DOC 5005W

MPhil in Fintech Thesis

"Online Platform for Deep Learning Education"

Student Number	Name	Submission Date
FRXAID001	Aidan Fourie	2019/01/27

Table of Contents

Table of Figures

Table of Tables

Introduction

Educating and future proofing in an ever-changing world

The world of work is on the precipice of irrecoverable change. Advances in technology in an unprecedented number of fields (such as artificial intelligence and 3D printing) pose an existential threat to the traditional labour force through the automation of processes and fields previously deemed 'future proof'. This is likely to cause a massive net contraction in the size of the labour market, and cause dramatic social upheaval. Simultaneously however, the advent of 'Fourth Industrial Revolution' will create lucrative, and plentiful, opportunities for those individuals with the scare skills needed to integrate and operate these advanced new technologies in fields as varied as logistics, healthcare and finance.

Therefore it is critical that educational institutions orient themselves to the forthcoming labour revolution and undertake to 'future proof' their students through education in the advanced technologies and theories that will underpin work of the future – irrespective of the students direct field of study. Furthermore, they also have the unprecedented opportunity to utilize the advent of online learning, a direct consequence of the 4th Industrial Revolution, to reach ever larger numbers of students, at lower cost, with potential greater efficacy than previously possible.

A fundamental requirement for successful 'future proofing' is access to easy to understand resources that facilitate learning and self study. To that end this thesis is an initial contribution to this challenge through the design and construction of an online platform for students and laypersons to learn about basic regression neural networks (an introductory and integral theoretical component of artificial intelligence, likely to be generalized across the industries of the future). This platform is intended to form part of a more formal online course organized between UCT and Get Smarter.

This thesis is therefore a small but crucial contribution to the education of artificial intelligence, as well as the future proofing of students, as it exists in the direct intersection of online learning, the labour and educational impact of the Fourth Industrial Revolution and basic neural network (artificial intelligence) theory.

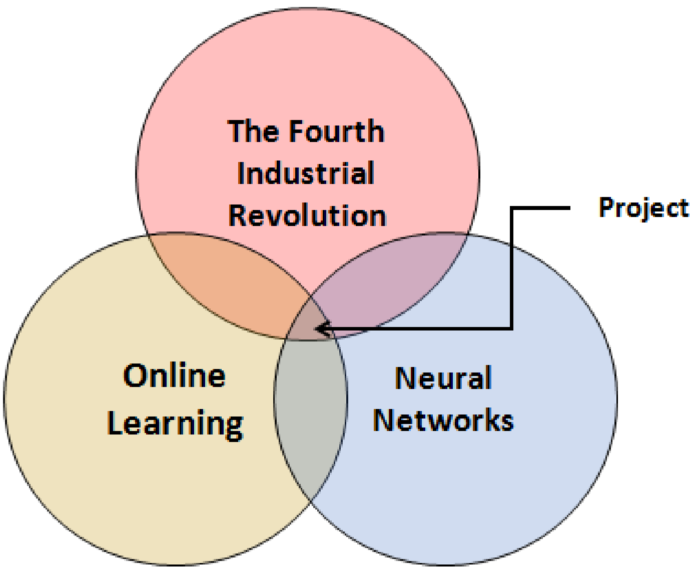

Figure 1: Project intersection of fields

Thesis Overview

This thesis consists of two deliverables, these being:

- This written document outlining relevant literature and the developmental process of the educational platform
- The developed stand alone platform

Overview of written submission

General chapter overview

The primary objective of this submission is to effectively frame the development of the online platform within the context of relevant literature, to overview the developmental journey, to describe the final deliverable, and thereafter comment as to its efficacy as an educational platform. This submission therefore consists of the following distinct segments:

Literature review

- An overview of the incoming 4th Industrial revolution, it's likely timeline, and its anticipated impact on the labour market and the broader society at large
- An overview of the advent of online learning and the current research on effective implementation techniques
- A historic, theoretical and computational overview of neural networks

Developmental overview and critiques

- A brief description of what would be considered a successful platform
- An analysis of the final HTML and JavaScript deliverable
- Conclusions regarding the efficacy and efficiency of the platform
- An analysis of the potential future work on the platform

Literature review overview

It was critical for this thesis to frame the development of the online platform in context of the impending upheaval of the labour and educational market through the 4th Industrial Revolution and the advent of online learning. Therefore, a literature review on the following topics was undertaken:

- **What is the incoming Fourth Industrial Revolution?** What impact is it likely to have on the nature of education and the labour market? And on what timeframe?
- **What is online learning?** Is it effective, and if so, how does one effectively implement it?
- **What are neural networks**? What is their underlying historical context, latest developments and computational theory?

This research would fundamentally be used to stress the importance of the development, to inform platform features and content, and finally be used to inform platform critiques as to its efficacy.

Developmental overview and critiques

The remainder of this thesis discusses the developmental process and overviews the final deliverable. This development was guided by literature as well as the specific technical and use case requirements of the problem statement. For example, specific concerns addressed were:

- How the platform addresses the potential multidisciplinary background of students
- How the platform utilizes effective online learning techniques (such as interactivity) uncovered in the research process

With these conditions in mind, an iterative design process was undertaken – with constant reference to the feedback of the thesis supervisor. This ultimate culminated in two distinct prototypes (one more illustrative, developed in HTML and JavaScript, the other more technical and developed in Flask). A fair critique process was carried out on both prototypes, and consequently the JavaScript implementation was selected for further development and submission alongside this thesis.

This development was ultimately concluded to be fit for purpose, with special acknowledge of its emphasis on interactivity, its multi-disciplinary approach and the general aesthetic. Failings of the review process, and therefore scope for further work, was **i)** a failure to obtain feedback from laypersons to incorporate in subsequent developmental versions and **ii)** a failure to interrogate the broader ecosystem of the online course likely running parallel to the platform to gauge the holistic effectiveness of the platform.

The incoming 4th industrial revolution

What is the 4th industrial revolution?

"We stand on the brink of a technological revolution that will fundamentally alter the way we live, work, and relate to one another. In its scale, scope, and complexity, the transformation will be unlike anything humankind has experienced before. We do not yet know just how it will unfold, but one thing is clear: the response to it must be integrated and comprehensive, involving all stakeholders of the global polity, from the public and private sectors to academia and civil society"

~ ***Klaus Schwab****, Founder and Executive Chairman, World Economic Forum*[1]

Society is in a constant state of flux, alternating between gradual and rapid technological advancements that radically alter our quality and way of life. These periods of rapid progress are often called 'Industrial Revolutions' and their impact on society can be unprecedented and fundamentally transformative. In our near past, for example, much of the world was stricken with famine, war and disease that had massive negative consequences for quality of life. The advent of the industrial, and thereafter the digital, revolution countered these global ills through the establishment of mechanisms to drastically improve our standards of living and life expectancy.[1]

Mankind has traversed three such revolutions as so far, these being[2]:

- **The First Industrial Revolution (1784):** the shift from biological power (such as human and animal labour) to fossil fuels and mechanical and steam power.
- **The Second Industrial Revolution (1870):** the establishment of electricity, wired and wireless communication and the establishment of new means of production and power generation
- **The Third Industrial Revolution (1950):** the advancements in computing power and the establishment of global connectivity

The impact of these technological advancements was not limited to a net improvement of quality of life. Rather, they fundamentally altered the manner in which we went about our daily lives, from what we wear, what we consume, how we communicate, and perhaps most importantly in context of this paper, what work we undertake. The most recent digital revolution, for instance, simultaneously shaped our shopping and entertainment habits via the establishment of online platforms, facilitated the rise of integrated multinationals and allowed the creation of platforms to facilitate mobile participants to engage in the "gig" economy – thus fundamentally disrupting traditional labour dynamics. [3]

We are now on the precipice of a new revolution – the rise of artificial intelligence, robotics, nanotechnology, 3D printing and biotechnology. This is also referred to as the rise of "cyber-physical systems", and it will facilitate unparalleled capabilities for humans and machines.[2]

Advances in machine learning and global interconnectivity are crafting a new future wherein "richly connected" global firms outsource their key decision making to the advanced (and perhaps autonomous) analysis of "big data", and laypersons will be able to access goods and services, from across the world, with unprecedented ease of access. [3] Simple examples include the opening of bank accounts with the use of a smart phone self portrait and the advent of cars that can drive themselves.[4][5] Others would be the ordering of goods via online platforms and subsequent autonomous delivery of goods via aerial drones or the creation of "designer babies" via genome editing. [6][2]

This revolution is unprecedented both in scope, and pace. Experts predict computers will be able to possess human intelligence by 2029, and should be able to surpass humanity (an event referred to as the Singularity) by 2045. Even the most conservative estimates put the achievement of a 'Super Intelligence' well within this century. [3] This has dramatic implications the scope of automation, and therefore unemployment, as machines slowly become more and more efficient and capable than their human counterparts in a wide variety of fields.

This rapid pace, and all encompassing scope, is concerning when one considers that there are severe apprehensions regarding the societal impacts of the 4th industrial revolution, specifically in regards to unemployment and the creation of a 'useless class'. [7] Never before have so many industries been fundamentally altered through the advent of new technologies - thus placing billions of individuals at risk of redundancy as their skill sets become automated. The 2015 World Summit on 'technological unemployment', for instance, concluded that there are a swathe of new, disruptive technologies that could contribute to a state of 'jobless growth', and that therefore that "technological unemployment will be one of the major societal issues facing the 21st century". This is a direct consequence of the increased automation and integration of the 4th Industrial Revolution. [1]

It is therefore uncertain whether the rise of artificial intelligence will contribute to a utopian, or dystopian, future – or somewhere in-between. The challenge is therefore how does society utilize the vast opportunities newly discovered, whilst avoiding the societal pitfalls of increased societal inequity, greater concentrations of wealth, and mass unemployment. [3]

Perspectives on the societal impact of the 4th Industrial Revolution

Literature refers to four broad camps of opinion in regards to the societal impact of the Fourth Industrial Revolution, these being [3]:

- **The Optimists**
- **The Pessimists**
- **The Pragmatists**
- **The Doubters**

Optimists are proponents that this unencumbered revolution will create an almost 'science fiction' like Utopia. The synthesis of genetics, nanotechnology and robotics would allow humans to artificially integrate the processing and memory capabilities of computers, manufactures would be able to create any product from raw schematics, and advances in medicine will allow humanity to greatly extent lifespan, while improving quality of life. The automation of labour, in this perspective, is a net positive as the massive creation of wealth facilitated by the revolution would allow humanity the ultimate freedom to indulge their passions.[3]

Pessimists view the advent of these new technologies as an existential threat to humanity. They stress that, in an ever more complex world, there will reach a point wherein humanity will let machines make broad reaching decisions (as these will most likely bring about better results due to ever increasing computational power and artificial intelligence). They view this progression somewhat of a Pandora's Box which will result in machines ultimately being in effective control of all critical decisions in global society, and humans afraid to challenge these decisions due to the potential consequences, thus effectively regulating the entirety of humanity to that of a 'useless class'. A simple example is that of driving – at a certain point technology would have progressed to a stage wherein autonomous vehicles would be infinitely safer than human driven alternatives. At that stage, it would be rationale to bar humans from driving due to the many lives it would save – thus eliminating an entire profession, and perhaps more importantly a fundamental skill and freedom, from humanity.[3]

The Pragmatists believe that AI can be controlled through effective regulation and technological interventions, thus ensuring that humanity maintains the ultimate control of society. They are thus opposed to unencumbered development of new technologies. They rather emphasise focused research on intelligence augmentation and AI safety prior to any new developments. In their view, AI ultimately serves to augment humanities natural abilities, and therefore our ability to control, and therefore exploit, these technologies is critical.[3]

The Doubters are fundamentally opposed to the idea that AI is possible; and therefore it can never pose a threat to humanity. Generally, they advocate that the true intricacies of the human mind cannot be captured by algorithms as, in their view, creativity is inherently anti-algorithmic. This means that all actions requiring innovation and creativity, such as strategic planning, entrepreneurship and risk taking, will remain the sole domain of humanity, thus severely curtailing the global impact of artificial intelligence and regulating to an augmentation of humanity (very much in line with the Pragmatists).[3]

In conclusion, perspectives on the advent of artificial intelligence are therefore generally split. Generally, optimists view it as a boom to the global economy, and a net contributor to the labour force or quality of life. They foresee a competitive, and ever growing, labour market wherein employees skilled in the new technologies compete in an adapted and enriched marketplace. Pessimists, on contrast, perceive it as inevitable that AI will cause a net reduction in jobs, with half of all jobs being automated by approximately 2050. There is only consensus between the camps on one thing - that developing the skills to work alongside these advanced technologies is critical for anyone “wishing to remain relevant in an increasingly automated workplace”. [8]

What is the impact of the 4th Industrial Revolution on the nature of labour and education?

