate them (Molin, 2016). The evaluators, using the heuristics, examine the interface and identify usability problems that users might encounter while interacting with an interface. Then they suggest recommendations to improve the usability of the interface. A set of comprehensive heuristics specific to medical devices was developed.

Through this method, the evaluators are able to judge whether an interface conforms to a set of usability principles. For this study, the goal of the usability testing was not to select a winner design alternative from the three design alternatives discussed earlier, but to get feedback from users regarding what they found as good or bad design features (identify requirements, listed functions, and operations the system must perform), which is necessary to refine the design. This is the time when a combination of end-users and heuristic evaluation experts were involved as evaluators of the designed artefacts whose valuable feedback helped refine the final prototype, as suggested by several researchers (P. Lilholt et al., 2015; Yen & Bakken, 2009). Since the designed artefacts are to be used by both novice and expert users, this testing tried to

find out how easy it was for users to accomplish the tasks the very first time they interact with the interfaces and whether it was easy to learn the operation of the artefacts after the initial instructions. This was necessary in order to keep the designed artefact easy to learn and spare users the effort of learning everything from the start every time they interact with the system as suggested by Wilkund et al (2013).

Table 5.13 Jacob Nielsen's usability heuristics

Heuristic	Question
H1 - visibility of system status	Are users informed about system progress through progress indicators or messages? The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.
H2 - Match between system and real world	Does the system cater for users with no prior experience of electronic devices? Does the system use user's language, with familiar words, phrases and concepts, rather than technical terms? The system should speak the users' language, with words, phrases, and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.
H3 - User control and freedom	Can users do what they want to do freely? Can the users navigate the screens with ease? Users often choose system functions by mistake and will need a clearly marked exit strategy to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.
H4 - Consistency and standards	Do the screen objects and actions have the same meaning and effect in different situations? Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.
H5 - Error prevention	Would users make mistakes they would easily avoid with a better system? Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.
H6 - Recognition rather than recall	Are the screen objects, actions, and options visible enough? Are the functionality of the screen objects obvious from labels? Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember

	information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.
H7 - Flexibility and efficiency of use	Does the system allow for both expert and novice users? Accelerators—unseen by the novice user—may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.
H8 - Aesthetics and minimalist design	Do the screens contain information which is irrelevant or rarely used? Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.
H9 - Help users recognise, diagnose and recover from errors	Are error messages expressed in simple language? Do they accurately describe the errors, and suggest solutions? Express error messages in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.
H10 - Help and documentation	Is appropriate help provided? Is the information easy for the user to find, understand and use? Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, be focused on the user's task, list concrete steps to be carried out, and not be too large.

#### 5.11.4 Evaluation procedure

The evaluation session began with a three-hour PowerPoint presentation and multimedia projector workshop in form of a seminar where the researcher welcomed the evaluators, read the test script that explained the objectives of the study, introduced the prototype, standard application of the Nielsen heuristics, the usability problems associated with each one of them to the evaluators, and the user's right to withdraw from the session at any time. It was also

explained to the evaluators that that they would be observed and that their screens would be recorded using screen capture software during the session. A 20-minute break was observed after each 45 minutes of presentation. It was then agreed that each screen be discussed (in the order of appearance) by each evaluator using one of the most popular, simple-to-use and low-cost methods, heuristic evaluation (de Carvalho, Évora, & Zem-Mascarenhas, 2016; Nielsen, 1993) for the detection of the heuristics violated, usability problems, their severity and the exact location of the problems. Solutions were also proffered for problems found. Where there were disagreements among the evaluators, it was discussed until an agreement was reached.

The evaluators were then asked to read and sign the consent form, after which a pre-test questionnaire was given to each user to fill out in order to obtain information regarding his/her background, expertise and experience. The evaluators went through the interface features in the application along with information about its functionality, objectives and standard terminology. This gave them a better understanding of what to evaluate. Each evaluator went through and analysed his/her prototype at least twice. The first run ensured that the evaluators were familiar with the interface, while the second run allowed the evaluators to search for usability problems using Jacob Nielsen's Heuristics (1993) and analysing severity on a scale of zero to four (Table 5.14). Each evaluator was given a Heuristic Evaluation Form (Appendix J) to enter their findings. A comment section is also included in the form.

Table 5.14 Usability scale based on usability evaluation

Severity	Meaning
0	I don't consider this a usability problem at all
1	A <b>cosmetic</b> problem only; need not be fixed unless there is time available
2	A <b>minor</b> usability problem; fixing this problem should be given low priority.
3	A <b>major</b> usability problem: it is important to fix and should be assigned a high priority.
4	Usability <b>catastrophe</b> : must be remedied prior to releasing the software

It is important to state that the severity of a usability problem is a combination of the three factors below (Joshi et a., 2009):

1. The **frequency** with which the problem occurs; whether it's rare or common

2. The **impact** of the problem if it occurs; whether it will be easy or difficult for the users to resolve?
3. The **persistence** of the problem; whether it's a one-time problem that users can resolve once they know about it, or repeatedly be concerned by the problem.

### 5.11.5 Recruitment of participants

Table 5.13 (graphically represented in Figure 5.9) summarises the panel of 62 clinicians and academic staff comprising both males and females who participated in the experiments to inspect the interfaces in search of problems and suggest ways to improve them. They were selected from twelve departments at the University of Benin and the teaching hospital for the evaluation of the developed system and the workflow. The selection was done using an intentional method, in which formal invitation was through emails and participation through a confirmation, as recommended in Dumas (1996).

Among these 62 evaluators, 56 reported to use PCs, laptops and tablets regularly while the remaining 6 reported that they use computers and laptops sporadically and with help from a third person, 6 were usability experts (experienced using different usability systems for at least six months on a monthly basis), 10 had domain knowledge in the medical and mental health with at least 5 years work experience. Considering the primary goal of the study, which is to develop predictive tool to support a heterogenous user population (University academic staff) in assessing their staff and students for depression, and that the artefacts would be used by all evaluators, such a high number and mix of panel of evaluators, consisting of both expert and novice users were deemed appropriate, as suggested by Wilkund et al (2013) To minimize the number of evaluators who would not turn up for the evaluation sessions, two approaches were taken: 1) recruitment and selection was conducted as close as possible to the time of the evaluation; and 2) evaluators were given incentives in the form of N1000 (one thousand naira) per hour or provision of lunch. This means that evaluators participated in the study by choice and could withdraw at any time.

Table 5.15 Panel of evaluators

<b>Evaluators</b>				
<b>Faculty</b>	<b>Department</b>	<b>Number</b>	<b>Male</b>	<b>Female</b>
Sciences	Computer science	5	4	1
	Zoology	5	2	3
Arts	Music	5	2	3
	French	5	2	3
Education	Science education	5	3	2
	Edu Adm	5	2	3
Social sciences	Marketing	5	2	3
	Economics	5	3	2
Health	Mental health	6	4	2
	Medicine	6	5	1
Engineering	Electrical engineering	5	5	0
	Petroleum engineering	5	5	0
<b>Total</b>		<b>62</b>	<b>39</b>	<b>23</b>

The University of Benin was selected for three reasons: 1) a pre-study survey had been done at the institution in which questionnaire was administered to staff, students and clinicians, as part of the motivation for the study 2) there had been prior contact with the respective heads of departments, lecturers and students at the University. Additionally, the study data was collected from the mental health hospital unit of the same University. The total number of evaluators depended on the inclusion and exclusion criteria composition of the required expertise, and availability of evaluators who confirmed their participation in the experiment.

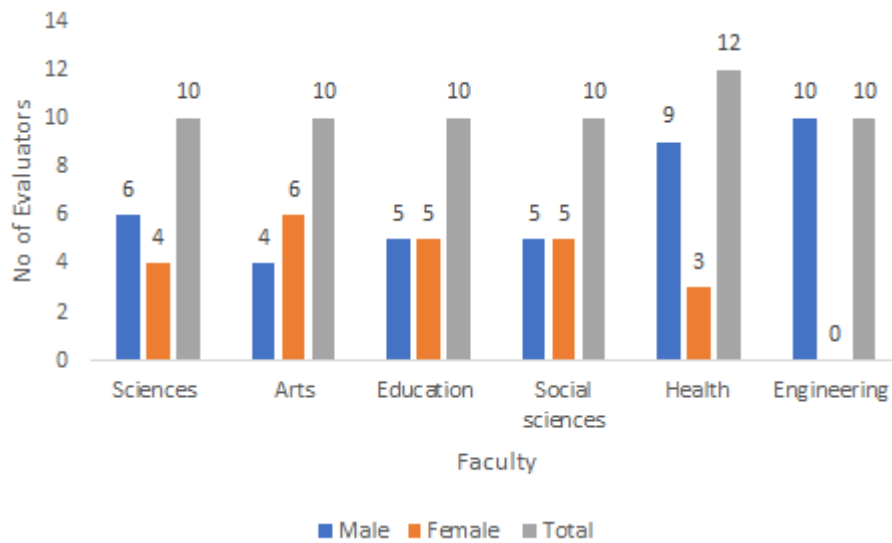


Figure 5.13 Graphical representation of panel of evaluators

### 5.11.6 Data collection method

As in many ICT4D projects, the research questions influenced the choice of data collection methods for evaluation. For instance, the fourth research questions (What machine learning techniques would be appropriate to screen for depression among University staff and students?), was addressed by testing and analysing the predictive strength of various machine learning techniques. The interesting result obtained, which agrees with those obtained in section 5.4.2, was evaluated with real clinical depression dataset and standard machine learning metrics, and published (Ojeme & Mbogho, 2016a). To address the second and third research questions, heuristic evaluation forms were administered to collect qualitative feedback from users on the suitability (convenience, ease-of-use) of the chosen evaluation method. A quantitative approach was used to assess the usability of the developed prototypes.

Having described the evaluation method that were used to assess the usability of the screening prototypes in the study, the usability problems that were identified and the suggestions that were made on how the usability of the prototypes could be improved upon are presented. These identified usability problems are summarised, categorised and explained. Recommendations and suggestions for each of these problems are also explained and presented together for each of the problems (Table 5.16).

Table 5.16 Usability problem description and location in the first prototype

User interface design principle	Violation	Score	Recommendation
<b>Visibility of system status:</b>	Some operations took a long time and no progress indicator was shown. The user does not have to be unsure as to whether or not anything was happening with the system or if he should try to click the function again	10	Include progress indicator to help users monitor the progress of some selected function
<b>Match between system and the real world</b>	Some of the words were not easy for me to understand because I don't have a good knowledge of depression and depression cases. The system doesn't say what happens after detecting a high risk for depression	7	Include an alert system that keeps popping for moderate and severe depression cases
<b>User control and freedom</b>	It's frustrating having to completely start over again to change an answer. Again, it should be possible to log off users' activities in the application No return and Home buttons on each page of the application making it difficult to navigate	7	-users should be able to change an answer once it has gone down the list of answered questions - The <i>return</i> to the <i>Home</i> page should be made visible on each page since they are used very often.
<b>Consistency and standards</b>	There was no consistency in the use of standard boxes and buttons and this made it somewhat difficult to navigate without any prior knowledge of the system	6	The consistency of the design should be improved upon

<b>User interface design principle</b>	<b>Violation</b>	<b>Score</b>	<b>Recommendation</b>
<b>Error prevention</b>	I found it difficult to retract a request. The features appear to be too close in colour thereby making it difficult to see the difference	7	Include a means whereby users can retract or cancel any request that has been made. Also separate the colour spectrum of the features so that it doesn't get confusing to users.
<b>Recognition rather than recall</b>	It was difficult for me to recognise the generate report of already screened students and staff.	9	Make the tab for 'generate report' much more recognisable so that the system can generate report from different users.
<b>Flexibility and efficiency of use</b>	The systems do not have any accelerators for expert users. The slowness of the system made it feel a little inefficient in some parts.	2	Include keyboard shortcut for all features
<b>Aesthetic and minimalist design</b>	The application design is simplistic and It's not attractive or appealing enough to impress the users -it is confusing having all questions in separate pages	7	Questions should be on same page but diagnosis should appear on a separate page. Use beautiful icons and graphics on the home page to capture users' attention. These should reflect the purpose of the application

<b>User interface design principle</b>	<b>Violation</b>	<b>Score</b>	<b>Recommendation</b>
<b>Help users recognize, diagnose, and recover from errors:</b>			
<b>Help and documentation</b>	A major problem is that no Help or documentation buttons	9	Make Help and documentation buttons more visible

Problems were categorised into nine sections according to nine usability problem areas and their corresponding sub-areas. Table 5.17 shows a summary of the number and severity of the heuristics violations in the first prototype.

Table 5.17 Number and severity of heuristic violations

Heuristic violated	severity					Total Count of usability problem
	0	Cosmetic	Minor	Major	Catastrophic	
System status visibility	2	2	3	3		10 (16%)
Match btw system interface and real world	1	3	1	2		7 (11%)
User control and freedom		2	3	2		7 (11%)
Consistency and standards		2	2	2		6 (9%)
Error prevention	1	2	3	1		7 (11%)
Recognition rather than recall	2	3	2	2		9 (14%)
Flexibility and efficiency of use			2			2 (3.1%)
Aesthetics and minimalist design	2	2	1	2		7 (11%)
Help and documentation	3	1	2	1	2	9 (14%)
Total by heuristic	11	17	19	15	2	64 (100)

## 5.12 Ethical considerations

Ethical considerations denote a moral attitude that are involved when conducting research to achieve not only high professional standards of technical procedures, but also respect and protection for the study participants consenting to be studied. Professional ethical standards should be maintained during all phases of research study (Cacciattolo, 2015; Lyon & Segal, 2013).

The works of Cacciattolo (2015) and Raudonis (1992) exhaustively described some fundamental ethical considerations in medical research. This include: confidentiality, anonymity, integrity, voluntary participation and consent requirement. Confidentiality concerns how the data, especially personal data, is handled and (Raudonis, 1992). In this study, the issue of confidentiality has mainly been achieved by not having personal information connected to the participant's responses, so even if access is accidentally gained to the data, it is impossible to deduce who the participant is. A second fundamental ethical concern in qualitative research is anonymity. The study had little ethical concerns regarding anonymity and it ties into what was previously stated, that staff and students identities anonymized immediately and only connected to a given ID. It was also made sure that the anonymity and confidentiality were clearly stated in the consent form (appendix B) and also explained to them by the researcher. Personal integrity is another ethical consideration in qualitative research, and it concerns the protection of the participant by not intruding on their privacy (Stevens, 2013). Due to the high level of anonymity, participant's privacy was protected. Again, the study did not target a vulnerable population and the topic did not concern anything sensitive which reduced the intrusion of the participant's privacy. The fourth ethical consideration is voluntariness. This has been addressed in the study by explicitly stating in the consent form that participation was voluntary and that participants had the option to withdraw from participation at any time. The last ethical consideration for the study is the consent. This principle states that those who participate in a study should be fully informed about the purpose and contents of the research (Raudonis, 1992). A consent form, as earlier mentioned, with general information about the study was presented prior to the session and it was mandatory to sign it in order to participate. Thus, no data was collected and shared without the consent of the participants.

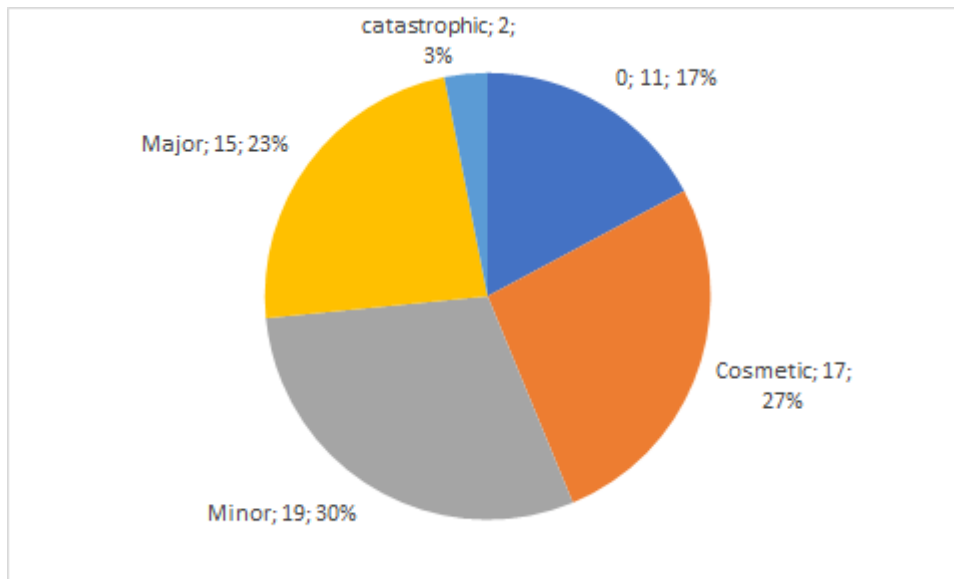


Figure 5.14 Count of usability problems and their severity

### 5.13 Results and analysis of first prototype evaluation

The results of Table 5.15 show that out of the ten (10) Jacob Nielsen (1993) usability heuristics, nine (9) violated heuristics and sixty four (64) usability problems with their corresponding severity were found. Table 5.14 shows the location of the problems and suggested solutions where necessary. The results show that “System status visibility” was the most often violated, with 10 (16%) usability problems found, followed by “Recognition rather than recall” and “Help and documentation”, with 9 (14%) usability problems found in each one. “Match between system interface and real world”, “User control and freedom”, “Error prevention”, and “Aesthetics and minimalist design” with 7 (11%) usability problems found in each one come next, followed by “Consistency and standards” with 6 (9%) usability problems. The least usability problems were found in “Flexibility and efficiency of use” with only 2 (3.1%) in the minor category. It is important to note that no problems were found in the “Help users recognise, diagnose and recover from errors” heuristics by the evaluators. It was only in the “Help and documentation” heuristics that severity grade 4 (catastrophic) was found. That was an indication of a serious violation that required serious attention. For visualisation purpose, these violations were graphically represented in simple pie chart in Figure 5.11. Other problems found by the evaluators were a few grammar errors, typos and insufficient explanation of some of the screens. These were corrected in the second version of the prototype application.

It is the belief of the researcher that all usability problems found in the study and the corresponding severity were important and could negatively affect the effectiveness of the final prototype. To that end, after discussing the problems identified and suggested solutions with the evaluators, they were all taken seriously, prioritised and implemented in the second prototype.

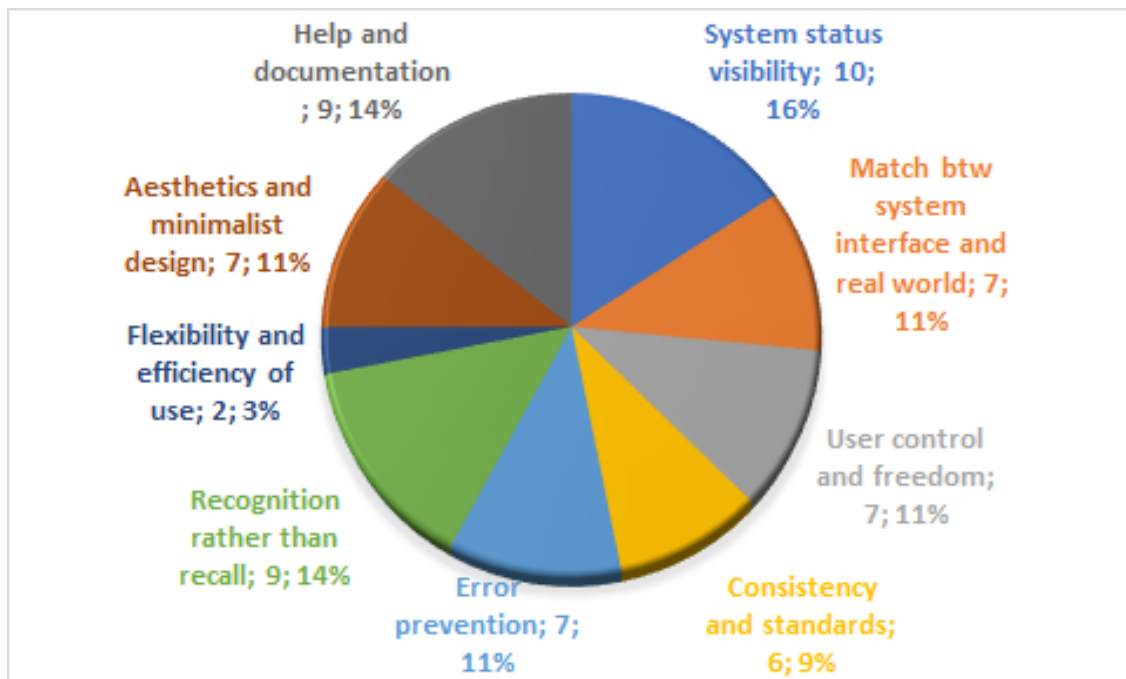


Figure 5.15 Graphical illustration of heuristic violation

## 5.14 Application interface: second prototype

In this section, a description of the second prototype showing some modifications, which resulted from feedback from the first prototype, are given. The details of the results that led to the modifications are discussed in section 5.10. The main interface, login, page, personal details of staff and student page, and depression assessment pages are still present as in the first prototype but the modifications to them are described in section 5.14.1.

