TRIBHUVAN UNIVERSITY

INSTITUTE OF SCIENCE AND TECHNOLOGY

Project Proposal

on

NEPSE STOCK FORECAST

Submitted To

Department of Computer Science and Information Technology

Soch College Of IT

In partial fulfillment of the requirement for the Bachelor Degree in Computer Science and Information Technology

Submitted By

Kushal Subedi (25085/076)

Niruta Bhattarai (25087/076)

Rachana Regmi (25089/076)

October 16,2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abbreviation

ML	**M**achine **L**earning
AI	**A**rtificial **I**ntelligence
NN	**N**eural **N**etwork
NEPSE	**N**epal **S**tock **E**xchange
API	**A**pplication **P**rogramming **I**nterface
LSTM	**L**ong **S**hort **T**erm **M**emory
RNN	**R**ecurrent **N**eural **N**etwork
SVM	**S**upport **V**ector **M**achine
NLP	**N**atural **L**anguage **P**rocessing
ANN	**A**rtificial **N**eural **N**etwork
CNN	**C**onvolutional **N**eural **N**etwork

CHAPTER 1
INTRODUCTION

1.1 Overview

In today's dynamic financial landscape, capital markets offer a compelling avenue for individuals to actively engage in economic activities, particularly through investments. Among the various investment options available, stocks stand out as a prominent choice. These securities represent ownership in companies, and the returns realised by stockholders are intricately tied to the financial performance of these entities. It's a simple equation: when companies thrive, so do their stockholders.[1]

Yet, the promise of significant profits in the stock market also comes with greater investment risk. This makes it increasingly important to predict future stock prices based on past data. This proposal aims to use the capabilities of machine learning to tackle this critical issue.[1]

The price movements of stocks are not isolated events; they are deeply intertwined with the broader economic landscape. Factors such as monetary policy, influenced by variables like the amount of money in circulation and interest rates, as well as fiscal policy, including taxation policies, play pivotal roles in shaping the behaviour of stock investors.[2]

In the not-so-distant past, various methodologies such as linear regression, time-series analysis, and even chaos theory have been applied to predict stock prices. Yet, with the advent of machine learning, we stand at the threshold of a transformative era in stock price prediction. This proposal explores how cutting-edge machine learning algorithms and techniques can revolutionise our ability to forecast stock prices accurately, empowering investors with invaluable insights and enabling them to make informed decisions in an ever-evolving financial ecosystem.

1.2 Problem Statement

The challenge at hand is to develop an accurate and reliable method for predicting future stock prices in a volatile financial market. The allure of potentially high returns in the stock market is counterbalanced by the inherent risks, making it imperative to find a robust solution that leverages historical data and machine learning techniques to make informed predictions.

The primary challenge is the development of an accurate and precise stock price prediction model that can effectively capture the complex interplay of market factors, including economic indicators, geopolitical events, corporate earnings, and investor sentiment.[3]

Moreover, it's widely recognized that managing risk plays a crucial role in investing.[2] With the stock market being so unpredictable, investors need tools to understand and reduce risks effectively. This involves thorough testing and refining the model to make it more accurate and reliable. Dealing with real-time predictions means considering factors like data delays, improving the model continuously, and strategies for using the latest information. Also, it's important to provide user-friendly charts and reports to help investors understand and act on the predictions. To follow the rules and regulations, we make sure we're in line with financial laws. In the end, our approach aims to help investors make well-informed decisions and manage risks in the ever-changing world of financial markets.

1.3 Aim and Objective

The primary goal is to provide investors and financial experts with reliable tools and strategies that improve decision-making, mitigate risks, adjust to market conditions, and offer trustworthy predictive models for stock investments, all while recognizing the inherent unpredictability in financial markets.

1.4 Scope and Limitations

1.4.1 Scope

1. **Web Based platform**:Our application will be accessible via the internet, providing a web-based platform.

2. **Educational Component**: In addition to prediction, the application may offer educational resources and insights to help users understand the factors affecting stock prices and improve their investment knowledge.

3. **Advanced Prediction Models**: Continuous improvement and integration of advanced prediction models to enhance the accuracy of stock price forecasts.

1.4.2 Limitations

1. **No Installable Application**:Since the application is not designed as a mobile app, it may not provide the convenience and real-time accessibility that some users prefer on their smartphones.

