A Tool for Student Grade Prediction, Revision Direction Assistance, and School Support Recommendation with the Assistance of Neural Networks

by Dylan Parry

Dylan R. Parry
961883

Submitted to Swansea University in fulfilment
of the requirements for the Degree of Master of Computer Science

Swansea University
Prifysgol Abertawe

Department of Computer Science
Swansea University

04/09/2023

Declaration

This work has not been previously accepted in substance for any degree and is not being con- currently submitted in candidature for any degree.

Signed .... (candidate)

Date 04/09/2023

Statement 1

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed .... (candidate)

Date 04/09/2023

Statement 2

I hereby give my consent for my thesis, if accepted, to be made available for photocopying and inter-library loan, and for the title and summary to be made available to outside organisations.

Signed .... (candidate)

Date 04/09/2023

Abstract

Student mental health is a problem that has been getting worse year on year in the UK. A part of this mental health problem is the huge amount of stress that is loaded upon students year in year out, much of which comes from exams and studying for them. To help one's stress and to improve one's grades, it is recommended to find a balanced studying habit that works for oneself, however this can be quite hard when suffering from mental health issues, leading to lowered grades, and spiralling mood. By using the concept of predicting student grades with AI, this paper proposes a pair of tools that students and teachers can use to predict their final grades based on their prior performance which also recommends to the students and teachers' ways for them to improve their final grade by the largest possible margin achievable by them. A pair of neural network models were created and implemented into user-facing python programs, which allow the users to discover their predicted grades, and points out ways that the school could help the student or which previous test they should revise in order to gain the most marks for their future final exam. The paper found that the models created are very accurate in their predictions and that current university students are interested in the capabilities of the program. With more detailed studies and an expanded training dataset, this tool could become a powerful part of the teaching toolkit in the UK, allowing students to improve their grades and take away the stress of figuring out what they should be revising.

Acknowledgements

Throughout the writing of this dissertation, I have received a range of help, support, and advice from multiple people.

I would like to first thank my supervisors, Professor Tom Crick, Dr Sofya Lyakhova, and Dr Tom Owen for approving my proposal for this paper and sending me down this rabbit hole.

Next, I would like to thank my dear friends Kyl and Palu for being there throughout the writing of this paper and the rest of the academic year. Their friendship has kept me happy and sane through many nights of writing and I cherish their companionship very much. You are both inspirations to me and I look forward to many more years of making up characters together.

I would also like to thank my friends Star, Millie, and Mortis. Getting to know you all better through D&D has been the highlight of my week for almost a full year now. Thank you for so passionately enjoying my creative writing, and for spending many a night chatting and watching anime together. I hope we can do so for many years to come.

Finally, I would like to thank my fellow cohort, for they welcomed me after I missed the start of the year due to covid. You are all amazing, and I wonder every day how I managed to be a part of a cohort full of such incredibly talented people.

Table of Contents

Chapter 1

Introduction

Every year, millions of university students across the country face assessments of many types, exams being one of the most popular. In 2022 alone there were more than 2,800,000 students enrolled in higher education in the UK (Higher Education Statistics Agency, 2023), many of whom would be facing these exams which could decide the future of their degree. It is well known that exams are a stressful experience for students, a survey conducted by NatWest discovered that more than a third of students experienced more stress than they expected during their summer exams, a similar number also reported being stressed over university itself (Natwest, 2022). Exam stress is such a common issue that student related charities and the NHS itself offers advice for those experiencing this kind of stress on its website (Student Minds, 2023) (NHS, 2023).

Stress is a natural feeling that can help oneself cope in challenging situations, it can be helpful or even motivating to some. However, prolonged, or too much stress can have a detrimental effect on a person reducing the helpful aspects of this response and making life much harder for the afflicted person. It can make someone irritable, cause them to lose sleep, binge eat, and even withdraw from their social circle (NHS, 2022). These kinds of symptoms can eventually be diagnosed as mental health issues such as Anxiety or Depression, and these mental health conditions are becoming more prevalent across the student population of the UK. In 2015 a parliamentary report suggested that 77% of students they surveyed had experienced feelings of anxiety and 69% had experienced feeling depressed (APPG, 2015). Meanwhile a survey on student mental health in English universities conducted by Student Minds reported that more than half of students self-reported a mental health issue and a quarter have a diagnosed mental health issue (Student Minds, 2023). It is clear from these figures that mental health issues are becoming a larger issue in Universities, the parliamentary statistics on student mental health in England reflects this, with a year-on-year rise in the number of students disclosing a mental health condition to their university. According to their numbers, the number of students with mental health conditions is now nearly seven times higher than it was a decade ago (Joe Lewis, 2023). It is also important to note that these are only the disclosed figures and that, as seen from the disparity in the figures from confidential and non-confidential responses, the real figure could be much higher.

Exam stress is just one contributor to the stress that a student can accrue while studying at university, but what if there was a way to reduce exam stress? Part of the stress of exams is the revision that goes into ensuring that a student can get the best grade possible. Revision is a tool that is encouraged by universities across the UK and the rest of the world as a way of solidifying knowledge and gaining a better understanding of the information. However, during the run-up to exams it can be difficult to know what exactly what topics to revise after months of learning so many different aspects of the module. Having to then balance this with revision for other exams too can make revision a stressful experience, having to decide what to and what not to revise to cover as much information as possible. This paper wants to raise the question of how we can make this revision less stressful and more effective for students. Is there a way for students to discover exactly the topics within a module that they should focus their efforts on, to give them the best chance at the best possible grade?

This paper will address this issue by proposing a project that will investigate the viability of using AI to both predict how a student will perform in an upcoming examination and personally recommend which topics they should try to focus on during their revision to raise this grade the most.

1.1 Aims

To eventually determine the success of this project, several aims have been developed.

1. Explore the background of student stress and mental health issues and the study habits that lead to a student's success.
2. Study how other papers have predicted student's grades and the possible effects this can have on their study habits and mental health.
3. Develop a program and AI model which can predict students' grades based on their prior scores in other exams and homework, which can also recommend what materials a student can study to best improve their grade.
4. Determine the viability of such an app in a real-world scenario through testing and interviews with real students or teachers.

Using these aims at the end of the project will help in the evaluation of whether this project was a success or not. These specific aims were chosen as they encapsulate the goals of the project at large: to research, identify, develop, and provide.

Chapter 2

Background

Before commencing development of the program element of this project, background research into the context of the problem and papers related to the planned solution, must be conducted. This section will consist of a discussion of the state of student stress and mental health issues, as well as a literature review into grade prediction methods and how they might affect a student's ability to study more effectively. This shall provide vital information on the importance of this subject and why the solution proposed may be effective in helping solve the issues at hand.

2.1 Stress and Mental Health

As discussed in the introduction, stress is the brain's natural response to challenging situations. It can be helpful, pushing us to work hard and achieve. However, stress can become a problem when it starts to negatively affect a person's life. When one is unable to escape from stress, physical and mental symptoms can begin to occur. Symptoms such as headaches, chest pain, difficulty concentrating, or feeling constantly worried can significantly reduce a person's quality of life. These symptoms can also cause changes in behaviour like eating too little or too much, sleeping too little or too much, or cutting off social interaction (NHS, 2022). Consistent and overwhelming stress can eventually lead to mental health issues such as anxiety and depression. One example of this is recorded in an article on the experiences of medical students during their first three years of study. Ruzhenkova et all reported that more than half of medical students were sleep deprived, with fewer than 6 hours of sleep, and more than 45% of students had experienced suicidal thoughts (Victoria V. Ruzhenkova*, 2018). Medical degrees are known to be very stressful, with high work loads and heavy topics. This paper shows the real problems that occur when students are placed under stress for long periods of time. Even on a hormonal level stress can be a major factor in the development of mental health issues like depression. For instance, a study discovered that serotonin and stress hormones change in response to sustained stress in the same way that may be observed in people with depression (Praag, 2004).

Depression and anxiety are big issues in the UK, with the country facing a mental health crisis that has been building since before the coronavirus pandemic. It is reported that 1 in 4 people will experience a mental health problem of some kind each year in England (McManus, 2009), and 1 in 6 people report experiencing a mental health problem (like anxiety or depression) each week in England (McManus S, 2014). These mental health issues have a crippling effect on a person's ability to get through everyday life and in the context of Students, can make studying effectively much more difficult. Not only this, but depression can lead to physical harm or even death in the worst cases. It is therefore important to mitigate the causes of depression and anxiety, whilst treating those who suffer from these illnesses in a meaningful and safe way. However, students may not be able to avoid stress in its entirety while studying in university due to the inherent stressful nature of higher education. Students must balance the demanding workload of their course, the pressure from deadlines, and their social lives, all of which can lead to considerable stress.