### 5.14.1 Modifications on the main interface

The main interface of the second prototype, as shown in Figure 5.15, contains four menus: *file*, *modules*, *utilities* and additional *Help* to help users search for any information on the application. The *file* and *utilities* menus contain the number of modules as in the first prototype but *modules* menu consists of four modules (as against three in the first prototype): the

*diagnostic, student, staff* and the additional *Appointment* module (Figure 5.17). Additionally, an audio icon for explaining every item on the main interface and the screening processes was provided on the task bar, for the visually impaired user. The task bar also houses shortcuts for the diagnostic module, staff module, student module, appointment module and the report module (Figure 5.18). Also included in the main interface are the *Return* and *the Home* buttons.



Figure 5.16 Four menus and shortcuts of the main application interface

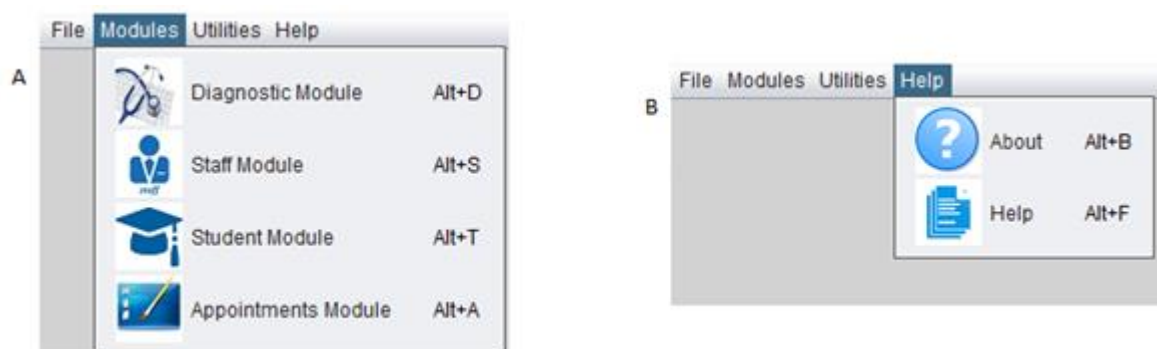


Figure 5.17 Modified diagnostic and file menus

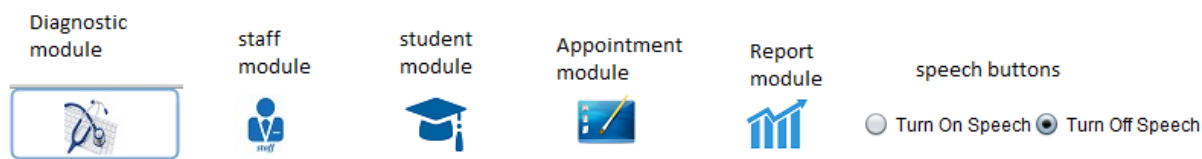


Figure 5.18 Shortcut of modules on the taskbar

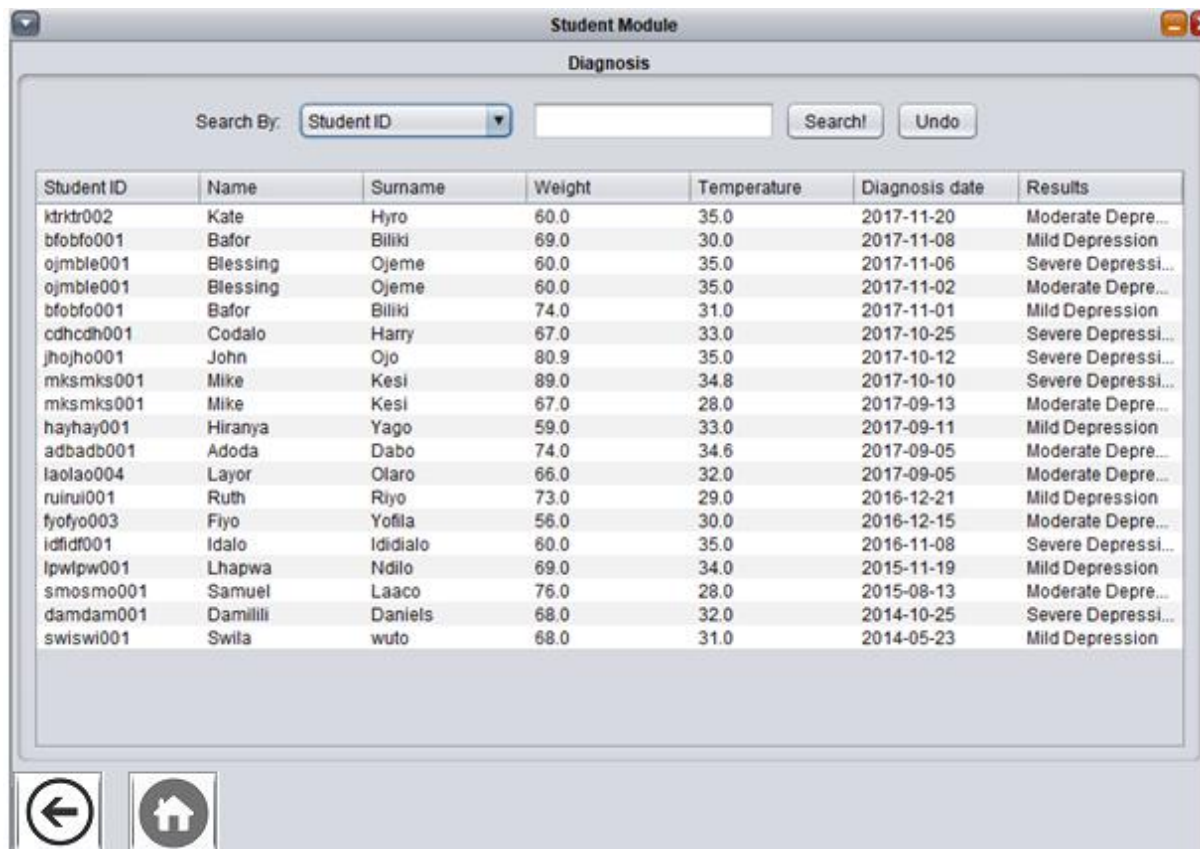
Staff and students make request for depression assessment using the request Appointment Form (Figure 5.19)

Figure 5.19 Request for assessment

### Report generation

Those authorized to generate the reports of already assessed staff and students are the HOD and management staff. Each HOD can generate report of staff or students who have been assessed in his/her department. A sample of such report showing student ID, name, surname, weight, temperature, date of assessment and the diagnosis is shown in figure 5.19. Aside the six months interval for the general depression assessment in the University, a staff or student can make a request for depression assessment whenever the need arises. Such request can be made using the Request Appointment Form (see Figure 5.18) in the student and staff module. It is the responsibility of HODs to attend to such pending requests from staff and student in his/her department fixing a date for the assessment and then notify the persons who made the requests. Meanwhile, the management staff is the only body entitled to generate report of every

staff and student in the University and send alert to the mental health unit of the University teaching hospital for further test for those screen with moderate and severe depression



Student ID	Name	Surname	Weight	Temperature	Diagnosis date	Results
ikrkr002	Kate	Hyro	60.0	35.0	2017-11-20	Moderate Depre...
bfobfo001	Bafor	Biliki	69.0	30.0	2017-11-08	Mild Depression
ojmble001	Blessing	Ojeme	60.0	35.0	2017-11-06	Severe Depressi...
ojmble001	Blessing	Ojeme	60.0	35.0	2017-11-02	Moderate Depre...
bfobfo001	Bafor	Biliki	74.0	31.0	2017-11-01	Mild Depression
cdhcdh001	Codalo	Harry	67.0	33.0	2017-10-25	Severe Depressi...
jhojho001	John	Ojo	80.9	35.0	2017-10-12	Severe Depressi...
mksmks001	Mike	Kesi	89.0	34.8	2017-10-10	Severe Depressi...
mksmks001	Mike	Kesi	67.0	28.0	2017-09-13	Moderate Depre...
hayhay001	Hiranya	Yago	59.0	33.0	2017-09-11	Mild Depression
adbadb001	Adoda	Dabo	74.0	34.6	2017-09-05	Moderate Depre...
laolao004	Layor	Olaro	66.0	32.0	2017-09-05	Moderate Depre...
ruirui001	Ruth	Riyo	73.0	29.0	2016-12-21	Mild Depression
fyofyo003	Fiyo	Yofila	56.0	30.0	2016-12-15	Moderate Depre...
idfidf001	Idalo	Ididialo	60.0	35.0	2016-11-08	Severe Depressi...
lpwlpw001	Lhapwa	Ndilo	69.0	34.0	2015-11-19	Mild Depression
somosmo001	Samuel	Laaco	76.0	28.0	2015-08-13	Moderate Depre...
damdmd001	Damilili	Daniels	68.0	32.0	2014-10-25	Severe Depressi...
swiswi001	Swila	wuto	68.0	31.0	2014-05-23	Mild Depression

Figure 5.20 Report of students already assessed

Following recommendations from evaluators on the first prototype, the diagnostic module was also modified to accommodate screening questions in one page (Figure 5.21, A, B, C). These questions had appeared on different pages in the first prototype.

Symptoms Symptoms Symptoms

Have you been depressed or down, most of the day, nearly every day, for the past two weeks?

No  Yes

In the past month, have you thought about killing yourself?

No  Yes

In the past two weeks, have you been much less interested in the things that you used to enjoy most of the time?

No  Yes

Did you have difficulty sleeping nearly every night?

No  Yes

Did you have difficulty waking up nearly every morning?

No  Yes

Was your appetite decreased or increased nearly every day?

No  Yes

Did you have difficulty sitting still or restless almost every day?

No  Yes

Did you move or talk more slowly than normal?

No  Yes

Did you feel tired or without energy almost every day?

No  Yes



 

Figure 5.21A Screening questions

Symptoms Symptoms Symptoms

Did you have difficulty concentrating, or be easily distracted almost every day?

No  Yes

Did you have difficulty making decisions almost every day?

No  Yes

In the past month, did you repeatedly think that you would be better off dead or wish you were dead?

No  Yes

Did you have difficulty maintaining relationship with friends or perform work efficiently for the past two weeks?

No  Yes

Did your weight increase without trying intentionally by about 5% or 8lb/3.5kg, for a 160lb/70kg person in a month?

No  Yes

Did your weight decrease without trying intentionally by about 5% or 8lb/3.5kg, for a 160lb/70kg person in a month?

No  Yes

Did these symptoms cause significant problems at home, at work, socially, at school, or in some other important ways?

No  Yes

Have you been under serious financial pressure for the past two weeks?

No  Yes

Does any of your family members suffer from depression for the past three months?

No  Yes

Did you feel worthless or guilty almost every day?



 

Figure 5.21B Screening questions

The screenshot shows a web form with a header containing three 'Symptoms' tabs. The main content area contains four questions, each with four radio button options: 'none', 'low', 'medium', and 'high'. The 'none' option is selected for all four questions. At the bottom, there is a 'Submit Symptoms' button and two navigation icons: a left arrow and a home icon.

Symptoms Symptoms Symptoms

What is your level of cigarette smoking?  
 none  low  medium  high

What is your level of alcohol or other drug consumption?  
 none  low  medium  high

What is your level of job satisfaction and/or academic performance?  
 none  low  medium  high

What is your level of family support and/or availability of accommodation?  
 none  low  medium  high

Submit Symptoms

Figure 5.21 C Screening questions

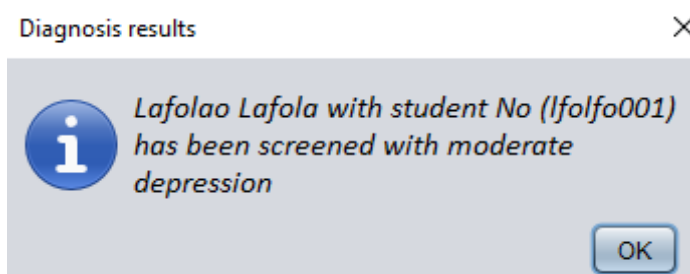


Figure 5.22 Screening result

### 5.15 Evaluation of the second prototype

It is interesting to note that the same number of evaluators were used during the second prototype. This was because of the mutual agreement during the training that all feedback from the first prototype would be corrected and included in the second prototype. Before the commencement of the evaluation the researcher had a one-hour training session with the users based on some of the changes that were made. At the completion of the given tasks, the

researcher collected and collated the feedback from the evaluators. These were summarized and presented in Table 5.18

Table 5.18 Summary of feedback from evaluators on second prototype

<b>User interface design principle</b>	<b>Results from validation</b>
<b>Visibility of system status</b>	The system had good information about what pages you were on. I was able to complete the tasks without difficulties. The simple and intuitive design made it easy to learn how to use the tool.
<b>Match between system and the real world</b>	The application was easy to understand even though I did not know anything about the field of study; the language was easy and similar to reality. Concepts appeared familiar to the users. The application is easy to use and generates report. - what the different menu lead to were mostly understood. -It was a good idea that students and staff could make request for assessment
<b>User control and freedom</b>	-It was good that the return and Home buttons were in all pages of the application.
<b>Consistency and standards</b>	-Consistent positioning of the navigation buttons was very good. -East to know where you are. -Good use of standard buttons and names made navigation easy even without any prerequisite knowledge of the study. - Terminologies and icons used were clear and easy to understand. -I think the prototype was consistent, easy to follow how the different parts were connected and easy to navigate through.
<b>Error prevention:</b>	I didn't encounter any errors, so I can only assume that the application serves the intended purpose.
<b>Recognition rather than recall:</b>	Instructions for use of the system were easy to find in the help menu accessible to anyone. Again, the speech buttons made explanation of the functionalities easy for visually impaired users
<b>Flexibility and efficiency of use:</b>	-When I've done the tasks more than once I easily learned the way of doing it. -If I had done different tasks one more time I think I would be going pretty much faster. It's so good!
<b>Aesthetic and minimalist design:.</b>	-Simple and beautiful design -Very clean design. It made me find the information I wanted to see very fast. - The design presents information in an aesthetically pleasing manner -A very simple design with great functionalities.
<b>Help users recognize, diagnose, and recover from errors:</b>	-Didn't encounter any errors -Did not experience any errors
<b>Help and documentation:</b>	-It was clear how to use it, and the Help button helped -I don't think extra documentation is needed if the user gets a proper introduction to the system and knows the specific terms. -The documentation and way to understand different functions is much easier in this system than in the first prototype.

The results and analysis from the first and second prototypes above underscore the importance of end-user evaluation of user interface in terms of strengths and weaknesses. In an efficient and effective manner, it has uncovered several usability violations that would have reduced the smooth interaction between users and the application and lead users of the application to incorrect interpretation of the action desired. The evaluation has also demonstrated that the application was functional and fits into the tight schedules of the University academic staff. The overall heuristic results suggest that timely and reliable feedback system of evaluation is a demonstration of the significance and relevance of end-user factors in the process of designing computer-aided diagnostic intervention, especially with respect to acceptance of the system for use in non-clinical environment. This was evident in the final prototype, which demonstrated much better ease of use and greater ease of navigation, while minimizing the errors earlier identified.

### **5.15.1 Practical implication of the second evaluation**

The practical implications from the evaluation of the second prototype is of particular interest to researchers, software developers, decision-makers, and innovation managers. With respect to perceived impact, majority of the participants were convinced that the developed depression screening tool was a good complement for the existing methodology as it could potentially resolve many of the challenges faced by the existing system. The feedback from many of the users also suggest that more consultations with users of the application within the mental health unit of hospitals during the development process could help designers and innovation managers to be able modify the innovations so that they are perceived as more impactful, which will make them much more likely to adoption. Though not part of the specification in the present design, the participants also suggested that decision-makers, designers and innovation managers must be proactive towards unforeseen events that may have negative effects on adoption intentions and decisions. This means that, for innovations that are currently on the innovation-decision process, it is important for designers and decision-makers to know how mental health users perceive and behave towards new applications or innovations. Regardless of whether an adoption-decision will certainly be made in some cases, this will still be an important implication for adoption decisions that are important to the user. By being familiar with the motivation behind the mental health users, decision-makers might respond appropriately in order for the adoption to be a success. For mental health unit, the implications are particular important since an application or innovation that has the potential to improve

screening or diagnosis of diseases in some way, might be rejected due to the way its attributes are perceived by the users. An application or innovation that is rejected also means a huge financial loss to both the designers, decision-makers and users.

### **5.15.2 Strengths and weaknesses of heuristics evaluation method**

Although heuristic evaluation has the potential to identify many usability problems, it is often criticised by many researchers for two reasons. The first is its heavy reliance on the expertise of the usability professionals who conduct the evaluation (P. H. Lilholt et al., 2015; Z. Tang et al., 2006). These experts may lack domain knowledge and could therefore overlook important domain-related usability problems which would be detected by end-users of the application. In this study, measures put in place to overcome this important limitation was to ensure that the heuristic evaluation was conducted by the same end-users, many of who have broad knowledge of HCI, machine learning systems and healthcare, as suggested by several usability studies (P. Lilholt et al., 2015; Ribeiro, Martins, Queirós, Silva, & Rocha, 2015; Z. Tang et al., 2006; Wilkund, M.E; Kendler, J;Yale, 2013; Yen & Bakken, 2009). For instance, 6 of the evaluators were usability experts, 10 had domain knowledge in the medical and mental health with at least 5 years working experience, and 56 of them were regular users of PCs, laptops and tablets. Another limitation of the heuristic evaluation method is its major focus on the weakness of the technology. Again, the use of end-users as evaluators in this study made it possible to identify some strengths of the application. For instance, despite the usability problems identified in the first prototype, almost all the evaluators in the second prototype confirmed its simplicity of design and ease of use, and that, generally, it was easy and convenient for those unfamiliar with technology to use the application for the purpose for which it was designed.

## Chapter Six

### Summary, achievements, conclusions and future work

#### 6.1 Summary

The motivation and driving for this study was to contribute towards finding workable local solutions to the growing threat of depression among University student population in Nigeria. The study started by establishing that depression is highly neglected (Abas et al., 2014) and is increasingly threatening public health in Nigeria, and that the major cause of the problem was the shortage of mental health professionals and facilities in Nigeria, and whose remedy was not anywhere at sight. The study also established through a survey with clinicians, staff and students of a Nigerian University that it was only reasonable and logical to seek alternative solutions to this problem since the current resources for managing depression hasn't made the desired impact in Nigeria.

Having established that the predictive strength of Bayesian networks was better than other supervised classification algorithms by testing the performance with standard metrics, the study proposed a collaboration of Bayesian networks models and HCI for use by University academic staff, as a first step, for screening of staff and students for depression, and where necessary, schedule appointment with the appropriate mental health authority. The combine Bayesian networks and HCI model was again evaluated through heuristics evaluation. Errors were identified by the team of evaluators, who were also the intended users of the tool, and the screening tool was refined until the performance reached an acceptable level. Though not tested with real clinical patients, the results from the study proved that it could complement the current method for managing detecting in Nigeria. The results also demonstrated the strength of such computer-aided diagnostic intervention, especially with respect to acceptance of the system for use in non-clinical environment.

#### 6.2 Strengths and limitations of study

##### 6.2.1 Strengths

The major strength of the study comes from the choice of methods and methodology. The study made use of several data collection techniques (questionnaire and interviews) and methods (machine learning and HCI) so as to obtain a firm basis for evaluation and add to the study's novelty. These are summarised thus:

- a) **Close involvement of researcher to all aspects of study:** Although the work of Denscombe (2010) notes the difficulty of replicating and finding similar results in mixed (quantitative and qualitative) study, the close involvement of the researcher in all aspects of the study ties into the reproducibility of the research. Great efforts were made to provide a fairly detailed description of the research processes in the research method chapter, in order to allow for other researchers to replicate the study as much as possible
- b) **Combined abilities:** Besides being easy to use by both experts and non-experts, the new system takes advantage of the strengths of the intended end-users, both in number, knowledge, and the reasoning capability of the system to make depression screening, quickly, thereby addressing the challenge of access to depression screening tools.
- c) **Knowledge transfer/exchange:** For the purpose of this study, knowledge transfer and exchange is conceptualised as the process by which the University staff and students' assessment system provides a virtual platform to share knowledge and decision models, and to facilitate the collaboration in a distributed way. This knowledge exchange and transfer from experts to non-experts is impossible with the traditional manual screening system. This is in line with the technology transfer, understood from the perspective of diffusion of innovation theory (Rogers, 2015)
- d) **Saved cost of consultation:** The new depression assessment system has the potential to save staff and students the extra cost of consultation and long waiting queues in hospitals if fully implemented and used.
- e) **Involvement of clinicians as confirmatory evaluation:** Another strength of the study comes from the use of clinicians as a complement in the heuristic evaluation to validate the methodology and accuracy of the processes. This was necessary because even though some of the evaluators are experienced in usability, it is complex to determine the accuracy of medical applications without subject matter expertise. This is supported by the work of Yuan et al (2013) where the study evaluated a clinical decision support system and found that combining evaluators with expertise in usability with evaluators with expertise in the specific domain was most effective in a heuristic evaluation of healthcare IT.