2. **Limited Data Source**: The application's data source is restricted to NEPSE. It does not consider global or external factors that might impact stock prices, limiting the comprehensiveness of its predictions.

3. **Accuracy**: The accuracy of the proposed system is highly dependent on that of algorithm used and Quality of Data.

4. **Market Volatility** : The stock market is subject to high volatility, and the application may not be able to account for extreme market conditions or sudden economic changes that can significantly affect stock prices.

5. **Historical Data Dependency**:The accuracy of predictions heavily relies on historical data. If there are data gaps or inaccuracies in the historical NEPSE data, it may impact the reliability of the predictions.

1.5 Motivation

Stock price prediction is a longstanding and crucial problem. Crafting a dependable model for predicting stock prices can provide valuable insights into how the market behaves over extended periods, helping us identify trends that might otherwise go unnoticed. With the increasing computing power of today's machines, machine learning emerges as an effective approach to address this challenge. Consequently, our inspiration arises from the aspiration to establish a

public service that harnesses historical data and user predictions to construct a more resilient model, ultimately benefiting a broader audience.

CHAPTER 2
BACKGROUND STUDY AND LITERATURE REVIEW

2.1 Background Study

Stock price prediction has been a longstanding challenge in financial markets. Accurate predictions are essential for investors, traders, and financial institutions to make informed decisions. Machine learning has gained significant attention in recent years as a promising approach for improving the accuracy of stock price predictions. This literature review discusses key studies in this domain[4].

2.1.1 Early Attempts at Predicting Stock Prices

In the early days of trying to predict stock prices, people mainly used traditional math and statistics. They looked at stock charts, used a method called time series analysis, and some math tools like moving averages. Then, in the 1960s and 70s, a person named Benoit Mandelbrot came up with the idea of using something called ”fractal analysis” to understand how stock prices move. But it's important to know that these early methods had problems because they couldn't really figure out the complicated and always-changing nature of the stock market.[5]

2.1.2 Machine Learning In Stock Forecasting

Machine learning algorithms have brought a fresh perspective to research on predicting stock prices. Studies have demonstrated several methods for forecasting stock prices as a result Some of them are below.

1. **Neural Networks**:Neural networks have gained popularity in recent years for various machine learning tasks, including stock price prediction [6]. Some neural network algorithms, such as Recurrent Neural Networks (RNNs) and Artificial Neural Networks (ANNs), are commonly practiced for predicting stock prices.

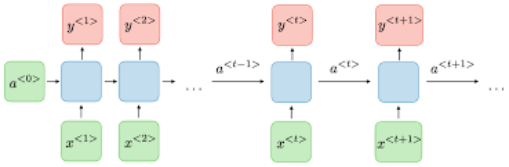

Figure 2.1: Recurrent Neural Network

2. **Random Forest Algorithm**: Random Forest is a machine learning algorithm introduced by Leo Breiman in 2001 [7]. It is an ensemble learning method that combines the predictions of multiple decision trees to improve accuracy and reduce overfitting.

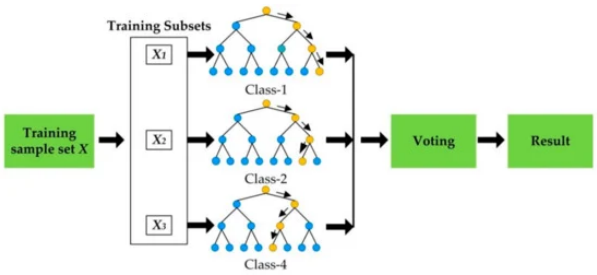

Figure 2.2: Random Forest

3. **SVM**:In the initial phases of applying machine learning to predict stock prices, the Support Vector Machine (SVM) was utilized. It employed the Support Vector Regression algorithm [8]

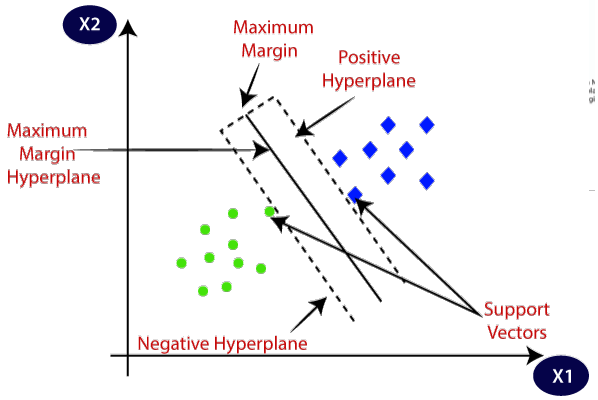

Figure 2.3: Support Vector Model

2.2 Literature Review

Literature survey is the process in which a complete and comprehensive review is conducted encompassing both the published and unpublished work from other alternative sources of information. This review is conducted in the domains of specific interest to the person or researcher. Further, the results of this process are documented.