2.2 Student Mental Health

As seen in the introduction, student mental health across the UK has been suffering, with many students reporting feelings of depression or anxiety, and a considerable number of students feeling stressed during their course. The statistics for the number of students experiencing mental health issues is worryingly high, whether one looks at the official figures or the charity reported figures. But why are so many students struggling with stress and mental health?

It is known that students are particularly vulnerable to the problems that come with stress, as they find themselves in a transitionary period of their life both individually and socially (K. Jayasankara Reddy, 2018). Because of this, they are vulnerable to also show signs of mental illnesses such as depression or anxiety. In one study on academic stress in Indian schools, more than a third of the students who participated showed high scores in the GHQ-28 Questionnaire, a common method of diagnosing mental illnesses (Sibnath Deb, 2015). Another paper followed first year students and their mental health in relation to their end of year academic performance. Whilst there was no link found between mental wellbeing at the start of the year and the academic performance of the student at the end of the year, they did discover that nearly a quarter of the sample reported almost-clinical levels of psychological distress and moderate to very severe anxiety (Phil Topham, 2011). This paper shows us that from the very beginning, university can be an incredibly stressful environment to adjust to. This makes sense, as this will be the first time many of the students will be living apart from their family and friends, a scary experience and change of situation for many people. But it is alarming to see just how many of these students might need clinical help with this experience. This shows just how vulnerable academic stress can make students to developing mental illnesses.

As for where this stress comes from, students frequently experience stress from many sources while at university. An article in the Biomedical and Pharmaceutical Journal investigated the sources of stress, specifically singling out these five pillars: Personal inadequacy, fear of failure, interpersonal difficulties with teachers, teacher pupil relationships and inadequate study facilities (K. Jayasankara Reddy, 2018). These five pillars are certainly contributors to stress levels among students, as discovered in the paper itself, but do not tell the whole story of every other contributor in universities. Many other papers try to pin-point the sources of stress for students, suggesting that things such as team work, placement, administration, cultural differences, accommodation, and economic factors are all factors that add to a student’s level of stress (Zeidner, 1992) (Purna Nandamuri, 2011). These additional areas of stress correlate with Phil Topham’s paper on the mental wellbeing of first year students, as discussed earlier. Accommodation, economic fears, and cultural differences are all worries that may come up immediately upon entering university for the first time.

Clearly, there is much to worry about when attending university. However, not only can this effect the mental wellbeing of a student, but this stress can have a knock-on effect on a student's grades (C. Ward Struthers, 2000). The unfortunate fact is that many students will have to bare the weight of this stress and push forward through their studies to attempt to gain the best mark they possibly can. Depending on the availability of resources in the student's local area, a student could have to wait upwards of 18 weeks to access mental health services through the NHS (NHS Digital, 2022). This simply isn't fast enough when the typical university study year is around 32 weeks. This long waiting time might put them off seeking help, but social stigma around mental illness might also be a factor (Thornicroft, 2011). Whilst social stigma around mental illness has been declining over the years, the fear of being diagnosed still exists within some people. This stigma could prevent students from seeking out help when things get too much for them.

That said, there are ways that students can help relieve stress while at university. This includes talking with others about one's problems, exercising, breathing exercises, and taking control of one's situation through things like time-management and planning ahead (NHS, 2022). The last tip in that list is an important one for students, as with exams and coursework piling up life can seem to spiral out of control. However, with proper management and effective revision, one can prepare for these situations.

2.3 Effective Revision and its Benefits

In this subchapter, we will discuss the benefits of effective revision habits and how this can positively affect a student's mental health and wellbeing. It is well known that revising material can help students better retain information (Garvin Brod, 2016), it is for this reason that so many teachers across the world recommend their students revise before exam season. However, revision itself can be a big challenge for students, some may procrastinate the activity, others may struggle to focus and take in the information that they are revising, some may simply not know where to begin with their studies.

That said, while it is well documented that better study habits result in better performance, it is not the case that simply spending more time studying will result in better grades. There are conflicting results showing that for some students, studying for more time has a positive correlation with their exam performance, whilst for other students the opposite is true (Sarath A. Nonis, 2010). This shows that there is much more to revision than simply re-reading over the same material for hours on end. In fact, one study suggests that it is much more effective to apply the knowledge that one already has in the field to new material in order to improve one's learning results (Monika Andergassen, 2014). Another study discovered that there is a correlation between students with academic self-efficacy and academic performance (Toni Honicke, 2016), this shows us that students who have the right mindset to achieve their goals and work towards them are more likely to achieve a better grade. From all of this we can see that there are many factors at play when it comes to revision and achieving better academic performance, but whatever benefits there might be for one's grade is only one half of the story, as there are benefits to a student's mental health from having better study habits.

One study found that students with lower anxiety tended to have more effective study habits and avoid procrastinating (WITTMAIER, 1972). It should be noted however that this study was conducted in 1972, which may have not taken into consideration other situations such as neurodivergence. But this study may be correct in assuming that those who are better prepared for their exams will feel less stress going into them. Another study found that students will spend about 20% of their active study time distracted and that distraction while studying has a negative effect on exam performance (Elise M. Walck-Shannon, 2021). Apply this school of thought to students with procrastination issues and you can see that these students are going to have a tougher time trying to achieve a higher grade. Procrastination itself is linked to higher levels of depression and anxiety (Gery Beswick, 1988), showing that students with poor study habits may suffer with worse mental health when compared to their peers. This makes sense as if one is unable to begin revising because of one's procrastination issues, the stress of the upcoming deadlines and lack of preparation completed for it will certainly weigh down on the mind.

From the evidence provided we can see that effective revision is an important part in best preparing a student for an exam both academically and mentally. Whilst the approach to revision will differ for each student, we can see that some revision is an important part in increasing the prospects of a student's academic performance. It can also be seen that students with these better study habits may show lower stress levels and signs of mental illness, making revision a strong part in helping a student stay stress-free during the exam period.

2.4 A Literature Review on Grade Prediction

This paper has suggested that in order to help a student choose what topics they should revise for, a program should be developed that predicts a student's future grade and then selects the material that, if improved upon, has the largest positive effect on this future grade. However, using AI to predict a student's grade comes with some questions that should be answered. Does knowing one's predicted grade influence a student's motivation? Can you accurately predict what a student will achieve using AI? To solve these problems, a literature review has been conducted to understand the field of grade prediction.

The first question we must answer is: "Is it possible to predict student grades accurately?" for that we can look at a myriad of papers who have studied different methods of predicting exam results and final grades. One such paper "Personalized

Grade Prediction: A Data Mining Approach" investigated the feasibility of predicting grades through a data-mining and mathematic solution. The paper predicted the final grade of each student based on classes that contained multiple homework quizzes through the year, and those that had a mid-term exam. What they specifically did different in this algorithm, however, is that the paper gives a personalised prediction based on the student's past performance in such a way that the prediction is only given when the accuracy of said prediction is sufficient. They then tested this algorithm on an under-graduate digital signal processing course over the past 7 years, this dataset contained the information on nearly 700 students. What they discovered is that they could indeed predict students grades with this algorithm much better than other machine-learning techniques. They also discovered that classes with mid-term exams are easier to predict, sometimes having as much as a 22% smaller cumulative prediction error over the quiz-based classes. From this exploration they recommend that using students' past data to predict their future grades could be an important tool in helping to intervene where students may be struggling (Yannick Meier, 2015).

Firstly, this paper shows promise as the algorithm involved can in fact accurately predict students grades at various points during their course. It does require time, usually the model is not accurate enough until after the mid-term exam to make a sufficient prediction for the student's grade, but this is the same with many kinds of machine learning techniques for this purpose. What is important for this paper, is that it can *very* accurately predict the final grade of the student, given all their previous homework's, projects, and mid-terms. What is interesting here is that they say when a mid-term exam was included in the course, the grade was much easier to predict. What this could be down to is the amount of effort a student is putting into the mid-term exam, compared to the amount of effort a student is putting into the homework quizzes. The paper states that "The weights of the performance assessments are given by: 20% homework assignments with equal weight on each assignment, 25% midterm exam, 15% course project and 40% final exam." (Yannick Meier, 2015) showing us that the different elements to the course are worth different amounts. There are 7 homework assignments, meaning that each quiz is worth about 2.9%, meanwhile the singular mid-term exam is worth a massive 25%. The lower accuracy of the course which does not have a mid-term exam may be down to how the students perceive the most effort-reward pay out. A student who is only going to get 2.9% of their final grade from a quiz may not put as much effort into it when compared to a mid-term exam which is worth 25%. Therefore, it is possible that this difference in error is down to the weighting of marks, rather than the nature of homework's and in-class exams. This is an important factor to take into future studies, as if the course is built around many smaller assessments and one final exam with a large weighting, grade prediction might not be appropriate to deploy in said course.