## 6.2.2 Study limitations

Though the developed systems in the study make it possible for non-medical personnel to do a first-line depression screening, it does not replace or duplicate many of the abilities of the currently used traditional manual diagnostic system model, as described in chapter four. In other words, the study has made vital contributions for managing depression in Nigeria through a local workable solution to complement the efforts of available resources in Nigeria. However, despite these strengths, there are some potential limitations and shortcomings in both the methods and how well they will be applied. These are highlighted thus:

- a) **The decision models were constructed with limited data:** The first study limitation of study was the performance of the knowledge-based system due to the availability of only a limited depression dataset ( $n = 1798$ ) for the construction of the models. Closely related to the first was that the dataset was unbalanced in distribution ('No depression', 372), ('mild depression', 419), ('moderate depression', 530) and ('severe depression', 477). Even though the stratified cross validation method was used to ensure a representation of each category, this could have affected the accuracy of the predictions made by the classifiers.
- b) **No trust mechanism in the knowledge transfer/exchange:** The new depression assessment system does not include any trust mechanism which can measure or control the transfer or exchange of unreliable knowledge.
- c) **Reluctance of mental health professionals to incorporate system into daily clinical sessions:** The interesting findings of the researcher during the pre-study survey about "knowledge and use of ICT tools in work place" by mental health professionals have been well discussed in section 1.3.2.3. This is in total agreement with compelling evidences in medical diagnosis literature, which clearly show that incorporating full computer-managed diagnostic system into routine clinical workflow was difficult. Among other reasons, clinicians argue that the process of medical decision making is an art rather than science and maintained that the reasoning ability of the new system cannot emulate the experience and skill of mental health professionals.
- d) The statistical population discussed in the study is only limited to the staff and students of one University in Nigeria. So, more future studies are needed to test the

reproducibility of the experiments with more data samples for the generalization of the findings to any other populations.

It is important to state that within the context of these limitations, the study achieved its key objectives, highlighted in section 6.3.

### **6.3 Key achievements of the study and contributions**

The goal of the study was to investigate the extent to which ICT4D frameworks can support the methodologies for screening for depression among Nigerian University staff and students. The specific objectives required to achieve the stated goal are summarised as follows:

1. To identify depression screening challenges faced by clinicians and then isolate those that that can be solved using ICT4D applications
2. To identify the extent to which ICT4D applications are used for routine disease screening in selected healthcare institution in Nigeria
3. Investigate the strengths of various classification techniques for the purpose of developing new models.
4. Address an important issue related to the most important predictors and relationships between the predictors
5. Integrate the knowledge-based framework into a graphical user application, which can assist University academic staff in the regular assessment of staff and students for depression
6. Investigate the effectiveness of supporting University academic staff in assessing their staff and students for depression using ICT4D application framework

In the course of the study, the researcher believed that the study goal and objectives, as summarised above, have been met. A breakdown of how this was done, leading to the key contributions of the study are thus described.

The contributions presented in this doctoral study provide compelling evidence of the huge potential that ICT4D framework (using machine learning and HCI methodologies) has for complementing available resources in the risk of depression at the level of the Nigerian University Community. As a common practice in computer science, some results of the study were published in scientific conference proceedings, resulting in the under-listed peer-reviewed publications, which are part of the contributions required for the completion of the

study. In particular, the predictive strength of Bayesian networks for diagnosis of depressive disorders was published (Ojeme & Mbogho, 2016a). An effort to understand the interactions between depression and co-existing physical disorders led to the development of various multidimensional machine learning models. The interesting results obtained were also published (Ojeme & Mbogho, 2016b). After more dataset was collected, another study revealing more reasoning ability of Bayesian networks in depression after all redundant features were removed using an unsupervised learning algorithm, the principal component analysis (PCA) was published (Ojeme et al., 2016).

1. **Paper A (Springer Publication).** Blessing Ojeme and Audrey Mbogho (2016). Predictive Strength of Bayesian Networks for Diagnosis of Depressive Disorders. *Proceedings of the 8th KES International Conference on Intelligent Decision Technologies (KES-IDT 2016)-Part II*, Tenerife, Spain, June 15-17, Vol. 56, Pp. 373-382. Available at:  
[http://link.springer.com/chapter/10.1007/978-3-319-39630-9\\_31](http://link.springer.com/chapter/10.1007/978-3-319-39630-9_31)
2. **Paper B (Elsevier Publication).** Blessing Ojeme and Audrey Mbogho (2016). Selecting Learning Algorithms for Simultaneous Identification of Depression and Comorbid Disorders. *Proceedings of the 20th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES-2016)*. York, UK, September 5-7, Vol. 96, Pp. 1294-1303. Available at:  
<http://www.sciencedirect.com/science/article/pii/S1877050916319846>
3. **Paper C (IEEE Xplore Publication).** Blessing Ojeme, Audrey Mbogho and Thomas Meyer (2016). Probabilistic Expert Systems for Reasoning in Clinical Depressive Disorders. *Proceedings of the 15th IEEE International Conference on Machine Learning and Applications (IEEE ICMLA'16)*, Anaheim, California, USA, December 18-20, Pp.599-604. Available at  
<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7838209>

However, the study takes advantage of the longitudinal and lengthy nature of the PhD study format to combine the foundations laid by the publications and expand their contents.

These contributions are summarized in twofold: with respect to predictive modelling, it is the development, evaluation and implementation of Bayesian networks to clinical datasets for

predictive modelling. With respect to real-life applications, it is the integration of computationally-managed diagnostic models into an easy-to-use screening tool in non-clinical settings and the interpretation of them. The contributions are as follows:

1. Empirical evidence of the workability of ICT4D framework as seen in the machine learning and HCI collaboration in building a depression screening tool. This was achieved in the literature and in chapter 5
2. Empirical evidence of which methods and techniques for revealing complex interdependencies and relations between (1) symptoms of depression (attributes), and (2) attributes and targets. This was achieved in section 5.5.
3. Extensive evaluation of the methods. machine learning methods were evaluated using real clinical datasets from a University population and standard metrics. The effectiveness of combining machine learning and graphical screening tool was evaluated using heuristic evaluation (sections 5.10.1 and 5.13), and the result obtained from users' feedback indicated ease of use, efficiency, acceptance, impact, and possible adoption. Within the timeframe of a PhD study, objective number six was achieved.
4. The artefact itself is a contribution to knowledge base in the sense that it provides an efficient, user-friendly solution to an existing viable depression problem. This was achieved by illustrating that existing HCI methods such as usability, user experience and heuristics can be used to build and evaluate systems in real and practical environment.
5. Empirical evidence of the effects of using subject matter expertise and experienced usability evaluators in evaluating a medical application. This was achieved during evaluation of prototypes in sections 5.10.1 and 5.13.
6. Empirical evidence of the knowledge and attitude of clinicians towards the use of ICT tools in work place. This was achieved in section 1.3.2.4.

### **6.3.1 Summary of research journey**

1. The study started by identifying an interesting researchable problem in Nigeria. This was done during the proof-of-concept study (section 1.3) conducted by the researcher through self-administered questionnaires to a group of 29 clinicians (Table 1.2) at the University of Benin Teaching Hospital, and 73 staff (teaching and non-teaching) and

students (Undergraduates, Masters and PhD) of the University of Benin, all in Nigeria. The users of the proposed system, what they want to do with it, and the environment in which it will operate were also discussed. The findings from the survey and from the literature helped to address research question one

2. The researcher, after the analysis of the data obtained in the proof-of-concept study, arrived at a workable local solution: the development of health ICT4D framework that is close to the users' skill, experience and understanding. This utilises the provisions of machine learning and HCI methodologies to address some of the peculiar difficulties screening for depression among University population. Research question two also addressed using these techniques during the study
3. A comparison of the performance of the proposed Bayesian network model with other commonly used probabilistic and non-probabilistic models was made, which found the Bayesian network models superior. The findings in section 5.5 helped to address research question four
4. Through constructed information gain and mutual information models, the study revealed the contributions of each symptom to the diagnoses-depression. The strengths of the synergistic combination of symptoms were also revealed. The findings from this analysis helped to address research question five
5. The study designed a screening tool and integrated it with the knowledge-based models. The performance was established through heuristic evaluation by a team of University staff who are also the intended end-users of the application, and complemented by clinicians. This was achieved in chapter five and helped to address research question three while also achieving objective number six

#### **6.4 Contribution to the field of Information and communication technologies for development**

Information and communication technologies for development (ICT4D) describes a multidisciplinary field, such as information systems, development studies, business and political science, that focuses on the application of information and communication technologies (ICT) to foster positive changes on the lives of the poor and marginalized individuals, communities or nations, by improving their economy, health, security, education and so on (Burrell & Toyama, 2009). In this study, the ICT techniques are the designed data analysis techniques, and the development aspect is in the contribution towards enhancing a skill in a complex subject such as contributing to reducing depression at the University levels.

In developing countries with limited mental health professionals and diagnostic facilities, the solution could be to use the devices that the universities already have and design applications that consider both the limitations of the available devices and users' needs. This study has shown that this is not only possible, it is workable. The prototypes developed in this study could be used in future studies that seek to understand the long-term impact of the use of desktop computers and machine learning tools in screening University staff and students for depression, within the context of a developing country.

It is important to note that ICT4D research focuses not only on the rural poor but also on the urban poor (Chepken, Mugwanya, Blake, & Marsden, 2012), who may experience resource constraints. In addition, research has shown that there is a gap in the literature that consider users who live in urban areas, with a lot more studies conducted with the rural poor (Chepken et al., 2012). The clinicians, staff and students who participated in this research were all from a University that was located in urban areas in a developing country, thus representing urban users who nevertheless may be in resource constrained environments. However, the tools developed are useful to users in both low-resource areas and areas without scarce resources. Therefore, this study contributes towards research that provides solutions to the urban poor or those who find themselves in resource-constrained situations while in urban areas.

The work of Khoja et al (2013) notes that health ICT research should be conducted using sound conceptual framework and proven theories. The development of the ICT4D methodologies in this study were based on a rigorous process using existing health machine learning guidelines. This has made the framework employed in the study easily replicated to design health machine learning techniques that support non-medical persons in other areas of healthcare settings.

#### **6.4.1 Novelty**

From our literature survey and to the best of our knowledge, this is the first time that the interaction between machine learning techniques and usability study has been considered to support a heterogenous user population in screening for depression among University population in Nigeria. Although there are several novel machine learning and HCI techniques in the literature, which could be used to support further simplification of the screening process for depression, this will be part of the focus of the study in the future to investigate how more benefits can accrue from these techniques. Past studies based on adoption and diffusion of innovation have been described on diffusions that have already occurred in the past, but the

present study is innovative since it promises to contribute to the adoption of assistive depression screening technology, in what Rogers described as “acceptability” research (Rogers, 2015). In addition to meeting the requirements for the award of PhD degree in computer science, the researcher intends to share the findings from the study with University of Benin teaching hospital (UBTH) and the Federal ministry of health in Nigeria. It is the hope of the researcher that these findings will inspire further studies and improve efforts to encourage possible faster adoption and utilisation, in the area of depression screening where adoption and diffusion of innovative technology promises to have great impact in improving health challenges in Nigeria.

## 6.5 Conclusions

With the increasing threat of depression among students of Nigerian universities, the consequences of turning a blind eye to the detrimental problem is grave. It is therefore only reasonable that a workable local solution that complements the efforts of mental health resources be sought. Complementing the existing method of depression screening with ICTs application is a necessary step in this direction. The tools designed in the study are two-folds; first is the task of building a knowledge base tool that is useful to discover the interactions between the variables in data. The second is integrating it into a user-friendly interface that help universities in Nigeria in the process of first-line screening for depression for their staff and students. The knowledge-based models were constructed using an established Bayesian network formalism with real-hospital dataset. The data was validated by a team of experienced mental health professionals. The experiments were performed with two machine learning software tools, Waikato environment for knowledge analysis (Weka) and BayesiaLab. The experimental results of the study have demonstrated, with practical application (and supported by theoretical analyses and justifications), the superiority of ICT4D framework over the existing system. The discovery of statistical relationships between symptoms and depression in a probability-based decision support system, and its integration into a graphical user interface is a definite step towards better methods in managing depression at the level of screening. The results also reveal the several number of unexplored possibilities that this knowledge can provide for building models and decision support systems. More importantly, the results demonstrate the strength of ICT-aided diagnostic intervention, especially with respect to acceptance and possible adoption of the system for use by non-clinicians in non-clinical environment.

## 6.6 Challenges faced in the study

Every PhD study is usually expected to come with a number of unique challenges. As such, this study was no exception. The challenges the study had, as explained in different sections of the work border mainly on setting the research questions right, answering them in the most professional manner thereby accomplishing the purpose for which the study was meant. These challenges are summarized as follows:

1. The first and the most daunting was the task of acquiring domain knowledge from the team of mental health professionals and then collecting real depression datasets from hospital and primary care centre. Presently, there are no combined expertise in both psychiatry and computer science in Nigeria, and as such, interdisciplinary collaboration was quite challenging.
2. Another challenge was gathering a team of evaluators comprising those knowledgeable in healthcare, machine learning and heuristic studies. These are professionals with very busy schedules. However, gathering was necessary because a screening tool, which represents the point of meeting between the expertise in both domains, needs their input to achieve the desired results (Wilkund, M.E; Kendler, J;Yale, 2013).

## 6.7 Future work

Based on the discussion in the previous sections, a number of areas can be identified for further research in the study, some relating to methodological advances, and others to the integration and maintenance of the screening tool. In the machine learning literature, there are compelling evidence that maintenance of clinical decision support system is a key aspect to its success. Since this is a responsibility of the domain expert, an interesting future direction is to follow up on this, exploring different strategies to accomplish this. For instance, the implementation of formal training for the users will be explored with other more powerful open source software (that would simplify the construction of the knowledge base).

Again, there is a persistent gap between medical research and quality of general health care, especially in the developing countries. As the study models may easily be modified to meet the needs of other settings, it will be interesting to extend the strength of knowledge engineering for health systems in this study to strengthen the inter-connection between data and knowledge flow from producers of health data research and the practitioners who need it. This measure has the potential to correct the imbalance.

Another future direction could be performed with another focus. Despite applying the user interface design principles specific for medical diagnosis-related decision support system, it will be interesting to redirect the study to other healthcare areas such as treatment planning and drug administration. Similarly, the study was focused on the implementation of Bayesian networks-based decision support systems for the screening of depression and make referrals for further diagnosis where necessary. The creation of computationally-managed clinical support systems for the implementation of other diseases and methods requires additional research.

Lastly, the study made use of only adoption and diffusion of innovation theory in its conceptual framework. It would a viable future agenda to extend the study to use combined theories.

## References

- Abas, M., Ali, G.-C., Nakimuli-Mpundu, E., & Chabanda, D. (2014). Depression in people living with HIV in Sub-Saharan Africa. *Tropical Medicine and International*, 19(12), 1392–1396. <https://doi.org/10.1111/tmi.12382>
- Abdulmalik, J., Kola, L., Fadahunsi, W., Adebayo, K., Yasamy, M. T., Musa, E., & Gureje, O. (2013). Country Contextualization of the Mental Health Gap Action Programme Intervention Guide: A Case Study from Nigeria. *PLoS Medicine*, 10(8). <https://doi.org/10.1371/journal.pmed.1001501>
- Acharya, C., Thimbleby, H., & Oladimeji, P. (2010). Human computer interaction and medical devices. In *BCS '10 Proceedings of the 24th BCS Interaction Specialist Group Conference* (pp. 168–176). Retrieved from <http://dl.acm.org/citation.cfm?id=2146329>
- Adeleke, T. (2015). ICT knowledge, Utilization and Perception among Healthcare Providers at National Hospital Abuja, Nigeria. *American Journal of Health Research*, 3(1), 47. <https://doi.org/10.11648/j.ajhr.s.2015030101.17>
- Adeyuya, A. O., Ola, B. A., Aloba, O. O., Mapayi, B. M., & Oginni, O. O. (2006). Depression amongst Nigerian university students: Prevalence and sociodemographic correlates. *Social Psychiatry and Psychiatric Epidemiology*, 41(8), 674–678. <https://doi.org/10.1007/s00127-006-0068-9>
- Adkins, D. E. (2017). Machine learning and electronic health records: A paradigm shift. *American Journal of Psychiatry*, 174(2), 93–94. <https://doi.org/10.1176/appi.ajp.2016.16101169>
- Afolabi, M. O., Abioye-Kuteyi, E. a, Fatoye, F. O., Bello, I. S., & Adeyuya, A. O. (2008). Pattern of Depression among Patients in a Nigerian Family Practice Population. *South African Family Practice*, 50(2). Retrieved from <http://www.safpj.co.za/index.php/safpj/article/view/858>
- Ahlan, A. R., & Isma'eel Ahmad, B. (2015). An overview of patient acceptance of Health Information Technology in developing countries: a review and conceptual model. *International Journal of Information Systems and Project Management*. <https://doi.org/10.12821/ijispm030102>
- Ahmed, K., & Bhugra, D. (2007). Depression across ethnic minority cultures : diagnostic issues. *World Cultural Psychiatry Research Review*, 2(3), 51. Retrieved from <http://www.wcpr.org/pdf/02-23/2007.0203.4756.pdf>
- Ahmed Kayode, O. (2014). Using e-psychiatry to Scale up Mental Health Services in Nigeria.

- Journal of Mass Communication & Journalism*, 04(4), 4–5. <https://doi.org/10.4172/2165-7912.1000184>
- Ajzen, I., & Fishbein, M. (1980). Theory of Reasoned Action / Theory of Planned Behavior. *Social Psychology*, 2007, 67–98. [https://doi.org/10.5771/9783845260341\\_1](https://doi.org/10.5771/9783845260341_1)
- Akande, T. (2004). Referral system in Nigeria: study of a tertiary health facility. *Annals of African Medicine*, 3(3), 130–133.
- Alhaj, T. A., Siraj, M. M., Zainal, A., Elshoush, H. T., & Elhaj, F. (2016). Feature selection using information gain for improved structural-based alert correlation. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0166017>
- Allgaier, A.-K., Pietsch, K., Frühe, B., Sigl-Glöckner, J., & Schulte-Körne, G. (2012). Screening for depression in adolescents: validity of the patient health questionnaire in paediatric care. *Depression and Anxiety*, 29(10), 906–913. <https://doi.org/10.1002/da.21971>
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association (Fifth Edition). Washington, DC: American Psychiatric Association. <https://doi.org/10.1176/appi.books.9780890425596.744053>
- Amershi, S., Cakmak, M., Knox, W. B., Kulesza, T., & Lau, T. (2013). IUI Workshop on Interactive Machine Learning. In *IUI'13 Companion, March 19–22, 2013, Santa Monica, CA, USA*. (p. 3). Retrieved from <https://dropline.net/wp-content/uploads/2012/02/p121-amershi.pdf>
- Anand, P., & Dolan, P. (2005). Equity, capabilities and health. *Social Science and Medicine*. <https://doi.org/10.1016/j.socscimed.2004.04.031>
- Ann, S., Dion, P., Christa, H., & Chojecki, D. (2015). *Optimizing adoption and diffusion of medical devices at the system level*. The Institute of Health Economics, Alberta, Canada. Available at <http://www.ihe.ca>.
- Ansell, D., Crispo, J. A. G., Simard, B., & Bjerre, L. M. (2017). Interventions to reduce wait times for primary care appointments: a systematic review. *BMC Health Services Research*, 17(1). <https://doi.org/10.1186/s12913-017-2219-y>
- Appari, A., Eric Johnson, M., & Anthony, D. L. (2013). Meaningful use of electronic health record systems and process quality of care: Evidence from a panel data analysis of U.S. acute-care hospitals. *Health Services Research*, 48(2 PART1), 354–375. <https://doi.org/10.1111/j.1475-6773.2012.01448.x>
- Armano, G., & Tamponi, E. (2018). Building forests of local trees. *Pattern Recognition*.