Automation of tasks is not new. However, the scope of application created by the 4th revolution is truly unprecedented. Studies have shown that as many as 47% of jobs in advanced economies are now at risk of redundancy. This is primarily due to the advent of new algorithms and techniques in machine learning and artificial intelligence, which have greatly expanded the possibility of automation far beyond its historic mandate of explicit and routine rule based tasks. Technological advances in robotics, for example, have created automatons with greatly enhanced senses and levels of dexterity, thus opening up a broader scope of manual tasks for automation, while advances in pattern recognition and artificial intelligence pose great threat to transportation and logistics companies.[9]

This will result in a massive redefinition of “the world of work” – with a predicted 65% of children entering primary schools today predicted to have occupations that do not yet exist. An analogy can be drawn to the between the impact of the First and Fourth Industrial Revolutions – just as the machines of the 19th Century disrupted the manual labour market, so too will the ‘Computer Revolution’ decimate many traditional, middle income, jobs. Areas such as logistics, law, accountancy, manufacturing and finance are particularly high risk for disruption, whereas more creative fields, such as media, marketing and industrial design have a more stable future outlook due to the more significant challenge of automating creativity. [9]

But the revolution is not without opportunity. New jobs will be created in fields relevant to the implementation and development of the new advanced algorithms, for example as data scientists and analysts, while those existing jobs that do survive automation, for example as business analysts, will also require new technical skill sets so as to incorporate, and interpret, the methods and results of the machine learning and artificial intelligence applied to their work. Therefore it is imperative that the global educational system orientate to the mantle laid before them, that is to future proof their students with the new skill sets required of them – or risk their ultimate redundancy [9].

The implications of the 4th Industrial revolution on the nature and focus of education is therefore vast – both in regards to how curricula must change to incorporate new skills for students, and on the requirement of educational institutions to vanguard the development of new techniques and methodologies [8]. Educational institutions therefore need to be aware as to the latest trends, both in the labour market and AI theory, and use this information to formulate and advance effective curricula so as to future proof their students. [9]

Lastly, there exists opportunity in the technological possibilities provided by the advancement of technology, with online learning and massive open online courses providing new platforms that allow Universities and educators to disseminate their work and courses to a far broader audience, at a lower cost, with learning toolsets and aids previously impossible. [9]

The advent of online learning

A challenge to the traditional brick and mortar classroom

The global education system is increasingly under pressure to increase student throughput, and to utilize ever fewer resources while doing so. This is particularly a concern for the United States due to ever-dwindling funding and ever-increasing international pressure. For example, between 2002 and 2012 US public university state funding decreased by 29 percent per student whilst enrolment increased by 28 percent, from 9.0 to 11.5 million full-time equivalent students. [10]

Traditionally these institutes have largely compensated for the decrease in public funding through above inflationary increases in tuition fees. However, this is becoming a politically inconvenient solution due to increasing resentment by students and society at large. This calls into question the viability of the current funding paradigm, and therefore the current operational model. [10]

A major contributing factor to this problem is the generally labour-intensive nature of education – with generally no easy means of increasing productivity or cost efficiency. However, the "brick and mortar" classroom is not immune to the information system revolution that has so dramatically influenced our personal and professional live. The internet has made online and mixed medium (i.e. partially online) learning a feasible alternative that could enhance student outcomes whilst combating the constrained resources of educational institutions. [10]

There is also increased demand for such platforms from the students themselves. This demand is driven, in part, by the realities of today's economic uncertainty and therefore the ever-present threat of unemployment. These harsh truths have driven many professionals to seek a lifestyle of lifelong learning to safeguard their professional success and stability. The digital revolution has facilitated the creation of online or dual medium institutions to cater to this market, creating flexible learning solutions at the touch of a screen.[11] It has been found that online learning is the chosen method of study for working students between the ages of 25 and 50 due to its convenience and flexibility – allowing them to balance their professional, private and scholastic lives however suits those best. [12]

There are currently over 150 million tertiary students globally. The number, and their demands upon a resource constrained educational system, is likely to keep growing exponentially. [11] Whilst it is naïve to think that online platforms and alternatives will completely replace brick and mortar institutes, it is equally naïve not to see their appeal.[10] Universities and other third parties have therefore been investing heavily in online learning due to its resource efficiency, and potentially academic benefits, over traditional methods of education. This has led to a new era of competition, and therefore democratization, of the educational marketplace due to increased international public and private competition.[11]

Is online learning as effective as traditional education?

The digital revolution has given rise to three broad categories of courses that will be discussed in this review [12]. These are:

1. **Traditional** – these are traditional classrooms wherein there is no real substantive online component
2. **Mixed medium** (**blended)** – these are courses that interweave online content and platforms with real world teaching methods and classrooms
3. **Online** – these are courses that are fully online with little to no real world interaction between staff and students

The rise of these alternatives methods have not been without critique. There are, for example, concerns online learning may depersonalize education. The Sloane Consortiums 2011 report reflected that while 51% of academic officers "believe that online instruction is comparable to face to face instruction", only 14% find that "it is superior". Additionally, only 63 % thought that student satisfaction was comparable across both platforms. This is predominantly due to the vastly different natures of either approach. For example, face to face instruction has historically allowed instructors to judge the students level of understanding from both verbal, and non verbal, cues. Online instruction, by its very nature, would remove this tool from the educator.[13]

But it is critical to recognize that "online learning" is not a singular, monolithic thing. It is a vast, and ever evolving, field of study ranging from the simple upload of lecture videos, the upload of materials such as tests, to the implementation of highly sophisticated, interactive platforms that make use of multiple feedback loops and integrate live traditional teaching. It can therefore actually provide a net increase tools to educator and students rather than limiting them. [10]

Furthermore, the opinions of the academic officers are not backed by empirical research. Many studies have, in fact, found statistically significant impact for student performance when compared to the traditional "brick and mortar" classrooms. These papers included metrics such as test scores, student engagement, self perception of learning and reduction failure rates. While reports on student satisfaction levels within these courses were mixed – with purely online courses suffering in comparison its traditional alternative - blended classrooms were in fact found to actually foster a stronger sense of community within the study body than traditional methods.[12]

Despite these positive findings, it is still unclear if this holds true across all studies and can therefore be treated as a fair and generalized conclusion. Analysis by the US Department of Education identified 45 studies between 1996 and 2008 that employed stringent methodologies and quantitatively measured student outcomes between the online and traditional format. These found that, on average, students in an online environment performed only modestly better than their traditional counterparts. Importantly however they also found that the number of studies that found mixed or negative impact on educational outcomes was dwarfed by those with neutral or positive findings. This implies that online courses were, at worst, on par with their traditional counterparts in regards to educational outcomes.[12]

Interestingly, this difference was most pronounced within the context of blended learning. These environments blended the traditional benefits of face to face instruction with the additional learning times and interactivity of the online platform. These, generally, were found to be a net improvement on traditional face to face instruction. A further unexpected finding was that publication year was also significant predicator as to the effectiveness of online education. This spoke to the evolution and increased sophistication of online platforms as they became viable and established alternatives to traditional mediums.[12]

The researchers ultimately concluded that there was not yet significant evidence that online education was superior, on average, to traditional instruction. Rather, there was found to be "no significant difference" between the two.[12]

This is significant. Online courses require far less resources and can be taught to a far broader audience of students with little or no geographic limitation or incremental cost. Therefore the finding of 'no significant difference' consequently dramatically increased the rate of investment in the field**.** For example, the number of universities who deemed online learning critical increased from below 50% to 77% between 2002 and 2017. [11]

In conclusion claims about the efficacy and utility of online learning are likely to be exaggerated. But equally, comments or perceptions that online systems cause sub standard results must be questioned. Research has consistently indicated that well crafted online learning solutions, especially the hybrid model, are able to achieve at least equivalent outcomes to the traditional method, whilst saving significant resources that could be deployed to more effective use elsewhere in the educational sector.[10]

MOOCS – The University of the Future?

Special mention must be made of massive open online courses, or MOOCs, which are generally free to access platforms that provide access to rich repositories of information online compiled by experts in their respective fields. Generally these platforms have massive numbers of enrolees in a vast array of fields, ranging from mathematics, to music, to software development. [11]

MOOCs are controversial. Their supporters see them as the "biggest innovation in education in 200 years", while their detractors view them as inefficient platforms with very low rates of student retention. For example, while MOOC providers intend to enrol 1 billion students in the next decade, only a small fraction of these students will likely complete their courses. [11]

This is backed up empirical research. It has been found that generally a student "funnel of participation" occurs wherein only a very small percentage of enrolees make it to course completion. Studies have shown that dropout rates within online courses can range anywhere from 10% to 50% higher, on average, than traditional alternatives – with some studies reporting completion rates as low as 7-10%. [13][11]

The reasons for this are vast. Badly designed courses can lead to student frustration, which can be compounded by the minimal level of peer to peer and peer to tutor contact. This lack of "community" can decrease motivation – which when combined with the low barriers to exit (i.e. many courses are free and not accredited) can cause high rates of student drop out.[14][11] Conversely another significant factor is the ease of access to online courses – many students may be simply tempted to sign up whilst "window shopping", and therefore be unmotivated to actually commit to the course. These students often drop out before the first assignment is due. [11]

There are however defining aspects to course design that can act to counter this. The adherence to principles of quality and close course proximity to individuals with knowledge in the field of study has been proven to be crucial.[11] It has also been shown that high levels of interaction, either via simulation or game-based learning approaches can also have a significant impact on student completion rates. Finally, emphasis on increased social integration can also act to decrease rates of student attrition as these interactions create a sense of "community", allow different avenues for the reinforcement of content knowledge, and create a sense of "social contract" between the participants that adds a social "cost" to the premature exit of a programme [13][12]

Lastly, while these platforms are challenging the exclusionary nature of traditional education enrolment, it is still unseen as to whether they will introduce formal accreditation to directly compete with traditional institutions in the formal student and labour market. Currently a collaborative approach has been pursued with formal institutions alternating between either encouraging the parallel use of these platforms, or alternatively formally integrating them into their traditional courses. There is no direct evidence that this is likely to change.[11]

In conclusion, MOOCs offer the promise of a direct democratization of education to vast numbers of people. However, they need to be cognizant of consistent and iterative improvement in course quality, student engagement and retention. Every effort must be made to ensure that students have an experience that is interactive, engaging and effective so as to combat the high rates of student attrition associated with the solution. Ultimately, regardless of student retention issues, it is still hard to deny the appeal of an education system that promises free or low cost tuition to traditionally excluded communities and peoples. [11]

The characteristic of a successful online course

A fundamental characteristic of successful course design is the alignment of course objectives, student expectations and course implementation and assessments. Objectives detail what students are expected to master at the end of the course, and thereby create student expectations. When courses become misaligned due to the objectives not matching either the teaching or the assessment methods, students start to become ambivalent towards the course as their perceptions of course quality falls.[13] As student motivation is intrinsically linked to student attrition, this has a severe knock on effect on the student "funnel of participation".[15]

When these are aligned however, great success can be had. A study of political science students using a mixed medium tool that enabled them to create, share and discuss multiple choice questions – with no instructor input – found that students generally had better learning outcomes, and perhaps more importantly in the context of online student retention, perceptions of learning and motivation. [12]

This finding corresponds with other research. It has been found that modern students greatly prefer interactive and dynamic – rather than passive or static – learning environments. This is most likely due to the fact that we live and engage in a highly interactive and dynamic world, thus shaping our expectations of our educational system. Therefore students thrive in the presence of multiple outlets for creativity, collaboration and competition. [16]

Fundamentally, research findings speak to two key learning models for successful online learners, these being: [13]

- **Exploratory model** – this model based on problem solving based learning methods aligned to the course objectives. This can be facilitated through online resources and multimedia.

- **Dialogical model** – this model focuses on learning through interactions – for example via group discussion forums, document sharing and collective reflection.

A vital component of this drive towards interactivity is the use of "gamification" - i.e. the application of gaming principles and virtual achievements to online learning. It was found that the implementation of a virtual achievement system on the previously mentioned political science platform, wherein students got rewarded with "badges" for certain actions and achievements, had "a significant positive effect on the quantity of student contributions" – with no associated loss of quality. [12]

Peer to peer, and peer to instructor, discussions are also another vital form of interaction. The importance of communication in learning mathematical principles (and therefore neural networks) is globally accepted. This does not change due to the medium of presentation shifting to the internet. [17]

Participants in platforms that facilitate these sorts of discussions gain the utility of collaborative knowledge sharing and construction. They are able to share ideas, learn from and reflect upon their peers, as well as their own, thoughts and musings in an environment that is conducive to learning. Research has consistently shown that the quality and quantity of these interactions are highly correlated with student satisfaction.[18] This is a massive boon for student motivation, and thus learning outcomes and student retention. [12]

There are two key formats for online communication, these being: [17]

- **Synchronous** (same time) e.g. voice and video conferences, shared whiteboards and live assessments.
- **Asynchronous** (out of time) – e.g. forum threads and e-mail chains

Both of these approaches have disadvantages and advantages.

