SYDE 660A SYSTEMS DESIGN GRADUATE WORKSHOP 1 - AI AND MACHINE LEARNING

Team 3 - Preliminary Project Report

Authors:
Shinong Mao (ID: 20570392)
Zhiliu He (ID: 21090731)
Beilin Ye (ID: 21053851)

Date: June 17, 2024

Contents

1 Introduction

In the financial trading market, the stock market has emerged as the most significant asset class. Developing prediction models for stock markets using artificial intelligence (AI) is a promising field of research[5]. Technically, stock prediction problems involve using computational models to forecast the price of stocks by training on historical data and testing the performance of models on unseen data to evaluate their predictive accuracy. In practical terms, stock prediction is the process by which investors and traders attempt to forecast the future price of a stock. This involves analyzing factors such as historical price movements, market indicators, etc., and applying models to make reliable predictions and profitable trading decisions.

Thus, accurate stock price predictions are critical for investors aiming to maximize returns and minimize risks. However, the complexity of stock markets poses significant challenges to stock forecasting. Compared to traditional statistical methods such as ARIMA and Moving Average, machine learning methods have demonstrated superior performance in stock price prediction.

Small and Medium-sized Enterprises (SMEs) are typically defined by their limited number of employees and revenue. They play a critical role in fostering new businesses and driving innovation, thereby significantly contributing to national economic growth[1]. According to a recent report by McKinsey &Company[12], SMEs play a crucial role in preventing monopolies and fostering a healthier economy by promoting competition. They challenge larger firms, which leads to improvements in product design, pricing, and efficiency. Additionally, SMEs support larger enterprises by supplying raw materials, components, and services, thereby enhancing the overall efficiency and productivity of the business ecosystem.

Among stock market sectors, Small and Medium-sized Enterprises (SMEs) exhibit relatively high uncertainty in stock price trends but also present significant investment potential due to their agility and growth opportunities. Despite their importance, there is a notable scarcity of research specifically focused on stock prediction for SMEs.

Therefore, this project aims to explore and optimize artificial intelligence models, including Long Short-Term Memory (LSTM) networks, Random Forest (RF), and Support Vector Machines (SVM), to address the stock prediction problem for SMEs. We will train and test the proposed models on stock data from SME-focused indices and exchanges such as the Russell 2000, S&P SmallCap 600, and TSX Venture Exchange. The performance of these stock price prediction models will be evaluated and compared using metrics such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

By identifying and verifying the most effective AI-based stock prediction model, we aim to provide a more accurate tool for investors and foster more informed investments in the SME sector.

2 Needs Assessment

2.1 Problem Space and Motivation

The situation of concern involves the limited research and tools available for predicting stock prices specifically for small-cap stocks. Most existing research in this field either focuses on large-cap stocks, which are more popular, or employs generalized models to address all stock prediction questions. For example, Wang et al.[20] used S&P 500 and Dow Jones Industrial Average datasets, Fischer and Krauss[7] utilized S&P 500 index constituents data, and Thakur and Kumar conducted research on stock indices from NASDAQ, Dow Jones, and S&P 500[10], which are dominated by large-cap stocks. These studies indicate that the datasets researchers commonly use to train and test their stock prediction models are largely focused on large-cap stocks.

In contrast, small-cap stocks (SMEs) exhibit unique characteristics that distinguish them from the general stock market or large caps, including high volatility, low liquidity (lack of popularity), thinner analyst coverage, and limited AI model research. These specialties necessitate the development of a model that can account for these factors, thereby making predictions more targeted, efficient, and accurate.

Investing in small-cap stocks can be highly profitable, as SMEs are often undervalued[16] and tend to outperform large caps, particularly after recessions[18]. Additionally, small caps typically peak and trough before large caps in market cycles and can outperform during young bull markets and economic recoveries[8]. Despite their higher growth potential, SMEs also present higher risks due to their narrow economic moats, which leads to increased competition and volatility[17]. Therefore, tools are needed to support personal investment decisions in small caps, helping investors navigate the associated risks and capitalize on the opportunities.

There are significant challenges and opportunities associated with investing in small-cap stocks. Small-cap stocks are more sensitive to factors such as interest rates, consumer spending, and market sentiment[8]. These insights from finance research can be quantified and incorporated into our AI model to enhance its predictive accuracy and utility.