<https://doi.org/10.1016/j.patcog.2017.11.017>

- Armero, C., Forné, C., Rué, M., Forte, A., Perpiñán, H., Gómez, G., & Baré, M. (2016). Bayesian joint ordinal and survival modeling for breast cancer risk assessment. *Statistics in Medicine*, 35(28), 5267–5282. <https://doi.org/10.1002/sim.7065>
- Arrieta, J., Aguerrebere, M., Raviola, G., Flores, H., Elliott, P., Espinosa, A., ... Franke, M. F. (2017). Validity and Utility of the Patient Health Questionnaire (PHQ)-2 and PHQ-9 for Screening and Diagnosis of Depression in Rural Chiapas, Mexico: A Cross-Sectional Study. *Journal of Clinical Psychology*. <https://doi.org/10.1002/jclp.22390>
- Asemahagn, M. A. (2016). Health professionals' challenge in using ICTs to manage their patients: The case of hospitals in Addis Ababa, Ethiopia. *Online Journal of Nursing Informatics*, 20(2).
- Asuke, S., Ibrahim, M. S., Sabitu, K., Asuke, A. U., Igbaver, I. I., & Joseph, S. (2016). A comparison of referrals among primary health - care workers in urban and rural local government areas in North - Western Nigeria, 93–97. <https://doi.org/10.4103/2276-7096.192233>
- Avgerou, C. (2017). Theoretical framing of ICT4D research. In *IFIP Advances in Information and Communication Technology* (Vol. 504, pp. 10–23). [https://doi.org/10.1007/978-3-319-59111-7\\_2](https://doi.org/10.1007/978-3-319-59111-7_2)
- Ayonrinde, O., Gureje, O., & Lawal, R. (2004). Psychiatric research in Nigeria: Bridging tradition and modernisation. *British Journal of Psychiatry*, 184, 536–538. <https://doi.org/10.1192/bjp.184.6.536>
- Ayoola, Y.-A. (2015). The Role of ICT in National Development - A Study of the Educational Sector in Lagos State. In *International Conference on African Development Issues (CU-ICADI)* (pp. 121–123).
- Baasher, T. A., Carstairs, G. M., Giel, R., & Hassler, F. R. (1975). Mental health services in developing countries. In *WHO seminar on the organisation of Mental Health Services*. World Health Organisation, Geneva. Retrieved from [http://whqlibdoc.who.int/offset/WHO\\_OFFSET\\_22\\_\(pt1-pt3\).pdf](http://whqlibdoc.who.int/offset/WHO_OFFSET_22_(pt1-pt3).pdf)
- Babalobi, O. (2010). *The challenge of changing trends in the development and use of information and internet web communication technologies for veterinary medicine education in Nigeria*. Department of Veterinary Public Health and Preventive Medicine, Faculty of Veterinary Medicine, University of Ibadan, Nigeria. <https://doi.org/10.1016/j.cardfail.2009.03.006>

- Baiyewu, O., Yusuf, A. J., & Ogundele, A. (2015). Depression in elderly people living in rural Nigeria and its association with perceived health, poverty, and social network. *International Psychogeriatrics*, 27(12), 2009–2015. <https://doi.org/10.1017/S1041610215001088>
- Bakare, A. (2012). Measuring the Income Inequality in Nigeria: the Lorenz Curve and Gini Co-efficient Approach. *American Journal of Economics*, 2(1), 47–52. <https://doi.org/10.5923/j.economics.20120201.06>
- Barber, D. (2011). *Bayesian Reasoning and Machine Learning*. <https://doi.org/10.1017/CBO9780511804779>
- Bardzell, J., & Bardzell, S. (2015). Humanistic HCI. *Synthesis Lectures on Human-Centered Informatics*, 8(4), 1–185. <https://doi.org/10.2200/S00664ED1V01Y201508HCI031>
- Barjis, J., Kolfshoten, G., & Maritz, J. (2013). A sustainable and affordable support system for rural healthcare delivery. *Decision Support Systems*, 56(1), 223–233. <https://doi.org/10.1016/j.dss.2013.06.005>
- Barnes, C., Claus, R., Driessen, P., Dos Santos, M. J. F., George, M. A., & Van Laerhoven, F. (2017). Uniting forest and livelihood outcomes? Analyzing external actor interventions in sustainable livelihoods in a community forest management context. *International Journal of the Commons*, 11(1), 532–571. <https://doi.org/10.18352/ijc.750>
- Batchelor, S., & Norrish, P. (2005). Framework for the assessment of ICT pilot projects: Beyond Monitoring and Evaluation to Applied Research. *InfoDev*, 20, 2009. Retrieved from [http://gamos.org/publications/Framework for the assessment of ICT Pilot Projects\\_infodev.pdf](http://gamos.org/publications/Framework%20for%20the%20assessment%20of%20ICT%20Pilot%20Projects_infodev.pdf)
- Bayes, T. (1763). An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. *Philosophical Transactions of the Royal Society of London*, 53(0), 370–418. <https://doi.org/10.1098/rstl.1763.0053>
- Beadnell, B., Baker, S. A., Gillmore, M. R., Morrison, D. M., Huang, B., & Stielstra, S. (2008). The theory of reasoned action and the role of external factors on heterosexual men's monogamy and condom use. *Journal of Applied Social Psychology*, 38(1), 97–134. <https://doi.org/10.1111/j.1559-1816.2008.00298.x>
- Beck, A. (1961). Beck Depression Inventory. *Depression*, 2006, 2–4. <https://doi.org/10.1093/ndt/gfr086>
- Bee, P., Brooks, H., Fraser, C., & Lovell, K. (2015). Professional perspectives on service user

- and carer involvement in mental health care planning: A qualitative study. *International Journal of Nursing Studies*, 52(12), 1834–1845. <https://doi.org/10.1016/j.ijnurstu.2015.07.008>
- Bennasar, M., Hicks, Y., & Setchi, R. (2015). Feature selection using Joint Mutual Information Maximisation. *Expert Systems with Applications*. <https://doi.org/10.1016/j.eswa.2015.07.007>
- Bergman, M. M. (2008). Method Mix, Technical Hex, and Theory Fix. *Advances in Mixed Methods Research*, 1–8. <https://doi.org/10.4135/9780857024329>
- Berndt, D. J., Fisher, J. W., Hevner, A. R., & Studnicki, J. (2001). Healthcare data warehousing and quality assurance. *Computer*, 34(12). <https://doi.org/10.1109/2.970578>
- Bhutkar, G., Konkani, A., Katre, D., & Ray, G. G. (2013). A review healthcare usability evaluation methods. *Biomedical Instrumentation and Technology*. <https://doi.org/10.2345/0899-8205-47.s2.45>
- Bing-Mei, C., Xiao-Ping, F., Zhi-Ming, Z., & Xue-Rong, L. (2011). Neural Network Structure Study In Child Mental Health Disorders Intelligent Diagnosis System\*. *Procedia Environmental Sciences*, 8, 669–678. <https://doi.org/10.1016/j.proenv.2011.10.103>
- Bligård, L. O., & Osvalder, A. L. (2014). CCPE: Methodology for a combined evaluation of cognitive and physical ergonomics in the interaction between human and machine. *Human Factors and Ergonomics In Manufacturing*, 24(6), 685–711. <https://doi.org/10.1002/hfm.20512>
- Bouckaert, R. R., Frank, E., Hall, M., Kirkby, R., Reutemann, P., Seewald, A., & Scuse, D. (2014). *WEKA Manual for Version 3-7-12*. University of Waikato, Hamilton, New Zealand. Retrieved from <papers3://publication/uuid/24E005A2-AA1B-4614-BAF5-4D92C4F37413>
- Boudreaux, E. D., Waring, M. E., Hayes, R. B., Sadasivam, R. S., Mullen, S., & Pagoto, S. (2014). Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. *Translational Behavioral Medicine*. <https://doi.org/10.1007/s13142-014-0293-9>
- Boylan, K. (2007). *Diagnosis Made Easier: Principles and Techniques for Mental Health Clinicians*. *Journal of the Canadian Academy of Child and Adolescent Psychiatry* (Vol. 16). <https://doi.org/10.1176/appi.ps.58.4.572>
- Braun, S. (2005). Usability for medical devices. In *2005 IEEE Symposium on Product Safety Engineering* (Vol. 2005, pp. 16–22). <https://doi.org/10.1109/PSES.2005.1529517>

- Bro, R., & Smilde, A. K. (2014). Principal component analysis. *Analytical Methods*, 6(9), 2812. <https://doi.org/10.1039/c3ay41907j>
- Burrell, J., & Toyama, K. (2009). What Constitutes Good ICTD Research? *Information Technologies and International Development*, 5(3), 82–94. <https://doi.org/10.1038/nm1110-1198>
- Cacciattolo, M. (2015). Ethical considerations in research. In *The Praxis of English Language Teaching and Learning (PELT): Beyond the Binaries: Researching Critically in EFL Classrooms* (pp. 55–73). [https://doi.org/10.1007/978-94-6300-112-0\\_4](https://doi.org/10.1007/978-94-6300-112-0_4)
- Cai, L., & Zhu, Y. (2015). The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. *Data Science Journal*, 14(0), 2. <https://doi.org/10.5334/dsj-2015-002>
- Cd, E., Cn, E., Me, O., & Ho, N. (2017). Giant Lipomas . A Report of Two Cases. *The Nigerian Journal of General Practice*, 15(2), 46–49. <https://doi.org/10.4103/NJGP.NJGP>
- CDC. (2009). *The Power of Prevention Chronic Disease... the public health challenge of the 21 st century*. National Center for Chronic Disease Prevention and Health Promotion. Available at: <https://www.cdc.gov/chronicdisease/pdf/2009-power-of-prevention.pdf>. Retrieved from <http://www.cdc.gov/chronicdisease/pdf/2009-Power-of-Prevention.pdf>
- Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers and Electrical Engineering*. <https://doi.org/10.1016/j.compeleceng.2013.11.024>
- Chang, S., Fan, T., Lo, T., Hung, C., & Yuan, M. (2014). Mobile cloud-based depression diagnosis using an ontology and a Bayesian network. *Future Generation Computer Systems*, 43–44, 87–98. <https://doi.org/10.1016/j.future.2014.05.004>
- Chattopadhyay, S. (2017). A neuro-fuzzy approach for the diagnosis of depression. *Applied Computing and Informatics*, 13, 10–18. [https://doi.org/y Elsevier Applied Computing and Informatics \(2017\) 13, 10–18 Saudi Computer Society, King Saud University Applied Computing and Informatics \(http://computer.org.sa\) www.ksu.edu.sa www.sciencedirect.com http://dx.doi.org/10.1016/j.aci.2014.01.00](https://doi.org/y Elsevier Applied Computing and Informatics (2017) 13, 10–18 Saudi Computer Society, King Saud University Applied Computing and Informatics (http://computer.org.sa) www.ksu.edu.sa www.sciencedirect.com http://dx.doi.org/10.1016/j.aci.2014.01.00)
- Chepken, C., Mugwanya, R., Blake, E., & Marsden, G. (2012). ICTD interventions. In *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development - ICTD '12* (p. 241). <https://doi.org/10.1145/2160673.2160704>
- Chetley, A., Trude, B., Ramirez, R., Shields, T., Drury, P., Kumekawa, J., ... Nyamai-Kisia, C. (2006). Improving health, connecting people: the role of ICTs in the health sector of developing countries A framework paper. *Infodev*, (7), 1–65. Retrieved from

<http://www.infobridge.org/asp/documents/3254.pdf>

- Chilisa, B., & Kawulich, B. (2015). Selecting a research approach: paradigm, methodology and methods. In *Doing Social Research: A global context, Chapter: Selecting a research approach: Paradigm, methodology and methods*, Publisher: McGraw Hill, Editors: C. Wagner, B. Kawulich, M. Garner, pp.51-61: <https://www.researchgate.net/publication/257944787> (pp. 1–22). *Doing Social Research: A global context, Chapter: Selecting a research approach: Paradigm, methodology and methods*, Publisher: McGraw Hill, Editors: C. Wagner, B. Kawulich, M. Garner, pp.51-61. Retrieved from [http://www.celt.mmu.ac.uk/researchmethods/Modules/Selection\\_of\\_methodology/](http://www.celt.mmu.ac.uk/researchmethods/Modules/Selection_of_methodology/)
- Chomutare, T., Fernandez-Luque, L., Arsand, E., & Hartvigsen, G. (2011). Features of mobile diabetes applications: Review of the literature and analysis of current applications compared against evidence-based guidelines. *Journal of Medical Internet Research*. <https://doi.org/10.2196/jmir.1874>
- Coast, E., Leone, T., Hirose, A., & Jones, E. (2012). Poverty and postnatal depression: a systematic mapping of the evidence from low and lower middle income countries. *Health & Place*, 18(5), 1188–1197. <https://doi.org/10.1016/j.healthplace.2012.05.004>
- Coast, J., Smith, R., & Lorgelly, P. (2008). Should the capability approach be applied in health economics? *Health Economics*, 17(6), 667–670. <https://doi.org/10.1002/hec.1359>
- Conrady, S., & Jouffe, L. (2015). *Bayesian Networks & BayesiaLab: A Practical Introduction for Researchers*. Bayesia USA (Vol. 1). Bayesia USA. <https://doi.org/10.1002/wics.48>
- Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1023/A:1022627411411>
- Creswell, J. (2014). *Research design: Qualitative, quantitative, and mixed method*. *Research design Qualitative quantitative and mixed methods approaches*. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Creswell, J. W. (2013). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. *Research design Qualitative quantitative and mixed methods approaches*. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Curiac, D.-I., Vasile, G., Baniias, O., Volosencu, C., & Albu, A. (2009). Bayesian Network Model for Diagnosis of Psychiatric Diseases. In *Proceedings of the ITI 2009 31st International Conference on Information Technology Interfaces* (pp. 61–66). <https://doi.org/10.1109/ITI.2009.5196055>

- Daimi, K., & Banitaan, S. (2014). Using Data Mining to Predict Possible Future Depression Cases. *International Journal of Public Health Science (IJPHS)*, 3(4), 231–240.
- Dang, A. T. (2014). Amartya Sen's Capability Approach: A Framework for Well-Being Evaluation and Policy Analysis? *Review of Social Economy*, 72(4), 460–484. <https://doi.org/10.1080/00346764.2014.958903>
- Das, J., Do, Q. T., Friedman, J., McKenzie, D., & Scott, K. (2007). Mental health and poverty in developing countries: Revisiting the relationship. *Social Science & Medicine*, 65(3), 467–480. <https://doi.org/10.1016/j.socscimed.2007.02.037>
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319. <https://doi.org/10.2307/249008>
- de Bruijne, M. (2016). Machine learning approaches in medical image analysis: From detection to diagnosis. *Medical Image Analysis*, 33, 94–97. <https://doi.org/10.1016/j.media.2016.06.032>
- de Carvalho, L. R., Évora, Y. D. M., & Zem-Mascarenhas, S. H. (2016). Assessment of the usability of a digital learning technology prototype for monitoring intracranial pressure. *Revista Latino-Americana de Enfermagem*, 24. <https://doi.org/10.1590/1518-8345.1054.2777>
- Dearden, A., Wright, P., Bowen, S., Rahman, F., Cobb, M., & Wolstenholme, D. (2010). User-centred design and pervasive health: A position statement from the User-Centred Healthcare Design project. *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2010 4th International Conference on-NO PERMISSIONS*, 8–11. <https://doi.org/10.4108/ICST.PERVASIVEHEALTH2010.8837>
- Deng, Z., Zhu, X., Cheng, D., Zong, M., & Zhang, S. (2016). Efficient kNN classification algorithm for big data. *Neurocomputing*, 195, 143–148. <https://doi.org/10.1016/j.neucom.2015.08.112>
- Denscombe, M. (2010). The good research guide for small-scale social research projects. *Psychological Science, 4th editio*, 356.
- DFID. (2008). DFID 's Sustainable Livelihoods Approach and its Framework. *Development*, 1–5.
- Diagnostic and Statistical Manual of Mental Disorders fifth edition*. (n.d.).
- Dixon, P. M. (2006). Nearest Neighbor Methods. In *Encyclopedia of Environmetrics*. <https://doi.org/10.1002/9780470057339.van007>
- Dubois, R. W., Lauer, M., & Perfetto, E. (2013). When is evidence sufficient for decision-

- making? A framework for understanding the pace of evidence adoption. *Journal of Comparative Effectiveness Research*, 2(4), 383–391. <https://doi.org/10.2217/cer.13.39>
- Duch, J., Rallo, R., & Guardiola, X. (2014). Determining the Diffusion of Innovations and the Dissemination Network in a Healthcare Organization. Retrieved from [http://www.biocent.cat/doc/Master\\_Thesis\\_Alex\\_Vinyals.pdf](http://www.biocent.cat/doc/Master_Thesis_Alex_Vinyals.pdf)
- Dumas, J. (1996). How many participants in a usability test are enough. *Common Ground, Newsletter of Usability Professionals' ...*, 1, 2–6. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:How+many+participants+in+a+usability+test+is+enough?#2%5Cnhttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:How+many+participants+in+a+usability+test+is+enough?%231>
- Eisenberg, D., Gollust, S. E., Golberstein, E., & Hefner, J. L. (2007). Prevalence and correlates of depression, anxiety, and suicidality among university students. *The American Journal of Orthopsychiatry*, 77(4), 534–542. <https://doi.org/10.1037/0002-9432.77.4.534>
- Endsley, M. R. (2016). *Designing for Situation Awareness: An Approach to User-Centered Design, Second Edition*. Design. <https://doi.org/10.1201/9780203485088>
- Engelbrecht, P., & Paquet, U. (2008). Training support vector machines with particle swarms. In *Proceedings of the International Joint Conference on Neural Networks* (pp. 1593–1598). <https://doi.org/10.1109/IJCNN.2003.1223937>
- Ethem, A. (2010). *Introduction to Machine Learning* (Second Edi). The MIT Press Cambridge, Massachusetts London, England.
- Ezema, I. J. (2016). Reproductive Health Information Needs and Access among Rural Women in Nigeria: A Study of Nsukka Zone in Enugu State. *The African Journal of Information and Communication (AJIC)*, 18(18), 117–133. <https://doi.org/10.23962/10539/21788>
- Fang, L., Zhao, H., Wang, P., Yu, M., Yan, J., Cheng, W., & Chen, P. (2015). Feature selection method based on mutual information and class separability for dimension reduction in multidimensional time series for clinical data. *Biomedical Signal Processing and Control*. <https://doi.org/10.1016/j.bspc.2015.05.011>
- Faust, O., Ang, P. C. A., Puthankattil, S. D., & Joseph, P. K. (2014). Depression Diagnosis Support System Based on Eeg Signal Entropies. *Journal of Mechanics in Medicine and Biology*, 14(03), 1450035. <https://doi.org/10.1142/S0219519414500353>
- Feingold, M. (1992). The Equivalence of Cohen's Kappa and Pearson's Chi-Square Statistics in the 2 x 2 Table. *Educational and Psychological Measurement*, 52(1), 57–61. <https://doi.org/10.1177/001316449205200105>

- FGN. (2012). *National Policy for Information Technology Development and Telecommunications (ICT4D)*. <https://doi.org/10.1016/j.cell.2011.02.013>
- First, M. B. (2015). Structured Clinical Interview for the DSM (SCID). *The Encyclopedia of Clinical Psychology*, 1–6. <https://doi.org/10.1002/9781118625392.wbecp351>
- Fisher, A. (2011). Depressive and other symptoms of emotional disorders in adolescents. *International Journal of Social Psychiatry*, 57(1\_suppl), 22–39. <https://doi.org/10.1177/0020764010396695>
- Fleiss, J., Spitzer, R., Cohen, J., & Endicott, J. (1972). Three Computer Diagnosis Methods Compared. *Archives of General Psychiatry*, 27, 643–650.
- Fontenot, F. (2014). The Affordable Care Act and Electronic Health Care Records: Can Technology Help Reduce the Cost of Health Care? *Physician Executive*, 40(1), 68–72. Retrieved from <http://library.ashford.edu/EzProxy.aspx?url=http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,cpid&custid=s8856897&db=bsh&AN=93541815&site=ehost-live>
- Foster, K. R., Koprowski, R., & Skufca, J. D. (2014). Machine learning, medical diagnosis, and biomedical engineering research - commentary. *Biomedical Engineering Online*, 13(1), 94. <https://doi.org/10.1186/1475-925X-13-94>
- Franklin, S., & Sridaran, S. (2012). Making Human-Computer Interaction Usable: Thinking from Healthcare Perspective. *International Journal of Computer Applications*, 58(1), 0975 – 8887. Retrieved from <http://research.ijcaonline.org/volume58/number1/pxc3883413.pdf>
- Freund, Y., & Schapire, R. E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, 55(1), 119–139. <https://doi.org/10.1006/jcss.1997.1504>
- Gagnon, M. (2015). Information and Communication Technology Adoption By Healthcare Professionals: an Overview of Theoretical Models and Their. *IADIS International Conference Applied Computing*, 65–72.
- Gambo, I. P., & Soriyan, A. H. (2017). ICT Implementation in the Nigerian Healthcare System. *IT Professional*, 19(2), 12–15. <https://doi.org/10.1109/MITP.2017.21>
- Ganju, K. K., Pavlou, P. A., & Banker, R. D. (2016). Does information and communication technology lead to the well-being of nations? A country-level empirical investigation. *MIS Quarterly*, 40(2), 417–430.