Synchronous communication has the advantage of allowing students and facilitators to create virtual 'classrooms' or 'offices' thus facilitating live feedback, community building and teamwork. A weakness of synchronous communication is the "my place and pace" dynamic. Scheduling convenient sessions for a global student base is administratively burdensome, moderation of fast paced and dynamic settings are difficult, and these discussions are often subject to complex interpersonal and social dynamics. [17]

Asynchronous communication has the advantage of fostering more comprehensive discussion between students (at their own convenience) – countering one of the major weaknesses of synchronous communication. Disadvantages include the absence of immediate feedback, time lag in the creation of mature and dynamic discussions, and a potential sense of social disconnect between students.[17]

Caution must be taken as not all online discussion is ranked equal. Quite often it can be found that these messaging systems often become divergent and fragmented as students either do not respond to, or build on, each other. Additionally, topics can veer off course rapidly and the depth of conversation can at times remain too shallow to stimulate student motivation. This corresponds with the findings that the most positive significant outcomes are reserved for students who participate actively, rather than passively, on these platforms – for example those who are actively creating content and engaging ideas robustly rather than those simply skimming other discussions for the answers to a tutorial question. In a sense, students get out what they put in. [18]

There are ways to counter act this however. One, for example, is to focus on an approach wherein the convener acts as the facilitator/mentor (who assists with key problem solving and critical thinking skills) while the teacher's assistants (TA's) function in a supporting mode – ensuring that the students are engaging one another in a constructive and robustly.[13]

Several studies have also found that immediate feedback is the perhaps the most critical factor in determining the effectiveness of an online course. One of the strongest explanations for the incremental value of online courses is the students can instantly determines if their answers are correct or incorrect. This allows students to quickly recognize their mistakes, and adjust their computational techniques according. In a traditional institutions there would be a time delay prior to receiving feedback, and as such, students have often moved on to new material by the time they it is received – thus harming their ability to cement proper computational technique. [19]

This is especially powerful when combined with randomized testing and homework. This enables students to take, and retake, assignments and tasks – each time with the problem, and therefore the computational technique required slightly altered. This gifts students with a broad exposure to implementation of different applications of theory, which they can immediately troubleshoot. In contrast, the traditional students remain exposed to singular problem definition, with delayed response time until feedback and therefore no effective means to troubleshoot.[19]

Authentic activities within online courses

It has also been found that "authentic activities" within the online learning paradigm can have significant benefits for learners. Authentic in this context generally refers to basing activities on real situations of sufficient complexity with an emphasis on the use of "conceptual knowledge and skills" – such as critical thinking. More formally, authentic activities can be defined as follows:[20]

- **"Authentic activities have real world relevance"**: Activities should mirror real world, and industry, challenges.
- **"Authentic activities are ill-defined, requiring students to define the tasks and sub-tasks needed to complete the activity"**: Problems should not be able to be easily via the mechanical implementation of taught techniques
- **"Authentic activities comprise complex tasks to be investigated by students over a sustained period of time"**: Tasks should require series time investment and mental effort.
- **"Authentic activities provide the opportunity for students to examine the task from different perspectives, using a variety of resources"**
- **"Authentic activities provide the opportunity to collaborate"**: Collaboration is a critical real world skill, and thus it should be fostered in any educational endeavour.
- **"Authentic activities provide the opportunity to reflect"**: Time should be dedicated to allow students to reflect on their decisions within the context of the project, and muse as to how they could have improved their performance.
- **"Authentic activities can be integrated and applied across different subject areas and lead beyond domain specific outcomes"**: Activities should aim to construct a broad and robust skill set rather than specialist and limited student base. This can be achieved by having students occupy different roles in the learning process.
- **"Authentic activities are seamlessly integrated with assessment"**: Assessments should be integrated naturally into the course of the work – as it is in the real world. Effectively, the emphasis is to avoid artificial and separated assessment processes that do not mirror how the real world operates.

- **“Authentic activities create polished products valuable in their own right rather than as preparation for something else”**:

The complexity of these activities allow create a diversity of outcomes, wherein students can provide their own competing solutions, rather than a singular correct answer that is found via an application of rule sets to a problem statement. This ultimately provides a more holistic, and motivating, educational experience for the student.[20]

A comprehensive pedagogy for online mathematical education

There are many parallels between the applications of machine learning, artificial intelligence and pure mathematics. Therefore any advances in the study of mathematics are relevant to interested in more formal machine learning. That said, the development of a comprehensive pedagogy for online learning specifically in the field of mathematics is still in its development phase, despite several studies indicating that students find web based mathematical learning more enjoyable. This lack of formal definition is primarily due most research conducted regarding online and distance learning being conducted ‘in general’ rather than discipline specific. Despite this, a few key principles of effective online mathematical education have been identified:[17]

- **Instructor facilitation**: Instructors must play the role of the “metronome” of the course, and guide how students interact with each other so as to ensure constructive, and on topic, discussions and communities. Students and teachers must engage consistently and electronically, and instructors need to prioritize the successful development and operational performance of the course – rather than presenting course contents.
- **Communication opportunities:** Students must be provided the opportunity to communicate with one another, either asynchronously and/or synchronously, and thus foster a sense of community and allow the construction of a ‘global knowledge base’. This interaction is critical in conceptual subjects, such as mathematics, as it allows students to engage and cement concepts independent of instructor feedback.
- **Internet resources:** Resources provided in the course need to comprehensive, thorough and make us of multi-media and visualizations where reasonable. Third parties should be integrated, for example online MOOCs and textbooks, that can be used in conjunction with the course material. This allows students to reinforce their knowledge and explore other, supplementary, teaching methods.
- **Appropriate interface**: The course interface needs to be simple, and one that students are comfortable with – i.e. in navigation, locating resources, engaging in discussions etc. Researchers must also be cognizant on the difference in perspectives and prior knowledge between themselves and students. For example, students require mathematics to be presented consistently in form throughout the course, whereas lecturers would generally be more flexible.
- **Online assessment**: The course should have assessments wherein students and instructors can measure performance and consequently course efficacy
- **Convenience, flexibility and accessibility**: Students should be able to access the material at their own convenience, and asynchronous discussions should be utilized to foster collaboration between students of varied geographies. Prospective students should be allowed to gain an overview of all course materials before enrolling formally in the course.
- **Dynamic learning environment:** Assignments, and feedback on performance, should be given to students as immediately as possible – and automated if possible. Additionally, in the course of

mathematical education students can sometimes get “hung up” on problems. While in traditional format courses there is generally no time to dwell on these, in the context of online courses provision should be made a full, asynchronous, exploration and exposition of these “problems of the day”

The emotional characteristics of a successful online learning student

The stereotypical profile of a successful online student is one who is self motivated, self directed and gifted with both above average executive functioning and technical skills.[13] These students are effective at self regulation – i.e. they possess the ability to set goals, plan ahead and are to endure challenging scholastic environments without significant loss of motivation. However, not all students who are successful in the context of online learning present this profile as studies have shown that the student’s level of technical ability – and their cognitive strategy in regards to learning – is not the primary factor in determining student success.[15]

Research has rather found that the student’s emotional experience within the context of a course is crucial to academic success. This is due to the extremely high levels of correlation and inter-dependency between the emotional experience, one’s view of self efficacy and therefore one’s academic motivation. [17]

Students who have positive emotions – for example hope, pride and enjoyment – will generally carry the positive traits of self believe, motivation and the ability to process complex negative emotions such as disappointment or frustration. Those who possess negative emotions, such as boredom, frustration or despair will generally suffer in regards to attention, motivation, self regulation and therefore ultimately suffer in academic performance. These emotional experiences often create a feedback loop as students with high levels of self regulation are able to attribute failure to a simple unsuccessful implementation of learning strategies – rather than an intrinsic failing of themselves. Therefore they corrective action by adjusting their approach, rather than compounding or create negative emotions. For students feeling frustrated, or hopeless, they attribute failure to personal failing, thus compounding their negative outlook.[15]

This emphasis on emotional perspective speaks to the one of the stark differences between online and traditional classrooms. Online courses are extremely learner focused, and such more demands are placed upon the student. The online learning requires students to act independently and consequently work at their own pace, whereas in the traditional classroom the teacher would be the pacemaker. Inherent assumptions have therefore been made regarding the learners ability to self direct their learning experience – and thus consequently students must be motivated to make adequate progress without significant guidance. [17]

It is therefore crucial to support and nurture students so as to gain a level of academic, and emotional, maturity beyond that of their relatively “spoon fed” traditional peers.[17] Course design, i.e. the correct implementation of communal interaction and interactivity, plays a massive role in this. For example students get frustrated with challenges in understanding when there is no option for immediate feedback. [13] These negative experiences can then be further compounded by ineffective design wherein students are isolated from one another – and therefore unable to receive social support from their peers. [17]

The ultimate conclusion of this research is that course supervisors should not solely pay attention to the traditional metrics of academic success. Attention should also be paid to student’s emotional journey. For example, if students emotional experiences could be improved (say via the redesign of a course, direct intervention via lessons in emotional self regulation or enhanced student interactions) then there should be a corresponding increase in traditional metrics of academic achievements due to reduced anxiety and increased enjoyment of course material. [17]

What are Neural Networks?

Encoding the human brain

A neural network consists of sequential layers of simple, non linear processing units (neurons). Input (i.e. first layer) neurons are "activated through sensors that perceive the environment", whilst the following layers of neurons in the system are "activated through weighted connections to previously activated neurons". These non linear, parallel chains of processing units can create complex and long chains of computational stages that allow the network to recreate complex behaviour, such as driving a car or object detection [21].

Neural networks are a key component in the "rise of artificial intelligence" and thus form the crux of thesis. The origins of neural networks are rooted in trying to mimic the human brain to perform tasks via the use of networks of non linear, simplified mathematical models. Fundamentally, they assume that the brain as a "highly complex and non linear parallel computing system" [22].

Historically, there have been three main categories of neural network models, these being [22]:

- **Supervised learning:** these models learn to map an input to an output. Effectively, they learn an underlying mathematical function to either predict correct values ('regression') or categories ('classification') for any assortment of inputs. In order to correctly train these models it needs a subset of matched inputs and outputs (i.e. training data) so as to learn the underlying function behaviour that will allow it to extrapolate to never before seen data.
- **Unsupervised learning:** these models are more descriptive – with the intention being to uncover new knowledge ('latent structures') within the data, and thus, enable parties to make more effective decisions or insights. An example of such a use case is the 'clustering' of retail customers into a finite number of archetypical customers that allow more effective marketing decision making.
- **Reinforcement learning:** these models are similar to supervised learning in regards to the fact that they wish to minimize some 'error' function, i.e. they intend to behave as accurately as possible. In the case of reinforcement learning, this function takes on the form of a reward/punishment signal that the algorithm attempts to maximize – i.e. the returns on a betting strategy in Poker for example.

Neural networks have their roots in the dawn of machine learning and artificial intelligence. In this era, work on artificial intelligence was done in parallel and collaboration with traditional work on neuroscience and psychology – often generating highly productive results – although at present the two fields have drifted apart due to the increasing complexity and specialisation of either field [23]. Within this historical context, the first neural networks were created with a dual objective, these being [22]:

- **To learn about biological processes** such as the nervous system via reverse engineering through computational models. These models allow hypothesis testing in the fields of psychology and neuroscience via simulation that provide results very similar to that of "in-vitro or in-vivo experimentation" – thus reducing requirements for invasive methods to gather empirical data.
- **To construct algorithms mimicking biological processes** – thus to take advantages of the evolutionary work undertaken in the development of these systems (i.e. bio mimicry)

There were numerous benefits to the emphasis on replicating and advancing biological systems and intelligence. They inspired new advances and approaches in machine learning both independent, and complimentary, of the traditional mathematical and logic based methods that dominated machine learning research. Furthermore, the discovery of any machine learning methodologies within biological systems, such as the brain, serves as a validation that that methodology is effective, and most likely an integral component of general, and human like, intelligence [23]. This is due to the fact that whilst computers are able to outperform the human brain in some tasks, they are nowhere yet near able to equal neither the brains flexibility, robustness nor its energy efficiency. [22]. Therefore the mimicry of any established biological process will most likely yield dividends in regards to the improvement of existing computational methods.

This approach has yield substantial real world results, with neural networks finding immensely influential use cases in fields as diverse as[23]:

- **"Language and understanding":** the ability to understand natural language, and consequently respond appropriately. Examples include voice recognition and instructions on cell phones, auto transcription of spoken statements, and chat bots.
- **"Perception":** the ability to recognize real world objects, and establish internal knowledge of various 'scenes'. Examples this include facial recognition software and the visual sensors of self driving vehicles.
- **"Modelling":** the ability to develop an internal schema and set of rules for the establishment of real world relationships between various agents. Examples of this include the prediction of weather patterns, identification of medical symptoms and the prediction of financial markets.
- **"Robotics":** this represents the integration of the aforementioned fields to create autonomous agents capable of navigating the real world, and consequently, perform a variety of tasks in an unsupervised fashion.