Another paper that may inform us on how students react to their predicted grades is "How Gaps between Target and Midcourse Grades Impact Undergraduates' Studying

Intentions and Grade Improvements”. This paper examines how students react to a gap between a student’s own target grade and the average grade they had achieved by the middle of the term. The paper studied 250 university students who voluntarily completed the surveys that were offered to the entire class of over 700. The paper discovered that, on average, most students set their initial target grade much higher than what they are currently capable of, with an average target of 78% against a midcourse average of 63% (H.F. (Herb) MacKenzie, 2020). The paper suggests that this is the effect of their expectations for high-school level grades where students commonly achieve higher marks. The paper also found that students would then re-evaluate their scores to be significantly lower, but still not within reach for their final grade, with average targets dropping to 69% and average final grades being only 61% (H.F. (Herb) MacKenzie, 2020). As one can see, the average grade dropped by two points from mid-term to final exam, but this may have little to do with the exposure to their predicted grades as the paper would go on to discover. The paper found that “…students with midcourse grades farther below their initial target grades had greater studying intentions.” (H.F. (Herb) MacKenzie, 2020) which shows that students who receive a lower-than-expected predicted grade may not be disheartened by this but in fact encouraged to improve upon their grade to achieve their new target. The paper would go on to discover through linear regression analysis that those students with midcourse grades lower than their target grades tended to get larger grade improvements (H.F. (Herb) MacKenzie, 2020).

The paper above shows good support for the work proposed in this paper, as it shows that students who are exposed to a lower-than-expected grade are not disheartened by this, which would go against our goal of making revision less stressful. However, there are many questions left un-answered by this paper, for instance we can see that the overall grades of the students fell after the mid-term, is this a common factor in university courses as the content gets more complex? Can this process be used only to soften this decrease in grades, or can it help improve on the average? The paper itself found out that among students with high growth mindsets, this method triggered a problematic response: a stronger intention to study, with no improvement in final grade (H.F. (Herb) MacKenzie, 2020). This may be down to the students involved in the study being first years, and therefore still un-experienced with the way university works, or it could be that this method pushes them to study more, but less effectively. Our implementation wishes to avoid this problem, by specifically pointing out where the student can improve upon, instead of letting them struggle on their own.

The last paper that we will be reviewing is: “A Preliminary Study of Grade Forecasting by Students” a paper that investigates how students respond to seeing their forecasted grades. By looking into this paper, we might find any overlap in the conclusions from the previous paper that we reviewed, strengthening our understanding of the field and motivation to continue our own work. The experiment in this paper used linear regression to determine the predicted grades of the students, using data from previous semester to provide the data. Once again, there were a few different assessments throughout the term: Two tests worth 15% each, two assignments worth 10% each, and a final exam worth 50%. The paper found that only the tests and the exam were statistically relevant (Armstrong, 2013), once again supporting the idea that at-home assignments worth less marks than tests are not treated the same by students and therefore have less impact on the prediction of the final grade. The 144 students who participated then calculated their forecasted grades and filled in a survey on how they felt afterwards. 47% reported that they were studying more than planned (Armstrong, 2013), which supports what the previous paper also found. The paper also discovered that most students experienced notable emotional reactions to the forecast, 31% reported positive feelings and 35% reported feeling negative feelings, 56% of students also reported increased motivation after the forecast, while only 7% said it decreased (Armstrong, 2013). These emotional responses are expected, as the paper states that 29% of students found that their forecasted grade was lower than expected (Armstrong, 2013), which might result in disappointment or other negative feelings. However, with 56% of students feeling motivated after this, it is not a feeling that lingers for long and seems to instead push the students to improve themselves further. As for how many students believed that this tool should be used in the future, a large majority of the students recommended it: 76% (Armstrong, 2013). This is a very important figure, as it shows that people who use these tools would like to use it again. Lastly, the paper reports that there is little influence from this tool on a student's grade improvement, hence they believe that the student's studying effort is not improved (Armstrong, 2013). This, however, may be because of a self-reported lack of lesser-performing students in this sample (Armstrong, 2013), meaning that those who did participate were already pushing the boundaries of what grade is possible for most students. That conclusion also does not totally fit with the conclusion given by the previous paper, which states that while there is not always an improvement in grade, their grades do not fall as much as other students do. With the larger sample size of the previous paper, it might be more reliable on this consideration.

2.5 Background Overview

We shall conclude this background research chapter with a quick summary of what has been discovered and discussed. We discovered from the papers that currently exist that mental health in students is a major issue that has been getting worse year on year, both in official figures from the government and from reported figures from mental health charities. Students who experience mental health issues such as anxiety can go on to struggle academically from the symptoms that plague them. We then also discovered that proper studying habits can help reduce stress within students, making them more prepared to face the challenges of their courses and achieve better grades. Our literature review considered three different papers on the subject of predicting grades in students, where we found that not only can it be accurately done by many different algorithms and methods, but the effects of finding out ones predicted grades are mostly not negative but in fact positive. Students experience stronger motivation and study more after seeing their predicted grades, and whilst this might not result in an improvement in their grades it can lessen any fall in their overall marks when compared to others in their course. This tool is also supported by the students that use them, with a strong number of students supporting the continued use of it within their course. We also discovered that it is important to think about what the assessment types involved in the course are and the weighting of their marks. High value tests are more significant in predicting a grade than low value homework.

With all of this in mind, we can move forward into the development of our implementation by ensuring that we keep in mind what we have seen. By developing a grade prediction tool that can also help a student focus their revision we can not only help the student feel more at ease about the revision process, but we can also help improve their academic performance, which other papers have failed to do. The framework and prototype we develop in this paper will hopefully lead towards a future experiment where this could be proven.

Chapter 3

Solution Design

This chapter discusses the features that were planned to be included in the final implementation.

3.1 Program Description

The application that we plan on developing is not too complicated in scale, as the task of predicting one's grade and then getting advice on what area one should focus their revision, should be a quick and easy one. For that reason, the tool will not have too many features. The main features of the application that are planned are a student focused grade predictor, a teacher focused grade predictor, a neural network model which will recommend areas to revise, and a way saving the information that has been entered into the application for quick re-entry.

3.1.1 Student Focused Predictor

The first feature of the program will be a student focused grade prediction service. Students will be able to input their grades from the various tests and assessments that they have completed as a part of their module and have a predicted grade given to them through a neural network-based AI. This will allow them to realistically understand what they might achieve given their current skills. An important part of this feature will be for the program to recognise when the prediction it is giving might not be accurate. Therefore, it shall detect when the prediction given is not an accurate grade, or if the input data is not similar in scope to the data it has been trained on. For instance, if the grades given wildly vairy in a way that few students in the dataset's information have, the model may not be able to give an accurate prediction. If this is the case, a warning will be given to inform the user that the prediction is not accurate. In addition, the predictor will come with fair warning on the average error of the model, so that a student is informed that the grade they are predicted to achieve may vary by some points.

This prediction will also come with a recommendation on what aspect of their previous tests they should revise to best improve their grade. This will allow students to stress less over what they should be studying and focus on the subjects that will most likely help them face the upcoming exam.

3.1.2 Teacher Focused Predictor

The next feature of the program is a teacher focused prediction tool for students' grades. This will consider the many different parts of a student that make up their school record, such as the number of days they have missed from school or their at-home situation. This feature is experimental, as there are ethical considerations to considered on taking other factors into account when predicting a student's grade. Largely it will work the same way as the student's grade predictor, however it will have many more inputs and be able to recommend if there are life-circumstances that, if changed, could help the student achieve a higher grade.

3.1.3 Save Feature

Both predictors will also feature a way to save the current inputs, so that a user can exit the program and come back later. This will especially be useful for the teacher side of the program as they will have many more inputs to make.

3.3 Final Specifications

After spending time to review and design these features, a final specification has been created to inform the makers of what is needed in the program to call it a completed framework. These specifications are as follows:

1	The ability for students to enter their previous grades or marks and receive a predicted grade for their final mark in an upcoming exam.
2	The ability for teachers to input the previous grades and other personal information about a student, to receive a predicted grade for an upcoming final exam.
3	The program should determine the one test that student has taken, or the aspect of a student's life, that if improved can have the biggest positive effect on the final grade's prediction. It should present this information to both the student and the teacher.

4	Both the student and the teacher must be able to save the information they have entered, and the information that they are given by the program, so that it is ready when they return.
5	The program must warn the user if any prediction it makes may be inaccurate and it must always inform the user that there is an average error rate to its predictions that the user must consider.
6	The program should not ask anything more than would reasonably be collected by the school when it comes to the teacher's prediction tool.
7	All data saved must be stored on the user's PC and not transferred to anywhere that might be insecure.
8	The app should be intuitive and simple to use.

Figure 3.1: A list of specification points, describing the features and principles that will inform the writers of a successful framework.

