<Your Project Title>

<Name>

<Student ID>

UXCFXK-30-3

Digital Systems Project



# Abstract

**Your Abstract**

# Acknowledgements

# Table of Contents

[Abstract 1](#_Toc83734960)

[Acknowledgements 2](#_Toc83734961)

[Table of Contents 3](#_Toc83734962)

[Table of Figures 4](#_Toc83734963)

[Introduction 5](#_Toc83734964)

[Literature Review 6](#_Toc83734965)

[Requirements 7](#_Toc83734966)

[Methodology 8](#_Toc83734967)

[Design 9](#_Toc83734968)

[Implementation 10](#_Toc83734969)

[Project Evaluation 11](#_Toc83734970)

[Further Work and Conclusions 12](#_Toc83734971)

[Glossary 13](#_Toc83734972)

[Table of Abbreviations 14](#_Toc83734973)

[References / Bibliography 15](#_Toc83734974)

[Appendix A: First Appendix 16](#_Toc83734975)

# Table of Figures

**No table of figures entries found.**

# Introduction

The rapid growth of retail trading in recent years has brought forth new challenges and opportunities in the investment landscape. As retail investors now account for nearly 25% of total equities trading volume, “Retail investors’ share of total equities trading volume is now approaching 25%, up from 20% in 2020 and 10-15% in the preceding decade.” ([link](https://www.bnymellonwealth.com/insights/the-rise-of-retail-traders.html)) large investment funds have started utilizing sentiment analysis on public forums to capitalize on this trend. Consequently, it is crucial to develop digital tools that empower retail traders by providing valuable insights into market sentiment and levelling the playing field against these large investment funds.

The objective of this digital system project is to design and implement a web-based platform that employs advanced machine learning algorithms for sentiment analysis of real-time news feeds and social media platforms. This investment tool aims to offer retail traders access to vital information on market sentiment, enabling them to make more informed investment decisions.

Throughout this project, we will delve into the technical aspects of designing and implementing the web-based platform. We will explore the integration of various data sources, such as news feeds and social media platforms, and the development of machine learning algorithms capable of accurately interpreting and analysing sentiment from these sources. Additionally, we will discuss the user experience design, ensuring the platform is accessible and user-friendly for retail traders.

Moreover, the project will consider the ethical implications and potential market impact of using sentiment analysis in investment decision-making. We will discuss the challenges and responsibilities associated with providing retail traders with access to sentiment analysis tools and the potential consequences on market dynamics.

Through the development of this digital system, we aim to contribute to the ongoing innovation and evolution of investment tools that leverage cutting-edge technologies, ultimately benefiting retail traders in an ever-changing financial landscape.

# Literature Review

# <https://www.grammarly.com/blog/literature-review/?gclid=CjwKCAiA9qKbBhAzEiwAS4yeDfLruDFL8ZYAdRpWOmBPnq8v27_VOsHj8kkKWoy6yI552qnhN19l9BoC6Q4QAvD_BwE&gclsrc=aw.ds>

**Introduction: Briefly explain the purpose of the Literature Review and its relevance to the project.**

This following chapter explores the problem this project is trying to address, along with background information about sentiment analysis and how it is currently being utilised within the financial sector. It will give insight into the reasons why this technology needs to be made more wildly available. Additionally, it will cover detailed research into the current attempts made with sentiment analysis and what models are available. finally, the topic of user experience (UX) will be covered, along with tools and languages used during APP development.

**Sentiment Analysis in the Financial News Domain: Discuss the importance of sentiment analysis in the financial news domain and its applications, including stock market prediction and trend analysis.**

The field of sentiment analysis is concerned with the computational identification and categorization of opinions expressed in text. The goal is to determine the writer's attitude towards a particular topic, product, etc., which can be positive, negative, or neutral.

In recent years, there has been a significant increase in the use of sentiment analysis due to the advantages it provides to companies in terms of gaining insights about their consumers. One such area of interest is the finance industry, where stock traders use sentiment analysis to analyse financial news feeds and make accurate stock market predictions.

The rise in the use of sentiment analysis is attributed to the exponential growth in data. According to a report by IDC, the global data sphere is expected to grow from “33 Zettabytes in 2018 to 175 Zettabytes by 2025” [1]. With such a massive increase in data, sentiment analysis provides a crucial tool for sieving through the information and gaining a competitive edge.