- Gao, Z., Xu, Y., Meng, F., Qi, F., & Lin, Z. (2014). Improved information gain-based feature selection for text categorization. In *2014 4th International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace and Electronic Systems, VITAE 2014 - Co-located with Global Wireless Summit*. <https://doi.org/10.1109/VITAE.2014.6934421>
- Gareth, J., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning with Applications in R*. <https://doi.org/10.1007/978-1-4614-7138-7>
- Gelaye, B., Williams, M. A., Lemma, S., Deyessa, N., Bahretibeb, Y., Shibire, T., ... Andrew Zhou, X. H. (2013). Validity of the patient health questionnaire-9 for depression screening and diagnosis in East Africa. *Psychiatry Research*, *210*(2), 653–661. <https://doi.org/10.1016/j.psychres.2013.07.015>
- Genie. (1999). *Open Genie reference Manual*. Retrieved from <http://download.opengenie.org/kits/doc/TheReferenceManual-2.0.pdf>
- Ghosh, J. K., & Valtorta, M. (1999). Probabilistic Bayesian Network Model Building of Heart Disease 1, 1999, 1–20.
- Gibson, E., Li, W., Sudre, C., Fidon, L., Shakir, D. I., Wang, G., ... Vercauteren, T. (2018). NiftyNet: a deep-learning platform for medical imaging. *Computer Methods and Programs in Biomedicine*, *158*, 113–122. <https://doi.org/10.1016/j.cmpb.2018.01.025>
- Globerman, S. (2013). *Reducing Wait Times for Health Care: What. Reducing Wait Times for Health Care*. Retrieved from <https://www.fraserinstitute.org/uploadedFiles/fraser-ca/Content/research-news/research/publications/reducing-wait-times-for-health-care.pdf#page=164>
- Google. (2017). Google Brain Team: Make machines intelligent. Improve people's lives.
- Gould, J. D., & Lewis, C. (1985). Designing for usability: key principles and what designers think. *Communications of the ACM*, *28*(3), 300–311. <https://doi.org/10.1145/3166.3170>
- Grudin, J. (2009). AI and HCI: Two Fields Divided by a Common Focus. *AI Magazine*, *30*(4), 48. <https://doi.org/10.1609/AIMAG.V30I4.2271>
- Grunfeld, H. (2007). Framework for Evaluating Contributions of ICT to Capabilities, Empowerment and Sustainability in Disadvantaged Communities. In *CPRsouth2 Conference* (pp. 1–21). Retrieved from [http://www.cprsouth.org/wp-content/uploads/drupal/Helena\\_Grunfeld.pdf](http://www.cprsouth.org/wp-content/uploads/drupal/Helena_Grunfeld.pdf)
- Gücin, N. Ö., & Berk, Ö. S. (2015). Technology Acceptance in Health Care: An Integrative Review of Predictive Factors and Intervention Programs. *Procedia - Social and*

- Behavioral Sciences*, 195, 1698–1704. <https://doi.org/10.1016/j.sbspro.2015.06.263>
- Guillemainault, L., Oukssel, H., Belleguic, C., Le Guen, Y., Germaud, P., Desfleurs, E., ... Magnan, A. (2017). Personalised medicine in asthma: From curative to preventive medicine. *European Respiratory Review*. <https://doi.org/10.1183/16000617.0010-2016>
- Gureje, O., Abdulmalik, J., Kola, L., Musa, E., Yasamy, M. T., & Adebayo, K. (2015). Integrating mental health into primary care in Nigeria: report of a demonstration project using the mental health gap action programme intervention guide. *BMC Health Services Research*, 15(1), 242. <https://doi.org/10.1186/s12913-015-0911-3>
- Gureje, O., Uwakwe, R., Oladeji, B., Makanjuola, V. O., & Esan, O. (2010). Depression in adult Nigerians: Results from the Nigerian Survey of Mental Health and Well-being. *Journal of Affective Disorders*, 120(1–3), 158–164. <https://doi.org/10.1016/j.jad.2009.04.030>
- Hall, M. A., & Witten, I. H. (2010). WEKA — Experiences with a Java Open-Source Project. *Journal of Machine Learning Research*, 11, 2533–2541.
- Haluza, D., & Jungwirth, D. (2015). ICT and the future of health care: Aspects of health promotion. *International Journal of Medical Informatics*, 84(1), 48–57. <https://doi.org/10.1016/j.ijmedinf.2014.09.005>
- Han, J., & Kamber, M. (2011). *Data Mining: Concepts and Techniques*. Morgan Kaufmann (Vol. 12). <https://doi.org/10.1007/978-3-642-19721-5>
- Harris, R. W. (2016). How ICT4D Research Fails the Poor. *Information Technology for Development*, 22(1), 177–192. <https://doi.org/10.1080/02681102.2015.1018115>
- Harvard Health publications. (2014). Using smartphone apps for heart health. <https://doi.org/10.1002/14651858.CD010013.pub2>
- Haux, R., Howe, J., Marschollek, M., Plischke, M., & Wolf, K. H. (2008). Health-enabling technologies for pervasive health care: On services and ICT architecture paradigms. *Informatics for Health and Social Care*, 33(2), 77–89. <https://doi.org/10.1080/17538150802127140>
- Heckerman, D. (1997). Bayesian Networks for Data Mining. *Data Mining and Knowledge Discovery*, 119(1), 79–119.
- Heuer, H. (2013). *On The Intersection Of Human-Computer Interaction ( HCI ) And Artificial Intelligence ( AI )*. KTH Royal Institute of Technology EIT ICT Labs Master School.
- Hewett, Baecker, Card, Carey, Gasen, Mantei, Perlman, S. & V. (2009). ACM SIGCHI Curricula for Human-Computer Interaction. *Human-Computer Interaction Series*.

Available at [Http://Sigchi.Org/Cdg/Cdg2.Html](http://Sigchi.Org/Cdg/Cdg2.Html).

- Hijazi, S., Page, A., Kantarci, B., & Soyata, T. (2016). Machine Learning in Cardiac Health Monitoring and Decision Support. *Computer*, 49(11), 38–48. <https://doi.org/10.1109/MC.2016.339>
- Hill, A., Douptcheva, N., Lerch, M., & Sauvain-dugerdil, C. (n.d.). IMPROVING HEALTH: HOW THE CAPABILITY APPROACH HELPS IN TWO WEST AFRICAN CASE STUDIES.
- HIMSS. (2009). Defining and Testing EMR Usability: Principles and Proposed Methods of EMR Usability Evaluation and Rating.
- Ho, M. R., Smyth, T. N., Kam, M., & Dearden, A. (2009). Human-Computer Interaction for Development: The Past, Present, and Future. *Information Technologies & International Development*, 5(4), 1–18. <https://doi.org/10.1145/1467247.1467265>
- Holloway, C. J., Suttie, J., Dass, S., & Neubauer, S. (2011). Clinical Cardiac Magnetic Resonance Spectroscopy. *Progress in Cardiovascular Diseases*, 54(3), 320–327. <https://doi.org/10.1016/j.pcad.2011.08.002>
- Holzinger, A. (2013). Human-Computer Interaction and Knowledge Discovery (HCI-KDD): What is the benefit of bringing those two fields to work together? In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 8127 LNCS, pp. 319–328). [https://doi.org/10.1007/978-3-642-40511-2\\_22](https://doi.org/10.1007/978-3-642-40511-2_22)
- Holzinger, A., & Jurisica, I. (2014). Knowledge discovery and data mining in biomedical informatics: The future is in integrative, interactive machine learning solutions. *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics*, 1–18. [https://doi.org/10.1007/978-3-662-43968-5\\_1](https://doi.org/10.1007/978-3-662-43968-5_1)
- Holzinger, A., Stocker, C., Ofner, B., Prohaska, G., Brabenetz, A., & Hofmann-Wellenhof, R. (2013). Combining HCI, natural language processing, and knowledge discovery - Potential of IBM content analytics as an assistive technology in the biomedical field. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 7947 LNCS, pp. 13–24). [https://doi.org/10.1007/978-3-642-39146-0\\_2](https://doi.org/10.1007/978-3-642-39146-0_2)
- Horný, M. (2014). *Bayesian networks: A Technical report*. Department of Health Policy and Management, Boston University. <https://doi.org/10.1145/1859204.1859227>
- Horsky, J., Schiff, G. D., Johnston, D., Mercincavage, L., Bell, D., & Middleton, B. (2012).

- Interface design principles for usable decision support: A targeted review of best practices for clinical prescribing interventions. *Journal of Biomedical Informatics*. <https://doi.org/10.1016/j.jbi.2012.09.002>
- Hosseinifard, B., Moradi, M. H., & Rostami, R. (2013). Classifying depression patients and normal subjects using machine learning techniques and non-linear features from EEG signal. *Computer Methods and Programs in Medicine*, 109(109), 339–345. <https://doi.org/10.1016/j.cmpb.2012.10.008>
- Hu, J., Perer, A., & Wang, F. (2015). Data driven analytics for personalized healthcare. In *Healthcare Information Management Systems: Cases, Strategies, and Solutions: Fourth Edition* (pp. 529–554). [https://doi.org/10.1007/978-3-319-20765-0\\_31](https://doi.org/10.1007/978-3-319-20765-0_31)
- Huang, S., LePendou, P., Iyer, S., Tai-Seale, M., Carrell, D., & Shah, N. (2014). Toward personalizing treatment for depression: predicting diagnosis and severity. *Journal of the American Medical Informatics Association : JAMIA*, 1–7. <https://doi.org/10.1136/amiajnl-2014-002733>
- Hugin. (2016). *Hugin api Reference Manual*. *Hugin Expert*. Retrieved from <http://download.hugin.com/webdocs/manuals/api-manual.pdf>
- IBM. (2017). IBM Watson Health - IBM Watson for Oncology.
- Ibrahim, A. K., Kelly, S. J., Adams, C. E., & Glazebrook, C. (2013). A systematic review of studies of depression prevalence in university students. *Journal of Psychiatric Research*, 47(3), 391–400. <https://doi.org/10.1016/j.jpsychires.2012.11.015>
- Ibrahim, A., Kelly, S., Adams, C., & Glazebrook, C. (2013). A systematic review of studies of depression prevalence in university students. *Journal of Psychiatric Research*. <https://doi.org/10.1016/j.jpsychires.2012.11.015>
- Idowu, P., Cornford, D., & Bastin, L. (2008). Health informatics deployment in Nigeria. *Journal of Health Informatics in Developing Countries*, 2(1), 16–23.
- Ilbeygi, M., & Shah-Hosseini, H. (2012). A novel fuzzy facial expression recognition system based on facial feature extraction from color face images. *Engineering Applications of Artificial Intelligence*, 25(1), 130–146. <https://doi.org/10.1016/j.engappai.2011.07.004>
- Isaković, M., Sedlar, U., Volk, M., & Bešter, J. (2016). Usability pitfalls of diabetes mHealth apps for the elderly. *Journal of Diabetes Research*. <https://doi.org/10.1155/2016/1604609>
- ISO. (2017). ISO 9241 Part 11: Guidance on usability. Retrieved from <https://www.userfocus.co.uk/resources/iso9241/part11.html>
- Jack-Ide, I. O., Uys, L. R., & Middleton, L. E. (2012). A comparative study of mental health

- services in two African countries: South Africa and Nigeria. *International Journal of Nursing and Midwifery*, 4(4), 50–57. <https://doi.org/10.5897/IJNM12.002>
- Jake-Schoffman, D., Silfee, V., Waring, M., Boudreaux, E., Sadasivam, R., Mullen, S., ... Pagoto, S. (2017). Methods for Evaluating the Content, Usability, and Efficacy of Commercial Mobile Health Apps. *JMIR MHealth and UHealth*, 5(12), e190. <https://doi.org/10.2196/mhealth.8758>
- Jeong, H., Ziemkiewicz, C., Ribarsky, W., & Chang, R. (2009). Understanding Principal Component Analysis Using a Visual Analytics Tool. Charl. In *Charlotte Visualization Center, UNC Charlotte* (pp. 1–10). Retrieved from papers3://publication/uuid/F74403D9-F5D5-420C-8A62-C44E90A43EBD
- Jimoh, L., Pate, M. A., Lin, L., & Schulman, K. A. (2012). A model for the adoption of ICT by health workers in Africa. *International Journal of Medical Informatics*, 81(11), 773–781. <https://doi.org/10.1016/j.ijmedinf.2012.08.005>
- Joshi, A., Arora, M., Dai, L., Price, K., Vizer, L., & Sears, A. (2009). Usability of a patient education and motivation tool using heuristic evaluation. *Journal of Medical Internet Research*, 11(4). <https://doi.org/10.2196/jmir.1244>
- Juan Fernandez del pozo, Pedro, Larranaga, C. bielza. (2013). *Technical Report on Stratified Cross-Validation in Multi-Label Classification Using Genetic Algorithms*. Computational Intelligence Group Universidad Politécnica de Madrid 1.
- Kamm, L., & Tretjakov, K. (2009). Construction of Bayesian Networks.
- Kane, J. M. (2014). Technology-based interventions in health care. *Epidemiology and Psychiatric Sciences*. <https://doi.org/10.1017/S2045796014000444>
- Karimi, M., Brazier, J., & Basarir, H. (2016). The Capability Approach: A Critical Review of Its Application in Health Economics. *Value in Health*, 19(6), 795–799. <https://doi.org/10.1016/j.jval.2016.05.006>
- Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine Learning and Data Mining Methods in Diabetes Research. *Computational and Structural Biotechnology Journal*, 15, 104–116. <https://doi.org/10.1016/j.csbj.2016.12.005>
- Khasawneh, M. M. F., & Ibrahim, H. B. H. (2008). TOWARD AN INFORMATION AND COMMUNICATION TECHNOLOGY DEVELOPMENT IN DEVELOPING COUNTRIES. *Innovation and Knowledge Management in Business Globalization: Theory & Practice, Vols 1 and 2*, 908–913.

- Khoja, S., Durrani, H., Scott, R. E., Sajwani, A., & Piryani, U. (2013). Conceptual framework for development of comprehensive e-health evaluation tool. *Telemedicine Journal and E-Health: The Official Journal of the American Telemedicine Association*, 19(1), 48–53. <https://doi.org/10.1089/tmj.2012.0073>
- Kim, Y., & Crowston, K. (2011). Technology adoption and use theory review for studying scientists' continued use of cyber-infrastructure. In *American Society for Information Science and Technology Annual Meeting* (pp. 1–14). <https://doi.org/10.1002/meet.2011.14504801197>
- Kleine, D. (2010). ICT4WHAT:-Using the Choice Framework to Operationalise the Capability Approach to Development. *Journal If International Development*, 22, 647–692. <https://doi.org/10.1002/jid>
- Kleine, D. (2015). ICT4D. In *The international encyclopedia of digital communication and society* (pp. 315–323). <https://doi.org/10.1002/9781118290743.wbiedcs161>
- Kleine, D., & Unwin, T. (2009). Technological revolution, evolution and new dependencies: What's new about ICT4D? *Third World Quarterly*, 30(5), 1045–1067. <https://doi.org/10.1080/01436590902959339>
- Koller, D., & Friedman, N. (2009). Probabilistic Graphical Models: Principles and Techniques. In *The MIT Press* (Vol. 51, p. 16). MIT Press. <https://doi.org/10.2307/478142>
- Korb, K., & Nicholson, A. (2004). *Bayesian Artificial Intelligence* (second edi, Vol. 2). Chapman & Hall.
- Kotsiantis, S. B. (2007). Supervised Machine Learning: A Review of Classification Techniques. *Informatica*, 31, 249–268. <https://doi.org/10.1115/1.1559160>
- Lacerda, T. C., Nunes, J. V., & Von Wangenheim, C. G. (2015). *Usability heuristics for mobile phone applications: A literature review. Emerging Perspectives on the Design, Use, and Evaluation of Mobile and Handheld Devices*. <https://doi.org/10.4018/978-1-4666-8583-3.ch007>
- Law, I., & Widdows, H. (2008). Conceptualising health: Insights from the capability approach. *Health Care Analysis*, 16(4), 303–314. <https://doi.org/10.1007/s10728-007-0070-8>
- Lawrence-Jones, J. (2010). Dual diagnosis (drug/alcohol and mental health): Service user experiences. *Practice*, 22(2), 115–131. <https://doi.org/10.1080/09503151003686684>
- Lieberman, H. (2009). User Interface Goals, AI Opportunities. *AI Magazine*, 30(4), 16–22. <https://doi.org/10.1609/AIMAG.V30I4.2266>
- Lilholt, P. H., Hæsum, L. K. E., & Hejlesen, O. K. (2015). Exploring User Experience of a

- Telehealth System for the Danish TeleCare North Trial. In *Studies in Health Technology and Informatics* (Vol. 210, pp. 301–305). <https://doi.org/10.3233/978-1-61499-512-8-301>
- Lilholt, P., Jensen, M., & Hejlesen, O. (2015). Heuristic evaluation of a telehealth system from the Danish TeleCare North Trial. *Int J Med Inform. Available at Http://Www.Sciencedirect.Com/Science/Article/Pii/S1386505615000155*, 84(5), 319–326. <https://doi.org/10.1016/j.ijmedinf.2015.01.012>
- Lim, G. Y., Tam, W. W., Lu, Y., Ho, C. S., Zhang, M. W., & Ho, R. C. (2018). Prevalence of Depression in the Community from 30 Countries between 1994 and 2014. *Scientific Reports*, 8(1). <https://doi.org/10.1038/s41598-018-21243-x>
- Limerick, H., Coyle, D., & Moore, J. W. (2014). The experience of agency in human-computer interactions: a review. *Frontiers in Human Neuroscience*, 8. <https://doi.org/10.3389/fnhum.2014.00643>
- Lluch, M. (2011). Healthcare professionals' organisational barriers to health information technologies-A literature review. *International Journal of Medical Informatics*, 80(12), 849–862. <https://doi.org/10.1016/j.ijmedinf.2011.09.005>
- Loh, Y. A. C. (2015). Approaches to ICT for development (ICT4D): vulnerabilities vs. capabilities. *Information Development*, 31(3), 229–238. <https://doi.org/10.1177/0266666913513198>
- Lorgelly, P. K., Lorimer, K., Fenwick, E. A. L., Briggs, A. H., & Anand, P. (2015). Operationalising the capability approach as an outcome measure in public health: The development of the OCAP-18. *Social Science and Medicine*, 142, 68–81. <https://doi.org/10.1016/j.socscimed.2015.08.002>
- Lu, C. (2005). *Probabilistic Machine Learning Approaches To Medical Classification Problems*. Departement Elektrotechniek, Katholieke Universiteit Leuven, Belgium.
- Lucas, P., Gaag, V. Der, & Abu-Hanna, A. (2004). Bayesian networks in biomedicine and health-care. *Artificial Intelligence in Medicine*, 30(3), 201–214. <https://doi.org/10.1016/j.artmed.2003.11.001>
- Luppa, M., Sikorski, C., Luck, T., Ehreke, L., Konnopka, A., Wiese, B., ... Riedel-Heller, S. G. (2012). Age- and gender-specific prevalence of depression in latest-life - Systematic review and meta-analysis. *Journal of Affective Disorders*. <https://doi.org/10.1016/j.jad.2010.11.033>
- Lyon, G. J., & Segal, J. P. (2013). Practical, ethical and regulatory considerations for the

- evolving medical and research genomics landscape. *Applied and Translational Genomics*, 2(1), 34–40. <https://doi.org/10.1016/j.atg.2013.02.001>
- Maaten, L. J. P., & Hinton, G. E. (2008). Visualizing high-dimensional data using t-sne. *Journal of Machine Learning Research*, 9, 2579–2605. <https://doi.org/10.1007/s10479-011-0841-3>
- Mandal, I. (2014). Developing new machine learning ensembles for quality spine diagnosis. *Knowledge-Based Systems*, 73(1), 298–310. <https://doi.org/10.1016/j.knosys.2014.10.012>
- Marco, G., Rebecca, F., Atau, T., Baptiste, C., Jérémie, G., Frédéric, B., ... Saleema, A. (2016). Human-Centered Machine Learning. In *Conference on Human Factors in Computing Systems (CHI 2016), May 2016, San Jose, California, United States*. (pp. 3558–3565).
- Martin, J. L. (2010). *Soft Computing based learning and Data Analysis: Missing Values and Data Complexity*.
- Martínez-Pérez, B., De La Torre-Díez, I., & López-Coronado, M. (2013). Mobile health applications for the most prevalent conditions by the world health organization: Review and analysis. *Journal of Medical Internet Research*. <https://doi.org/10.2196/jmir.2600>
- Mazibuko, S. (2013). Understanding underdevelopment through the sustainable livelihoods approach. *Community Development*, 44(2), 173–187. <https://doi.org/10.1080/15575330.2012.683798>
- McCoy, A. B., Waitman, L. R., Lewis, J. B., Wright, J. A., Choma, D. P., Miller, R. A., & Peterson, J. F. (2012). A framework for evaluating the appropriateness of clinical decision support alerts and responses. *Journal of the American Medical Informatics Association*, 19(3), 346–352. <https://doi.org/10.1136/amiajnl-2011-000185>
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115–133. <https://doi.org/10.1007/BF02478259>
- McCurdie, T., Taneva, S., Casselman, M., Yeung, M., McDaniel, C., Ho, W., & Cafazzo, J. (2012). mHealth consumer apps: the case for user-centered design. *Biomedical Instrumentation & Technology / Association for the Advancement of Medical Instrumentation, Suppl*, 49–56. <https://doi.org/10.2345/0899-8205-46.s2.49>
- McKim, C. A. (2017). The Value of Mixed Methods Research: A Mixed Methods Study. *Journal of Mixed Methods Research*, 11(2), 202–222. <https://doi.org/10.1177/1558689815607096>

- Mehdi, D., & Maryam, A. (2014). Cross-cultural Study of the Prevalence of Depressive Symptoms and Depressive Severity: The Cases of Iranian and Indian Adolescents. *International Journal of Psychology*. Available at [Http://Www.Sid.Ir/En/VEWSSID/J\\_pdf/5062520140103.Pdf](Http://Www.Sid.Ir/En/VEWSSID/J_pdf/5062520140103.Pdf), 8(1), 43–63.
- Mertens, D. M. (2008). *Transformative Research and Evaluation*. [ *Evaluation*. Retrieved from <http://books.google.com/books?hl=en&lr=&id=q2fy6hriphYC&oi=fnd&pg=PA1&dq=Transformative+research+and+evalua&ots=nljbAjsQy&sig=AP5e0KTZTQrGbOAdybTSAvbHE2g%5Cnhttp://books.google.com/books?hl=en&lr=&id=q2fy6hriphYC&o>
- Miche, Y., Oliver, I., Ren, W., Holtmanns, S., Akusok, A., & Lendasse, A. (2017). Practical estimation of mutual information on non-euclidean spaces. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. [https://doi.org/10.1007/978-3-319-66808-6\\_9](https://doi.org/10.1007/978-3-319-66808-6_9)
- Mitchell, P. M., Roberts, T. E., Barton, P. M., & Coast, J. (2017). Applications of the Capability Approach in the Health Field: A Literature Review. *Social Indicators Research*, 133(1), 345–371. <https://doi.org/10.1007/s11205-016-1356-8>
- Miyamoto, D., Hazeyama, H., & Kadobayashi, Y. (2009). An evaluation of machine learning-based methods for detection of phishing sites. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 5506 LNCS, pp. 539–546). [https://doi.org/10.1007/978-3-642-02490-0\\_66](https://doi.org/10.1007/978-3-642-02490-0_66)
- Molin, E. (2016). *Evaluation of medication-related clinical decision support using user interface design principles: A qualitative study of Sfinx and NjuRen*. Kalinska institute of Health Informatics, University of Stockholm.
- Morrison, K. M., Shin, S., Tarnopolsky, M., & Taylor, V. H. (2014). Association of depression & health related quality of life with body composition in children and youth with obesity. *Journal of Affective Disorders*, 172, 18–23. <https://doi.org/10.1016/j.jad.2014.09.014>
- Morse, J. M. (2016). *Essentials of Qualitatively-Driven Mixed-Method Designs*. *Essentials of Qualitatively-Driven Mixed-Method Designs*. <https://doi.org/10.4324/9781315543406>
- Moustakis, V., & Herrmann, J. (1997). Where do machine learning human computer interaction meet? *Applied Artificial Intelligence*, 11, 595–609. Retrieved from <https://www.tandfonline.com/doi/pdf/10.1080/088395197117948?needAccess=true>