This specification will be used to help in the evaluation of the project, these points will be used to understand whether the final creation matches what was planned.

Chapter 4

Methodology

In this chapter the tools and methods that have been used to construct the program are discussed and justified.

4.1 Python

Python is a general-purpose programming language frequently used in data analysis tasks but can also be used for a wide range of different projects. Python is especially useful in the case of AI thanks to a wealth of different libraries designed specifically for machine learning and deep learning. Python is also a very popular language with a massive amount of documentation and help available online from both official and community sources, making it an easier language to use when exploring a task that may be new to the programmer. Another part of what makes Python a great choice for this project is the readability of the code, which makes it more interpretable. This is very useful as the code from the program is planned to be presented in later chapters of this paper, and making this easier to read is an important part of explaining how it works. Finally, the writer of this paper is experienced in Python programming, meaning that it will take less time to produce the implementation than if they were to explore other languages. This will also lead to less errors and a more solid program overall. For all these reasons, Python was chosen as the language to program the implementation of this project.

4.2 TensorFlow

TensorFlow is a free and open-source library for machine learning and artificial intelligence. It was developed by the Google Brain team for internal research and development (Abadi M, 2016), before being released to the public in 2015. TensorFlow is an immensely popular option for building and training models thanks to its prestige from the developers of Google and its integration with the Keras API which makes the process of building and training significantly easier and more understandable. For these reasons, TensorFlow is the right choice of library to use for training our AI for the implementation.

4.3 Student Performance Dataset

In order to train our AI, we need to have data, unfortunately there are very few datasets of student grades over a period of time available for free on the internet, and due to time restrictions we are unable to gain access to a proper dataset to explore the task to the extent that is possible. Fortunately, there is one dataset available that may allow us to present the concept of our theory. The “Student Performance Dataset” uploaded to the UCI Machine Learning Repository shows the student achievement in secondary education of two Portuguese schools. It collects the three scores of students in their maths and Portuguese language classes: year 1, 2 and 3. It also collects a large amount of data about the individual students, including at-home situations and how many days of school they have missed. With this information we will be able to train an AI to predict the final grade of these students, using the two prior grades that are given. This is not optimal for our project but given the lack of data it is the best that we can do. A proper discussion about what the optimal dataset we would require will be discussed in the future work section of the evaluation.

4.4 Google Collaboratory

The process of training a neural network is experimental and usually involves many adjustments to the number of layers. As such it is much easier to train a model in an environment where you can run only parts of the code that you have written and edit parts of the code on the fly. For this, Google Collaboratory is a great choice. Collab is a browser-based notebook for python-based code, and it is particularly suited for machine learning, data analysis, and education. Specifically, Google Collab is a hosted Jupyter Notebook service, allowing one to use the tools of Jupyter Notebook with no set-up and providing access to computing resources such as GPU’s free of charge. Jupyter notebooks allow you to sperate code into “cells” that can be run independently. This allows you to run only chunks of your program, without affecting everything else within it. This is very useful if you want to make small changes to variables without having to run the entire code. It also has strong markdown capabilities, letting you create detailed instructions for others to read and understand the code that is within the notebook. Thanks to this service, we will be able to much more efficiently train our neural network, and provide a notebook for others to read about the training process of the notebook upon completion of the project.

Chapter 5

Implementation

The following chapter discusses how the prediction programs were implemented, it includes screenshots of both programs and explanations of how the underlying code works. Special attention is drawn to interesting parts of the code that have high significance to how the program works. Also included are details on the accuracy statistics of the neural network model and the code on how it was constructed in a notebook.

5.1 Neural Network Notebook

The first part of the implementation to be discussed is the construction of the neural network. This was conducted in a Google Collaboratory notebook and was the first part of the program to be built.

5.1.1 Data Cleaning

The first task in training the neural network was to clean up the data that we have found. The original dataset contained information that was not relevant to the task at hand and inputs that were not numerical in their nature. For a neural network to be trained on this information, those inputs would need to be converted into a numerical type. So, after the irrelevant information was dropped from the dataset, the following code was used to convert any non-numerical inputs into a numerical alternative.

```
df["sex"] = df["sex"].apply(lambda x:1 if x =='F' else 0)
df["address"] = df["address"].apply(lambda x:1 if x =='U' else 0)
df["famsize"] = df["famsize"].apply(lambda x:1 if x =='GT3' else 0)
df["Pstatus"] = df["Pstatus"].apply(lambda x:1 if x =='T' else 0)
df["guardian"] = df["guardian"].apply(lambda x:1 if x =='mother' else
(0 if x == 'father' else 2))
df["schoolsup"] = df["schoolsup"].apply(lambda x:1 if x =='yes' else 0)
df["famsup"] = df["famsup"].apply(lambda x:1 if x =='yes' else 0)
df["paid"] = df["paid"].apply(lambda x:1 if x =='yes' else 0)
df["activities"] = df["activities"].apply(lambda x:1 if x =='yes' else
0)
df["nursery"] = df["nursery"].apply(lambda x:1 if x =='yes' else 0)
df["higher"] = df["higher"].apply(lambda x:1 if x =='yes' else 0)
df["internet"] = df["internet"].apply(lambda x:1 if x =='yes' else 0)
df["romantic"] = df["romantic"].apply(lambda x:1 if x =='yes' else 0)
```

Figure 5.1: The code above removes non-numerical inputs and replaces them with numerical alternatives.

After this has been done, the data has been cleaned to the point where it is ready for being used for training. For this, it was split into two sets: one where all data other than grades were dropped, and a second set where the full information available was kept. This was done so that a model based on only a student's performance could be created, and another model which is based on all aspects of a student.

5.1.2 Neural Network Design

With the data now clean the neural network was designed. Using TensorFlow and Keras we experimented with a number of different layers to find a model that had the best statistics when tested with the test portion of the dataset. The model that was found to be most effective is shown below.

```
model = Sequential()
model.add(Dense(units=29,activation='relu',input_dim=len(X_train.column
s)))
model.add(Dense(units=64,activation='relu'))
model.add(Dense(units=128,activation='relu'))
model.add(Dense(units=64,activation='relu'))
model.add(Dense(units=29,activation='relu'))
model.add(Dense(units=1, activation='linear'))

model.compile(loss='mean_squared_error', optimizer='adam',
metrics='accuracy')

model.fit(X_train, Y_train, epochs=200, batch_size=32)
```

Figure 5.2: The code above shows the design, compilation, and training of the neural network model.

The model takes in a certain number of inputs, depending on whether it is the teacher model or the student model, and returns a number between 0 and 20. We use "relu" activation functions to keep the output steered towards a positive number, however with extreme inputs it can produce small negative numbers. A training time of 200 epochs was found to be the best length of time to train the model to avoid under or overfitting.

5.1.3 Neural Network Statistics

The model was able to achieve fairly impressive results despite its simplicity. There were two metrics by which we could judge the effectiveness of the model: the r2 score and the root mean squared error. For the student model, which only took two inputs, the scores were:

- R2_score: 0.83
- Root Mean squared Error: 1.35

An R2 score of 1 is considered a perfect predictive model, so a score of 0.83 is very strong, indicating very good predictive abilities. Meanwhile the Root Mean Squarred Error gives us an average error of 1.35, meaning that the average difference between prediction and actual grade is 1.35, which gives us a 7% error variation. This is a very small variation value, meaning that the model will generally give us a prediction within 1.35 points of the actual grade a student will achieve. This is important as it means that students, in most cases, can trust the judgement of the program.

The "complex" model with many more inputs had slightly different scores, which makes sense as it has much more information to use in its prediction. The scores for the teacher model were as follows:

- R2_Score: 0.87
- Root Mean Squared Error: 1.26

It is clear from the only slight improvement in scores that the most important part of the prediction model is the first- and second-year grades, which tell much more about a student's future success than their personal circumstances. That said, the model has narrowed its error variation by 1 percentage point down to 6%. The R2 score has also improved by 0.04 points, meaning that this model is *slightly* more accurate in its predictions that the student model. However, for the amount of information needed to be inputted for the increase in accuracy, it makes sense for students to only use the first model and accept the slightly less accurate prediction.

5.2 Student Prediction Tool

With the model complete and tested, it was ready to be placed into a user-facing program, this is where the development changed into IDLE, a Python IDE which is bundled with the Python download. The first tool to be created was the student prediction tool.

The student prediction tool is made up of a single window, containing two input boxes, a single output box, and three buttons. The input text boxes are where a student may enter their first and second year grades whilst the output box displays the predicted grade and revision advice. The three buttons available are "Save" "Load" and "Predict".