Specific real world use cases of neural networks include the use of artificial intelligence to assist radiologists in the identification of lesions, and the extent of illness, as a form of automated 'second opinion'.[23] These algorithms open up the possibility of the entire of field of radiology eventually becoming automated as the accuracy of these automated systems completely outperforms the humans due to our inherent fallibility. A more mundane, yet just as powerful, use case is the integration of neural networks into accounting databases. Humans are easily overwhelmed with the large volumes of numeric data presented in such systems – the implementation of effective artificial intelligence can assist decision makers with classifying and interpreting transactions – for example in the identification of fraud or mismanagement. [23]

That is not to say that neural networks are without their flaws. Neural networks are generally considered a "black box" - i.e. they are able to reproduce almost any function with great accuracy (if given sufficient time), but the later study of the trained network itself reveals little to nothing regarding the system properties or decision making process. To phrase it plainly, one will not be able to explain the rationale regarding a neural networks output – it must simply be accepted. This is concerning in applications or use cases wherein failure or inaccuracy can have catastrophic results.[22]

Furthermore, we have yet to establish a singular "universal" model that can be applied generically to problem solving and information processing. Generally, models are framed within the context of a specific problem – with this problem determining the network architecture and complexity. These models tend to excel at their area of design, but fail miserably when transposed to a new field of use. The direct consequence of this is that a multitude of model types and variations have been developed to deal with various niches of problems experienced in the real world, and as such, any data scientist or engineer must possess a great level of knowledge, and exhaust a fair amount of time in optimization, in order to efficiently and effectively solve any problems presented.

Neural networks also face more mundane, technical problems when implemented in new systems. Examples of these problems include[22]:

- **Insufficient or incomplete data to effectively train the models** leading to low model accuracy
- **Imbalanced training data** - i.e. if we are training a model to identify images as either being that of a cat, or a dog, it is problematic if 99% of the images we provide it to train are of cats. The network will quickly discover it is most effective to assume every image is that of a cat, thus not truly learning the underlying structure of the task. It is therefore imperative to 'balance' any training data provided to the network
- **Deluge of high dimensionality data** – due to the ever expanding data requirements, and dimensionality, of the world, it is becoming more and more time prohibitive to train models directly on raw source data. Rather, effort must be made in the pre-processing phase to extract any key features or patterns within the raw data, which can thereafter be passed to more advanced machine learning algorithms.

Despite these concerns, recent research in artificial intelligence has been dramatic and unprecedented. Artificial intelligence can now directly challenge, and outperform, humans in object recognition, and complicated 'games' such as DOTA, GO or chess.[23] They are further able to generate synthetic art and complicated visuals of artificial people, faces and never before seen landscapes in a manner that is almost indistinguishable from the real world. [24] Another example is the creation of a full 3 dimensional avatar of a human face from a static 2 dimensional image in almost real time – for integration into video calls, gaming and newer cinematic experiences [25].

However, for now, these systems still trail far behind humans in generalized, non specific, intelligence. Human intelligence is still distinguished by its ability to rapidly learn new concepts with minimal training, through the use of logic and the leveraging of prior knowledge to ensure inductive inferences. Infants, for example, still carry well more developed neural systems and processes in this regard than even the most advanced artificial intelligences. It is hoped that advances in neuroscience and biology will narrow this gap between biological and artificial systems, and in many experts opinion it is only a matter of time until this is achieved. [23]

Computational theory of dense feed forward multi-layered neural networks

Neurons and neural network construction

A **neural network** is a transformative function that converts an input observation of ***X*** dimensions into a ***Y*** dimensional output. These networks are generally trained to as to be able to make predications based on some provided dataset, for example predicting a stock price based or identifying images of cats. [26]

In context of regression neural networks ***Y*** will always possess a dimensional value of one – i.e. the network has a single output for any input. This is so as to allow the network to make single numeric predication (say for example a stock price) irrespective of the amount of data fed to it (e.g. the previous 6 months closing prices, the trading volume over the last month, index's of public sentiment, commodity prices etc) [26]

A neural network consists on ***N*** neurons, arranged in ***M*** layers. **Neurons** are stand alone computational units, which transform an input of ***Z*** dimensions into a single output value. It is effective to think of each neuron as a stand-alone ***Z*** dimensional linear model, whose output is used as an input into a internal function ('the activation function') which transforms this model so as to provide the final neuron output. Formally, a neuron consists of the following components [26][27]:

- **Inputs**: This is either the raw data if the neuron is in the first layer of the neural network, or the outputs of **all** neurons in the preceding layer if the neuron if in the 2nd layer or beyond. Each feature/dimension of any input can be thought of as an input dimension of the linear model (e.g. X, Y, Z) and the actual value merely a position along that axis.
- **Weights:** These are the coefficients of the linear model – i.e. they scale the magnitude of the corresponding input by their value. There are as many weights as there are inputs per observation, and therefore the dimensionality of the weight vector is a direct function of either the number of training features per observation or the number of neurons in the preceding layer. These weights are initially randomly generated and are "trained" (incrementally adjusted) through a process of "**back propagation**" so as to improve predictive accuracy.
- **Bias:** A bias is a special weight that exists independent of the weight vector and can be thought of as the intercept of the linear model – i.e. it shifts the linear plane that results from the multiplication of the weights and inputs up or down by its own value.
- **Activation:** The output of the linear model is feed through an activation function – for example a sigmoid function or a rectified linear function – thus transforming it. This allows for non-linearity in the neural network that in turn compounds as the input observations works its way through all layers and neurons. This allows the model to model to "learn" complex mathematical behaviours despite being constructed from relatively simply components.

Consider the following simple example of a neuron with a **sinusoidal** activation function, weights of values **[3, 4, 5],** inputs of values **[2, 2, 2]** and a bias of value **[10]**. We could calculate the output of the neuron as follows:

- The product of the weights is **[6,8,10]**
- The sum of this is **24 –** this would be the linear output ignoring the bias term
- The bias is **10** – therefore the final linear output would be **34,** which is fed to the activation function
- This is then transformed via a sin function i.e**. sin (34)** – giving a final neuron output **of 0.52.**

Layers refer to the arrangement of these neurons relative to each other, which fundamentally impacts the computational process of the network (as inputs to any layer within the network are the outputs of the neurons in the layer that precedes it). The dimensionality of the network output is therefore determined by the number of neurons in the final layer – for example if the layer contains 2 neurons, the output will have 2 dimensions, hence why regression networks always have a single neuron in the final layer. The image below outlines the typical depiction of a 3x2x1 neural network (i.e. 6 neurons arranged in a pyramid layout over 3 layers) [26].

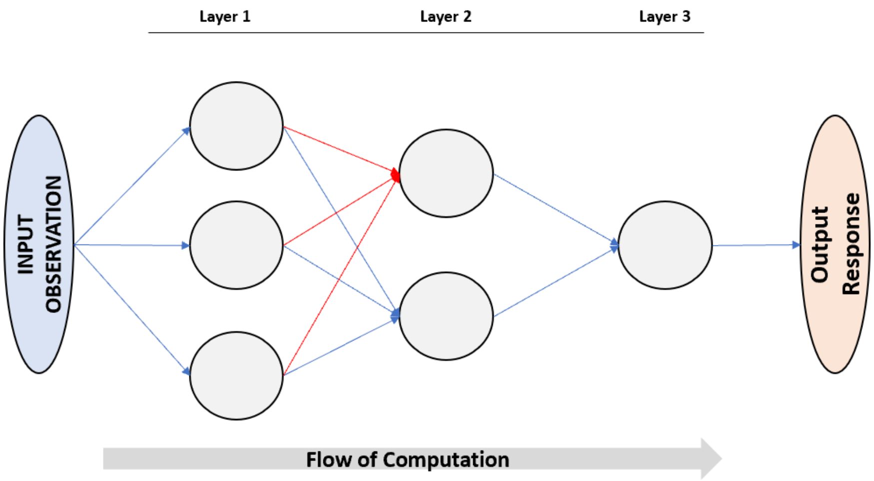

Figure 2 : An arrangement of Neurons in a Neural Network

Compounding complexity

The plot depicts the simple use of an activation function, in this case the "sin" function, in a simple 1x1 neural network. Plotted are **i)** the linear outputs of the first neuron prior to activation **ii)** the final output of the first neuron and **iii)** the final output of the second neuron. As can be seen, the complexity compounds as the information passes through the output [26].

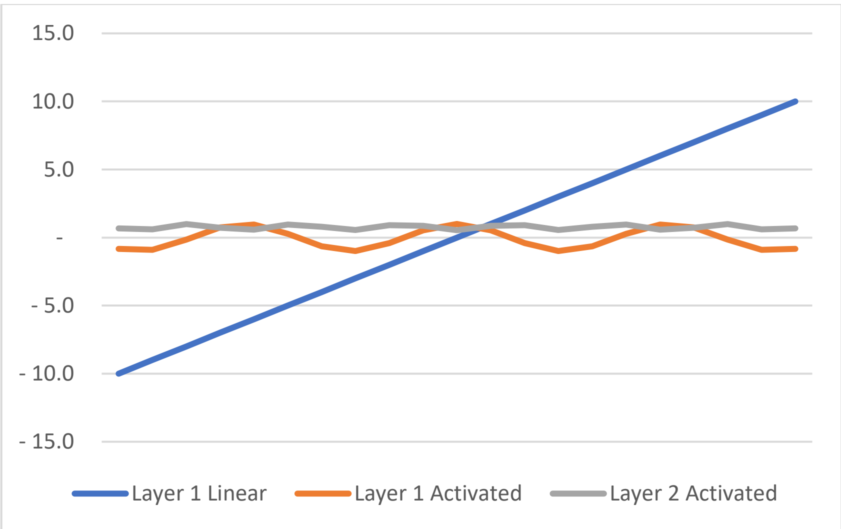

Figure 3 : Transformation of linear neuron outputs via activation function

Mathematical description of a neural network

Mathematically, each neuron could be written as [26][28]:

$$Neuron_{output} = f_{activation}(I_1 * W_1 + I_2 * W_2 + \cdots + I_n W_n + W_{bias})$$

Where:

$$I_n = Input\ Feature$$
$$Wn = Neuron\ Weight$$

These "nest" in subsequent layers, thus allowing for greater complexity and non-linearity. For example, in the second layer of a 2x1 neural network the equation of the output would be:

$$\begin{gathered} Network_{output} = f_{activation}(\\ f_{activation}(I_1 * W_{11} + I_2 * W_{12} + \cdots + I_{1n} W_{1n} + W_{1bias}) * W_{31} \\ + f_{activation}(I_1 * W_{21} + I_2 * W_{22} + \cdots + I_n W_{2n} + W_{2bias}) * W_{32} \\ + W_{3bias} \\ ) \end{gathered}$$

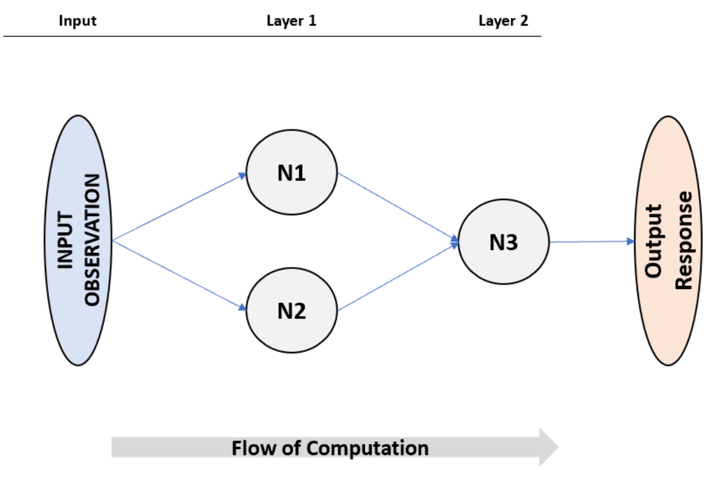

Figure 4: Figure and example of nested equation

Therefore it is apparent that appropriate choice of network arrangement and activation functions is crucial for modelling success as, generally, these define the range of mathematical possibilities that the neuron can output – with training merely adjusting weights to find the most appropriate possibility.

Training a neural network via back propagation

We have until this point in the thesis discussed neurons, both conceptually and mathematically, and explained how these are arranged in 'layers' to create complex outputs. But we have yet to discuss how we can harness these structures to train models that allow us to create powerful and accurate predictions [26][29].

This process is best first explained conceptually. Every training observation we pass to the neural network will generate a single output as it runs through the final layer. But each observation is also associated with a training response – i.e. the "true" output that it is trying to emulate (e.g. the stock price). The variance between these two values is the direct input into our error term for the model – i.e. the measure of how accurate we are [26] [29].

The exact mathematical definition of error varies by use case. In the cast of regression, it is most often mean squared error, which is calculated as follows [30]:

$$E_{mse} = \frac{\Sigma\left(Y_{i\_true} - Y_{i\_prediction}\right)^2}{n}$$

Where:

$$i = values\ associated\ with\ an\ observation$$

$$n = the\ number\ of\ observations\ over\ which\ error\ is\ computed$$

We can use the above term to determine whether our model is accurate – the lower the number the better – and thereafter use it optimize and improve our model using a method called "back propagation". This is a methodology that allows us to find a local minimum of this error/loss term for the entire network. It does this by finding the partial derivative of error term relative to each weight and bias term in the network – and thereafter uses this information to increment each of these terms by a small amount, in many cycles, so as to minimize the error [29] [26]. Put simply, it adjusts the output of the network (via manipulation of the weights and biases) so as to better represent the true outputs.