- Muja, M., & Lowe, D. G. (2014). Scalable Nearest Neighbour Algorithms for High Dimensional Data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(11), 2227–2240. <https://doi.org/10.1109/TPAMI.2014.2321376>
- Murphy, K. (2012). *Machine Learning: A Probabilistic Perspective*. The MIT Press, Cambridge. [https://doi.org/10.1007/SpringerReference\\_35834](https://doi.org/10.1007/SpringerReference_35834)
- Nadaraya, E. A. (1964). On Estimating Regression. *Theory of Probability & Its Applications*, 9(1), 141–142. <https://doi.org/10.1137/1109020>
- Narasimha, S., Agnisarman, S., Madathil, K. C., Gramopadhye, A. K., Welch, B., & McElligott, J. (2016). An investigation of the usability issues of home-based video telemedicine systems with geriatric patients. In *Proceedings of the Human Factors and Ergonomics Society*. <https://doi.org/10.1177/1541931213601412>
- NASS. (2012). Nigerian Disability Act. *Nationa Assembly Journal*, 8(48), 23. Retrieved from <http://placng.org/wp/wp-content/uploads/2016/04/DisabilityBill.pdf>
- Ndukwe, E. (2011). *The Telecommunication Revolution in Nigeria*. Lagos Business School's Centre for Infrastructure, Policy, Regulation and Advancement (CIRPA).
- Negnevitsky, M. (2005). *Artificial intelligence: A Guide to Intelligent Systems* (Second edi). Addison Wesley: Pearson Education.
- Nendaz, M., & Perrier, A. (2012). Diagnostic errors and flaws in clinical reasoning: Mechanisms and prevention in practice. *Swiss Medical Weekly*. <https://doi.org/10.4414/smw.2012.13706>
- Netica. (2010). *Netica-J Reference Manual*. Retrieved from [http://www.norsys.com/netica-j/docs/NeticaJ\\_Man.pdf](http://www.norsys.com/netica-j/docs/NeticaJ_Man.pdf)
- Nielsen, J. (1993). *Usability Engineering*. *Usability Engineering* (Vol. 44). <https://doi.org/10.1145/1508044.1508050>
- Nielsen, J. (2010). What is Usability? In *In User Experience Re-Mastered* by Wilson, Chauncey (pp. 4–7). Morgan Kaufmann. <https://doi.org/10.1016/C2009-0-20682-9>
- Nielsen, J., & Molich, R. (1990). Heuristic Evaluation of user interfaces. *CHI '90 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (April), 249–256. <https://doi.org/10.1145/97243.97281>
- Nigeria Federal Ministry of Health. (2014). National Health Information System STRATEGIC PLAN 2014 - 2018 Federal Ministry of Health Department of Health Planning , Research and Statistics. *National Health Information System STRATEGIC PLAN*, 1–43. Retrieved from <http://www.health.gov.ng/doc/National> HIS Strategic

- Plan.pdf%0Ahttps://dhis2nigeria.org.ng/dhis/api/documents/QCdRZMqmoF2/data
- NPC. (2013). Nigeria demographic and health survey 1999. ... , *MD: National Population Commission and ORC* ..., 59. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Nigeria+Demographic+And+Health+Survey#0>
- Nussbaum, M. C. (2005). Capabilities as fundamental entitlements: Sen and social justice. In *Capabilities Equality: Basic Issues and Problems* (pp. 44–70). <https://doi.org/10.4324/9780203799444>
- Nussbaum, M. C. (2011). Capabilities, entitlements, rights: Supplementation and critique. *Journal of Human Development and Capabilities*, 12(1), 23–37. <https://doi.org/10.1080/19452829.2011.541731>
- O'Connor, E., Rossom, R. C., Henninger, M., Groom, H. C., & Burda, B. U. (2016). Primary Care Screening for and Treatment of Depression in Pregnant and Postpartum Women. *JAMA*, 315(4), 388. <https://doi.org/10.1001/jama.2015.18948>
- Obadeji, A., Oluwole, L. O., Dada, M. U., Ajiboye, A. S., Kumolalo, B. F., & Solomon, O. a. (2015). Assessment of Depression in a Primary Care Setting in Nigeria using the PHQ-9. *Journal of Family Medicine and Primary Care*, 4(1), 30–34. <https://doi.org/10.4103/2249-4863.152246>
- Oche, M., & Adamu, H. (2013). Determinants of patient waiting time in the general outpatient department of a tertiary health institution in north Western Nigeria. *Annals of Medical and Health Sciences Research*. <https://doi.org/10.4103/2141-9248.122123>
- OECD. (2003). Papers on Official Development Assistance (ODA). *OECD Journal on Development*, 3(4), 5–35. [https://doi.org/10.1787/journal\\_dev-v3-art27-en](https://doi.org/10.1787/journal_dev-v3-art27-en)
- Ojeme, B., & Mbogho, A. (2016a). Predictive Strength of Bayesian Networks for Diagnosis of Depressive Disorders. In *Proceedings of the 8th KES International Conference on Intelligent Decision Technologies (KES-IDT 2016)-Tenerife, Spain, June 15-17, Vol 56, Pp.373-382* . Available at: [http://link.springer.com/chapter/10.1007/978-3-319-39630-9\\_31](http://link.springer.com/chapter/10.1007/978-3-319-39630-9_31) (pp. 373–382). Tenerife, Spain: Springer International Publishing. [https://doi.org/10.1007/978-3-319-39630-9\\_31](https://doi.org/10.1007/978-3-319-39630-9_31)
- Ojeme, B., & Mbogho, A. (2016b). Selecting Learning Algorithms for Simultaneous Identification of Depression and Comorbid Disorders. In *Proceedings of the 20th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES2016)*. York, UK, September 5-7, Vol. 96, Pp 1294-1303.

- Available at: <http://www.sciencedirect.com/science/article/pii/S18770509163198> (Vol. 96, pp. 1294–1303). The Author(s). <https://doi.org/10.1016/j.procs.2016.08.174>
- Ojeme, B., Mbogho, A., & Meyer, T. (2016). Probabilistic expert systems for reasoning in clinical depressive disorders. In *Proceedings of the 15th IEEE International Conference on Machine Learning and Applications, Anaheim, California, USA, December 18-20, Pp 599-604*. Available at <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7838209> (pp. 599–604). Retrieved from <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7838209>
- Olatokun, W. M. (2009). Information and Communication Technology Use By Reproductive Health Workers in Nigeria : State of the Art , Issues , and Challenges. *Environments*, 5(November), 181–207.
- Orrù, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., & Mechelli, A. (2012). Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. *Neuroscience and Biobehavioral Reviews*. <https://doi.org/10.1016/j.neubiorev.2012.01.004>
- Osemene, K. P., & Lamikanra, A. (2012). A study of the prevalence of self-medication practice among university students in southwestern Nigeria. *Tropical Journal of Pharmaceutical Research*, 11(4), 683–689. <https://doi.org/10.4314/tjpr.v11i4.21>
- Oteniya, L. (2008). *Bayesian Belief Networks for Dementia Diagnosis and Other Applications: A Comparison of Hand-Crafting and Construction using A Novel Data Driven Technique. A PhD Thesis*. Department of Computing Science and Mathematics, University of Stirling.
- Oteniya, L., Cowie, J., & Coles, R. (2005). Diagnosis of Dementia and Its Pathologies Using Bayesian Belief Networks. In *Proceedings of the 22nd HealthCare Computing Conference* (pp. 289–297).
- Othieno, C. J., Okoth, R. O., Peltzer, K., Pengpid, S., & Malla, L. O. (2014). Depression among university students in Kenya: Prevalence and sociodemographic correlates. *Journal of Affective Disorders*, 165, 120–125. <https://doi.org/10.1016/j.jad.2014.04.070>
- Otieno, O. C., Liyala, S., Odongo, B. C., & Abeka, S. (2016). Theory of Reasoned Action as an Underpinning to Technological Innovation Adoption Studies. *World Journal of Computer Application and Technology*, 4(1), 1–7. <https://doi.org/10.13189/wjcat.2016.040101>
- Oyegoke, L. (2013). *Adoption and Utilization of ICT in Nigeria Hospitals (Government*

- Owned*). Degree programme in Business IT, Haaga-Helia university of Applied Sciences. Retrieved from [https://www.theseus.fi/bitstream/handle/10024/66898/Oyegoke\\_Thesis.pdf;sequence=1](https://www.theseus.fi/bitstream/handle/10024/66898/Oyegoke_Thesis.pdf;sequence=1)
- Panir, M. J. H. (2011). Role of ICTs in the health sector in developing countries: a critical review of literature. *Journal of Health Informatics in Developing Countries*, 5, 197–208. Retrieved from <http://jhidc.org/index.php/jhidc/article/view/61/96%5Cnhttp://shibboleth.ovid.com/secure/?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=cagh&AN=20113197964%5Cnhttp://sfx.kcl.ac.uk/kings?genre=article&atitle=Role+of+ICTs+in+the+health+sector+in+developing+countries:+a>
- Patel, M. J., Khalaf, A., & Aizenstein, H. J. (2016a). Studying depression using imaging and machine learning methods. *NeuroImage: Clinical*. <https://doi.org/10.1016/j.nicl.2015.11.003>
- Patel, M. J., Khalaf, A., & Aizenstein, H. J. (2016b). Studying depression using imaging and machine learning methods. *NeuroImage: Clinical*, 10, 115–123. <https://doi.org/10.1016/j.nicl.2015.11.003>
- Payne, T. H., Hines, L. E., Chan, R. C., Hartman, S., Kapusnik-Uner, J., Russ, A. L., ... Malone, D. C. (2015). Recommendations to improve the usability of drug-drug interaction clinical decision support alerts. *Journal of the American Medical Informatics Association*, 22(6), 1243–1250. <https://doi.org/10.1093/jamia/ocv011>
- Paz, F., & Pow-Sang, J. A. (2015). Usability Evaluation Methods for Software Development: A Systematic Mapping Review. In *2015 8th International Conference on Advanced Software Engineering & Its Applications (ASEA)* (pp. 1–4). <https://doi.org/10.1109/ASEA.2015.8>
- Peischl, B., Ferk, M., & Holzinger, A. (2013). Integrating user-centred design in an early stage of mobile medical application prototyping a case study on data acquisition in health organisations. *ICETE 2013 - 10th Int. Joint Conf. on E-Business and Telecommunications; 4th Int. Conf. DCNET 2013, - 10th Int. Conf. on ICE-B 2013 and OPTICS 2013 - 4th Int. Conf. on Optical Communication Systems*, 185–195. <https://doi.org/10.5220/0004493901850195>
- Peltzer, K., Pengpid, S., Olowu, S., & Olasupo, M. (2013). Depression and Associated Factors Among University Students in Western Nigeria. *Journal of Psychology in Africa*, 23(3), 459–466. <https://doi.org/1433-0237>

- Peng, L. K., Ramaiah, C. K., & Foo, S. (2004). Heuristic-based user interface evaluation at Nanyang Technological University in Singapore. *Program: Electronic Library and Information Systems*, 38(1), 42–59. <https://doi.org/10.1108/00330330410519198>
- Petersen, E. K., & Pedersen, M. L. (2010). The Sustainable Livelihoods Approach. From a psychological perspective. *Knowledge Solutions*, (June), 21–26. [https://doi.org/10.1007/978-981-10-0983-9\\_5](https://doi.org/10.1007/978-981-10-0983-9_5)
- Pevnick, J. M., Li, X., Grein, J., Bell, D. S., & Silka, P. (2013). A Retrospective Analysis of Interruptive versus Non-interruptive Clinical Decision Support for Identification of Patients Needing Contact Isolation. *Applied Clinical Informatics*, 4(4), 569–582. <https://doi.org/10.4338/ACI-2013-04-RA-0021>
- Pollino, C., & Henderson, C. (2010). *Bayesian networks : A guide for their application in natural resource Management and Policy. Integrated Catchment Assessment and Management Centre, Fenner School of Environment and Society, Australian National University, Canberra. A Technical Report.* Retrieved from [http://www.utas.edu.au/\\_\\_data/assets/pdf\\_file/0009/588474/TR\\_14\\_BNs\\_a\\_resource\\_guide.pdf](http://www.utas.edu.au/__data/assets/pdf_file/0009/588474/TR_14_BNs_a_resource_guide.pdf)
- Preece, J., Rogers, Y., & Sharp, H. (2015). Interaction Design: Beyond Human-Computer Interaction. *Book*. <https://doi.org/10.1162/leon.2005.38.5.401>
- Qian, W., & Shu, W. (2015). Mutual information criterion for feature selection from incomplete data. *Neurocomputing*. <https://doi.org/10.1016/j.neucom.2015.05.105>
- Quinlan, J. R. (1992). *C4.5: Programs for Machine Learning. Morgan Kaufmann San Mateo California* (Vol. 1). [https://doi.org/10.1016/S0019-9958\(62\)90649-6](https://doi.org/10.1016/S0019-9958(62)90649-6)
- Quinlan, J. R. (2006). Bagging, boosting, and C4.5. *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, 5(Quinlan 1993), 725–730.
- Ramírez, J., Górriz, J. M., Salas-Gonzalez, D., Romero, A., López, M., Álvarez, I., & Gómez-Río, M. (2013). Computer-aided diagnosis of Alzheimer's type dementia combining support vector machines and discriminant set of features. In *Information Sciences* (Vol. 237, pp. 59–72). <https://doi.org/10.1016/j.ins.2009.05.012>
- Ramlo, S. (2016). Mixed Method Lessons Learned From 80 Years of Q Methodology. *Journal of Mixed Methods Research*, 10(1), 28–45. <https://doi.org/10.1177/1558689815610998>
- Raudonis, B. M. (1992). Ethical considerations in qualitative research with hospice patients. *Qualitative Health Research*, 2(2), 238–249. <https://doi.org/10.1177/104973239200200207>

- Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., & Yang, G. Z. (2017). Deep Learning for Health Informatics. *IEEE Journal of Biomedical and Health Informatics*, 21(1), 4–21. <https://doi.org/10.1109/JBHI.2016.2636665>
- Razzouk, D., Sharan, P., Gallo, C., Gureje, O., Lamberte, E. E., de Jesus Mari, J., ... Saxena, S. (2010). Scarcity and inequity of mental health research resources in low-and-middle income countries: A global survey. *Health Policy*, 94(3), 211–220. <https://doi.org/10.1016/j.healthpol.2009.09.009>
- Read, J., Reutemann, P., Pfahringer, B., & Holmes, G. (2016). MEKA: A Multi-label / Multi-target Extension to WEKA. *Journal of Machine Learning Research*. Available at: <Http://Jmlr.Org/Papers/Volume17/12-164/12-164.Pdf>, 17(21), 1–5.
- Reolon, M., Lacerda, T. C., Krone, C., Gresse, C., & Wangenheim, V. (2016). Usability Heuristics for Evaluating Health-care Applications for Smartphones: A Systematic Literature Review Computação na Escola View project Evaluation of Educational Games for Computing Education View project. <https://doi.org/10.13140/RG.2.2.34854.73282>
- Ribeiro, V. S., Martins, A. I., Queirós, A., Silva, A. G., & Rocha, N. P. (2015). Usability Evaluation of a Health Care Application Based on IPTV. In *Procedia Computer Science* (Vol. 64, pp. 635–642). <https://doi.org/10.1016/j.procs.2015.08.577>
- Robeyns, I. (2005). The Capability Approach: a theoretical survey. *Journal of Human Development*, 6(1), 93–117. <https://doi.org/10.1080/146498805200034266>
- Rogers, E. (2003). *Diffusion of innovations*. *Diffusion of Innovation*. <https://doi.org/citeulike-article-id:126680>
- Rogers, E. (2015). Evolution: Diffusion of Innovations. In *International Encyclopedia of the Social & Behavioral Sciences: Second Edition* (pp. 378–381). <https://doi.org/10.1016/B978-0-08-097086-8.81064-8>
- Ruxwana, N. L., Herselman, M. E., & Conradie, D. P. (2010). ICT applications as e-health solutions in rural healthcare in the Eastern Cape Province of South Africa. *The HIM Journal*. <https://doi.org/10.1109/ISMICT.2015.7107509>
- Saito, T., & Rehmsmeier, M. (2015). The Precision-Recall Plot is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PloS One*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- Sakthivel, N. R., Nair, B. B., Elangovan, M., Sugumaran, V., & Saravanmurugan, S. (2014). Comparison of dimensionality reduction techniques for the fault diagnosis of mono block centrifugal pump using vibration signals. *Engineering Science and Technology, an*

- International Journal*, 17(1), 30–38. <https://doi.org/10.1016/j.jestch.2014.02.005>
- Salihu, A. S. (2015). Impact of Somatic Symptoms on Identification of Depression among General Outpatients by Family Physicians in. *Open Journal of Psychiatry*, 5(July), 278–284.
- Sato, J. R., Moll, J., Green, S., Deakin, J. F. W., Thomaz, C. E., & Zahn, R. (2015). Machine learning algorithm accurately detects fMRI signature of vulnerability to major depression. *Psychiatry Research - Neuroimaging*, 233(2), 289–291. <https://doi.org/10.1016/j.psychres.2015.07.001>
- Schaakxs, R., Comijs, H. C., van der Mast, R. C., Schoevers, R. A., Beekman, A. T. F., & Penninx, B. W. J. H. (2017). Risk Factors for Depression: Differential Across Age? *American Journal of Geriatric Psychiatry*, 25(9), 966–977. <https://doi.org/10.1016/j.jagp.2017.04.004>
- Schmid, W., Bronisch, T., & Von Zerssen, D. (1982). A comparative study of PSE/CATEGO and DiaSiKa: Two psychiatric computer diagnostic systems. *British Journal of Psychiatry*, 141(3), 292–295.
- Schnall, R., Cho, H., & Liu, J. (2018). Health Information Technology Usability Evaluation Scale (Health-ITUES) for Usability Assessment of Mobile Health Technology: Validation Study. *JMIR MHealth and UHealth*. <https://doi.org/10.2196/mhealth.8851>
- Schnall, R., Rojas, M., Bakken, S., Brown, W., Carballo-Diequez, A., Carry, M., ... Travers, J. (2016). A user-centered model for designing consumer mobile health (mHealth) applications (apps). *Journal of Biomedical Informatics*. <https://doi.org/10.1016/j.jbi.2016.02.002>
- Sein, M. K., Hatakka, M., Thapa, D., & Sæbø, Ø. (2016). What theories do we need to know to conduct ICT4D research? In *SIG GlobDev Ninth Annual Workshop* (pp. 1–19).
- Seixas, F. L., Zadrozny, B., Laks, J., Conci, A., & Muchaluat Saade, D. C. (2014). A Bayesian Network Decision Model for Supporting the Diagnosis of Dementia, Alzheimer's Disease and Mild Cognitive Impairment. *Computers in Biology and Medicine*, 51, 140–158. <https://doi.org/10.1016/j.combiomed.2014.04.010>
- Semrau, M., Evans-Lacko, S., Alem, A., Ayuso-Mateos, J. L., Chisholm, D., Gureje, O., ... Thornicroft, G. (2015). Strengthening mental health systems in low- and middle-income countries: the Emerald programme. *BMC Medicine*, 13(1), 79. <https://doi.org/10.1186/s12916-015-0309-4>
- Sen, A. (1997). From Income Inequality to Economic Inequality. *Southern Economic Journal*,

- 64(2), 384–401. <https://doi.org/10.2307/1060857>
- Sen, A. (2000). *SOCIAL EXCLUSION: CONCEPT, APPLICATION, AND SCRUTINY*. Amartya Sen. *Social Development*. <https://doi.org/10.1.1.100.1010>
- Sen, A. (2001). Introduction “Development as freedom.” In *Development as freedom* (pp. 3–11). <https://doi.org/10.1215/0961754X-9-2-350>
- Sen, A. (2007). Capability and well-being. In *The Philosophy of Economics: An Anthology* (pp. 270–294). <https://doi.org/10.1017/CBO9780511819025.019>
- Serrat, O. (2017). The Sustainable Livelihoods Approach. In *Knowledge Solutions* (pp. 21–26). [https://doi.org/10.1007/978-981-10-0983-9\\_5](https://doi.org/10.1007/978-981-10-0983-9_5)
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*. <https://doi.org/10.1145/584091.584093>
- Sharma, P., & Dhir, S. (2016). Functional & non-functional requirement elicitation and risk assessment for agile processes. *International Journal of Control Theory and Applications*, 9(18), 9005–9010.
- Sharp, H., Rogers, Y., & Preece, J. (2011). Interaction Design: Beyond Human-Computer Interaction. *Book*. <https://doi.org/10.1162/leon.2005.38.5.401>
- Shneiderman, B. (2002). Inventing discovery tools: Combining information visualization with data mining. *Information Visualization*. <https://doi.org/10.1057/palgrave.ivs.9500006>
- Shojaei Estabragh, Z., Riahi Kashani, M. M., Jeddi Moghaddam, F., Sari, S., Taherifar, Z., Moradi Moosavy, S., & Sadeghi Oskooyee, K. (2013). Bayesian network modeling for diagnosis of social anxiety using some cognitive-behavioral factors. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 2(4), 257–265. <https://doi.org/10.1007/s13721-013-0042-x>
- Shortliffe, E., & Lindberg, D. (1976). *Computer-Based Medical Consultations: Mycin*. *Computer-Based Medical Consultations: Mycin*. <https://doi.org/10.1016/B978-0-444-00179-5.50005-6>
- Silipo, R., Aday, I., Hart, A., & Berthold, M. (2014). Seven Techniques for Dimensionality Reduction. *KNIME.Com*, 1–21.
- Simon, J., Anand, P., Gray, A., Rugkåsa, J., Yeeles, K., & Burns, T. (2013). Operationalising the capability approach for outcome measurement in mental health research. *Social Science & Medicine*, 98, 187–196. <https://doi.org/10.1016/j.socscimed.2013.09.019>
- Sluga, D., & Lotrič, U. (2017). Quadratic mutual information feature selection. *Entropy*. <https://doi.org/10.3390/e19040157>