2.2 Existing Solutions and Competitors

Current solutions include investment analyst reports, which are conducted by humans and tend to be expensive and slow. Generalized AI models for stock predictions do exist, but they often do not account for the specific characteristics of SMEs, thereby limiting their effectiveness for small-cap stock prediction. We aim to fill this gap by developing a specialized AI model that considers the unique features of small-cap stocks, making investment in these stocks more informed and potentially more profitable.

2.2.1 General AI-based Stock Prediction Models

To address the low training efficiency of traditional neural networks, the paper[15] introduced the Time-series Recurrent Neural Network (TRNN) model for predicting stock prices. The model was applied to predict the Dow Jones Index from 1990 to 2019. The results indicated that TRNN is more time-efficient than LSTM but less accurate. Comparative experiments with RNN, LSTM, ARIMA, and GARCH demonstrated TRNN's efficiency advantage.

The paper[3] explores the application of Long Short-Term Memory (LSTM) networks for time-series prediction of stock prices, focusing on companies such as Apple Inc., Amazon Inc., Google LLC, Tesla Inc., Netflix Inc., and BEXIMCO Pharmaceutical. The results revealed that LSTM models achieved a mean absolute error (MAE) reduction of 23.4% and an average prediction accuracy of 89.7%, significantly outperforming baseline models like the moving average.

To enhance the predictive ability for interval-valued stock prices, the paper[11] introduces a dual convolutional neural network (Dual-CNN) model. Applied to six randomly selected stocks from the Chinese market, the Dual-CNN method outperformed several popular models, including CNN, LSTM, GRU, MLP, MSVR, Holt-MSVR, and IDE, demonstrating superior predictive capability.

To improve stock prediction accuracy, this paper[22] proposes a two-stage model that first decomposes the stock price time series using Variational Mode Decomposition (VMD). The decomposed sub-series are then predicted using Support Vector Machine Regression (SVR), Extreme Learning Machine (ELM), and Deep Neural Network (DNN). The preliminary predictions are aggregated using an ELM-based nonlinear ensemble strategy, resulting in superior performance on daily closing prices of four stocks from the Shanghai Stock Exchange.

Similarly, the paper[21] aims to enhance stock prediction accuracy by proposing a two-stage deep ensemble paradigm that uses Singular Spectrum Analysis (SSA) and Variational Mode Decomposition (VMD) to decompose closing prices. Applied to the Shenzhen Stock Index (SZI) and Shanghai Stock Exchange (SSEC), the model integrates Bidirectional Gated Recurrent Unit (BIGRU) for final predictions, achieving better performance in capturing hidden information and providing relevant data for accurate stock price forecasts compared to other models.

Some hybrid methods combine the AI model with text sentiment analysis for stock prediction. To capture stock price fluctuations and the impact of public sentiment more effectively, this paper[6] combines Ensemble Empirical Mode Decomposition (EEMD) and ensemble convolutional neural networks (CNN) with Twitter sentiment analysis. Using public comments and Nifty50 market prices from Yahoo Finance, the model outperformed baseline models, including various LSTM and GRU variants, in terms of MSE, MAE, and sMAPE metrics.

The study[14] investigates the joint impact of market data and financial media on stock prices by introducing the Multi-View Learning Support Vector Machine (MVL-SVM) classifier for stock price trend prediction. Applied to the Shanghai Stock Exchange 50 index (SSE 50 index), the MVL-SVM model demonstrated higher accuracy compared to ARIMA and classic SVM models, highlighting the importance of integrating diverse data sources.

The research[4] proposed the AI Stock Valuation Index (AI-SVI) and the AI in Finance Index (AI-FINX) to reflect the sensitivity of stock prices and returns to AI technology development. It utilized Google Trends to analyze the search volume for several AI-related terms. It examined the correlation between these AI terms' search volumes and the stock returns of companies including Nvidia, Microsoft, and Google. The results indicated that the AI-SVI and AI-FINX are effective matrices for analyzing the sensitivity of a stock to the AI field. According to the research, a further step would be to investigate how to apply the framework to other companies.

2.2.2 AI-based Stock Prediction Models for SMEs

Innovative SMEs have significantly impacted the economies of emerging countries, with their stock price volatility closely tied to economic development and investor behaviors. The study[19] explores the factors affecting the performance of Vietnamese small and medium-sized enterprises (SMEs) using a quantitative method known as the Generalized Method of Moments (GMM). The research verified six statistically significant variables positively influencing SME performance: profitability lag, firm size, leverage ratio, revenue growth, gross domestic product (GDP) growth, and the quality of national governance. This non-AI model emphasizes the importance of financial and governance factors in enhancing SME performance in Vietnam.