**Previous Studies: Summarize previous studies that have applied sentiment analysis to financial news headlines or similar data sources, highlighting their methodology, results, and limitations.**

Nicolas Pröllochs et al [2] propose enhancing sentiment analysis of financial news by detecting negation scopes. In the case of Machine learning, it utilizes different variants of Hidden Markov models using both supervised and unsupervised learning, to negation parts of a sentence. Their paper revealed that a rules-based algorithm led to superior results in both applications, these made the prediction forecasts accuracy of up to 89.97%, with an improvement in correlations between sentiment value and stock market return of 9.80. showing the impact between negation is far more significant than previously thought.

**Sentiment Analysis Approaches: Discuss the different approaches to sentiment analysis, including rule-based, machine learning-based, and deep learning-based methods.**

**Rules based approach:**

Rule-based sentiment analysis, also commonly known as lexicon-based sentiment analysis, is one of the simplest and earliest forms of sentiment analysis. In this approach, sentiment is determined by matching words or phrases in the input text against a predefined list of sentiment-bearing words and phrases known as a sentiment lexicon.

A sentiment lexicon is a database of words and phrases annotated with sentiment scores, such as positive, negative, or neutral. Sentiment scores can be assigned in various ways, such as by manual annotation or using automated algorithms that aggregate word associations from large datasets. Common lexicons used in rule-based sentiment analysis include SentiWordNet and WordNet-Affect. These sentiment lexicons can be adjusted towards certain text, for the case of this study Finacial news where phases may be unique to the financial reporting industry.

The process of rule-based sentiment analysis involves mapping words and phrases in the input text to their corresponding sentiment scores in the lexicon, a basic example of this would be “Good, Great, positive” as a dictionary of simple sentiment words, and with phrases “record growth, increased support” they would compute a sentiment score for each document (Kennedy and Inkpen, 2006; Taboada et al., 2006, 2011). With the use of this approach, each expression would influence the sentiment score. With a starting sum of 0 this would define a neutral sentiment score; each positive expression would provide a +1 and a negative –1 after the given text has been passed through the model it would have a final score and that score indicating whether the text was overall negative or positive.

One of the main advantages of rule-based sentiment analysis is its simplicity and interpretability. It is easy to understand the reasoning behind the sentiment classification, as it is based on a predefined set of rules. However, rule-based sentiment analysis has limitations in terms of accuracy and scalability, as it is difficult to capture subtle nuances and context-specific sentiments in the text such as "This stock is on fire, but not in a good way”, a phrase like this could misguide this module as its contains both positive and negative words. Additionally, sentiment lexicons may be limited in scope and coverage, leading to missed sentiment annotations and errors in sentiment classification.

**Machine learning based approach:**

Taking a Machine learning based is more complex than

**Previous Models: Summarize previous models used for sentiment analysis in the financial news domain, including their architecture, performance, and limitations.**

**FinBERT Model: Discuss the finBERT model and its potential for sentiment analysis in the financial news domain. Explain its architecture, pre-training, and fine-tuning, and compare its performance to other models.**

**React and Python: Briefly mention the use of React as the front-end library and Python as the back-end programming language in the project. Discuss the benefits of these technologies and how they will be used in the development of the mobile application.**

**User Experience (UX) Design: Briefly mention the importance of UX design in the development of mobile applications and its role in creating a user-friendly experience. Discuss any previous studies or projects that have used similar approaches to UX design in the development of mobile applications.**

**Conclusion: Summarize the findings of the Literature Review, highlighting the strengths and limitations of existing models, and explaining why finBERT was selected for the project. Discuss how the use of React, Python, and UX design will contribute to the overall user experience of the mobile application.**