" Accurately predicting the change of stock price can reduce the investment risk of stock investors and effectively improve the investment return". [9] Because of the stock market's inherent volatility, predicting stock prices typically involves complex nonlinear time series forecasting. Stock prices are influenced by a multitude of variables, making accurate prediction challenging with simplistic models.

Selecting a model for stock price prediction has become more accessible, despite the dynamic behavior of the stock market. Recurrent Neural Networks (RNN) excel in this context, primarily due to their ability to retain and utilize prior states in current state calculations. Unlike independent layers, the hidden layers in RNN are interconnected, incorporating not just input layer outputs but also outputs from preceding hidden layers. This characteristic empowers RNN to perform well with sequential data. RNN's advantage lies in its capacity to consider the contextual information within the data during training, making it particularly suitable for stock and Forex scenarios. This is because fluctuations at a specific time often bear a connection to the previous trends. [10]

Another Survey on *"Chaos theory"*, a branch of mathematics and physics, has found intriguing applications in the field of finance, offering a fresh perspective on understanding and predicting market dynamics. This theory posits that complex systems, such as financial markets, are inherently sensitive to initial conditions and can exhibit chaotic, unpredictable behavior, even though they may appear random at first glance. [11]

One of the fundamental concepts of chaos theory is the "butterfly effect." It suggests that a

small event, like the flapping of a butterfly's wings, can trigger a cascade of effects that ultimately lead to major, seemingly unrelated events. In the context of finance, this means that minor changes or perturbations in market conditions can have far-reaching consequences. Chaos theory's recognition of unpredictability and sudden shifts in financial markets underscores the importance of robust risk management strategies. By acknowledging the potential for chaotic behavior, investors and institutions can better prepare for unforeseen market events.

The stock market is significantly impacted by the behavior and trends on social media. Conducting sentiment analysis on social media content can provide valuable insights into trader sentiment and market trends. This, in turn, aids in risk management by allowing for proactive analysis of real-time news and market behavior. [12]

Long Short-Term Memory (LSTM) networks have brought a revolutionary shift in the landscape of stock analysis. LSTM networks stand out as the best choice in stock analysis due to their versatility in capturing complex temporal patterns, handling non-linearity, and automatically learning significant features within financial time series data. They empower traders and analysts with an advanced tool for understanding market dynamics and making informed decisions. As research in this field continues to grow, the significance of LSTMs in stock analysis becomes increasingly evident, and their influence in the financial world is set to expand, providing traders with a powerful ally in navigating the intricacies of the stock market. [13]

2.2.1 Ongoing Research For Stock Price Prediction

In addition to Recurrent Neural Networks (RNNs), Artificial Neural Networks (ANNs), and linear regression models, ongoing research in stock price prediction includes investigating areas such as sentiment analysis of the market, employing Natural Language Processing (NLP) for analyzing social media data [14], utilizing high-frequency trading techniques, and applying Convolutional Neural Networks (CNNs), Quantum Computing[15] to gain insights into and make predictions about the stock market.

2.3 Study Of Existing Application

2.3.1 Chukul NEPSE

Chukul is an AI-based algorithm Application that aims to predict the trend of Nepal stock market. It is a tool that facilities a trader to make decisions without having to do any technical and fundamental research. Chukul provides a novice user a safe landing in this unpredictable environment of rumor filled share market. It saves a lot of analysis time regarding the rise and fall of the share value[16].

2.3.2 Nepal Stock Exchange

Nepal Stock Exchange is a mobile Application and web based application that provides updates of daily ups and down in NEPSE.The app provides real-time access to market data, news, and insights pertinent to the Nepal Stock Exchange. The app also provide Comprehensive Analysis of NEPSE Data and news related to Nepal Share Market. [17]

CHAPTER 3
DEVELOPMENT METHODOLOGY

3.1 System Development Model

3.1.1 Agile Model

The Agile model is an iterative and incremental approach to software development. It emphasizes collaboration, flexibility, and customer feedback throughout the development process. **Why we choose Agile ?**:

1. **User-Centric Approach**: Agile places a strong emphasis on customer feedback and involvement

2. **Flexibility**: Agile allows for changes and adaptations as the project progresses. This flexibility is valuable for a stock prediction application because financial markets are dynamic and frequently updating.