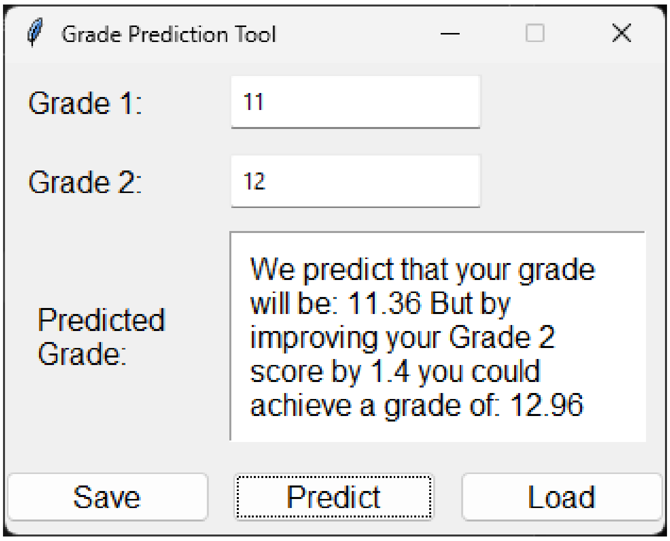

Figure 5.3: The Student Grade Prediction tool, the output box currently displays the predicted grade and revision advice for the student.

The save button saves all text currently written in both the input and output boxes to a text file, allowing a student to then load this information with the load button at a later time. There is only space for a single saved output, so a student will over-write their past save if they save a new prediction, however this was justified as this is designed for a single class.

```
#Saves the current inputs to a text file.
def onSave():
    saveFile = open("studentSave.txt","w")
    L =
[entry1.get()+"\n",entry2.get()+"\n",outputField.get("1.0",END)+"\n"]
    saveFile.writelines(L)
    saveFile.close()

    outputField.delete(1.0, tk.END)
    entry1.delete(0, tk.END)
    entry2.delete(0, tk.END)
    print("Save Sucessful")

#Load the saved inputs to the program.
def onLoad():
    loadFile = open("studentSave.txt","r")
    L = loadFile.readlines()
    loadFile.close()

    outputField.delete(1.0, tk.END)
    outputField.insert(tk.END, L[2].strip())

    entry1.delete(0, tk.END)
    entry1.insert(tk.END, L[0].strip())

    entry2.delete(0, tk.END)
    entry2.insert(tk.END, L[1].strip())
```

Figure 5.4: The saving and loading code. All inputs and outputs are saved to a text file and separated by a newline command. They are then able to be loaded line by line into the program again at a later date.

Upon clicking the "Predict" button, the program takes the input grades and checks them for errors. So long as the inputs are correct and of the right type, the program will then continue into the core of its purpose. The program first decides a grade that the student *could* aim to achieve if they revised the topic more. At first this was a flat increase of 5 points above their current grade, however it was decided that it is more realistic that the higher grade a student has, the harder it would be to achieve a higher grade due to the difficulty of getting a perfect 20 out of 20. Similarly, it was found that students with lower grades tend to improve more on resits, that those with higher grades (T Sutch, 2013). Therefore a function was created that would give those with lower grades, a higher "Achievable" grade, while those with higher grades would have a smaller increase to make to their grade.

This function is estimated as $10 \cdot 0.85^{x}$ where x is the current grade. This is applied to both grade 1 and 2 to find how much a student could improve on that grade.

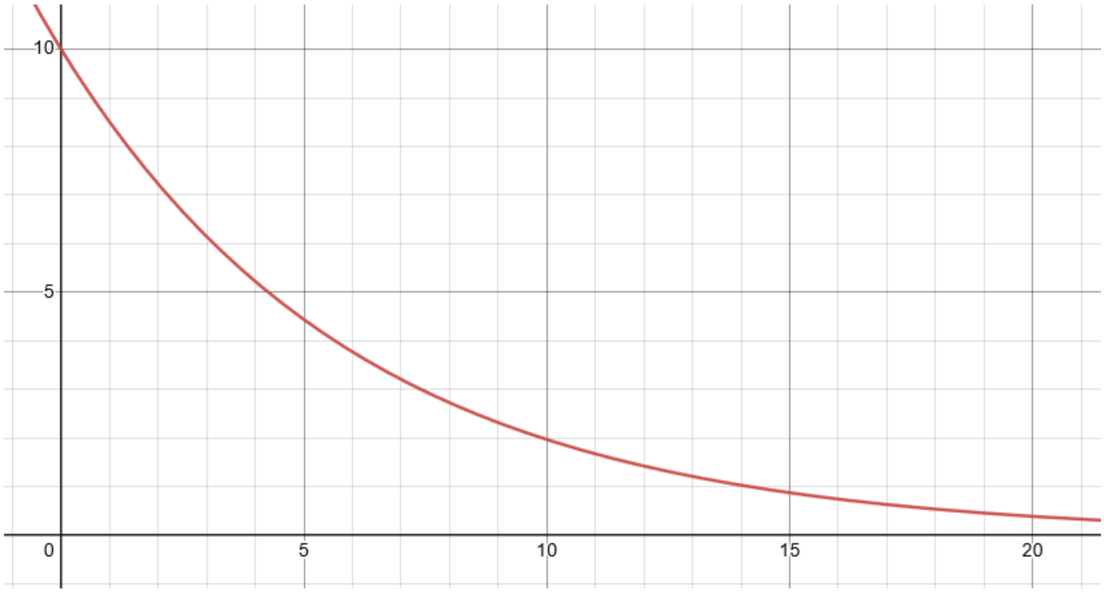

Figure 5.5: The function $10 \cdot 0.85^x$ gives a curve that estimates the potential achievable improvement that a student can make to their current grade.

This potential improvement is added to their original grade to give the “Potential Grade”, this is done for both the first-year grade and the second-year grade to give two potential new grades that a student could achieve with revision. These grades are saved as two tuples, where the new grade is paired with the current grade of the other year. The program must now decide which of these grades gives the largest increase to their predicted grade.

The model uses the three pairs of grades, the original pair and two potential improvements, to predict three final grades. The first is the current grade that the student may achieve, the second is the grade a student could achieve if they focused on their grade 1 revision, and the third is the grade a student could achieve if they focused on their grade 2 revision. The program decides which of the improvement predictions has the largest positive difference to the original predicted grade, and selects that grade as the one which will be recommended for revision.

```
#If the prediction is less than 1 mark, the prediction will be deemed
as innacurate as the liklihood of getting no marks in an exam is
exceptionally low.

if originalResult < 1:
    output = "The predicted grade is out of the accuracy range.
Therefore it has been withheld."
#Otherwise, we check which alternate result has a larger improvement to
the final grade and select that one as the advised revision area.

elif alternateResult1[0,0]-originalResult[0,0] > alternateResult2[0,0]-
originalResult[0,0]:
    output = "We predict that your grade will be: {0:.2f} But by
improving your Grade 1 score by {1:.1f} you could achieve a grade of:
{2:.2f}"
    .format(originalResult[0,0],(10*pow(0.85,grade1)),
alternateResult1[0,0])
else:
    output = "We predict that your grade will be: {0:.2f} But by
improving your Grade 2 score by {1:.1f} you could achieve a grade of:
{2:.2f}"
    .format(originalResult[0,0],(10*pow(0.85,grade2)),
alternateResult2[0,0])
#If the gap between grades is more than or equal to seven, a warning is
given.

if abs(grade1-grade2) >= 7:
    append = " WARNING This result may have higher innacuracy due to
the sizeable difference in input grades."
    output = output + append
```

Figure 5.6: Code for selecting an output. Whichever grade has the highest effect on the original output is recommended as revision advice.

During this time, the program also ensures that any inaccurate results are either thrown out or informed to the user. If a result is lower than 1, it is considered inaccurate, and an error is shown. Also, if the difference between the first and second year grades is larger than seven, which is the largest gap in the dataset, a warning is shown that the prediction given could be inaccurate. This is because the model has no information to train on when it comes to grade gaps of this size, leading to some strange results. Finally, the result is sent to the output box so that the user can read and save it.

5.3 Teacher Prediction Tool

The second tool that was created was the "Teacher Prediction Tool", a more complex version of the student prediction tool that uses many more inputs than just grade 1 and 2. For this tool, it was decided that it would focus on telling a teacher if there is anything that the school itself can do or adjust for the student to help them achieve a better grade. These options are:

- Shortening their travel time.
- Advising them to increase their study time.
- Grant them school support.
- Advising them to cut their extracurricular activities.
- Granting them access to the internet.
- Asking their parents to give them more free time.
- Advising them to spend less time going out.

There are many more options that are included in the program, with 26 different inputs available. However, these seven were considered the most realistic options for a teacher to be able to intervene in.