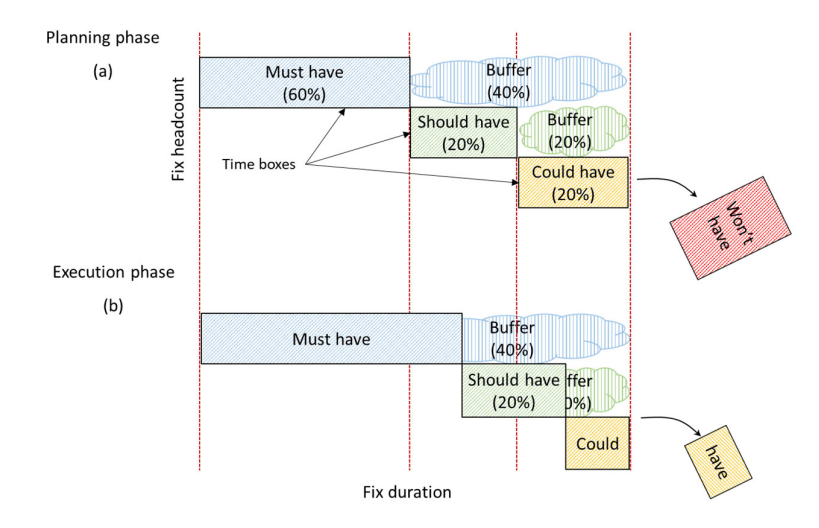
# Requirements

# Chapter introduction

This Section of the report will cover the requirements analysis within this project. There is a wide range of needs that to meet for this project and since the time frame for the digital systems project is limited, it’s essential to prioritise the requirements efficiently. In the chapter It will be making use of the Moscow Method (M - *Must have*, S - *Should have*, C - *Could have*, W - *Won't have*.)

# Moscow Method

This Moscow method for prioritisation commonly down any gather requirements both function and non-function of a project then separates them into four categories: Must have, S - Should have, C - Could have, W - Won't have. (Eduardo Miranda, December 2021). Eduardo states that “*. Each of the first three categories is allocated a fraction of the development budget, typically 60, 20 and 20 percent, and features assigned to them according to the preferences1 of the product owner until the allocated budgets are exhausted by subtracting from them, the development effort estimated for each feature assigned to the category.*” This is very useful when working not with a typical monetary budget but a one based on time as it enables the projects requirements to be allocated a time allowance for that feature.



# Functional Requirements

In the section of the report, the functional requirements of this digital systems project will be looked at and broken down via Moscow Method.

# User Interface programme functional requirements

|  |  |  |
| --- | --- | --- |
| ID | Moscow | description |
| FR-1 | must | The interface program must be able to display if the article is positive or negative |
| FR-2 | must | The interface program must be able to display current news articles |
| FR-3 | must | The application must be secured via user authentication |
| FR-4 | must | The application must allow a user to login, create accounts. |
| FR-5 | must | The application must allow a user to add a preference of stock to be shown. |
| FR-6 | must | The application must be able to link to the sources of the articles shown in the newsfeed |
| FR-7 | must | The application must be easy to navigate. |
| FR-8 | must | The user's data must be secured and protected. |
| FR-9 | must | The application must be able to load a screen in under a second |
| FR-10 | should | The user be able to update account settings |
| FR-11 | should | The user to be able to share articles |
| FR-12 | Won't | Have a way for a user to invest via the application |
| FR-13 | Won't | Have a way for a user to upload their own articles. |

# Sentiment analysis models requirements

|  |  |  |
| --- | --- | --- |
| ID | Moscow | description |
| FR-14 | must | Be able to detect if a news headline is positive or negative with a accuracy over 65% |
| FR-15 | must | The model must be able to work with real time data. |
| FR-16 | must | The model must be able to label any unlabelled data |

# Non-Functional Requirements

# In the section of the report, we will cover the non-functional requirements of our project, these requirements are to define the useability, accessibility, reliability, and performance of the digital systems project.

# User Interface programme Non-Functional requirements

|  |  |  |
| --- | --- | --- |
| ID | Moscow | description |
| NFR-1 | must | This news feed should update with real time data. |
| NFR-2 | must | The interface must be accessible to people suffering with colour blindness. |
| NFR-3 | must | It must be quick and |

# Sentiment analysis models Non-Functional requirements

|  |  |  |
| --- | --- | --- |
| ID | Moscow | description |
| NFR-1 | must | Must be able to read a headline of over 50 characters |
| NFR-2 | must | Be able to return an analysis in under 5 seconds |
| NFR-3 | must | It must reject any headline not in English. |

# Methodology

# Chapter introduction

This Section of the report will cover the methodology that was used during this project. Then a breakdown of the chosen methodology and how it was implemented during the duration of the project.

# Agile Methodology

An agile methodology is a modern approach to managing a large project its commonly used within software development and the broader IT industries. The reason it's so commonly used within these industries is that it’s flexible and collaborative approach enables continuous feedback on current tasks and facilitates rapid adaptation to any changes within the requirements that stakeholders might produce.