3. **Research and Development**: Agile encourages a culture of continuous improvement. Which encourages our team to regularly reflect on their processes and seek ways to enhance efficiency and with continious improvement approach.[18]

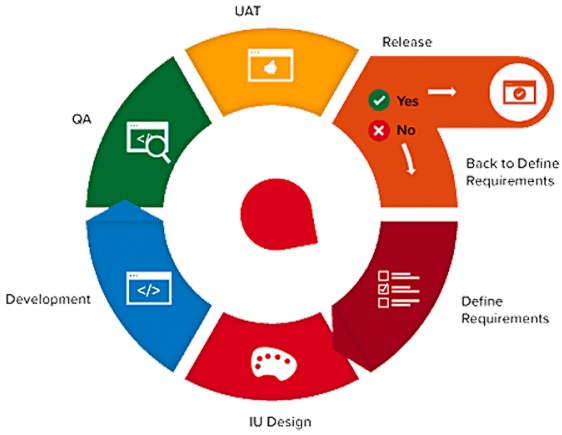

Figure 3.1: Agile Development

3.2 Requirement Engineering

Based on our background study and literature review (refer to Chapter 2), we've identified several projects similar to ours that employ different methodologies and algorithms. Although our project shares similarities with these endeavors, it's essential to clarify that our primary objective does not revolve around the development of a commercial application. Instead, due to time constraints, we will not be incorporating multiple algorithms.

As a result of this, we have compiled the following set of requirements for our Stock Price Prediction Application.

3.2.1 Functional Requirements

Functional requirements specify the specific functions, features, and capabilities a system or software must have to meet user needs and perform its intended tasks.

1. **Data Collection and Integration:**
    - The system should gather historical stock price data from NEPSE .
2. **Data Prepossing:**
    - The system should clean and preprocess raw data, handling missing values and outliers.
    - It should normalize or scale data appropriately for modeling.
3. **Prediction Models:**
    - The system should implement LSTM model for prediction .
4. **Training and Testing:**
    - The system should split data into training and testing datasets.
    - It should provide options for cross-validation and model evaluation.
5. **Visualization:**
    - The system should offer interactive charts and visualizations of historical and predicted stock prices.

- Users should be able to customize the time range and parameters displayed.

3.2.2 Non-Functional Requirements :

1. **Accuracy and Precision:**
    - The system should achieve a specified level of accuracy in stock price predictions.
    - Precision requirements should be defined to minimize false positives in alerting.
2. **Scalability:**
    - The system should handle a growing volume of data and users without performance degradation.
    - It should support parallel processing or distributed computing for scalability.
3. **Reliability:**
    - The system should be available and reliable 24/7 to support real-time trading decisions.
    - Redundancy and failover mechanisms should be in place to minimize downtime.
4. **Response Time:**
    - The system should provide real-time or near-real-time predictions and alerts.
    - Response time requirements should be specified for different functions.
5. **User Interface:**
    - The user interface should be intuitive and user-friendly, catering to both novice and experienced users.
    - Accessibility and usability guidelines should be followed.

3.3 Feasibility Study

A feasibility study is a comprehensive analysis conducted to assess the practicality and viability of a proposed project, system, product, or business idea. It serves as a critical tool for decision-making, helping stakeholders determine whether a project should be pursued, modified, or abandoned. The primary objectives of a feasibility study are to evaluate the project's potential for success, identify potential challenges, and provide a basis for informed decision-making.

3.3.1 Technical Feasibility

In the technical aspect of our system, which relies on algorithmic and machine learning models, we will utilize **Python** as our primary programming language. We'll leverage essential libraries such as **Pandas, Numpy, Keras, and scikit-learn** to support our core functionalities. As the system grows and evolves, there may be a need to incorporate additional programming languages like JavaScript, and markup languages like HTML and CSS to facilitate further expansion and enhance the user interface.

3.3.2 Operational Feasibility

For our stock prediction system, operational feasibility means making sure we have the right teammates who understand Python, machine learning, and web technologies. We also need to check that our computers and internet systems can handle the work. We must make sure our data is managed well, and we follow the rules about data privacy. We'll need plans to keep the system running, and handle any future growth. If we can do all this, our stock prediction system will work smoothly and give us the information we need.