The paper[13] introduces a machine learning approach optimized by Bayesian techniques to forecast the stock prices of innovative SMEs listed on China's SSE STAR market. The model incorporates 34 determinants from historical trading data, stock price-related indices, and exchange rates. It utilizes four machine learning models: Random Forest (RF), Deep Neural Network (DNN), Gradient Boosting Decision Tree (GBDT), and Adaboost models. The K-fold method and Bayesian optimization enhance the model's accuracy and robustness, demonstrating its superior predictive ability in the dynamic market environment of innovative SMEs.

2.3 Users Analysis

2.3.1 User Segmentation

Our primary users are individual investors who frequently use our tool. These investors favor shorter-term trading and are less risk-averse, prioritizing returns and willing to take moderate risks.

Secondary users are those who occasionally use our solution, such as those looking for signal indicators for merger and acquisition opportunities. These users will conduct thorough research on investment decisions following initial indications from our tool.

Our anti-users, who are not our target audience and should not use our solution, are institutional investors. These investors require maximum accuracy, typically have a longer-term investment horizon, and prefer fundamental investment over technical analysis.

2.3.2 What our users care about

Drawing from our targeted user demographics and insights gleaned from informal interviews with potential users, we've identified key priorities driving user preferences.
Our users place a premium on cost-effectiveness and operational efficiency. They value predictions with a moderate level of accuracy and prioritize transparency to ensure accessibility for non-technical users. While they prefer limited customization options, they demand frequent, real-time updates and alerts to facilitate informed decision-making.
This framework also serves as our "Performance Evaluation Matrix," guiding our product design process and quantifying user satisfaction metrics.

2.3.3 How prevalent this problem is to our users

Based on our analysis and user feedback, the identified problem holds significant prevalence among our users. The shortage of tailored solutions for specifically predicting small-cap stock performance with all SME stocks characteristics considered, is a widespread concern. Existing research predominantly focuses on large-cap stocks, leaving a gap in the market for tools that cater specifically to the unique characteristics and needs of small-cap investors.
Our user base, comprising individual investors and institutional users seeking merger and acquisition opportunities, regularly encounters challenges in accessing efficient predictive tools tailored to give them hind of the market opportunities and support their investment decisions.
Therefore, our project addresses a prevalent need within the investment community, offering good solutions to enhance decision-making and improve profit outcomes for our users.

3 Impact Assessment

The session discusses the economic, social, technological, environmental, and ethical implications of our project, the AI-based stock prediction model for SMEs.
As for economic and social impacts, studies have shown that AI can significantly enhance the accuracy and efficiency of stock price predictions. Our proposed AI-based stock prediction tool for SMEs can help users quickly identify patterns that human analysts might miss, enabling better investment decisions in SME stocks. This increases diversity in the stock market by bringing attention to small and medium-sized companies, allowing SMEs to obtain more capital for growth. According to a report by The

World Bank[2], SMEs represent about 90% of businesses and more than 50% of employment worldwide. Therefore, the growth of SMEs can contribute to job creation and global economic development. Additionally, SMEs are essential for the creation of new industries and the diversification of the economic base, reducing dependence on a few large firms, which contributes to a healthy economic structure. SMEs also account for a significant portion of technological innovation. Regarding potential risks, the tool could negatively impact users if they overly rely on the stock prediction tool for their investment decisions. This is particularly true for less experienced investors, who might make poor decisions and incur losses if they depend too much on the tool.
From a technological perspective, a new AI-based stock prediction model will be developed and specifically trained to address the prediction challenges for SME stocks.
For environmental considerations, using the AI tool will have some indirect effects. The energy consumption of AI models is a growing concern. Therefore, optimizing the AI tool to require fewer computational resources can be more environmentally friendly. Additionally, accurate predictions can help direct funds towards environmentally friendly SMEs, indirectly supporting sustainability.
For ethical considerations, the use of AI in finance should adhere to regulations. It is crucial to consider data privacy and protection in the AI tool. The need for transparency in AI systems is emphasized by organizations such as the IEEE[9], which advocates that users should understand how decisions are made by AI models.