- Smarsly, K., Dragos, K., & Wiggenbrock, J. (2016). Machine learning techniques for structural health monitoring. *Proceedings of the 8th European Workshop On Structural Health Monitoring (EWSHM 2016)*, (July), 5–8. Retrieved from <http://www.ndt.net/search/docs.php3?showForm=off&id=19828Machine>
- Snaith, R. P., & Zigmond, A. S. (1986). The hospital anxiety and depression scale. *British Medical Journal (Clinical Research Ed.)*. <https://doi.org/10.1136/bmj.292.6516.344>
- SPitzer, R., & Endicott, J. (1969). DIAGNO II: Further Developments in a Computer Program for Psychiatric Diagnosis. *American Journal of Psychiatry*, 125(7S), 12–21. <https://doi.org/10.1176/ajp.125.7S.12>
- Spitzer, R. L., & Endicott, J. (1968). DIAGNO: A Computer Program for Psychiatric Diagnosis Utilizing the Differential Diagnostic Procedure. *Archives of General Psychiatry*, 18(6), 746–756. <https://doi.org/10.1001/archpsyc.1968.01740060106013>
- Spitzer, R. L., Kroenke, K., & Williams, J. B. W. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ Primary Care Study. *Journal of the American Medical Association*, 282(18), 1737–1744. <https://doi.org/10.1001/jama.282.18.1737>
- Stein, D. J. (2012). Dimensional or categorical: different classifications and measures of anxiety and depression. *Medicographia*, 34, 270–275. Retrieved from <http://www.medicographia.com/2013/01/dimensional-or-categorical-different-classifications-and-measures-of-anxietyand-depression/>
- Stevens, M. (2013). Ethical issues in qualitative research Research ethics. *Social Care Workforce Research Unit, King's College London*, 1–41.
- Su, C., Andrew, A., Karagas, M. R., & Borsuk, M. E. (2013). Using Bayesian networks to discover relations between genes, environment, and disease. *BioData Mining*, 6(1), 6. <https://doi.org/10.1186/1756-0381-6-6>
- Suhasini, A., Palanivel, S., & Ramalingam, V. (2010). Multi decision support model for Psychiatry Problem. *International Journal of Computer Applications*, 1(25), 74–84. <https://doi.org/10.5120/456-760>
- Sumathi, M. R., & Poorna, B. (2015). A Bayesian Framework for Diagnosing Depression Level of Adolescents. *International Conference on Computing Intelligence Systems*, 4(March), 1350–1354.
- Suzuki, K. (2017). Overview of deep learning in medical imaging. *Radiological Physics and Technology*. <https://doi.org/10.1007/s12194-017-0406-5>
- Tang, J., Deng, C., & Guang, G.-B. (2015). Extreme learning machine for multilayer

- perceptron. *IEEE Transactions on Neural Networks and Learning Systems*, 27(4), 809–821. <https://doi.org/10.1109/TNNLS.2015.2424995>
- Tang, Z., Johnson, T., Tindall, D., & Zhang, J. (2006). Applying Heuristic Evaluation to Improve the Usability of a Telemedicine System. *Telemedicine Journal and E-Health*. Available at: <Http://Dx.Doi.Org/10.1089/Tmj.>, 12(1), 24–34. <https://doi.org/10.1089/tmj.2006.12.24>
- Taniar, D., & Chen, L. (2011). Integrations of Data Warehousing, Data Mining and Database Technologies. *Innovative Approaches*. <https://doi.org/10.4018/978-1-60960-537-7>
- Teufel, R. J., Kazley, A. S., Andrews, A. L., Ebeling, M. D., & Basco, W. T. (2013). Electronic medical record adoption in hospitals that care for children. *Academic Pediatrics*, 13(3), 259–263. <https://doi.org/10.1016/j.acap.2013.01.010>
- Thapa, D., & Hatakka, M. (2017). Understanding ICT in ICT4D: An Affordance Perspective. In *50th Hawaii International Conference on System Sciences (HICSS)* (pp. 2618–2626). <https://doi.org/10.24251/HICSS.2017.316>
- Thomas, M. A., Li, Y., & Oliveira, T. (2017). Nuances of development contexts for ICT4D research in least developed countries: An empirical investigation in Haiti. *Telematics and Informatics*, 34(7), 1093–1112. <https://doi.org/10.1016/j.tele.2017.05.001>
- Toyama, K. (2010). Human–Computer Interaction and Global Development: Introduction. *Foundations and Trends® in Human–Computer Interaction*, 4(1), 1–79. <https://doi.org/10.1561/11000000021>
- Toyama, K., & Dias, M. B. (2008). Information and communication technologies for development. *Computer*, 41(6), 22–25. <https://doi.org/10.1109/MC.2008.193>
- Twomey, C., & O'Reilly, G. (2016). Effectiveness of a freely available computerised cognitive behavioural therapy programme (MoodGYM) for depression: Meta-analysis. *Australian and New Zealand Journal of Psychiatry*, (August), 0004867416656258. <https://doi.org/10.1177/0004867416656258>
- UNCTAD. (2014). *A Framework For Information and Communications Technology Policy Reviews: Helping Countries Leverage ICT for Development*. United Nations, New York and Geneva. [http://unctad.org/en/PublicationsLibrary/dtlstict2013d6\\_en.pdf](http://unctad.org/en/PublicationsLibrary/dtlstict2013d6_en.pdf). Retrieved from [http://unctad.org/en/PublicationsLibrary/dtlstict2013d6\\_en.pdf](http://unctad.org/en/PublicationsLibrary/dtlstict2013d6_en.pdf)
- UNDP. (2015). *Human Development Report 2015: Work for human development*. Undp. Retrieved from [http://hdr.undp.org/sites/default/files/2015\\_human\\_development\\_report.pdf](http://hdr.undp.org/sites/default/files/2015_human_development_report.pdf)

- Usability.gov. (2012). Types of Evaluations.
- Usher-Smith, J., Emery, J., Hamilton, W., Griffin, S. J., & Walter, F. M. (2015). Risk prediction tools for cancer in primary care. *British Journal of Cancer*, *113*(12), 1645–1650. <https://doi.org/10.1038/bjc.2015.409>
- Valtonen, T., Kukkonen, J., Kontkanen, S., Sormunen, K., Dillon, P., & Sointu, E. (2015). The impact of authentic learning experiences with ICT on pre-service teachers' intentions to use ICT for teaching and learning. *Computers & Education*, *81*, 49–58. <https://doi.org/10.1016/j.compedu.2014.09.008>
- Varkey, P., Reller, M. K., & Resar, R. K. (2007). Basics of Quality Improvement in Health Care. *Mayo Clinic Proceedings*, *82*(6), 735–739. <https://doi.org/10.4065/82.6.735>
- Varshney, U. (2014). Mobile health. In *Proceedings of the 4th ACM MobiHoc workshop on Pervasive wireless healthcare - MobileHealth '14* (pp. 37–42). <https://doi.org/10.1145/2633651.2633656>
- Vincent, C., & Blandford, A. (2011). Designing for Safety and Usability: User-centred Techniques in Medical Device Design Practice, 793–797. <https://doi.org/10.1177/1071181311551164>
- Viola, P., & Jones, M. (2004). Robust real-time face detection. *International Journal of Computer Vision*, *57*(2), 137–154. <https://doi.org/10.1023/B:VISI.0000013087.49260.fb>
- Wang, L., Wang, Y., & Chang, Q. (2016). Feature selection methods for big data bioinformatics: A survey from the search perspective. *Methods*. <https://doi.org/10.1016/j.ymeth.2016.08.014>
- Wang, Y., Kung, L. A., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*, *126*, 3–13. <https://doi.org/10.1016/j.techfore.2015.12.019>
- Watson, G. S. (1964). Smooth regression analysis. *The Indian Journal of Statistics*, *26*(4), 359–372. <https://doi.org/10.2307/25049340>
- WHO. (2005). Health and the Millennium Development Goals. *World Health Organization Press*, 84. Retrieved from [http://www.who.int/hdp/publications/mdg\\_en.pdf?ua=1](http://www.who.int/hdp/publications/mdg_en.pdf?ua=1)
- WHO. (2012). *Depression: A global public health concern*. Available at [http://www.who.int/mental\\_health/management/depression/who\\_paper\\_depression\\_wfmh\\_2012.pdf](http://www.who.int/mental_health/management/depression/who_paper_depression_wfmh_2012.pdf). WHO Department of Mental Health and Substance Abuse. WHO Department of Mental Health and Substance Abuse. Retrieved from [http://www.who.int/mental\\_health/management/depression/who\\_paper\\_depression\\_wfmh](http://www.who.int/mental_health/management/depression/who_paper_depression_wfmh)

h\_2012.pdf

- WHO. (2013). mHealth: New horizons for health through mobile technologies. *Observatory*, 3(June), 66–71. <https://doi.org/10.4258/hir.2012.18.3.231>
- Wilkund, M.E; Kendler, J;Yale, A. S. (2013). *Usability Testing of Medical Devices. Journal of Chemical Information and Modeling*. <https://doi.org/10.1017/CBO9781107415324.004>
- Willis, A. (2014). UX for Healthcare: What You Need to Know Before You Start. *User Experience Magazine*, Available at <Http://Uxpamagazine.Org/Ux-for-Healthcare/>, 14(3).
- Wilson-Shaw, L., Pistrang, N., & Herlihy, J. (2012). Non-clinicians' judgments about asylum seekers' mental health: How do legal representatives of asylum seekers decide when to request medico-legal reports? *European Journal of Psychotraumatology*, 3. <https://doi.org/10.3402/ejpt.v3i0.18406>
- Wing, J. K., Babor, T., Brugha, T., Burke, J., Cooper, J. E., Giel, R., ... Sartorius, N. (1990). SCAN. Schedules for Clinical Assessment in Neuropsychiatry. *Archives of General Psychiatry*, 47(6), 589–593. <https://doi.org/10.1001/archpsyc.1990.01810180089012>
- Wisdom, J. P., Chor, K. H. B., Hoagwood, K. E., & Horwitz, S. M. (2014). Innovation adoption: a review of theories and constructs. *Administration and Policy in Mental Health*, 41(4), 480–502. <https://doi.org/10.1007/s10488-013-0486-4>
- Wiseman, S., Cox, A. L., & Brumby, D. P. (2013). Designing Devices With the Task in Mind: Which Numbers Are Really Used in Hospitals? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 55(1), 61–74. <https://doi.org/10.1177/0018720812471988>
- Witten, I., Frank, E., & Hall, M. (2011). *Data Mining : Practical Machine Learning Tools and Techniques* (Third Edit). Morgan KaufmannPublishers. [https://doi.org/10.1002/1521-3773\(20010316\)40:6<9823::AID-ANIE9823>3.3.CO;2-C](https://doi.org/10.1002/1521-3773(20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C)
- Wosley, J. M. (2011). Comparison of Contemporary Technology Acceptance Models and Evaluation of the Best Fit for Health Industry Organizations . *International Journal of Computer Science Engineering and Technology*, 1(11), 709–717.
- World Bank. (2013). *World Development Indicators 2013*. Vasa. <https://doi.org/10.1596/978-0-8213-9824-1>
- World Health Organization. (1990). Composite International Diagnostic Interview (CIDI). *WHO, Division of Mental Health*. <https://doi.org/10.1002/mpr.168>
- World Health Organization. (1992). *The ICD-10 Classification of Mental and Behavioural Disorders. Mental Health Division, World Health Organization* (Vol. 10).

[https://doi.org/10.1002/1520-6505\(2000\)](https://doi.org/10.1002/1520-6505(2000))

- World Health Organization. (2011a). *Technical Report on mhGAP: Mental Health Gap Action Programme*. World Health Organization, Department of Mental Health and Substance Abuse (MSD). Available at: [www.who.int/mental\\_health/mhgap/en/](http://www.who.int/mental_health/mhgap/en/). WHO, Department of Mental Health and Substance Abuse.
- World Health Organization. (2011b). World Health Organization Mental Health Atlas, Geneva. *World Health Organization*, 1–81. <https://doi.org/10.1093/bja/aes067>
- Worldometers. (2016). Nigeria Population (2016) -. *Nigeria Population (LIVE)*. Retrieved from <http://www.worldometers.info/world-population/nigeria-population/>
- Wright, P., Blythe, M., & McCarthy, J. (2006). User Experience and the Idea of Design in HCI BT - Interactive Systems Design. In *Interactive Systems Design* (Vol. 3941, pp. 1–14). <https://doi.org/10.1007/11752707>
- Wu, X., Kumar, V., Ross, Q. J., Ghosh, J., Yang, Q., Motoda, H., ... Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14(1), 1–37. <https://doi.org/10.1007/s10115-007-0114-2>
- Wyllie, D., & Davies, J. (2015). Role of data warehousing in healthcare epidemiology. *Journal of Hospital Infection*. <https://doi.org/10.1016/j.jhin.2015.01.005>
- Xuesong, Y., Wei, L., Qinghua, W., & Victor, S. (2016). A Double Weighted Naive Bayes for Multi-label Classification. In *In: Li K., Li J., Liu Y., Castiglione A. (eds) Computational Intelligence and Intelligent Systems. Communications in Computer and Information Science, vol 575. Springer, Singapore (pp. 382–389)*. [https://doi.org/https://dx.doi.org/10.1007/978-981-10-0356-1\\_40](https://doi.org/https://dx.doi.org/10.1007/978-981-10-0356-1_40)
- Yap, M., & Yu, E. (2016). Operationalising the capability approach: developing culturally relevant indicators of indigenous wellbeing – an Australian example. *Oxford Development Studies*, 44(3), 315–331. <https://doi.org/10.1080/13600818.2016.1178223>
- Yen, P.-Y., & Bakken, S. (2009). A comparison of usability evaluation methods: heuristic evaluation versus end-user think-aloud protocol - an example from a web-based communication tool for nurse scheduling. *AMIA ... Annual Symposium Proceedings. AMIA Symposium*.
- Yoon, S., & Kim, S. (2008). AdaBoost-based multiple SVM-RFE for classification of mammograms in DDSM. In *Proceedings - 2008 IEEE International Conference on Bioinformatics and Biomedicine Workshops, BIBMW* (pp. 75–82). <https://doi.org/10.1109/BIBMW.2008.4686212>

- Yu, W., Liu, T., Valdez, R., Gwinn, M., & Khoury, M. J. (2010). Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes. *BMC Medical Informatics and Decision Making*, *10*(1), 16. <https://doi.org/10.1186/1472-6947-10-16>
- Yuan, M. J., Finley, G. M., Long, J., Mills, C., & Johnson, R. K. (2013). Evaluation of user interface and workflow design of a bedside nursing clinical decision support system. *Journal of Medical Internet Research*, *15*(1). <https://doi.org/10.2196/ijmr.2402>
- Zaidi, N., & Cerquides, J. (2013). Alleviating Naive Bayes attribute independence assumption by attribute weighting. *The Journal of Machine ...*, *14*, 1947–1988. Retrieved from <http://dl.acm.org/citation.cfm?id=2567725>
- Zakaria, S., & Ghani, M. K. A. (2013). The Impact of EMR User Interface Design on Doctor Satisfaction. *E-Proceeding of Software Engineering Postgraduates Workshop (SEPoW)*, 94.
- Zayyad, M. A., & Toycan, M. (2018). Factors affecting sustainable adoption of e-health technology in developing countries: an exploratory survey of Nigerian hospitals from the perspective of healthcare professionals. *PeerJ*. <https://doi.org/10.7717/peerj.4436>
- Zhang, X., Yu, P., Yan, J., & Ton, S. I. (2015). Using diffusion of innovation theory to understand the factors impacting patient acceptance and use of consumer e-health innovations: A case study in a primary care clinic. *BMC Health Serv Res*, *15*, 71. <https://doi.org/10.1186/s12913-015-0726-2>
- Zheng, T., Xie, W., Xu, L., He, X., Zhang, Y., You, M., ... Chen, Y. (2017). A machine learning-based framework to identify type 2 diabetes through electronic health records. *International Journal of Medical Informatics*, *97*, 120–127. <https://doi.org/10.1016/j.ijmedinf.2016.09.014>
- Zhou, L., & Lai, K. K. (2016). AdaBoost Models for Corporate Bankruptcy Prediction with Missing Data. *Computational Economics*, pp. 1–26. <https://doi.org/10.1007/s10614-016-9581-4>
- Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361–370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>

## Appendices

### Appendix A: Letter sent to clinicians, University staff and students

Dear Dr/ Ms,

I am Blessing Ojeme, a PhD research student in the Department of Computer Science at the University of Cape Town in South Africa. The aim of my PhD study is to investigate the effectiveness of the collaboration of machine learning methodologies and human computer interaction (HCI) in creating screening tools for depression in Nigerian universities. As one of the major stakeholders in the healthcare industry, I would like to offer you the opportunity of being included in my study, which will involve acquiring statistical usage data on the use of computing technologies in clinical practice. If you agree to partake in the survey, then I would undertake to keep all data collected as any data referenced in my thesis would be anonymous. This study is a great opportunity for your hospital to obtain very useful depression data from Nigerian University students so, I do hope that you will accept this offer. If you require more information, please contact, either myself or my supervisors.

Your participation in this survey is voluntary and you can withdraw your consent to participation at any time without stating any particular reason

Thank you.

Yours sincerely,

Signed by candidate

Blessing Ojeme

PhD Research Student, Department of Computer Science, University of Cape Town, Rondebosch, Cape Town, Tel +27725574409, Email: bojeme@cs.uct.ac.za

#### Supervisors:

Professor Thomas Meyer

Dr Audrey Mbogho

Department of Computer Science

Department of Computer Science

University of Cape Town,

Pwani University

Rondebosch, Cape Town

Kilifi, Kenya

Tel +27828064114

Tel +254703455867

Email: [tmeyer@cs.uct.ac.za](mailto:tmeyer@cs.uct.ac.za)

[a.mbogho@pu.ac.ke](mailto:a.mbogho@pu.ac.ke)

## Appendix B: Consent Form for clinicians, staff and students

UNIVERSITY OF CAPE TOWN



Department of  
Computer Science

Private Bag  
Rondebosch  
7701  
Phone: +27.21.650.2663  
Fax: +27.21.689.9465



### Personal Information of Researcher

Blessing Ojeme, PhD candidate, Department of Computer Science, University of Cape Town.

### Purpose of study:

To create a computer system that models clinician-patient conversations for predicting the risk of depression for University staff and students

**Email:** [bojeme@cs.uct.ac.za](mailto:bojeme@cs.uct.ac.za)

Please tick (✓)

I confirm that the purpose of the questionnaire/interview, as explained by the Researcher, is well understood by me.

I agree to participate in the study.

I agree that my participation in the study is voluntary and that I can withdraw at any time, without giving reason.

I agree to the interview / observation/ being **audio**-recorded.

I agree that the results of this interview will be used in publications.