It is important to note that we can do this as neural network is fundamentally just a continuous mathematically expression, with several constants (i.e. the weights and bias). Therefore, we can always compute a derivative of this equation – and thereafter use this to conduct back propagation. Differing network structures, neuron types and activation functions will only alter the final form of the derivatives – not the fundamental fashion in which we approach the optimization of our networks [29] [26].

With a basic understanding of back propagation, we still need to define the following terms so as to fully understand model optimization [26][31]:

- **Batch –** A batch is a group of training observations feed to the neural network before we updated the weights. This number can vary between 1 and the size of the training set. Whilst larger batch sizes are computationally more efficient (due to the lower number of updates required) they may prevent the system from learning localized features in the data. For example, if we fed the entire training set to the model prior to updating the weights then the only minima the model would be able to identify would be the global mean. Therefore, for each data set, there is a trade off between computational efficiency and sufficient granularity so as to be able to truly capture the intricacies of the feature space.

- **Epoch –** As previously discussed we tend to segment the training data into batches. An epoch is defined as a full cycle through all training observations – i.e. a full cycle through all batches within the training data.
- **Learning Rate** – In the process of training, after each batch, we will identify a partial derivative of the loss function relative to each of the weights in the network. This rate (between 0 and 1) thereafter controls how materially we increment each of the weights – effectively scaling the partial derivative. Generally, we want the model to learn slowly as this allows it to capture delicate behaviour in the feature space. If the learning rate is too high we either **a)** could reach converge but have an overall lower training accuracy as the ,model as failed to capture local details of the feature space or **b)** fail to converge as the model "jumps" too dramatically from training epoch to epoch.

With these now terms now defined, we can mathematically define how we update the network weights every training batch [26]:

$$W_{new} = W_{old} - \frac{dError}{dW} * \alpha$$

Where:

$$W_{new} = The\ weight\ after\ updating$$

$$W_{old} = The\ weight\ prior\ to\ updating$$

$$\alpha = The\ learning\ rate\ of\ the\ model$$

$$Error = The\ global\ error\ term$$

It can therefore be seen that the weights with the largest partial derivative, and thus impact on the model, will have the greatest adjustment undertaken per update cycle. The method of gradient descent can be illustrated on the following curve. This curve represents the error term of the model given the adjustment of a single weight in the network. It is important to note that in practice the network does this simultaneously on all weights and biases.

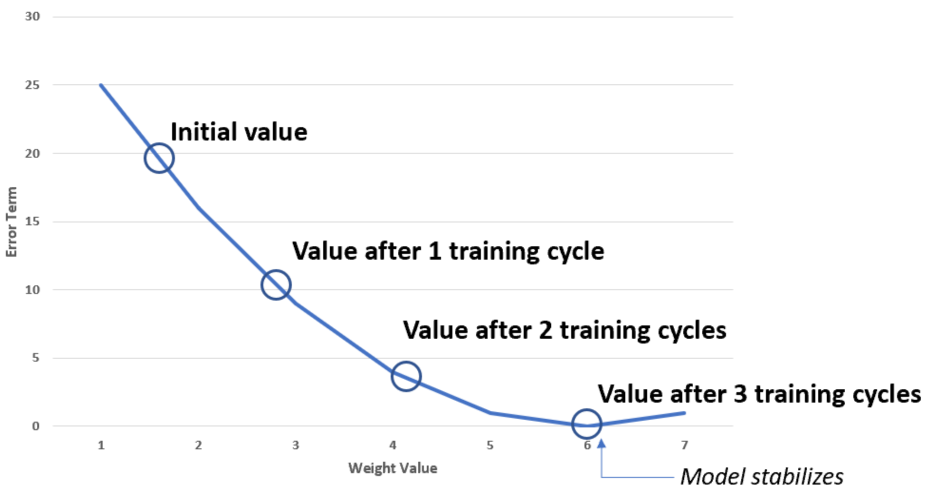

Figure 5: Weight optimization through gradient descent

Generally, we will train the model until one of the following conditions is met [26]:

- We exceed some predetermined number of training epochs or computational time
- We reach some predetermined training error target
- Our cross validation error starts to increase – i.e. we start over fitting our model

These ending conditions can be illustrated on the figure below:

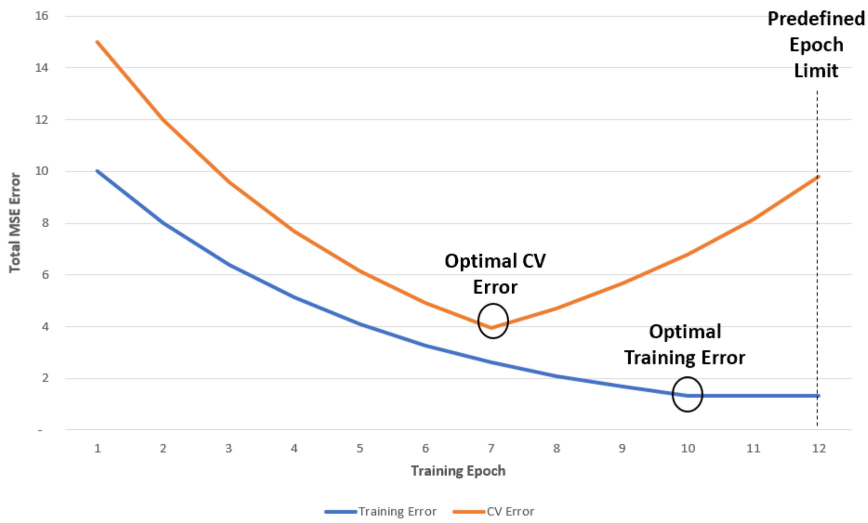

Figure 6: Training end conditions

Summary of network building and training

The general process of training creating, and training, a basic regression network could therefore be best summarized as follows. Please note that this excludes any data preparation or feature engineering steps of the process (e.g. splitting the data into test and train sets) [26].

Step 1 – Determine Network Architecture

- Determine how many layers the neural network requires
- Determine how many neurons are required per layer – remembering that the output layer requires a single neuron in the instance of regression

Step 2 – Determine Neuron Architecture

- Decide which activation function the neurons will use

Step 3 – Determine Learning Parameters

- Decide on the loss function for the network – generally mean squared error in the instance of regression
- Decide on the learning rate of the network
- Decide on the stopping criteria of the network

The following is a very simple and high-level visualization of how the process of training would generally work. Please note that this excludes the splitting of data into test, validation and train and the setting of the end criteria for the network training cycle.

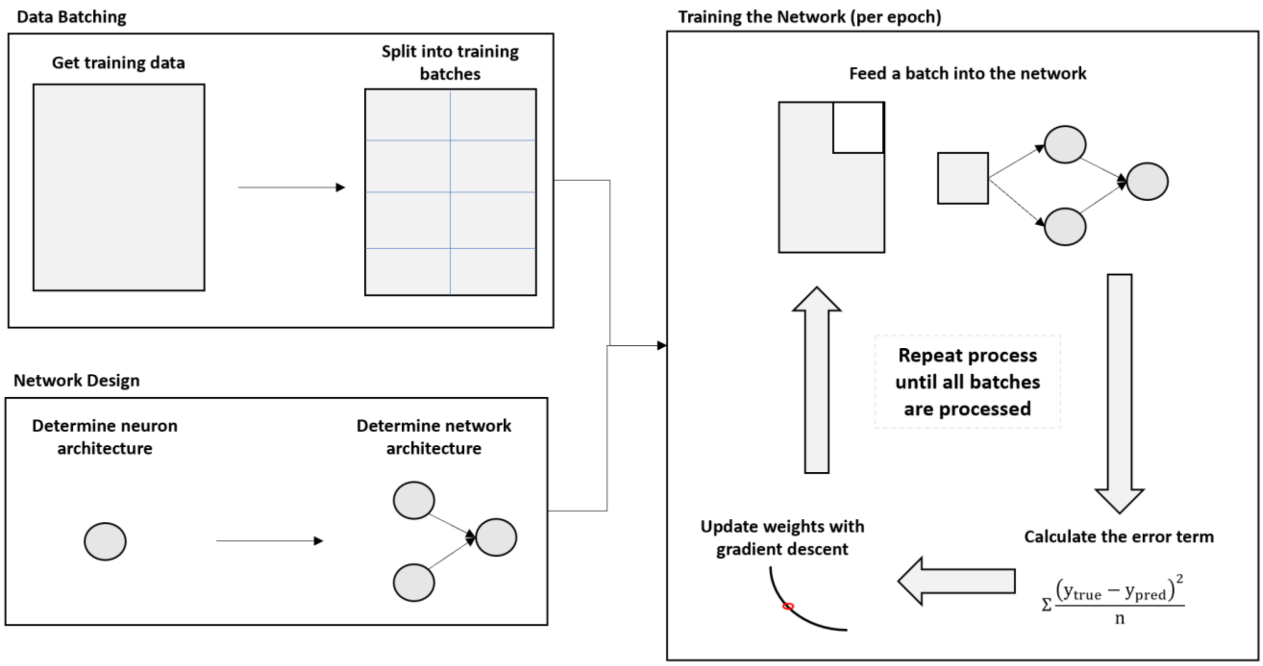

Figure 7: Basic network training

What is an effective platform design?

As discussed in the introduction to this thesis, an iterative design process was undertaken so as to maximise the likelihood of platform success. Each of these successive designs was bound by boundary conditions (sourced from literature and the problem statement), including but not limited to:

- The platform should summarize the theoretical content covered in the "Computational Theory of Neural Networks" section of this thesis
- The platform should approach the problem statement from both a theoretical, mathematical and coding perspective to allow a broad scope of learning for prospective students for disparate educational background and levels
- Simple examples should be used with explicit calculations clarify explanations as the pre-existing knowledge set of the participants is unknown (e.g. no complicated mathematical symbols should be used)
- There is to be a broad emphasis on "interaction" and "visualization" in the final platform, so that students can develop an intuitive feel for neural networks
- A more visual and fun platform would serve to capture the user's imagination and encourage student interaction and improve retention rates
- The platform needs to be fully web based, and therefore it would be ideal if all functionally were either embedded in HTML and JavaScript or some alternative (such as Flask and Python)

Evaluation of each design/prototype

An efficient, and consistent, evaluation matrix is required so as to ensure fair critique between potentially disparate prototypes. This will allow for a – somewhat – objective evaluation of the minimum theoretical requirements of each prototype developed, and ideally would provide an indication of the strengths and weaknesses of each. It also guided development through the standardization the information that each platform will attempt to convey.

The evaluation matrix ultimately used for this thesis was as follows:

Table 1: Evaluation matrix of platform

Metric	Score
Does the platform explain what a neuron is and how it works?	Yes / No
Does the platform explain that a neural network is effectively a collection of interlinked neurons?	Yes / No
Does the platform effectively convey how a neural network learns through back propagation and gradient descent?	Yes / No
Does the platform allow interactivity?	Yes / No
Is the platform effective?	Yes / No
Is the platform visually appealing?	Yes / No

JavaScript and HTML Platform

Technical approach

It was apparent after the first Flask and Python iteration (located in the **Appendix)** that the aesthetics and interactivity of the website is absolutely crucial. After some research it was determined that the following technical approach would be most likely to yield a successful development:

- A standard **JavaScript and HTML** based web development would both allow for a more aesthetic product while greatly improving the accessibility of the website (i.e. it could be run client side and could be reviewed by simply opening a file in a browser rather than launching a Flask server)
- To that end, a free to use HTML template would be sourced to provide the general aesthetic of the website. This is primarily to save a significant amount of time of front end “beautification” through CSS and HTML development. That said, any template used would be dramatically altered from its initial form in the process of development. [32]
- **P5.js** is a JavaScript library that allows the easy creation of interactive canvas HTML objects and would be the base for any interactive mechanisms on the website[33]
- **Tensorflow.js** is a powerful JavaScript library for the development and implementation of neural networks. It would form the core of any neural networks used on the website[34]

Descriptive walkthrough of platform

This iteration was a single scrolling website, consisting of the following sections:

- **Navbar** - A navigation bar that remains top of page as the user scrolls down. It consists of an AIFMRM logo on the left-hand side and links to the relevant sections on the right-hand side. Aesthetics are enhanced by having the currently active section (and any hovered over section link) in the header underlined in blue.

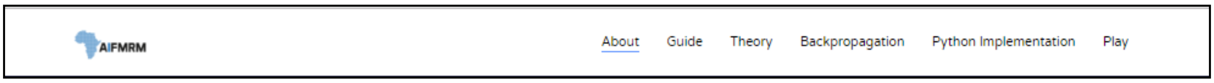

Figure 8: Website Navbar

- **About** - This banner introduces the user to the purpose of the site, i.e. to learn how neural networks work. The blue colour scheme is continued through the dynamic JavaScript background. The background itself is a **Particles.js** script that interacts dynamically with the user’s cursor – they particles actively avoid the cursor and highlight themselves if the user left clicks. The purpose of this is to create a sense of intrigue and interest in the site.