4 Project Objectives

After identifying the problem space and target users, we came up with the objectives of this project as follows: 1) We will propose a new stock price prediction model aiming small and medium-sized companies; 2) The new model should perform better than the general stock price prediction model; 3) The overall cost of model training should be affordable for personal users; 4) We will have a prototype to demo; 5) Each member will gain hands-on experience in AI model training.
Because inadequate studies of AI stock price predictions were performed on small and medium-sized companies. We will train a new prediction model based on the datasets we collected to fill that gap. Since we are a small team and we have limited resources, we narrow our scope to 1-2 small companies and we will try to improve our model on more general companies depending on our capabilities. And our target users are personal investors who are sensitive to the cost. So add the low model training cost as one of the requirements.
We will start with literature reviews and build models with the general stock price algorithms. We will first do the parallel comparisons of 3 algorithms that are studied individually by each member. The best-performed model will be chosen for further improvement and eventually adopted in our tool. We will try to improve the performance by adding data inputs or adjusting training parameters.

Due to the size and diversity of our team, we could not make a fancy application in the end. But we are not going to just give a model. Instead, we will wrap our trained model inside a backend with well-documented APIs as well as provide a simple UI interface. That makes our tool easy to use and accessible to users with little technical knowledge.
There will be lots of challenges as we work on the project. First, there are many general stock price prediction algorithms in the prior studies. Picking up the best algorithm is challenging since there are various factors affecting the performance. We need to have a deep understanding of the algorithms, narrow our choices of methods, and compare their performances. Another challenge is to improve the algorithm we choose. We care about not only the accuracy but also the overall cost and speed of the predictions, and even more factors. We need to choose a balance point among them.
The whole project was divided into executable tasks and they are summarized in a GANTT chart attached in the appendix. The roles and contributions of each member are as follows:
Contributions so far: 1) We all get together to brainstorm project topics; 2) All members did literature reviews on existing stock price prediction solutions; 3) (Beilin & Zhiliu) Created the team contract; 4) (Beilin & Shinong) Discuss the problem space and user needs; 5) All members join either TA sessions, office hours, or even brief chats after class to gain feedback; 6) The report was divided into 3 parts, and each member took one section: (Beilin) Project plan, and searching evidence for user needs and impacts; (Shinong) Project objectives, design specifications; (Zhiliu) Introduction, need assessment, impact assessment.
Task assignments for the project: 1) We need to test 3 algorithms and choose the one with the best overall performance, so, each member will select a potential algorithm from our prior review publications; 2) (Shinong) create a backend with API that integrates our model; 3) Each member will test on model performance and the final tool; 4) Each member will contribute to the engineering analysis of the model we created; 5) Same as the initial report, we will divide the report into 3 sections and assign them to each member.

5 Design Specifications

After identifying the user requirements, we discussed the design requirements and constraints: 1) The tool should be easy to use: Should be accessible to people without any knowledge; Should contain straightforward UI; Should have simple API and clear documentation; 2) The model should be accurate and correct in prediction: Should predict the trend (either increase or decrease); Should give a predicted price within the tolerance; 3) The model should give the result rapidly: Should give the result in seconds. They are detailed in the Quality Function Deployment (QFD) in the appendix.
From the QFD chart, it is clear that the accuracy of the model on stock price prediction is the most important factor. That is obvious because personal inventors will rely on

our tool to make revenues from the stock market, so they need a reliable prediction. If the model is not accurate, it will lead to wrong decisions and losing money as a result. However, since our target users are personal investors, accuracy is not the only consideration. The cost of the model is also critical. The QFD chat also reveals that fact, where the low computational resource consumption is the second most important factor. Because personal investors are sensitive to the cost compared with large institutions. They usually have small amount of investments, if the revenues earned won't cover the cost. They would not consider using the tool.

6 Project Plan

In this session, we present our preliminary plan for completing the project with more technical detail.

6.1 Engineering Analysis and Modeling Plan

6.1.1 Analysis Techniques

In our engineering analysis and modeling plan, we employ various analysis techniques to ensure a comprehensive understanding of the problem space. We start by selecting the problem space through a SWOT analysis, identifying the strengths, weaknesses, opportunities, and threats relevant to our project.

To determine the needs, users, and potential impacts, we conduct extensive research. This includes reviewing financial papers, AI research papers on stock predictions, reports from independent investment research companies, and articles from news and magazines. These sources provide a well-rounded perspective on the market and user requirements.

Next, we identify the models we will focus on, which will be detailed in subsequent sections. We collect data accordingly, utilizing reliable sources such as Nasdaq Data Link, Yahoo Finance, and Kaggle. These platforms provide trustworthy data that align with our problem space and chosen models.