\_\_\_\_\_  
Name of Participant

\_\_\_\_\_  
Date & Signature

Blessing Ojeme

Researcher

\_\_\_\_\_  
Date & Signature

### Appendix C: Questionnaire for clinicians for pre-study data survey

	Study variables	Responses	value
1	Age	31-40	9 (34.6%)
		41-50	11 (42.3%)
		51-60	6 (23.1%)
	Gender	Male	16 (61.5%)
		Female	10 (38.5%)
	Professional category	Doctors	6 (23.1%)
		Psychiatrists	3 (11.5%)
		Psychologists	3 (11.5%)
		Nurses	12 (46.2%)
		Social workers	2 (7.7%)
	Working experience	= 5 years	7 (26.9%)
		6- 10 years	12 (46.2%)
		11-15 years	5 (19.2%)
		=16 years	2 (7.7%)

	<b>Study variables</b>	<b>Responses</b>	<b>value</b>
2	ICT awareness	Yes	26 (100%)
		No	
	Computer literacy	Yes	26 (100%)
		No	
	Reasons for computer illiteracy	Absence of computers	
		Financial problems	
		Less attention to ICTs	26 (100%)
	Have you heard about practical applications of machine learning and HCI in medical diagnosis	Yes	14
		No	12 (46.2%)
	Have you or your colleagues been involved in a machine learning or HCI research project, aiming to identify new patterns or finding new rules for patient diagnostics, prediction of treatment results.	Yes	16 (61.5%)
		No	10 (38.5%)
	If machine learning and HCI methods have been used, was your experience successful? Please comment	Yes	10 (38.5%)
		No	16 (61.5%)
	Please specify which clinical specialties could benefit by using machine learning and HCI methods on collected patient clinical data in your hospital	MIS	20 (76.9%)
		Diagnostic	6 (23.1%)
		Treatment planning	

	<b>Study variables</b>	<b>Responses</b>	<b>value</b>
3	Job satisfaction	<b>Yes</b>	<b>18 (69.2%)</b>
		No	8 (30.8%)
	Reasons for job dissatisfaction	Poor salary	
		Poor training opportunity	2 (7.7%)
		Facility related problems	5 (19.2%)
	How many years has patient data been collected in IT systems in your organization	Management problem	1 (3.8%)
		Little experienced	5 (19.2%)
		None	21 (80.8%)
		= 2 years	5 (19.2%)
		2-5 years	
Does the hospital management provide a personal computer for every clinician	5-10 years		
	=10 years		
If yes, which type	<b>Yes</b>	<b>26 (100%)</b>	
	No		
	Desktop		
How often do you use a personal computer to complete tasks related to your work a day	Laptop		
	Both	26 (100%)	
	Less than 2 hours		
	2 – 4 hours	8 (30.8%)	
	More than 4 hours	18 (69.2%)	

<b>Study variables</b>	<b>Responses</b>	<b>value</b>
Do you think ICT has a role in your profession	Yes	26 (100%)
	No	
Do you think ICT tools can improve evidence-based medical practice?	Yes	21 (80.8%)
	No	5 (19.2%)
Do you think the adoption rate of ICT tools in medical practice is still very low in Nigeria?	Yes	20 (76.9%)
	No	6 (23.1%)
Given the shortages of mental health professionals, do you think there is need for a local workable solution to mental health problems for the benefit of society	Yes	15 (57.7%)
Are you or your colleagues potentially interested in the benefits machine learning and HCI technologies could provide to you?	No	11 (42.3%)
	Yes	17 (65.4)
	No	9 (34.6%)
What is the most common mental disorders you come across	Anxiety	4 (15.4%)
	Depression	19 (73.1%)
	Schizophrenia	3 (11.5%)
	Dementia	
What age bracket would you say are the most commonly depressed	3 – 20 years	2 (7.7%)
	21 – 40 years	15 (57.7%)
	41 – 60 years	5 (19.2%)
	= 61 years	4 (15.4%)

**Appendix D: Questionnaire for staff and student for pre-study data survey**

	<b>Study variables</b>	<b>Responses</b>	<b>value</b>
1	Age	18- 28	35 (55.6%)
		29- 39	16 (25.4%)
		= 40	12 (19.0%)
	Gender	Male	43(68.3%)
		Female	20 (31.7%)
	Students	Undergraduate	26 (41.3%)
		Masters	11 (17.5%)
		PhD	6 (9.5%)
	Staff	Academic staff	10 (15.9%)
		No-academic staff	10 (15.9%)

	Study variables	Responses	value
2	Accommodation	Parents home	10 (15.9%)
		With a relative	5 (7.9%)
		Rented room	48 (76.2%)
	Level of family support	None	29 (46.0%)
		Low	24 (38.1%)
		medium	7 (11.1%)
		High	3 (4.8%)
	Family background	Poor	42(66.7%)
		Well off	14 (22.2%)
		Wealthy	7 (11.1%)
	Cigarette smoking	Non-smoker	4 (6.3%)
		Low	6 (9.5%)
		medium	25 (39.7%)
		High	28 (44.4%)
	Alcohol consumption	Non-drinker	9 (14.3%)
		Low	3 (4.8%)
		medium	21 (33.3%)
		High	31 (49.2%)
	Academic performance	Poor	21 (33.3%)
		Good	11 (17.5%)
		Excellent	9 (14.3%)
		Not applicable	22 (35.0%)
	Do you feel depressed because of the above problems	Yes	35 (55.6%)
		No	28 (44.4%)
	How frequently do you feel depressed	Weekly	4 (6.3%)
		Monthly	31 (49.2%)
		Yearly	28 (44.4%)
Do you often feel the need to see a clinician to examine your health status	Yes	45 (71.4%)	
	No	18 (28.6%)	
If yes, were you satisfied with the outcome	Yes	16 (25.4%)	
	No	47 (74.6%)	

<b>Study variables</b>	<b>Responses</b>	<b>value</b>
Reasons for your dissatisfaction	Long waiting times to receive medical services	
	Long waiting times to retrieve my medical history	
	Costs	
	All of the above	63 (100%)
Given the challenges in getting medical attention, do you think there is need for a local workable solution to mental health problems for the benefit of society	Yes	54 (85.7%)
	No	9 (14.3%)
Do you think combining human expertise and ICT can reduce the cost and waiting time of medical services	<b>Yes</b>	<b>49 (77.8%)</b>
	No	14 (22.2%)
Does the university provide a personal computer for staff and students	<b>Yes</b>	<b>63 (100%)</b>
	No	
If yes, which type	Desktop	
	Laptop	
	Both	63 (100%)

	Study variables	Responses	value
3	Computer literacy	Yes	63 (100%)
		No	
	Reasons for computer illiteracy	Absence of computer center	
		Financial problems	
		Less attention to ICTs	
		Time shortage	
	Job satisfaction	Yes	23 (36.5%)
		No	40 (63.5%)
	Reasons for job dissatisfaction	Poor salary	19 (30.2%)
		Poor learning environment	10 (15.9%)
		Facility related problems	14 (22.2%)
		Management problem	20 (31.7%)
	How often do you use a personal computer to complete tasks related to your work a day	Less than 2 hours	
		2 – 4 hours	10 (15.9%)
		More than 4 hours	53 (84.1%)
	Have you heard about practical applications of machine learning and HCI in solving real-world problems including medical diagnosis	Yes	46 (73.0%)
		No	17 (27.0%)
	Have you or your colleagues been involved in a machine learning or HCI research project, aiming to identify new patterns or finding new rules for real-word problem	Yes	31 (49.2%)
		no	32 (50.8%)
	If machine learning and HCI methods have been used, was your experience successful? Please comment	Yes	25 (39.7%)
No		38 (60.3%)	

## Appendix E Faculty of science, UCT, ethical approval

Faculty of Science  
University of Cape Town  
RONDEBOSCH 7701  
South Africa

E-mail: richard.hill@uct.ac.za  
Telephone: + 27 21 650 2786  
Fax: + 27 21 650 3456



10 March 2015

Mr Blessing Ojeme  
Department of Computer Science

EXPERT SYSTEMS FOR DIFFERENTIAL DIAGNOSIS OF DEPRESSION DISORDERS IN  
LOW-RESOURCE SETTINGS OF AFRICA

Dear Mr Blessing Ojeme

I am pleased to inform you that the Faculty of Science Research Ethics Committee has approved the above-named application for research ethics clearance, subject to the conditions listed below. You are required to:

- implement the measures described in your application to ensure that the process of your research is ethically sound; and
- uphold ethical principles throughout all stages of the research, responding appropriately to unanticipated issues: please contact me if you need advice on ethical issues that arise; and
- ensure that any conceptual models and expert systems developed are sent to the mental health professionals who were interviewed in the research, asking for their comments on, and approval of, these models and systems, and that such comments be incorporated in the dissertation.

Your approval code is: FSREC 06– 2015

I wish you success in your research.


Yours sincerely

A handwritten signature in cursive script, appearing to read 'Richard Hill'.

Dr Richard Hill  
Chair: Faculty of Science Research Ethics Committee

Cc: Dr Geoff Nitschke, Supervisor

**Appendix F University of Benin Teaching Hospital Ethical approval**



**UNIVERSITY OF BENIN TEACHING HOSPITAL**  
**P.M.B. 1111 BENIN CITY NIGERIA**

Telephone: 052-600418  
 Telegram: UNITECHOS, BENIN  
 Telex: 41120 NG  
 Website: ubth.org

**CHAIRMAN:** GEN. A.B. MAMMAN (RTD)  
 mni, FSS, psc, OFR  
 E-mail: [ashoha2@yahoo.com](mailto:ashoha2@yahoo.com);  
[genabmamanabdulahi@yahoo.com](mailto:genabmamanabdulahi@yahoo.com)

**CHIEF MEDICAL DIRECTOR:** PROF.M.O. IBADIN  
 MBBS (Benin), FMCP, (Paed) M.S, (IMMUNOLOGY & IMMUNOCHEM)  
 E-mail: [mikobadin@yahoo.com](mailto:mikobadin@yahoo.com);  
[mikobadin@ubth.org](mailto:mikobadin@ubth.org)

**CHAIRMAN, MEDICAL ADVISORY COMMITTEE:** PROF. G.E. OFOVWE  
 B.M, B.ch, FWACP (Paed)  
 E-mail: [ofovwegabriel@gmail.com](mailto:ofovwegabriel@gmail.com)

**DIRECTOR OF ADMINISTRATION:** A.P. OMOREGIE (MRS)  
 B.Sc, MSc, MHPH  
 E-mail: [officeoftheda@ubth.org](mailto:officeoftheda@ubth.org)

**ETHICS AND RESEARCH COMMITTEE  
 CLEARANCE CERTIFICATE**

PROTOCOL NUMBER: ADM/E 22/A/VOL. VII/1234

**PROJECT TITLE:** "HYBRID SOFT COMPUTING SYSTEMS FOR DIFFERENTIAL DIAGNOSIS OF DEPRESSIVE DISORDERS IN LOW-RESOURCE SETTING OF NIGERIA"

**PRINCIPAL INVESTIGATOR(S)** BLESSING OJEME

**DEPARTMENT/INSTITUTION:** DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF CAPE TOWN, SOUTH AFRICA

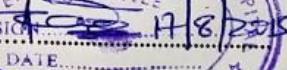
**DATE CONSIDERED** AUGUST 17<sup>TH</sup>, 2015

**SUPERVISOR(S):** DR. GEOFF NITSCHKE, DR. KINGSLEY AKHIGBE

**DECLARATION DECISION OF THE COMMITTEE:** APPROVED

**REMARK:**

**CHAIRMAN:** PROF. A.N. ONUNU


**DECLARATION BY INVESTIGATOR(S):** SIGNATURE & DATE:  17/8/2015

**PROTOCOL NUMBER** (please quote in all enquiries)

To be completed in four and three copies returned to the secretary, Ethics and Research committee, Clinical Services and Training Division, University of Benin Teaching Hospital Benin City.

I/We fully understand the conditions under which I am/we are authorized to conduct the above mentioned research and I/We undertake to resubmit the protocol to the Ethics and Research Committee.

Signature..... Date.....



## Appendix G Mutual information and contribution of symptoms to depression

S/N	Parent/ Target/depression	Symptom	Mutual information	Overall contribution (%)
1	Depression	Impaired function	0.6309	17.1944
2	Depression	Loss of energy	0.5621	15.318
3	Depression	Worthlessness	0.3650	9.9484
4	Depression	Weight loss	0.3445	9.3879
5	Depression	Lack of thinking	0.3101	8.4521
6	Depression	Alcohol or other drug consumption	0.3098	8.3152
7	Depression	Indecisiveness	0.3069	8.3627
8	Depression	Loss of appetite	0.2022	5.5093
9	Depression	Recurrent thought of death	0.1898	5.1730
10	Depression	Family support and availability of accommodation	0.1774	5.1689
11	Depression	Psychomotor retardation	0.1896	5.1674
12	Depression	Loss of pleasure	0.1618	4.4101
13	Depression	Job satisfaction or academic performance	0.1310	3.8041
14	Depression	Insomnia	0.1179	3.2135
15	Depression	Sad mood	0.0863	2.3532
16	Depression	Suicide attempt	0.0680	1.8526
17	Depression	Weight gain	0.0352	0.9602
18	Depression	Depression in family	0.0277	0.7540
19	Depression	Hypersomnia	0.0214	0.5825
20	Depression	Stressful life events	0.0202	0.5503
21	Depression	Psychomotor agitation	0.0171	0.4654
22	Depression	Financial pressure	0.032	0.0873
23	Depression	Cigarette smoking	0.0152	0.0550
24	Depression	Employment status	0.0097	0.0257

## Appendix H Synergistic combination of symptoms to depression

S/N	Symptom	Symptom	Mean synergy (%)	S/N	Symptom	Symptom	Mean synergy (%)
1	Loss of appetite	Sad mood	6.9968	23	Lack of thinking	Weight loss	0.8029
2	Sad mood	Weight loss	6.7256	24	Depression in family	Impaired function	0.7392
3	Insomnia	Loss of pleasure	4.5318	25	Job satisfaction or academic performance	Sad mood	0.6845
4	Loss of appetite	Loss of pleasure	4.3210	26	Impaired function	Cigarette smoking	0.6705
5	Alcohol or other drug consumption	Impaired function	4.1431	27	Recurrent thought of death	Weight loss	0.6430
6	Insomnia	Loss of energy	4.0504	28	Psychomotor retardation	Weight loss	0.6206
7	Impaired function	Sad mood	3.6677	29	Impaired function	Psychomotor agitation	0.5298
8	Loss of energy	Loss of pleasure	3.2593	30	Depression in family	Indecisiveness	0.5031
9	Insomnia	Loss of appetite	3.1744	31	Depression in family	Lack of thinking	0.4873
10	Loss of appetite	Loss of energy	2.5662	32	Impaired function	Suicide attempt	0.4427
11	Impaired function	Worthlessness	2.1636	33	Depression in family	Worthlessness	0.4084
12	Impaired function	Indecisiveness	1.7308	34	Job satisfaction or academic performance	indecisiveness	0.3392
13	Impaired function	Lack of thinking	1.6686	35	Cigarette smoking	Weight loss	0.3210
14	Impaired function	Recurrent thought of death	1.4990	36	Depression in family	Psychomotor retardation	0.2645
15	Family support and availability of accommodation	Job satisfaction or academic performance	1.3994	37	Depression in family	Weight loss	0.2423

S/N	Symptom	Symptom	Mean synergy (%)	S/N	Symptom	Symptom	Mean synergy (%)
16	Impaired function	Psychomotor retardation	1.3568	38	Suicide attempt	Weight loss	0.2318
17	Sad mood	Alcohol or other drug consumption	1.2224	39	Depression in family	Recurrent thought of death	0.2210
18	Weight loss	Worthlessness	1.0756	40	Psychomotor agitation	Worthlessness	0.2161
19	Impaired function	Weight loss	1.0288	41	Loss of pleasure	Stressful life events	0.2081
20	Insomnia	Sad mood	0.8742	42	Psychomotor agitation	Weight loss	0.2027
21	Indecisiveness	Weight loss	0.8229	43	Financial pressure	Alcohol or other drug consumption	0.1786
22	Weight gain	Family support and availability of accommodation	0.8125	44	Hypersomnia	Weight gain	0.1684
45	Lack of thinking	Psychomotor agitation	0.1622	68	Loss of appetite	Stressful life events	0.0611
46	Depression in family	Psychomotor agitation	0.1608	69	Hypersomnia	Indecisiveness	0.0569
47	Loss of pleasure	Job satisfaction or academic performance	0.1600	70	Psychomotor agitation	Suicide attempt	0.0541
48	Cigarette smoking	Psychomotor agitation	0.1573	71	Family support and availability of accommodation	Recurrent thought of death	0.0517
49	Family support or availability of accommodation	Alcohol or other drug consumption	0.1552	72	Hypersomnia	Weight loss	0.0510
50	Family support and availability	Weight loss	0.1541	73	Job satisfaction or academic performance	Lack of thinking	0.05024
51	Hypersomnia	Impaired function	0.1532	74	Lack of thinking	Weight gain	0.0490
52	Psychomotor agitation	Recurrent thought of death	0.1446	75	Psychomotor agitation	Suicide attempt	0.0482
53	Indecisiveness	Psychomotor agitation	0.1434	76	Insomnia	Alcohol or other drug consumption	0.0444

S/N	Symptom	Symptom	Mean synergy (%)	S/N	Symptom	Symptom	Mean synergy (%)
54	Family support and availability of accommodation	Impaired function	0.14008	77	Sad mood	Family support and availability of accommodation	0.0442
55	Alcohol or other drug consumption	Suicide attempt	0.13414	78	Weight gain	Weight loss	0.0441
56	Impaired function	Weight gain	0.1271	79	Financial pressure	Worthlessness	0.0440
57	Depression in family	Cigarette smoking	0.1252	80	Financial pressure	Impaired function	0.0417
58	Hypersomnia	Worthlessness	0.1199	81	Psychomotor retardation	Weight gain	0.0407
59	Hypersomnia	Lack of thinking	0.1180	82	Psychomotor agitation	Weight gain	0.0378
60	Hypersomnia	Psychomotor retardation	0.1178	83	Financial pressure	Recurrent thought of death	0.0328
61	Depression in family	Hypersomnia	0.1090	84	Alcohol or other drug consumption	Weight loss	0.0315
62	Hypersomnia	Recurrent thought of death	0.1025	85	Financial pressure	Lack of thinking	0.0309
63	Insomnia	Stressful life events	0.1022	86	Hypersomnia	Psychomotor agitation	0.0307
64	Indecisiveness	Weight gain	0.0885	87	Financial pressure	Indecisiveness	0.0295
65	Depression in family	Weight gain	0.0851	88	Financial pressure	Psychomotor retardation	0.0260
66	Loss of energy	Stressful life events	0.0781	89	Psychomotor agitation	Psychomotor retardation	0.0259
67	Depression in family	Suicide attempt	0.0731	90	Weight gain	worthlessness	0.0258

S/N	Symptom	Symptom	Mean synergy (%)	S/N	Symptom	Symptom	Mean synergy (%)
91	Recurrent thought of death	Weight gain	0.0252	99	Financial pressure	Hypersomnia	0.0039
92	Financial pressure	Cigarette smoking	0.0225	100	Financial pressure	Weight gain	0.0023
93	Family support and availability of accommodation	Suicide attempt	0.02147	101	Hypersomnia	Suicide attempt	0.0020
94	Financial pressure	Weight loss	0.0187	102	Hypersomnia	Cigarette smoking	0.0014
95	Financial pressure	Suicide attempt	0.01402	103	Employment status	Job satisfaction or academic performance	0.010
96	Depression in family	Financial pressure	0.0134	104	Cigarette smoking	Weight gain	0.0009
97	Financial pressure	Psychomotor agitation	0.0090	105	hypersomnia	Family support and availability of accommodation	0.0006
98	Job satisfaction or academic performance	Lack of energy	0.0077	106	Suicide attempt	Weight gain	0.0003

## Appendix I Consent form for staff and students

University staff and students' depression assessment system

### Staff and student consent for depression assessment

**By signing this form, I confirm that this consent form has been explained to me in terms which I understand**

I consent for depression assessment to be made of me. I understand that the information may be used in my medical record, for the purposes of medical teaching, or for publication in medical textbooks or journals as I have designated below. Refusal to consent to this depression assessment will in no way affect my work or studies in the university. I also understand that I can withdraw my consent to this depression assessment at any time.

#### If illiterate

A literate consentor must sign. Consentors who are illiterate should include their thumb-print as well.

Name of student/staff \_\_\_\_\_ ND    Thumb print of illiterate consentor

Signature student/staff \_\_\_\_\_

Date \_\_\_\_\_

Day/month/year

Consentor Number

#### Statement by the person taking consent

I confirm that the consentor was given an opportunity to ask questions about the assessment, and all the questions asked by the consentor have been answered correctly and to the best of my ability. I confirm that the consentor has not been coerced into giving consent, and the consent has been given freely and voluntarily.

Name and number of person taking the consent \_\_\_\_\_

Signature of person taking the consent \_\_\_\_\_

Date \_\_\_\_\_

Day/month/year

## Appendix J Consent form for staff and students

### Heuristic Evaluation Form

I am a PhD student in Computer Science at the University of Cape Town, South Africa, writing a thesis “Adoption of ICT4D framework to support screening for depression in Nigerian universities”. I have designed a prototype for the assessment and will appreciate your help in evaluating the prototype. The evaluation consists of two parts; Performing some simple tasks by following the current practises and the designed prototype; and evaluating the prototype following Nielsen's heuristics where you will be asked to rate the prototype using a scale of 0-4 where 0 = No problem; 1 = Cosmetic usability problem; 2 = Minor usability problem; 3 = Major usability problem; and 4 = Catastrophic usability problem. Comments can be added under each heuristic.

#### H<sub>1</sub>: Visibility of system status

0	1 Cosmetic	2 Minor	3 Major	4 Catastrophic

Comments

#### H<sub>2</sub>: Match between system and the real world

0	1 Cosmetic	2 Minor	3 Major	4 Catastrophic

Comments

**H3: User control and freedom**

0	1 Cosmetic	2 Minor	3 Major	4 Catastrophic

Comments

--

**H4: Consistency and standards**

0	1 Cosmetic	2 Minor	3 Major	4 Catastrophic

Comments

--

**H5: Error and prevention**

0	1 Cosmetic	2 Minor	3 Major	4 Catastrophic

Comments

--

**H6: Recognition rather than recall**

0	1 Cosmetic	2 Minor	3 Major	4 Catastrophic

Comments

--

**H7: Flexibility and efficiency of use**

0	1 Cosmetic	2 Minor	3 Major	4 Catastrophic

Comments

**H8: Aesthetic and minimalist design**

0	1 Cosmetic	2 Minor	3 Major	4 Catastrophic

Comments

**H9: Help users recognise, diagnose and recover from errors**

0	1 Cosmetic	2 Minor	3 Major	4 Catastrophic

Comments

**H10: Help and documentation**

0	1 Cosmetic	2 Minor	3 Major	4 Catastrophic

Comments