The program once again consists of a single window, however this time it contains 26 input boxes and then the same output box and three buttons. These 26 inputs are complex in their nature, which is why a README text file was included with this program to instruct the user on how to fill out the information. The available 26 inputs are as follows:

- Sex: The gender of the student.
- Age: The age of the student.
- Address: Whether the student is from a rural area, or urban area.
- Family Size: Whether the student's family is smaller or larger than 3.
- Parent Cohabitation Status: Whether the student's parents are together or apart.
- Mothers Education: The level of study the student's mother has achieved.
- Fathers Education: The level of study the student's father has achieved.
- Guardian: The kind of guardian the student has. Mother, Father, Other.
- Travel Time: The time it takes for the student to get to school.
- Study Time: The amount of time the student spends studying.
- Failures: The number of failed classes.
- Receives School Support: Whether the student receives school support.
- Receives Family Education Support: Whether the student receives family education support.
- Receives Paid Tutorship: Whether the student receives paid tutorship.
- Attends Extra Activities: Whether the student attends extra-curricular activities.
- Attended Nursery: Whether the student attended nursery.
- Aiming for Higher Education: Whether the student is aiming to attend higher education.
- Has Internet: Whether the student has access to the internet.
- Is in a relationship: Whether the student is in a relationship.

- Quality of Family Relationship: A rating of the student's family relationship.
- Amount of free time: A rating of how much free time the student has.
- Time spent with Friends: A rating of how much time the student spends going out with friends.
- Health Status: A rating of how healthy the student is.
- Number of absences: The number of absences the student has had.
- Grade 1: The first year grade.
- Grade 2: The second year grade.

As there are many inputs, the window is much larger in this version of the program.

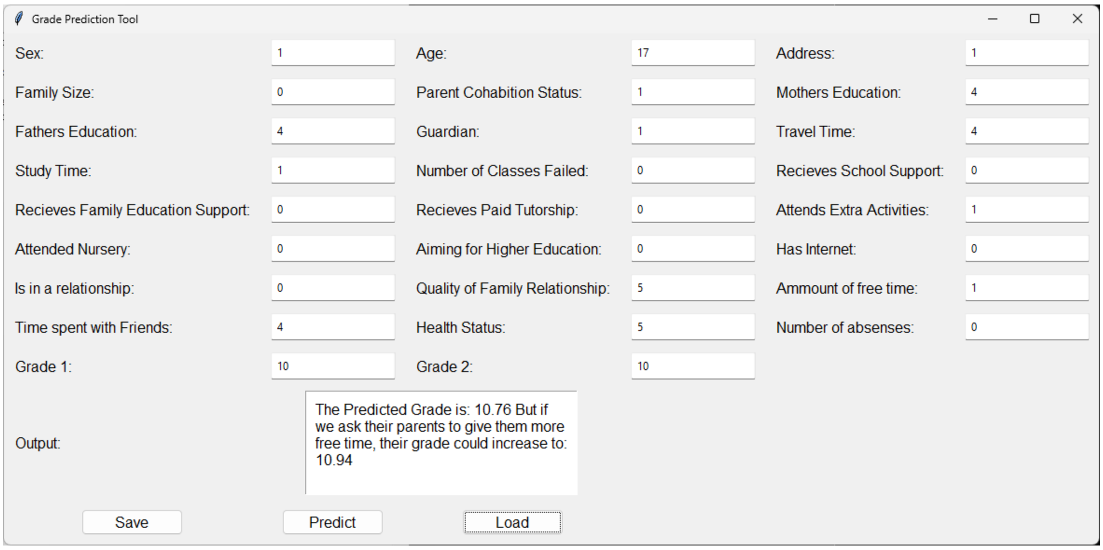

Figure 5.7: The teacher prediction tool, showing the results of a prediction for a student.

Much of the code is similar to the student prediction tool, such as the saving and loading buttons, however the prediction button is slightly different. When the prediction button is clicked the program checks if any of the seven areas for school support are in need of improvement. For instance, if the student's travel time could be cut or if they need internet access. Each time one of these variables is found to be improvable, a deep-copy of the original student record is made and the value is altered by one increment. That is, true to false, false to true, or on a scale reducing or increasing by 1. The program then uses this new student record to make a prediction of the student's grade if this change is made.

```
travelResult = 0
if travelTime > 1:
    travelTimeAlter = copy.deepcopy(originalStudentRecord)
    travelTimeAlter["traveltime"][0] =
travelTimeAlter["traveltime"][0]-1
    dfTravelTimeAlter = pd.DataFrame(originalStudentRecord)
    travelResult = complexModel.predict(dfTravelTimeAlter)[0,0]

studyResult = 0
if studyTime < 4:
    studyTimeAlter = copy.deepcopy(originalStudentRecord)
    studyTimeAlter["studytime"][0] = travelTimeAlter["studytime"][0]+1
    dfStudyTimeAlter = pd.DataFrame(studyTimeAlter)
    studyResult = complexModel.predict(dfStudyTimeAlter)[0,0]
```

Figure 5.8: An example of how the program checks, alters, and predicts an area where the school could help a student improve.

When the program has collected all the possible changes' predictions, they are compared to find which prediction has the largest difference to the original student record's grade predictions. The largest one is selected as the output for the teacher's advice.

```
results = [travelResult, studyResult, schoolSupportResult,
extraActResult, internetResult, freeTimeResult,friendResult]

output = "The Predicted Grade is: {0:.2f}".format(originalResult[0,0])

if max(results) > originalResult[0,0]:
    index = pd.Series(results).idxmax()
    match index:
        case 0:
            output = output + " But if we shortern their travel time,
their grade could increase to: {0:.2f}".format(travelResult)

        case 1:
            output = output + " But if we advise them to increase their
study time, their grade could increase to: {0:.2f}".format(studyResult)

        case 2:
            output = output + " But if we give them school support,
their grade could increase to: {0:.2f}".format(schoolSupportResult)

        case 3:
            output = output + " But if we ask them to reduce their
extracurricular activities, their grade could increase to:
{0:.2f}".format(extraActResult)

        case 4:
            output = output + " But if we give them access to the
internet, their grade could increase to:
{0:.2f}".format(internetResult)

        case 5:
            output = output + " But if we ask their parents to give
them more free time, their grade could increase to:
{0:.2f}".format(freeTimeResult)

        case 6:
            output = output + " But if we ask them to spend less time
going out with friends, their grade could increase to:
{0:.2f}".format(friendResult)
```

Figure 5.9: Code that selects which attribute of the student record should be changed to gain the largest increase in grade.

This advice is added to the student grade prediction and outputted as a single message in the output box (figure 5.7). This can then be saved to a text file and loaded again later.

Chapter 6

Results

After the tool had been completed and was ready to be demonstrated to others, several interviews with current university students were planned to gather the thoughts and opinions of the public on whether they would use such a tool.

6.1 Interview Details

Six students were to be interviewed both in-person and over internet video calls. These students would be shown how the tool works in a demonstration, and then asked questions in a free-form discussion on what they think about the tool, predicting grades, and artificial intelligence. Whilst a study would have been preferred, with the lack of resources and time, there was no chance of one being conducted in this initial paper. Time will be dedicated to discussing how a user-study would be conducted in the evaluation chapter.

Each interview was planned to take less than 10 minutes but would extend onwards if the student had much to say. No rewards or incentives were given for attendance, to gain a less biased judgement.

6.2 Interview Results

The interviews came with many different insights that were useful to overall understanding on the public's perspective on the tool and AI in general. There were areas where the vast majority of interviewees agreed and others where there were differing opinions. Firstly, the interviewees were conflicted on the usefulness of the tool to them and the rest of the student population. All were quick to say that this is a useful tool that they would use, but some of the students had concerns about its usefulness for certain kinds of students. One student brought up that they have struggled with studying in the past and might not be motivated to use this tool to help improve their grades, whilst another pointed out that they might be worried about checking their predicted grade if they were currently doing bad, like "being tight on money and not wanting to check your bank account." Another student mentioned that they would be more likely to use the tool if it was on their smartphone. Many noted that it seemed like a good way of finding areas for improvement, or to inform them where they should be spending more of their time on revising.

On the question of whether they would use this over talking to a teacher, most said that they would prefer that a teacher gave them advice on what to study. However multiple students pointed out that this would be a great tool for introverted or shy students who struggle with talking to teachers. One student discussed how teachers could have bias's and that this would be a good way to get an objective opinion. Discussing this we approached the topic of trust when it comes to AI. Many of the students were sceptical about the tool at first, one student would prefer to know how the tool was trained and on what data before they would be comfortable trusting it. However, many of the other students were more comfortable trusting the tool after being given the R2 and Root Mean Square Error values.