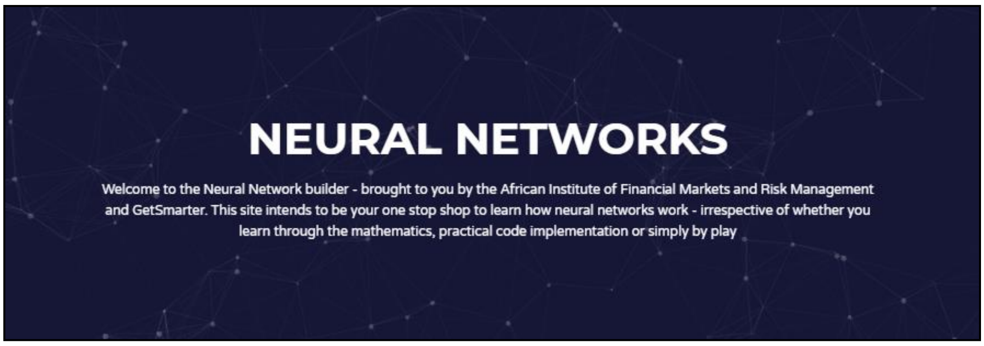

Figure 9: Website About Section

- **Navigation centre** - A navigation centre contains brief descriptions, and links, to the four core sections of the website. The blue colour scheme is continued, and icons are used to represent each of the main sections of the website. The colour scheme of each navigation block is dynamic – they are normally white, but when a user places their cursor over a block, they turn black. This serves the purpose of enhancing visual intrigue and creating a sense of professionalism regarding the site.

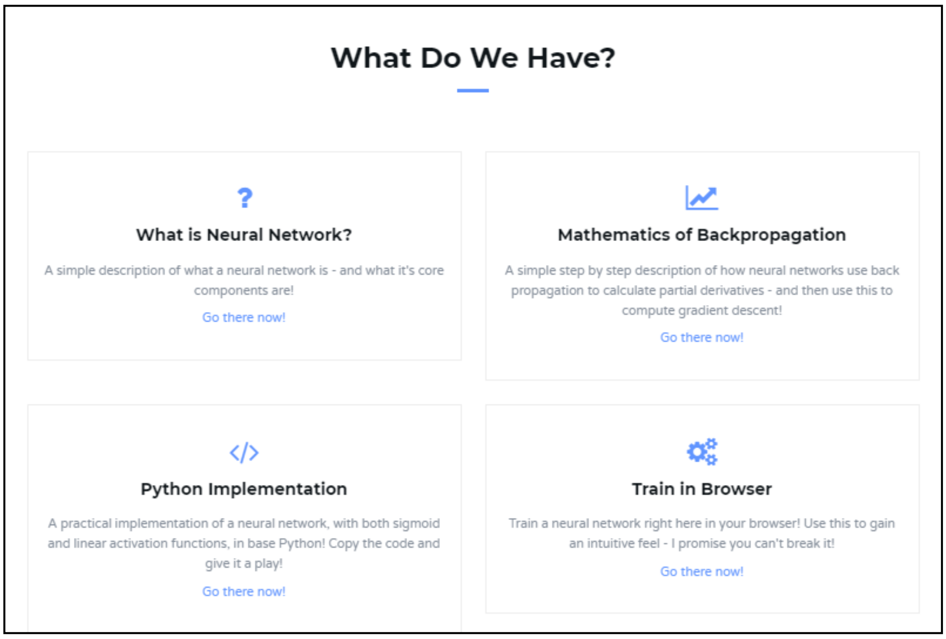

Figure 10 : Website Guide

- **Theory** – This is a brief section explaining the mechanics of how a neuron operates, and how they are interlaced to form a neural network. It makes use of simple and effective diagrams and example calculations. Its intention is to be short, punchy and effective at conveying the basic theory of how a network works.

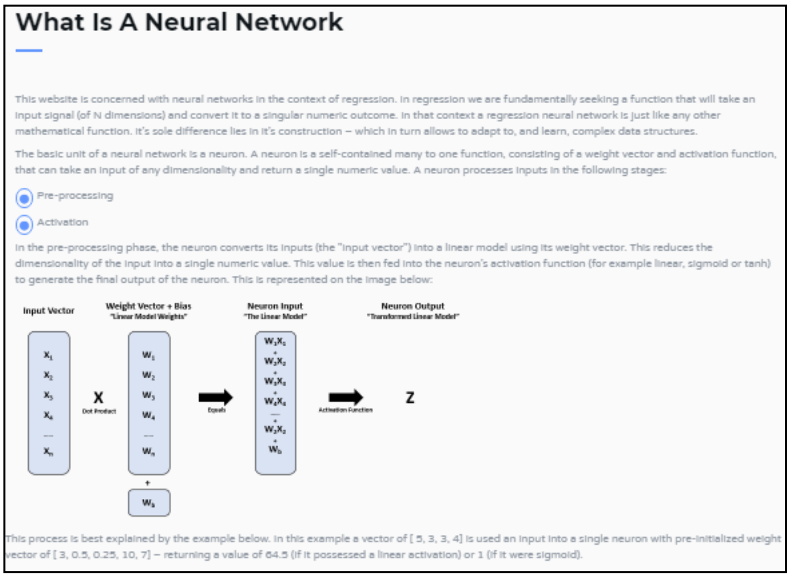

Figure 11: Highlight of website theory section

- **Back propagation** – This section discusses the basic principles of how a network is randomly initialized and how we thereafter train (i.e. adjust the weights) using the partial derivative of the network. This is done via an illustrative calculation of a 2x1 networks actual partial derivative. Importantly this section contains our first dynamic JavaScript canvas to help illustrate the realities of how a network learns via the updating of its weights and bias terms. It is hoped that the interactivity of the canvas reinforces the lesson on back propagation that immediately precedes it. The canvas consists of the following parts:
  - The left-hand side of the canvas consists of a white straight line, and two equations – one of the straight line and another for the soon to be created neural network. Once the user clicks "Reset" a randomly generated 2x1 linearly activated neural network is generated. This has the impact of updating the equation on the lower section of the canvas, whilst drawing a red line representing it in the upper half. When the user clicks "train 1 epoch", the network goes through one training cycle – updating both the equation and the line representing it. This allows the user to see how the network incrementally solves a simple problem, slowly orientating to its training objective.
  - The right-hand side of the canvas contains the weights and biases of all neurons within the network – and tracks them over time. This allows the user to see how each weight tends to solution as the network trains. It also reveals that there are infinitely many configurations of weights and biases that allow the network to emulate the provided target. This is illustrated by the fact that, despite how many times the network resets, a full history is kept since the inception of the platform. Therefore, different equilibrium positions can be viewed at a glance.

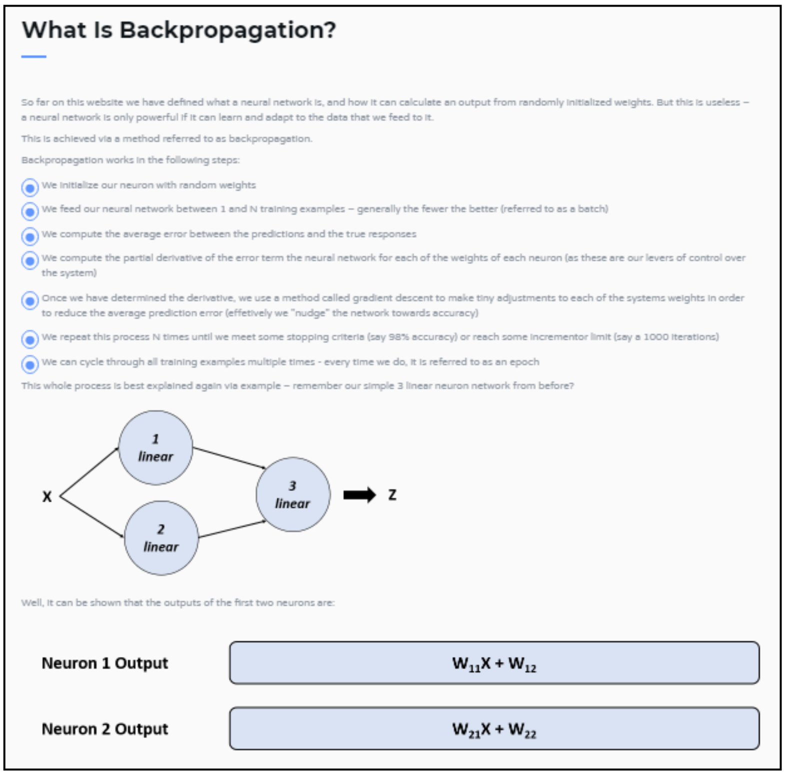

Figure 12: Highlight of back propagation section

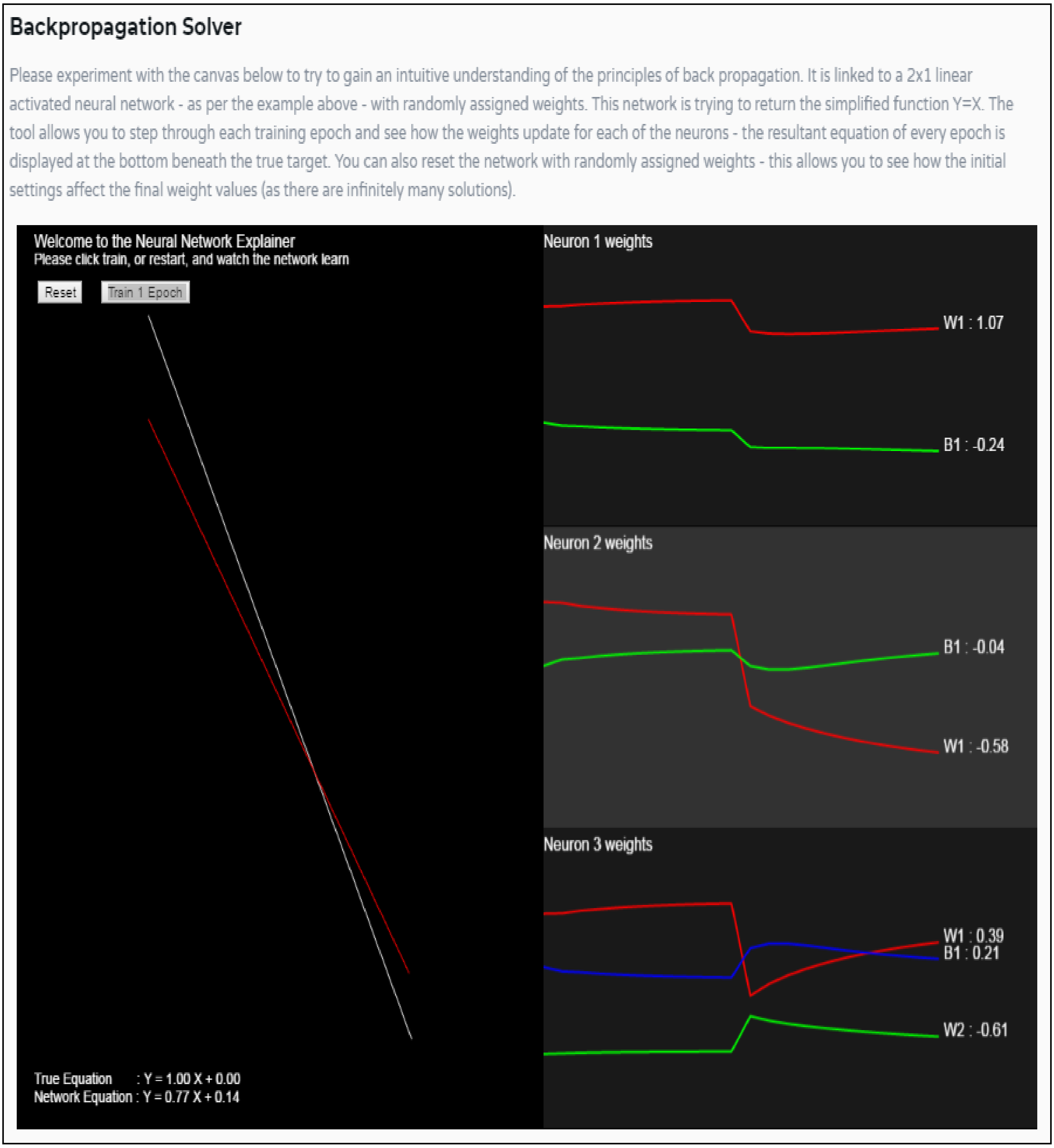

Figure 13: Back propagation interactive solver

- **Python Implementation** – This section contains an independent brief theoretical overview of what a neural network is, and how it works, followed by a fully commented implementation of both a neural network and it's training in base Python. This is achieved via a static embedded IPython notebook and is fully commented to allow the user to read through and digest it in chronological pieces. This section therefore serves two core functions. Firstly, it allows users who are code savvy to interrogate the code, copy it, and use it for their own purposes. It also provides an up skilling opportunity for users who are not. Secondly, it allows a granular step through of how neural networks works that provides a level of understanding that basic theoretical explanations may not provide. To put it plainly, users can finally "see the cogs turning" – which has immense utility on an educational platform.