Data preparation is a crucial technique in our process. We begin by splitting the data into training and testing sets with an 80/20 ratio. Following this, we engage in data cleaning, feature selection, and normalization to ensure the data is suitable for modeling. During the training phase, we use cross-validation to validate the models' performance comprehensively.

To select the optimal model for our final product, we perform a comparative analysis of several models. This technique allows us to choose the model that best meets our criteria. Model fine-tuning is then conducted using sensitivity analysis, ensuring the model performs optimally under various conditions.

Our performance evaluation metrics focus on achieving lower costs and higher efficiency while maintaining moderate accuracy. We emphasize high transparency so that non-technical users can easily understand the outcomes. Although customization is limited, we ensure frequent or real-time updates and alerts are available.

Finally, we use performance visualization techniques, such as graphs and plots, to present the results clearly and effectively. This visualization aids in communicating the findings to both technical and non-technical stakeholders.

6.1.2 Modelling Techniques

Through our research, we have identified several popular models commonly used for stock price predictions. Each model has its strengths and weaknesses. In the papers we reviewed, these models were either proposed or investigated for general application across all stocks, with the test cases mainly focusing on popular large-cap stocks. We aim to explore these models to determine their performance specifically on SME stocks. The models we will investigate include Support Vector Machine (SVM), Multilayer Perceptron (MLP), Naive Bayes, Random Forest, Logistic Regression, and Long Short-Term Memory (LSTM).

To assess these models, we will conduct pilot programs, testing each model with a selection of 10 stocks. These stocks are chosen to accurately represent our problem space. We will evaluate the performance of each model using our success metrics.

Based on the results of these evaluations, we will select the model that demonstrates the best performance. This model will be integrated into our final design, ensuring we utilize the most effective approach for stock price prediction.

6.1.3 Software Platforms

Our project leverages several software platforms to enhance efficiency and collaboration. Given our specialization in AI and Machine Learning, we will focus on optimizing the AI components, emphasizing backend development rather than the user interface. All of our group members have experience in Python. For these purposes, we will use Python for backend development and testing.

We will utilize GitHub for version control purposes, ensuring that all changes to our codebase are tracked and managed effectively. For project management, we will rely on Jira to coordinate tasks and monitor progress.

All documents, including deliverables, project research, meeting notes, and collected datasets, will be stored on Google Drive for easy access and sharing. We will use Zotero for literature management, facilitating collaboration on research sources. Finally, Overleaf will be employed for report formatting, enabling efficient and collaborative document preparation.

6.2 Prototyping Plan

6.2.1 Level of fidelity

In our prototyping plan, we've opted for a medium-fidelity approach. This decision is grounded in several key factors. Firstly, given our focus on AI and Machine Learning, a significant portion of our time will be dedicated to optimizing these components within the final product. Consequently, we've determined to allocate limited resources towards developing a highly user-friendly interface. Additionally, our design prioritizes meeting user needs over extensive user-friendliness. As such, iterative back-and-forth user testing isn't a requirement for our design process. Instead, we'll concentrate on backend testing to evaluate the efficiency, resource utilization, and accuracy of our design. Given our tight project timeline and small team size, this medium-fidelity approach strikes a balance between functionality and resource allocation. While we won't conduct extensive user research, our design aims to address core user needs by providing essential functionality for quick stock price prediction.

6.2.2 Designed Elements

We've chosen to utilize off-the-shelf designed elements for our prototype. This decision stems from the abundance of existing stock prediction algorithms across various markets and applications. Recognizing our team's limitations in AI and machine learning expertise, we've selected to focus on a niche problem space: small-cap stocks in the US market. Despite limited existing research in this area, we're leveraging established algorithms, adapting them to our specific use case, and evaluating their effectiveness. Our approach involves iterative model development, where we test alternatives, make evidence-based adjustments, and select the most successful algorithm to form the core of our tool. Success is gauged by how well the model addresses user needs, as outlined in the user needs section of our report.

6.2.3 Purpose of prototype

Our prototype serves two primary purposes. Firstly, it acts as a demonstration tool for stakeholders, showcasing our design concept and functionality. Secondly, it serves as a validation mechanism, confirming that our idea can indeed be transformed into a useful tool. We aim to gather user feedback to ensure that our prototype effectively addresses their needs. Time permitting, we'll incorporate this feedback to iteratively improve the functionality of our design.

6.2.4 Tools and technologies

Our project entails the design of a software tool, for which we'll utilize Python and essential AI/ML libraries. Time permitting, we aim to augment this tool with a straight-

forward user interface developed using TypeScript, React, and Ant Design.
In formulating our design requirements, we're employing various techniques such as whiteboarding, mind mapping, QFD chart analysis, and other methodologies outlined in this report.