Eventually the interview shifted to a discussion on how teachers might use this tool, after an explanation of the teacher's side of things many of the students were much more excited about the tool. More than half of the students believed that this tool would work better as a teacher support tool, they liked how it could help them identify struggling students. There was one student however who saw this as a possible risk, that if a teacher started to rely on this tool, they would stop trying to get to know their own students. They saw this as a reason for the tool not being used by teachers. Others, meanwhile, wanted to see the tool expanded for students, flagging where students may be succeeding in certain kinds of work but failing in others, such as coursework vs exams. Another student suggested that this should purely be a teacher tool, and students are given their predicted grades and advice from the teacher who might be using this tool in the background. In general, the opinion leaned towards this tool being more useful for teachers than students.

The discussion finally ended on a talk about what they would like to see from a full implementation of this tool. Unanimously the students would want it to be implemented into their online learning environment, as this would allow them to not have to search for their grades or enter them manually. One student believed that it being integrated into the school's platform would make it more trustworthy than a downloadable tool. One student worried that their ADHD would affect how much they would be able to use this tool, as their grades can fluctuate wildly, so if it was able to handle fluctuating grades it could be of more use to them. Another student wanted the tool to run automatically on the online learning platform and for it to inform them when any of their modules are falling below a certain threshold. However, another student rejected this idea, wanting it to only run when they manually ask it to check their grades. This was out of a sense of control of their data and to avoid the stress of being informed about their final grade predictions on a weekly basis. Finally, a student wanted the tool to be able to recognise when a student is struggling with their grades and inform them of any official help they can get, such as contacting a lecturer or visiting extra classes for basic skill improvement.

Overall, this series of interviews allowed the researcher to build a basis of understanding on what real students think of this concept and how they would see it changed. It was considered to be a useful tool in the hands of some students, but it might be more effective as a tool for teachers. There are also edge cases that need addressing such as students who have fluctuating grades, and additional features that could be included to help teachers identify struggling students.

Chapter 7

Evaluation

Now that the tools have been created, tested, and exposed to real students, the next step of this paper is to evaluate the tools and reflect on what happened during the process of creating them.

7.1 Specification Comparison

To aid in the creation of the tools, an 8-point specification was created (figure 3.1). This specification is what we will use to evaluate the success of the project.

The first point of the specification was "The ability for students to enter their previous grades or marks and receive a predicted grade for their final mark in an upcoming exam." This was the core of this project and was achieved almost perfectly. The tool that we created is very accurate in its predictions and is able to provide students with a final grade given they enter their previous two grades. This was *almost* perfect, as there is one edge case where this is not quite accurate. When the difference in grades is equal to or greater than 7 the program informs the user that it cannot make an accurate prediction, while still providing a number. So in this case, a reliable grade cannot be given to the student. Whilst this is an unfortunate circumstance that will require rectifying in future work, this is an edge case which rarely occurs and given that a later specification point requires the tool to warn students of inaccurate predictions, it is an acceptable shortcoming. This specification point can be considered completed.

The second point in the specification is "The ability for teachers to input the previous grades and other personal information about a student, to receive a predicted grade for an upcoming final exam." The second tool was created to serve this purpose, it allows teachers to enter 26 different aspects of a student's life and use it to predict their final grade on an exam. It was able to predict a student's grade using this information more accurately than the student-facing tool, but only by a small margin. However, this tool should still only be considered as a thought experiment, as much of the information included on it is information that a school would not likely keep on a student. For the sake of this project, we can consider this point achieved, but in the long-term process this tool would need strong redevelopment to fit GDPR regulations and ethical guidelines.

The third point on the specification was “The program should determine the one test that student has taken, or the aspect of a student’s life, that if improved can have the biggest positive effect on the final grade’s prediction. It should present this information to both the student and the teacher.” This specification point was the second core aspect of this project and has been achieved thoroughly. The tools are able to discern the right grade or aspect of a student record that should be changed to achieve the highest possible grade. The student facing tool can even simulate the room for improvement that lower-mark students have when compared to higher-performing students. Through a process of improvement and checking, a student could certainly improve their mark with this tool. As for the teacher tool, it is able to advise on up to seven different areas of a student’s life that could be changed to improve their grades. This grade difference is small, but this is because of how important the grades are when compared to their environment. With more experimentation a more effective version of this could be found, but for now this works as a concept of what could be achieved.

The fourth point on the specification is “Both the student and the teacher must be able to save the information they have entered, and the information that they are given by the program, so that it is ready when they return.” This point has also been implemented into both tools. Save and load buttons are prominently displayed and can both effectively save what has been inputted and displayed. Only one grade can be saved at this time, which is a limitation that would need changing, but in the context of this tool being used by one student for a single subject, it works effectively. This specification point is a success for these reasons.

The fifth point on the specification is “The program must warn the user if any prediction it makes may be inaccurate and it must always inform the user that there is an average error rate to its predictions that the user must consider.” Both tools are packaged with a README file which details the statistics on how accurate the models are, and the student tool informs the user if the prediction is likely to be inaccurate due to the range of the grades inputted. The teacher model does not have this ability, but it is intended for higher-level users and therefore is intended to have less guard rails. However, both tools could certainly use some more warnings in prominent areas of the program, or ones which pop-up when the program detects it rather than allowing the user to still see the predicted grade. For this reason, we can consider this specification point to be half-complete. There is room for improvement here in future developments.

The sixth point on the specification is “The program should not ask anything more than would reasonably be collected by the school when it comes to the teacher’s prediction tool.” The teacher tool takes 26 inputs from across the range of information that was available in the dataset, however much of this information is not information that a school would realistically collect on a student. As this tool was seen as more of an experiment than a true working tool, it was considered fine to make full use of all the information available. But, from a specification standpoint, this point is a failure and one that needs to be rectified in any future studies or updates.

The seventh point of the specification is “All data saved must be stored on the user’s PC and not transferred to anywhere that might be insecure.” This point was simple to complete, as all information saved is written to a text file on the user’s PC. This is a very simple way of saving information and one that would need to be changed in a full release, but in terms of the rules set out in this specification it is acceptable. This saving and loading system gets the job done and keeps the information securely on the user’s PC. For that reason, this specification point can be considered complete.

The final point on the specification is “The app should be intuitive and simple to use.” Both of the tools are very simple in their design, to a fault. In a real release these tools would need to be made to look much more interesting to catch the attention of the user. The student tool is easy to understand so long as one knows that the mark they are inputting is between 0-20. However, the teacher tool is not at all intuitive, as one must reference a README file on what to input for it to work. The teacher tool would need to be redesigned to use dropdown lists rather than text boxes for this to be easier to use. For this reason, we can only consider this specification point to be half complete.

If each specification point is worth 1 point, the implementation of this project would score 6 out of 8 points, meaning that we achieved 75% of what we set out to achieve in our specification. This is an ok score, but one that could certainly be improved upon with more time and a fresh start. More on what we would do with more time will be written in the Future Work section of this chapter.

7.2 Problems in Development

Development of this project was fairly smooth with few problems. There were a few small issues with the development, however. The first was the discovery of how strong the model weighted the second-grade mark in its predictions. This is likely because the second grade has more fundamental crossover with the final grade, or because it is simply of a similar difficulty level. This meant that the tool would always recommend that a student revise grade 2 unless there was a major gap between grade 1 and 2. After some research, we discovered the theory that students who take resits stand to gain more when they have lower marks than higher marks, therefore the graph function was developed to ensure that the room for improvement was dynamic in relation to the current marks the student has. This balanced the revision recommendation, allowing it to advise a student to revise more-easier content, rather than trying to learn the few marks they are missing from one grade.

Additional problems occurred when developing the teacher tool, which had so many inputs that it became a long chore to create all of the widgets. There is most likely a more efficient way to create all the widgets, but the researcher could not think of one at the time of development, leading to some very repetitive code.

Another issue was the installation of TensorFlow, as the PC being used by the researcher was brand-new and had not been set up for virtualisation, nor was PIP installed in the right location during the installation of Python, leading to much troubleshooting to find how to change the root in the command prompt to include PIP.

Overall, the development of these tools was mostly hindered by the lack of proper dataset for the task. There was only one student dataset available on the internet that could fulfil the needs of this project, but it was lacking in the number of tests that it contained. An ideal dataset would have many more of these tests, 5 or more would be a good starting point. Due to the lack of time for arranging legal matters and background checks, this project could not collect a dataset of this kind itself, and therefore had to rely on this publicly available one.

7.3 Future Work

With time there are many places that this project can be taken and improved upon. The following sub-chapter discusses the additional changes that could be made given a fresh start and more time to collect a proper dataset.

7.3.1 Future Tool Work

The first changes to be detailed are those to the tools themselves. Firstly, they would be integrated into an online system so that students and teachers are able to pull grades from their records, rather than having to input them manually. This would make the tools much easier to use and less daunting for teachers especially. Additionally, it would allow for students to use this on their mobile devices, as most of these services are web-based.