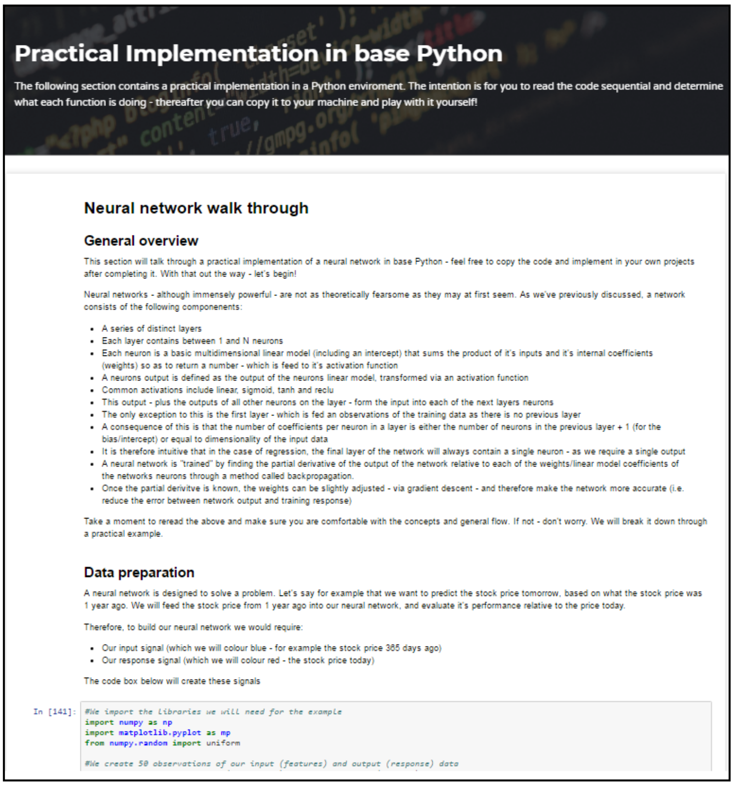

Figure 14: Extract of website Python section

- **Build in Browser –** This is the final educational section of the site and is the second, and last, interactive canvas developed. The intention of this section is to intrigue the user by providing them a "game" to play – one that simultaneously teaches them how a network iteratively learns whilst you play. The canvas is constructed as follows:
    - The upper area of the canvas is a "drawing board" – the user and either click or drag their mouse to draw curves and establish patterns. Each data point drawn is represented back to the users as a white dot. This section then links back to a **tensorflow.js** neural network that will attempt to "learn" the curve once the "train" button has been pressed. At the end of every training epoch the networks current predictions are plotted as a solid red line on the canvas. This allows the user to dynamically see how the network learn, and perhaps answer some questions such as:

        - Which aspects of the feature space does the network initially learn?
        - Which aspects of the feature space does the network struggle with?
        - Does the network oscillate around a local error minimum?
        - How long does it take the network to stabilize?
        - How well does the network handle “noise” in the data?
        - Can the particular network configuration provided handle all shapes provided? Are there some it is unable to cope with?
    - The bottom section of the canvas contains a plot of the training error per epoch within the network. This is useful for the following reasons:
        - The ever-decreasing absolute gradient of the curve shows the user that the network tends to discover major patterns before fine tuning the data. A vast majority of all learning therefore occurs in the first few epochs
        - Oscillations in the curve show that not every iteration of the network is more effective than its predecessors – in fact some of far worse. But it can also be seen how the network can self-correct rapidly. It can also show if the network is oscillating around a local minimum due to an overly aggressive learning rate
        - It is now very easy to identify when the network has plateau in learning progress

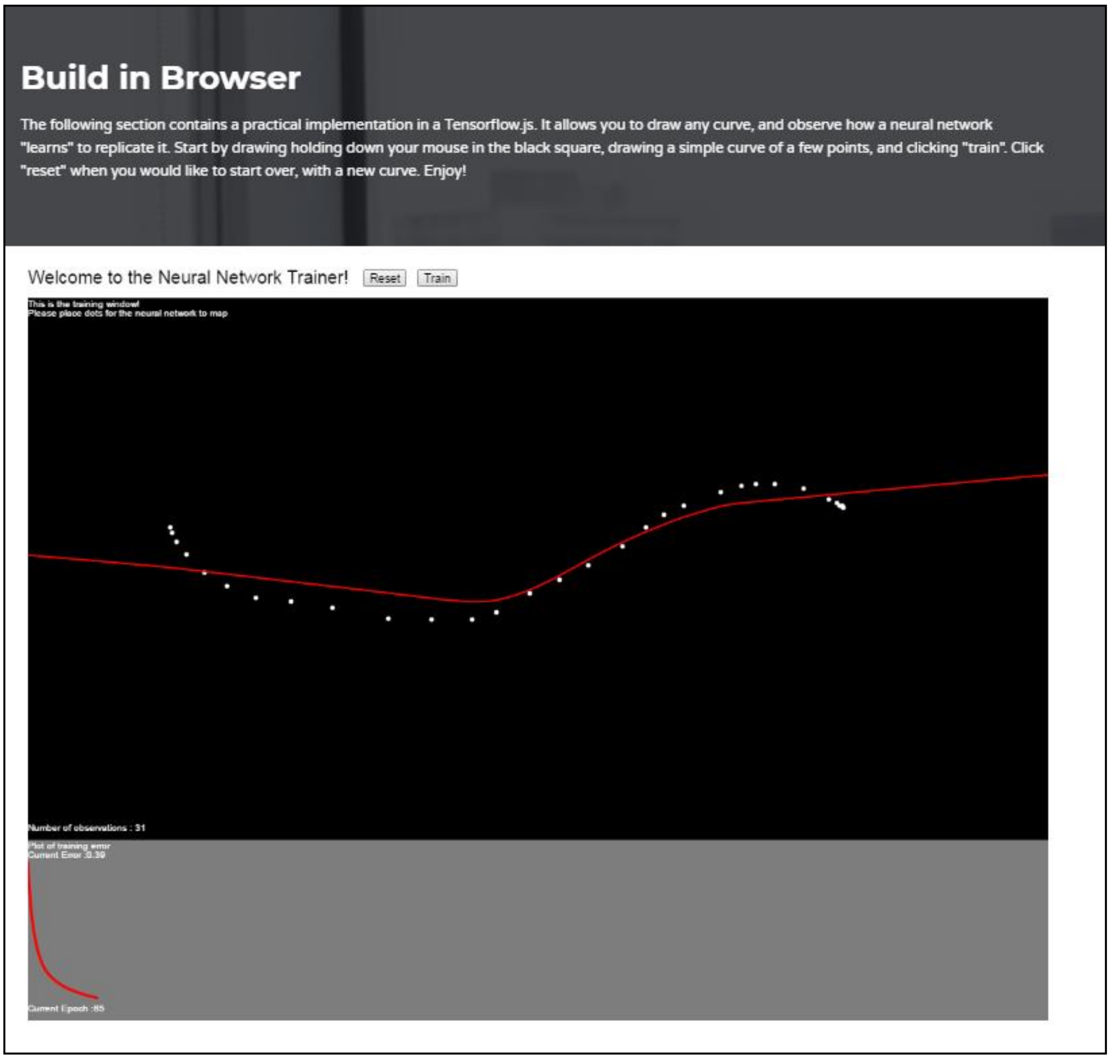

Figure 15: Extract of build in browser section

- **Website Footer –** the final section of the website is a footer highlighting the African Institute of Financial Markets logo, and giving accreditation to the original HTML template used to develop the website.

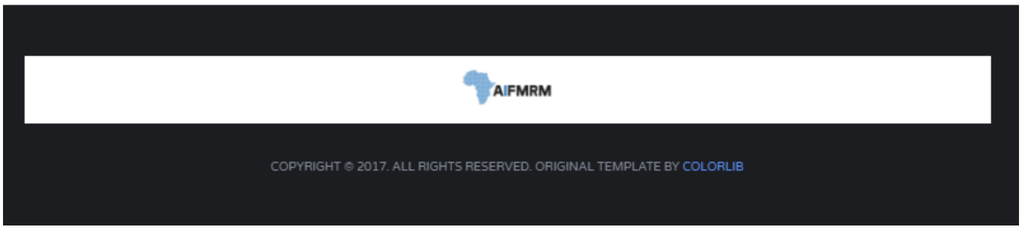

Figure 16: Website footer

General comments relative to Flask and Python development

This platform represents significant improvement over the initial design, particularly in the following areas:

- Aesthetically, the platform is significantly more pleasing. The use of an HTML template and JavaScript have facilitated dynamic feedback to user actions, the inclusion of interactive applet, and an improved overall all visual scheme. This is a significant improvement upon the initial draft.
- The incorporation of a dynamic JavaScript applets both greatly reinforces the material taught through step by step visualization, but also should dramatically increase user retention as it incorporates an element of “gamification”.
- The use of a “single scroll” template is highly effective – it gives the impression that the website is a single resource and that users don’t have to navigate far in order to access the material. This is reinforced by the ever present NavBar at the top of the browser – allowing easy and immediate access to any required section.
- The different areas are simply and effectively demarcated via the use of “header banners” – whose images and text contribute to the overall aesthetic – and differing coloured backgrounds.
- The “Python Implementation” section of the new report effectively encapsulates the entirety of the theoretical content of the previous iteration of the design. Therefore, the new development represents a net increase in content from a theoretical perspective.
- The only significant downside to this iteration is the removal of the ability to build custom neural networks on real world financial data. This has instead been replaced by the interactive canvases. In the opinion of the author this does not significantly weaken the design for the following reasons:
    - The interactive nature of the newly developed canvases is far superior to the “single click” functionality of the initial version. They introduce the concept of gamification, whilst further illustrating, in an iterative fashion, the principles of how a network leanings through its training epochs.
    - The ability to train on real data, whilst impressive, is not critical to the reinforcement of the underlying principles of neural networks. In fact, it could be harmful to the learning experience as real-world datasets containing noise and are needlessly very complex for the purpose of creating theoretical understanding of core principles.

- The ability to customize the neural network that was being trained, whilst very interesting, is again not critical. In the opinion of the author the initial platform provided was not structured sufficiently as to allow the users to truly understand the impact of the various inputs. Therefore, the actual mechanics of effectively designing and optimizing a neural network for a particular dataset should be left to a separate tutorial that can integrate and interweave simple training examples, the impact of various network inputs, and complex real-world data sets.

In conclusion, this iteration therefore represented a net improvement over the original, beyond what is implied by the evaluation matrix below. Given the constrained time of thesis development, it was therefore selected as the final prototype of the development process.

Table 2 : Evaluation matrix of JavaScript and HTML platform

Metric	Score
Does the platform explain what a neuron is and how it works?	Yes
Does the platform explain that a neural network is effectively a collection of interlinked neurons?	Yes
Does the platform effectively convey how a neural network learns through back propagation and gradient descent?	Yes
Does the platform allow interactivity?	Yes
Is the platform effective?	Yes
Is the platform visually appealing?	Yes

Conclusions on platform efficacy

The efficacy of the implemented platform must be gauged on the intersection of its design brief (i.e. the development of a deep learning platform for users of varied backgrounds) and the principles of effective online learning uncovered in literature. Examples of such principles include the use of interaction and gamification, the creation of sense of social inclusion and interaction, the facilitation of immediate feedback, and the moderation of the student's emotional experience.

There was difficulty in applying many of these metrics to the development platform due to its limited scope, and its intended inclusion in a far broader online learning ecosystem. For example, the social dynamics of the course would be independent of the platform as it effectively would just function as a study reference for these students.

There is a powerful, and obvious, second avenue of evaluation, direct iterative user feedback, which not used in the course of this thesis. Using humans as a direct data source requires substantive ethics clearance, and due to the limited development time of this thesis, it was unable to be obtained. This therefore would be undertaken as "future work" for this thesis prior to finalizing the current developmental iteration

That said, we can make the following statements regarding the proposed development:

- The platform covers the intended material in an accessible fashion
- The developed platform caters for different learning strategies (i.e. mathematical, coding and conceptual) and utilizes simple illustrative examples and calculations where possible

- The platform utilizes interactivity and gamification through the implementation of JavaScript applets to both reinforce theory, and encourage play
- The platform is visually appealing, and makes effort to promote a sense of excitement and quality using dynamic and effective styling and colour coding
- The platform is web-based and can be simply operated on any browser via drag and drop if online

Probably the most material omission from the platform is the ability for students to measure their level understanding and gain immediate feedback as to such. An example of a feature that would enable this would be an implementation of a somewhat randomized problem set section (containing both mathematical and conceptual questions) that students can attempt.

Irrespective of that omission, the author would argue that the platform, as is, represents a fair and effective solution to the problem statement of this thesis as it allows students, from a variety of backgrounds and skill sets, to learn about neural networks in a stimulating environment that facilitates gamification and interactivity.