The Agile development process is based on a fixed set of guiding principles, those have been spoken about in a tech target article being *“Individual interactions are more important than processes and tools.”, and “A focus on working software rather than thorough documentation. Collaboration instead of contract negotiations”* and *“A focus on responding to change.”. (Brush and Silverthorne, 2019).* These four values lay the foundation for the methodology.

# Gannt charts

# 

# Sprint design

# The most straightforward way to manage time and project allocations is by breaking down the work into sprints. This allowed for the Decomposition of the work into smaller more manageable chunks with clearly defined goals and outcomes, the order of tasks we’re so the project had accomplished the Must haves, that were defined during the requirements sectuion of the report.

# At the start of a sprint, the gannt chart , would be used to track goals and general progression through each task, this was achieved by the use of progress bars, the use of tracking each major task durinng a given sprint was to enable accountability and moreover reduced the overall chance of missing a key function

# At the end of any given sprint, each task would be reviewed and make for any adjustments for example if a given project required more time It could move into the next sprint if it was deemed a must have, via the aforementioned MOSCOW method.

# Sprint 1- backend development (Django and model)

# The goal of the first sprint was to lay the foundation for project. The reason for this to be prioritised was that the tasks during this sprint were critical to the feasibility of project, havig access to a working API for the news, and handling of the data. Would make training and developing the sentiment analysis model far easier.

# Sprint 2- frontend development

# In this sprint, the primary aim was to meticulously design and construct an intuitive and aesthetically pleasing front-end interface for the project, ensuring seamless interaction between the users and the website. Emphasis was placed on the development of an accessible news feed, fostering a gratifying user experience that promotes effortless exploration and consumption of content.

# Sprint 3- Testing and bug fixes

# The goal for this sprint will be to ensure the website is free from any major bugs/errors. Any time saved during the earlier sprints will be used to add more features that were considered “could haves” that the timeline enabled implementing. This is also where the website would be given an Alpha testing to friends and family that have an interest in the project to provide feedback and make any suggestions to changes to the product that could be implemented during the future works section of the report.

# Design

# Dealing with a complex project that uses multiple libraires, that also takes advantage of A.P.I ‘s m the planning/design stage is critical as it enables both hgigh level design to understand the broad strokes of the project. While also showing how a user would interact at a low level, this is to help plan out what pages are required and also the tests that would meed to be developed in order to make sure the site as minimal as possible errors.

# High-level design

# System architecture

# 

# Fig xyzv system architecture diagram

# FIG XYZ Shows the overall system architecture design. The architecture consists of website that communicates t to the database via l queries, the database will be populated by a news feed API that is passed through a sentiment analysis model. These interactions will be hidden from the user, so they will only witness the news feed being updated.

# User interface design

# 

# Fig xyzv : User interface diagram-discovery

# FIG XYZ overall shows the wireframe layout of the website, the design was inspired by the thoughts of simplicity and readability at the heart website. The website's way of interacting is via a scrolling field of news articles, with the ability to to “like” certain companies.

# 

# Fig xyzv User interface diagram-custom

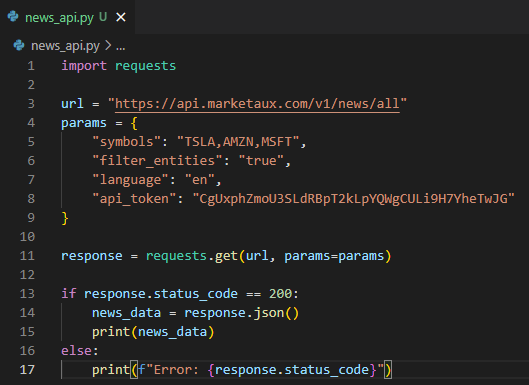
# FIG XYZ Shows the wire frame layout of the changed to custom feed design, it's intended to be identical to the discovery view, the articles here will be limited to ones by the companies that the user has “liked”.

# Database design

# Class diagram

# Sequence diagram

# Implementation



# Project Evaluation

# Further Work and Conclusions

# Glossary

# Table of Abbreviations

# References / Bibliography

# Appendix A: First Appendix

# Miranda, E. (2022) Moscow Rules: A Quantitative Exposé (Accepted For Presentation at Xp2022). *Moscow Rules: A Quantitative Exposé* [online]. [Accessed 11/03/2023].

Brush, K. and Silverthorne, V. (2019) *What is Agile Software Development (Agile Methodologies)?* *SearchSoftwareQuality*.November 2019 [online]. Available from: https://www.techtarget.com/searchsoftwarequality/definition/agile-software-development.