Another change would be to the kind of information the teacher tool uses in a future tool it would strictly only use the information that is currently available on the school database. This will allow it to pass ethical reviews, as well as GDPR regulations. Another change that should be made is to how the models are trained. They both require more data from a properly sourced dataset that includes a wide variety of grades from different kinds of learners. Students with disabilities such as ADHD and Dyslexia should be included. The teaching tool should be able to notice when a student is struggling and inform the teacher, pointing out that a student is consistently late or is strong in one kind of assessment but weak in another.

Due to the nature of how a model needs a certain amount of information to make a prediction, a weekly predicting tool as suggested by one of the interviewee's is unlikely to be a useful feature. However, offering the tool earlier than just during final exam season is a reasonable idea, and should be implemented as a way for students to understand how they are currently standing. This would most likely be a possibility after half of the year is completed.

Another important part of the tool will be making it enticing to students of all capabilities. It should be seen as a tool that can help anyone, no matter how much they are struggling. For this reason, the tool will need to be properly re-designed to look much more professional and give less clinical answers. Animation and art could be a strong part of the tool, making sure that students know that this tool is not here to criticise, but instead to help.

7.3.2 Future Study Work

With a proper study, a future paper could discover the effectiveness of this tool on a scientific scale, ensuring that this tool is a useful part of the revision process. The study would need to follow one or more classes of students who are allowed to use the tool as they enter the final exam season of the school year. The students would be given questionnaires to figure out what kind of student they are at the start of the exam season, and a second question at the end of the exam season to find out what they thought of the tool. By studying the marks of the students who used the tool, against those who didn't, we will be able to see if there is any effect on improving grades and reducing stress for these students.

There are important things to keep in mind while conducting this test, a wide range of students should be selected, so that the results aren't skewed by students who are more likely to get involved in a study who may also have higher grades on average. Additionally, it would be interesting to conduct this test across different years, such as GCSE students and A-level students, to see if it is more suited to certain age-brackets. Students and teachers should also be brought on board to gain their opinion on the design process of the tool itself, as involving others in the creation of the tools is important in making sure that what is being made is what they want and need.

Chapter 8

Reflection

The final step in this paper is to reflect upon the project as a whole and the lessons that have been learned from it. The aims that this project set out in chapter 1 were as follows:

1. Explore the background of student stress and mental health issues and the study habits that lead to a student's success.
2. Study how other papers have predicted student's grades and the possible effects this can have on their study habits and mental health.
3. Develop a program and AI model which can predict students' grades based on their prior scores in other exams and homework, which can also recommend what materials a student can study to best improve their grade.
4. Determine the viability of such an app in a real-world scenario through testing and interviews with real students or teachers.

Of these aims this paper has managed to achieve the first three aims. The paper was able to thoroughly explore the background of student stress and mental health issues, as well as showing how proper study habits can lead to greater success and lower stress levels but do not necessarily guarantee higher marks.

We followed up this information with research on how other papers predicted student grades and how this might affect student's mental health, discovering that it is both possible to predict grades and that the effect on student's is mostly negligible. With this information we were able to commit to developing a program that would both predict a student's grade *and* give them solid advice on how to improve that grade.

We developed the program and ai model which could both predict student's grades and recommend what they should do to best improve their grades. The model performed very well, with strong results and statistics showing that it is both very accurate and has a low error range.

However, we were unable to determine the viability of such an app in a real-world scenario, only being able to record what people *think* of the tools and not their real thoughts of using the tools.

With this, we can say that this project can be considered a success, providing a strong starting point for this tool to be developed further and to be studied in real world situations. The tool works effectively and has captured the interest of those who have seen it. With time and resources, these grade prediction tools could become a useful part of student's and teachers toolkits.

References

Abadi M, A. A. (2016). *Tensorflow: Large-scale machine learning on heterogeneous distributed systems.*

APPG. (2015, December 7). *Mental Health Poll Nov 15.* Retrieved from All-Party Parliamentary Group on Students: http://appg-students.org.uk/wp-content/uploads/2016/03/Mental-Health-Poll-November-15-Summary.pdf

Armstrong, M. J. (2013). A Preliminary Study of Grade Forecasting. *Journal of Innovative Education*.

C. Ward Struthers, R. P. (2000). An Examination of the Relationship Among Academic Stress, Coping, Motivation, and Performance in College. *Research in Higher Education*, 581-592.

Elise M. Walck-Shannon, S. F. (2021). To What Extent Do Study Habits Relate to Performance? *CBE—Life Sciences Education*.

Garvin Brod, U. L. (2016). Knowledge Acquisition during Exam Preparation Improves Memory and Modulates Memory Formation. *Journal of Neuroscience*, 8103-8111.

Gery Beswick, E. D. (1988). sychological antecedents of student procrastination. *Australian Psychologist*, 207-217.

H.F. (Herb) MacKenzie, M. J. (2020). How Gaps between Target and Midcourse Grades Impact Undergraduates' Studying Intentions and Grade Improvements. *Journal of Innovative Education*, 549-567.

Higher Education Statistics Agency. (2023, January 19). *Higher Education Student Statistics: UK 2021/22 released*. Retrieved from HESA: https://www.hesa.ac.uk/news/19-01-2023/higher-education-student-statistics-uk-202122-released

Joe Lewis, P. B. (2023, May 30). *Student mental health in England: Statistics, policy, and guidance*. Retrieved from House of Commons Library: https://commonslibrary.parliament.uk/research-briefings/cbp-8593/

K. Jayasankara Reddy, K. R. (2018). Academic Stress and its Sources Among University Students. *Biomed Pharmacol Journal*.

McManus S, B. P. (2014). *Mental health and wellbeing in England: Adult psychiatric morbidity survey.* NHS Digital.

McManus, S. M. (2009). *Adult psychiatric morbidity in England, 2007: results of a household survey.* NHS Digital.

Monika Andergassen, F. M. (2014). Practice and Repetition during Exam Preparation in Blended Learning Courses: Correlations with Learning Results. *Journal of Learning Analytics*, 48-74.

Natwest. (2022, August). *Student Living Index 2022*. Retrieved from Natwest: https://www.natwest.com/life-moments/students-and-graduates/student-living-index.html

NHS. (2022, Novermber 22). *Stress*. Retrieved from NHS: https://www.nhs.uk/mental-health/feelings-symptoms-behaviours/feelings-and-symptoms/stress/

NHS. (2023, March 14). *Student stress*. Retrieved from NHS: https://www.nhs.uk/mental-health/children-and-young-adults/help-for-teenagers-young-adults-and-students/student-stress-self-help-tips/

NHS Digital. (2022). *Psychological Therapies: reports on the use of IAPT services, England, June 2022 Final including a report on the IAPT Employment Advisers pilot.* NHS. Retrieved from https://digital.nhs.uk/data-and-information/publications/statistical/psychological-therapies-report-on-the-use-of-iapt-services/june-2022-final-including-a-report-on-the-iapt-employment-advisers-pilot/waiting-times

Phil Topham, N. M. (2011). New students' psychological well-being and its relation to first year academic performance in a UK university. *Counselling and Psychotherapy Research*, 196-203.

Praag, H. v. (2004). *Can stress cause depression?* Progress in Neuro-Psychopharmacology and Biological Psychiatry.

Purna Nandamuri, C. G. (2011). Sources of Academic Stress – A Study on Management Students. *Journal of Management and Science*, 31-42.

Sarath A. Nonis, G. I. (2010). Performance of College Students: Impact of Study Time and Study Habits. *Journal of Education for Business* , 229-238.

Sibnath Deb, E. S. (2015). Academic Stress, Parental Pressure, Anxiety and Mental Health among Indian High School Students. *International Journal of Psychology and Behavioral Sciences*, 26-34.

Student Minds. (2023, 07 14). *Exam stress*. Retrieved from Student Minds: https://www.studentminds.org.uk/examstress.html

Student Minds. (2023). *Student Minds Research Briefing - Febuary '23.* Student Minds.

T Sutch, F. W. (2013). *Reaching for the A*: Exploring the extent and effect of.* Cambridge Assessment .

Thornicroft, G. (2011). Stigma and discrimination limit access to mental health care. *Epidemiology and Psychiatric Sciences*.

Toni Honicke, J. B. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review*, 63-84.

Victoria V. Ruzhenkova*, V. A. (2018). Academic stress and its effect on medical students' mental. *Drug Invention Today*.

WITTMAIER, B. C. (1972). Test Anxiety and Study Habits. *THE JOURNAL OF EDUCATIONAL RESEARCH*, 352-254.

Yannick Meier, J. X. (2015). Personalized Grade Prediction: A Data Mining Approach. *IEEE International Conference on Data Mining.* Atlantic City, NJ, USA: IEEE.

Zeidner, M. (1992). Sources of academic stress: the case of first year Jewish and Arab college students in Israel. *Higher Education*, 25-40.