Future work

Future work for this thesis would include (but not be limited to):

- A substantive, and iterative, user review process – with special emphasis on obtaining user feedback from individuals of diverse educational backgrounds. This would allow a more thorough critique of the established platform, and therefore, allow for improved iterative design. The diversity of background is crucial as the intended audience for the platform is very broad, and thus, its efficacy must be tested for a broad range of parties. The review would consist of a questionnaire asking the user for their broad opinion, suggestions for improvement as well as broad theoretical questions so as to test their level of knowledge.
  - This was not possible in this implementation of the thesis due to the narrow timeframe of development, and the substantive ethics clearance process required using humans as a primary data source.
- A thorough analysis of the online course habitat in which this platform would exist. Most of the theory researched could not be directly applied to the platform as developed due to its limited scope. For example, the level of social engagement provided by the course or the moderation of the student’s emotional journey exists beyond the reach of the platform. But broadening the scope of investigation would allow a holistic review of the efficacy of implementation.
- A randomized problem and answer section would be implemented. This would include multiple choice theoretical questions and randomized calculation problems (e.g. a network and input will be given, and the user will have to calculate the outputs of every neuron). This would allow users to immediately gain feedback as to their level of understanding, and hopefully allow them to reinforce this knowledge using randomized problem sets.
- Formal integration into the online course would occur. This is most likely just a technical exercise but crucial as without this development of the site would have served no purpose.

Appendix

FLASK and Python Implementation

The first iteration developed was intended to be a simple, modular website. It effectively aimed to combine two resources into a single site. These being:

- An explanatory IPython notebook that would take the user through a theoretical, and practical, walkthrough of neural networks
- An interactive interface that would allow users granular control over a network operating on real data.

Ultimately, the developed website consisted of the following pages:

- An introductory splash page over viewing the intention of the website
- A code and theory IPython walkthrough of the principles of deep learning. This would simply explain the principles of neural networks through a “tutorial” of both theory and code snippets. These code snippets would include a full neural network library, coded in base Python, that would:
    - Allow developers to directly copy code and modify as required whilst
    - allowing laymen to follow the concepts of neural networks in a granular fashion
- A build page that would allow users to deploy their very own neural networks to real world financial data. Some of the controls offered to the user include:
    - The financial ticker
    - The time period of analysis
    - The training and test split of the data
    - The number of neurons per layer in the neural network
    - The activation functions per layer of the neural network
    - The number of epochs the network would train for
    - The batch size of the network
    - The stopping condition – i.e. would it stop when cross validation error starts to increase or when the epoch limit is hit

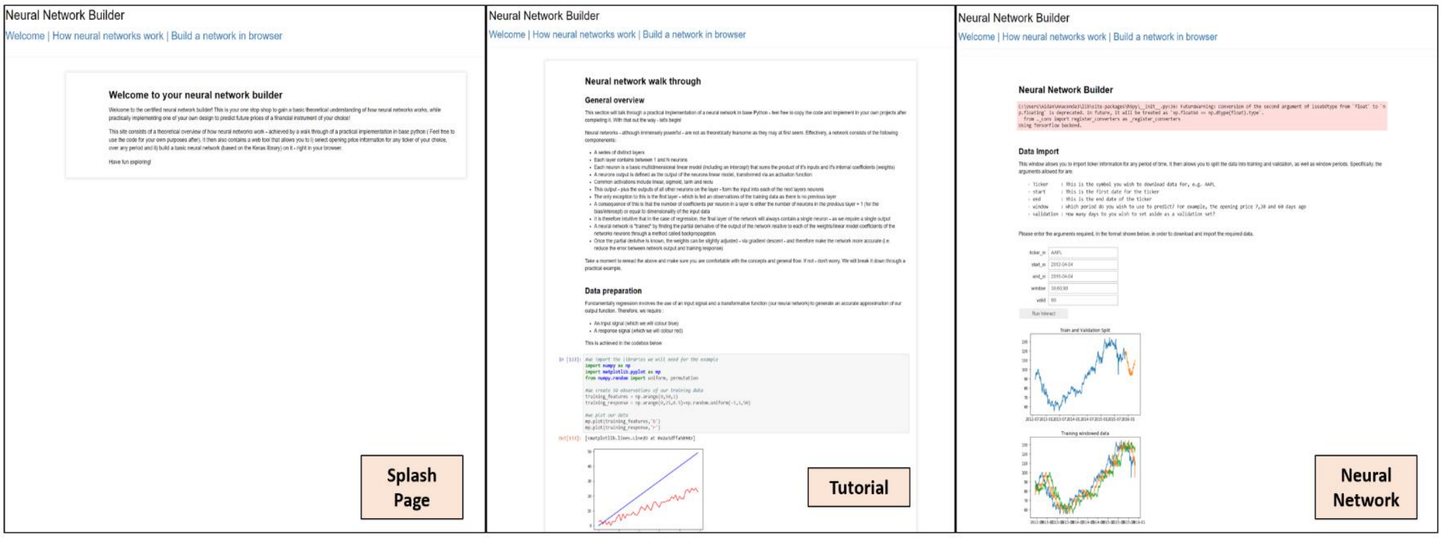

Figure 17 : Extracts from platform pages

General comments on platform efficacy

This platform – in a way – represents an "engineers' solution". It is extremely functional, and simple to update, but ultimately lacks any effective management of the user experience. There is little regard given to aesthetics, with all pages rather barren and bare. In the context of an educational platform, in the opinion of the author, this is a significant failing as visual appeal is a key factor in the retention of attention.

The most effective, and powerful, component of this iteration was the neural network builder. This allowed very granular control of both the network itself, and the real-world data set, on which it would operate. However, it too had problems. The kernel upon which the neural network is running can take several minutes to start up in a new session – which is prohibitively long given the context of the platform. Furthermore, there is little "interactivity" or learning provided by this network – it is very much just "plugs and play". This, to a degree, inhibits the extent to which a user will engage with the platform – once they've run it once, they've seen all functionality.

This iteration was abandoned partway through development (for the reasons mentioned above) after consultation with the thesis supervisor. Key takeaways from the review session indicated the need for an improvement of the general user interface and aesthetics. Ironically, it was also decided by the author that the intention to use real world data, whilst novel, was inhibiting the development of an effective platform. Illustrative examples could be far more effective at facilitating learning. Lastly, a concentred effort would be made in the future iteration to improve the use of interactivity and "play" to foster intrigue and learning.

Table 3 : Evaluation matrix of FLASK and Python iteration

Metric	Score
Does the platform explain what a neuron is and how it works?	Yes
Does the platform explain that a neural network is effectively a collection of interlinked neurons?	Yes
Does the platform effectively convey how a neural network learns through back propagation and gradient descent?	Yes
Does the platform allow interactivity?	Yes (somewhat)
Is the platform effective?	No
Is the platform visually appealing?	No

Installation/Run instructions (Windows)

- Ensure you have Python 3.6.5 installed – 3.7 is not supported yet by some of the libraries used on the site
- Install the following packages on your machine
    - **Flask** – for website hosting
    - **FlaskWTF** – for dynamic website forms to allow the submission of data
    - **Pandas_datareader –** for access to real world financial information
    - **Matplotlib** – for the plotting of the various required charts
    - **Keras** - for the creation of neural networks
    - **Tensorflow** – an engine for the Keras system

- Navigate to the folder in your cmd window
- Type the following command "*set FLASK_APP = microblog.py*"
- Type the command "*flask run"* to begin hosting the website
- The website should now be hosted at ***http://127.0.0.1:5000/***
- Please note that the "build" tab may take a few minutes to initialize once you activate widgets – you can confirm progress by viewing the console in the developer tab. When the kernel is connected it will be ready to begin processing requests.

Works Cited

[1] M. A. Peters, "Technological unemployment: Educating for the fourth industrial revolution," *Educational Philosphy and Theory.*

[2] N. Davis, "What is the fourth industrial revolution?," World Economic Forum, 19 January 2016. [Online]. Available: https://www.weforum.org/agenda/2016/01/what-is-the-fourth-industrial-revolution/. [Accessed 1 January 2019].

[3] S. Makridakis, "The Forthcoming Artificial Intelligence (AI) Revolution : It's Impact on Society and Firms".

[4] "Switch With A Selfie," First National Bank, 2018. [Online]. Available: https://www.fnb.co.za/ways-to-bank/digital/switchWithaSelfie.html. [Accessed 01 01 2018].

[5] Tesla, "Autopilot," Tesla, 2018. [Online]. Available: https://www.tesla.com/autopilot. [Accessed 1 1 2019].

[6] "Amazon and UPS are betting big on drone delivery," Business Insider, 18 March 2018. [Online]. Available: https://www.businessinsider.com/amazon-and-ups-are-betting-big-on-drone-delivery-2018-3?IR=T. [Accessed 1 1 2019].

[7] Y. N. Harari, Homo Deus, Harvill Secker, 2016.

[8] J. B. Adams, "The Fourth Industrial Revolution," *South African Journal of Science,* vol. 114, no. 5.

[9] M. A. Peters, "Smart Work: The transformation of the labour market due to the fourth industrial revolution," *International Journal of Business and Economic Sciences Applied Research.*

[10] W. G. Bowen, M. M. Chingos, K. A. Lack and T. I. Nygren, Interactive Learning Online at Public Universitie: Evidence from Randomized Trials, Ithaka S+R.

[11] S. d. Freitas, J. Morgan and D. Gibson, "Will MOOCs Transform Learning and Teaching in Higher," *British Journal of Educational Technology,* vol. 46, no. 3, pp. 455-471, 2015.

[12] T. Nguyen, "The Effectiveness of Online Learning: Beyong No Significant Difference and Future Horizons," *MERLOT Journal of Online Learning and Teaching ,* vol. 11, no. 2, 2015.

[13] H. Kauffman, "A review of predictive factors of student success in and satisfaction," *Research in Learning Technology,* 2015.

[14] M. Cole, D. Shelley and L. Swartz, "Online Instruction, E-Learning, and Student Satisfaction: A Three Year Study," *The International Review of Research in Open and Distance Learning,* vol. 15, no. 6, 2014.

[15] M.-H. Choa and M. L. Heron, "Self-regulated learning: the role of motivation, emotion, and use of learning strategies in students' learning experiences in a self-paced online mathematics course," *Distance Education,* vol. 36, no. 1, pp. 80-99, 2015.

[16] C. Dziuban, P. Moskal, J. Thompson, L. Kramer, G. DeCantis and A. Hermsdorfer, “Student Satisfaction with Online Learning: Is it a Psychological Contract,” Research Initiative for Teaching Effectiveness.

[17] J. Engelbrecht and A. Harding, “Teaching undergraduate mathematics on the internet,” University of Pretoria, 2004.

[18] C. Kent, E. Laslo and S. Rafaeli, “Interactivity in online discussions and learning outcomes,” *Computers & Education,* vol. 97, pp. 116-128, 2016.

[19] D. Nguyen and G. Kulm, “Using Web-based Practice to Enhance Mathematics Learning and Achievement,” *Journal of Interactive Online Learning,* vol. 3, no. 3, 2005.

[20] J. Herrington, R. Oliver and T. C. Reeves, “Patterns of engagement in authentic online learning enviroments,” *Australian Journal of Educational Technology,* vol. 19, no. 1, pp. 59-71, 2003.

[21] J. Schmidhuber, *Neural Networks,* no. 16, pp. 85-117, 2015.

[22] B. P. E. M. O. E. R. F. P. J. O. I. R. Alberto Prieto, “Neural networks: An overview of early research, current frameworks and new challenges,” *Neurocomputing,* no. 214, pp. 242-268, 2016.

[23] D. K. C. S. M. B. Demis Hassabis, “Neuroscience-Inspired Artifical Intelligence,” *Neuron Review.*

[24] “AI Based Video To Video Synthesis,” [Online]. Available: https://www.youtube.com/watch?v=GRQuRcpf5Gc. [Accessed 1 1 2019].

[25] “AI Learns Real-Time 3D Face Reconstruction,” [Online]. Available: https://www.youtube.com/watch?v=m9XyXiL6n8w. [Accessed 1 1 2019].

[26] U. D. o. Statistics, *MSc Data Science Supervised and Unsupervised Courses,* Cape Town, 2018.

[27] V. Gupta, “Understanding Feedforward Neural Networks,” 9 October 2017. [Online]. Available: https://www.learnopencv.com/understanding-feedforward-neural-networks/. [Accessed 1 January 2019].

[28] T. Babs, “The Mathematics of Neural Networks,” 14 July 2018. [Online]. Available: https://medium.com/coinmonks/the-mathematics-of-neural-network-60a112dd3e05. [Accessed 1 January 2018].

[29] “What is back propogation really doing?,” [Online]. Available: https://www.youtube.com/watch?v=Ilg3gGewQ5U. [Accessed 1 1 2019].

[30] “Backpropagation calculus,” 3 November 2017. [Online]. Available: https://www.youtube.com/watch?v=tIeHLnjs5U8. [Accessed 1 January 2019].

[31] J. Brownless, “What is the difference between a batch and a epoch in a neural network?,” 20 July 2018. [Online]. Available: https://machinelearningmastery.com/difference-between-a-batch-and-an-epoch/. [Accessed 1 January 2019].

[32] “HTML & Bootstrap Webtemplates,” Colorlib, [Online]. Available:

https://colorlib.com/wp/templates/. [Accessed 1 January 2019].

[33] “P5.js Documentation,” Processing Foundation, [Online]. Available: https://p5js.org/. [Accessed 1 January 2019].

[34] “TensorFlow.js Documentation,” [Online]. Available: https://js.tensorflow.org/. [Accessed 1 January 2019].