6.2.5 Iterative Design Process

We will apply the Iterative Design Process. Since it is an AI software, we could test it as much as we can on our computer for the efficiency and accuracy of the core part, and modify our design if the result is not satisfying. This helps us to catch problems early.
Due to the time limits, users will have limited participation in testing. But we will still incorporate users, as we want to ensure our final product successfully addresses their needs. We want to have continual, incremental improvements.
Constant research will be in this project as we are new on implementing this kind of model, and we wish the design method could provide us with flexibility.
Resources available to our group: professor's and TA's feedback, we will consistently listen to the suggestions and incorporate them into our design.

6.2.6 Roles

With only three members in our small group, assigning specific roles becomes challenging. Consequently, each of us will shoulder multiple responsibilities.
Given our specialization is in AI and machine learning, our primary focus for this project will be on model development and refinement. We've adopted a collaborative approach wherein each team member selects an AI model for training. Subsequently, we convene to compare the merits and drawbacks of each model, ultimately deciding on the most suitable one to proceed with.
Beilin Ye, leveraging her background in finance, will spearhead the research efforts, ensuring we have the necessary background information to inform our project. Shinong Mao, with a software engineering background, will lead the development and design aspects, overseeing software functionality design, virtual environment setup, and coding tasks. Meanwhile, Zhiliu He will take on the roles of resource coordinator and additional researcher, facilitating efficient resource allocation and contributing to the research endeavors alongside Beilin. This collaborative and multifaceted approach ensures that each team member contributes meaningfully to the project's success, despite the absence of rigidly defined roles.

6.3 Testing Plan

6.3.1 Goal of the testing plan

The primary goal of our testing plan is to validate the feasibility and utility of our software tool. We aim to assess whether our design effectively meets user needs and gather

feedback to identify any areas for improvement. Time permitting, we will refine the functionality based on the results of these tests. These objectives will help us evaluate how well we have met our project goals and how effectively our prototype adheres to the established requirements and specifications.
Building upon the previously discussed user needs, we will measure the success of our prototype using a performance evaluation matrix rather than focusing solely on improving accuracy. This matrix will serve as our benchmark for assessing the performance of each design iteration, allowing us to make informed adjustments and enhance our prototype accordingly through rigorous testing.

6.3.2 Physical testing methods

To conduct our physical testing, we will use the following resources: 1) Software: Python, PyTorch, and other relevant AI/ML libraries. 2) Equipment: Development and testing computers with Python installed.
Determine Test Protocol: 1) Split the initial dataset into an 80/20 ratio, with 80% allocated for training and 20% reserved for testing purposes. 2) Set up a virtual environment with all necessary libraries installed. 3) Quantify the success matrix measurement to evaluate performance. 4) Establish the benchmark metrics we aim to surpass.
Perform Test: 1) Run the algorithms on the testing data. 2) Interpret the results to assess the performance and identify areas for improvement.

6.3.3 User Testing methods

In a user-centered design process, we should have involved users extensively. However, due to time constraints, we will limit user participation in our project. Instead of holding participatory design sessions and coding all their input, we will focus on the key feedback they provide about our prototype and whether it successfully addresses their needs.
Steps: 1) Recruit Participants: Identify and recruit participants from our primary user group, as previously defined. Secondary users, such as institutional investors, will not be involved due to limited access. 2) Schedule User Testing: Arrange sessions for user testing. 3) Perform User Testing: Gather feedback on the functionality of our tool, listening to users' impressions and suggestions. 4) Interpret Feedback: Analyze the feedback and translate it into updated design requirements. 5) Update Our Tool: Implement necessary updates to our tool based on the interpreted feedback.

7 Declaration

Used the generative AI only for the purpose of correcting grammar mistakes.

References

[1] Abdulaziz M Abdulsaleh and Andrew C Worthington. Small and medium-sized enterprises financing: A review of literature. *International Journal of Business and Management*, 8(14):36, 2013. pages 3

[2] World Bank. Small and medium enterprises (smes) finance, 2020. Accessed: 2024-06-07. pages 8

[3] Md Masum Billah, Azmery Sultana, Farzana Bhuiyan, and Mohammed Golam Kaosar. Stock price prediction: comparison of different moving average techniques using deep learning model. *Neural Computing and Applications*, pages 1–11, 2024. pages 5

[4] Yosef Bonaparte. Artificial intelligence in finance: Valuations and opportunities. *Finance Research Letters*, 60:104851, 2024. pages 6

[5] Fatima Dakalbab, Manar Abu Talib, Qassim Nassir, and Tracy Ishak. Artificial intelligence techniques in financial trading: A systematic literature review. *Journal of King Saud University-Computer and Information Sciences*, page 102015, 2024. pages 3

[6] Nabanita Das, Bikash Sadhukhan, Susmit Shekhar Bhakta, and Satyajit Chakrabarti. Integrating eemd and ensemble cnn with x (twitter) sentiment for enhanced stock price predictions. *Social Network Analysis and Mining*, 14(1):1–18, 2024. pages 5

[7] Thomas Fischer and Christopher Krauss. Deep learning with long short-term memory networks for financial market predictions. *European journal of operational research*, 270(2):654–669, 2018. pages 4

[8] The Motley Fool. When to buy small-cap stocks, 2024. Accessed: 2024-06-07. pages 4

[9] IEEE. Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems, 2020. Accessed: 2024-06-07. pages 8

[10] J. Jiang, E. Chen, X. Yan, and Y. Feng. Applications of artificial intelligence in stock market prediction: A survey. *Journal of Big Data*, 7(1):1–30, 2020. pages 4

[11] Manrui Jiang, Wei Chen, Huilin Xu, and Yanxin Liu. A novel interval dual convolutional neural network method for interval-valued stock price prediction. *Pattern Recognition*, 145:109920, 2024. pages 5

[12] David-Yue Lin, Satya N. Rayavarapu, Karim Tadjeddine, and Rajen Yeoh. Beyond financials: Helping small and medium-sized enterprises thrive. `https://www.mckinsey.com/industries/public-sector/our-insights/beyond-financials-helping-small-and-medium-size-enterprises-thrive`, January 26 2022. pages 3

[13] Wei Liu, Yoshihisa Suzuki, and Shuyi Du. Forecasting the stock price of listed innovative smes using machine learning methods based on bayesian optimization: Evidence from china. *Computational Economics*, pages 1–34, 2023. pages 6

[14] Wen Long, Jing Gao, Kehan Bai, and Zhichen Lu. A hybrid model for stock price prediction based on multi-view heterogeneous data. *Financial Innovation*, 10(1):48, 2024. pages 5

[15] Minrong Lu and Xuerong Xu. Trnn: An efficient time-series recurrent neural network for stock price prediction. *Information Sciences*, 657:119951, 2024. pages 5

[16] MorningStar. Where we see opportunities as june stocks recover losses, 2023. Accessed: 2024-06-07. pages 4

[17] MorningStar. 2024 outlook: Stock market economy, 2024. Accessed: 2024-06-07. pages 4

[18] MSCI. Small caps have been a big opportunity, 2021. Accessed: 2024-06-07. pages 4

[19] Nguyen Kim Quoc Trung. Determinants of small and medium-sized enterprises performance: The evidence from vietnam. *Cogent Business & Management*, 8(1):1984626, 2021. pages 6

[20] Jujie Wang, Junjie He, Chunchen Feng, Liu Feng, and Yang Li. Stock index prediction and uncertainty analysis using multi-scale nonlinear ensemble paradigm of optimal feature extraction, two-stage deep learning and gaussian process regression. *Applied Soft Computing*, 113:107898, 2021. pages 4

[21] Jujie Wang and Jing Liu. Two-stage deep ensemble paradigm based on optimal multi-scale decomposition and multi-factor analysis for stock price prediction. *Cognitive Computation*, 16(1):243–264, 2024. pages 5

[22] Jun Zhang and Xuedong Chen. A two-stage model for stock price prediction based on variational mode decomposition and ensemble machine learning method. *Soft Computing*, 28(3):2385–2408, 2024. pages 5

Appendix

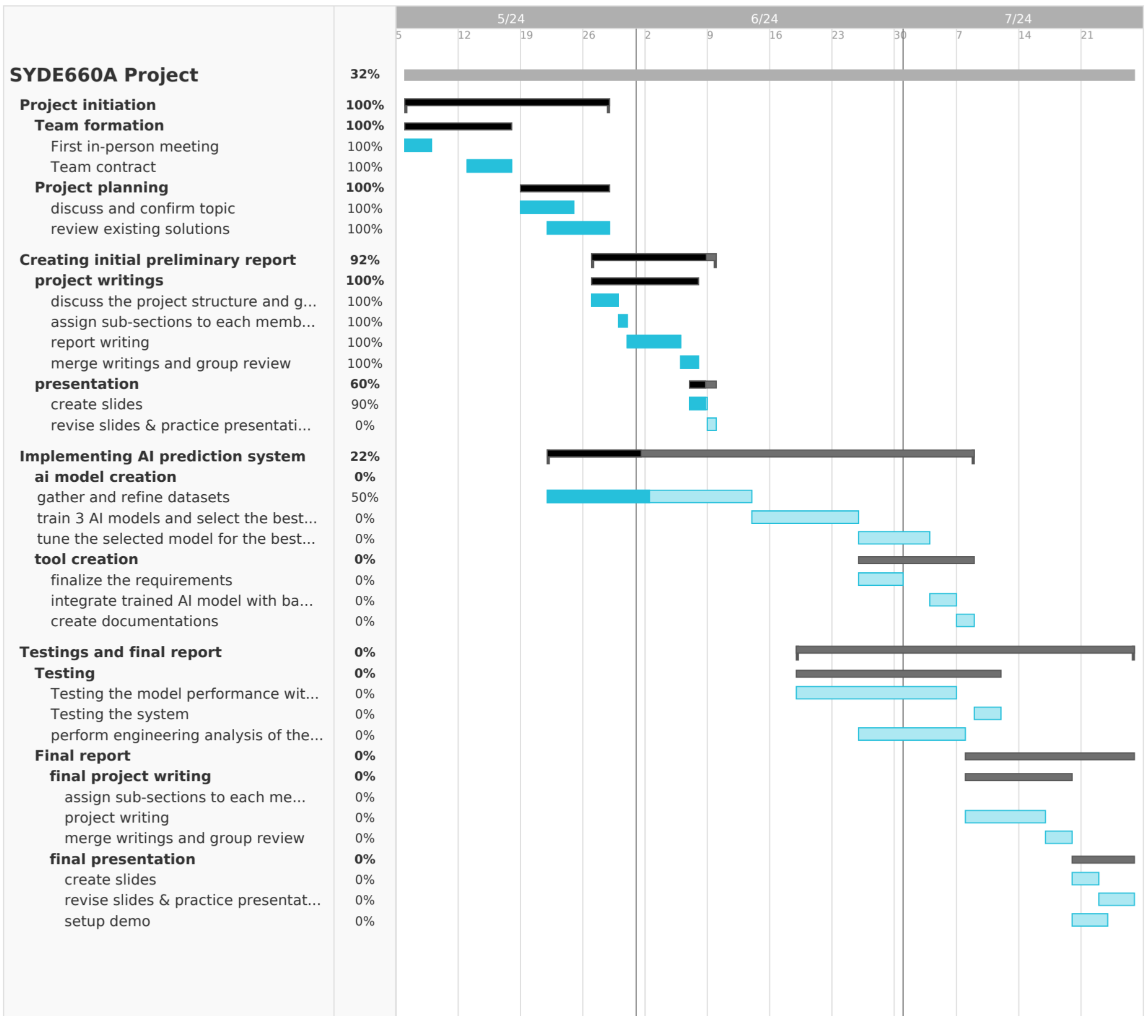
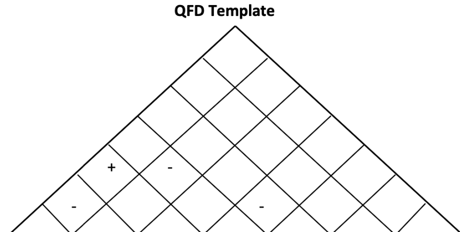


Figure 1: GANTT chart

QFD Template



User Needs/Functions	Importance (What)	Engineering Requirements							Competitive Products		
		Should be easy to use	Should not be customizable	Should give prediction result rapidly	Should consume low computation resources	Should output accurate price prediction			Existing AI for stock price prediction	Financial investment analyst	
Know the future trend of the stock price	15			1		9			4	3	
Making revenues from stock markets	35			3	1	9			2	3	
Support personal investors to make decision	20	1		3		9			3	5	
Required few technical knowledge	10	9	6						4	5	
Affordable	20				9				2	1	
Unit		likert	likert	second	CDN	%					
Importance (HOW)		110	60	180	215	630	1195				
Importance (%)		9	5	15	18	53					
Minimum Threshold		1	1	1	1	60					
Desired Target		8	6	10	10	90					
Maximum Threshold		10	10	600	50	99					

Figure 2: QFD chart