3.3.3 Economical Feasibility

The primary expenses associated with developing the system revolve around ensuring that we have adequately equipped devices and a reliable internet connection. Our economic viability also encompasses costs related to travel, wages for staff during system launch, as well as expenses for meals. Additionally, it includes spending on online

courses and books for research and study purposes. Deployment costs are another aspect of our economic feasibility plan, along with potential expenses for accessing third-party system APIs.

Break-Even Analysis

Cost/Revenue Item	**Amount (NRs:)**
Initial Investment	5,000
Internet and Device Costs	25,000
Travel Expenses	4,500
Staff Wages (Launch Phase)	15,000
Research Materials (Courses, Books)	7,500
Deployment Costs	3,000
Third-Party API Expenses	1,200
Total Costs	61,200
Projected Sales/Income	20,000
Break-Even Point (BEP)	41,200

3.4 Use Case Diagram

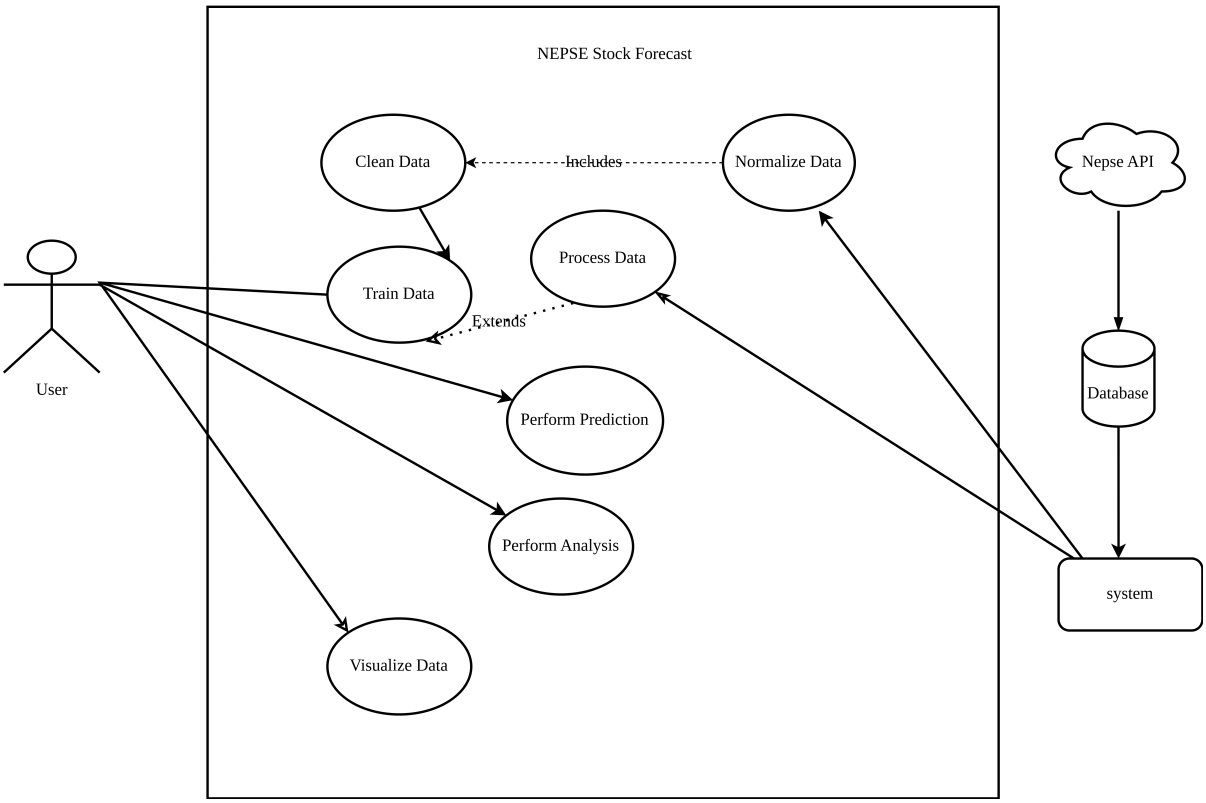

Figure 3.2: Usecase Diagram

3.5 Gantt Chart

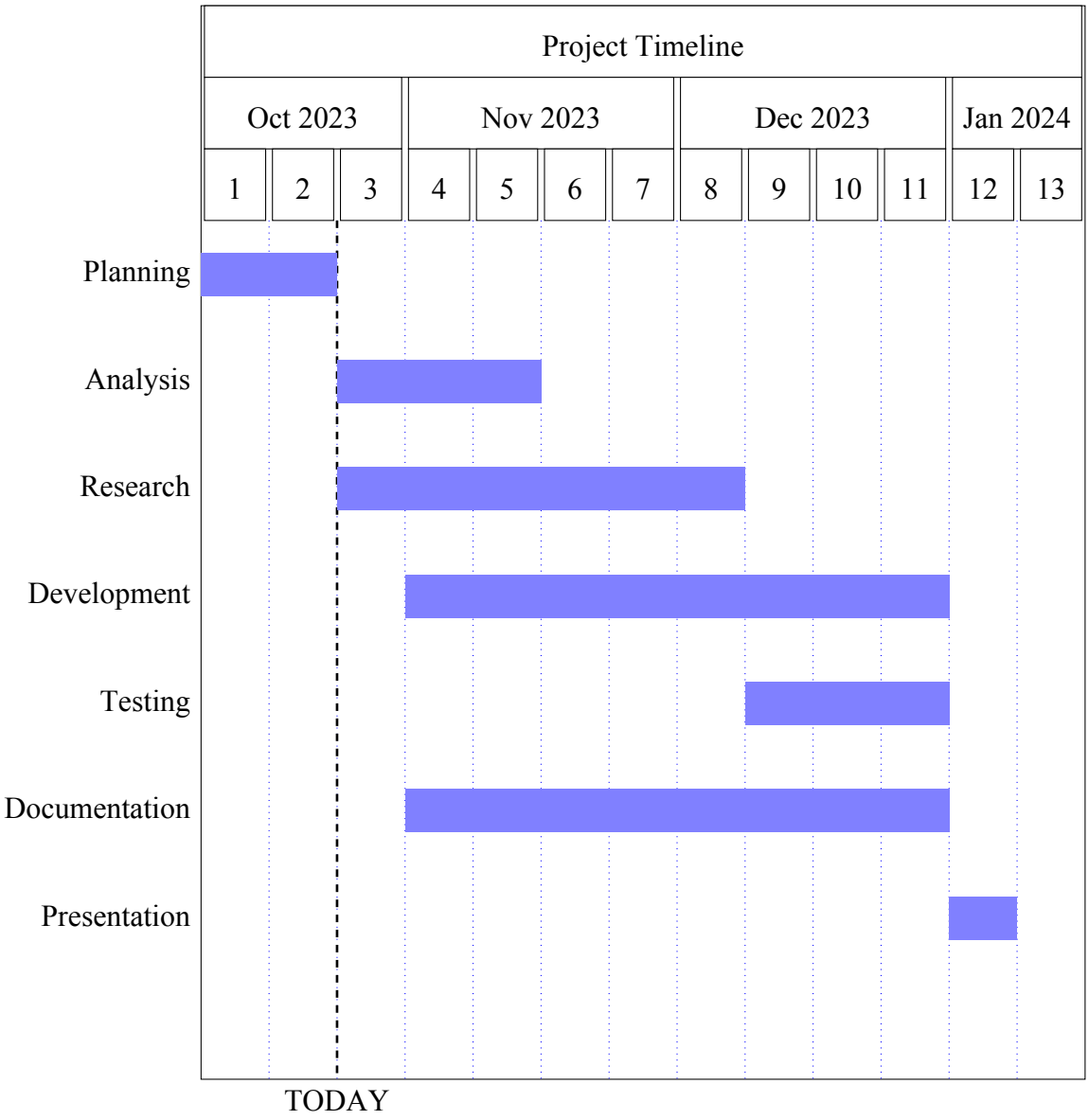

Figure 3.3: Gantt-Chart

CHAPTER 4
SYSTEM ANALYSIS AND DESIGN

This chapter gives an overview of the proposed system along with defining the system architecture and UML diagrams such as use-case and class diagrams. A UML diagram is a diagram based on the UML (Unified Modelling Language) with the purpose of visually representing a system along with its main actors, roles, actions, artifacts or classes, in order to better understand, alter, maintain, or document information about the system.

4.1 Overview of System Design

We are implementing a system that will create a predictive model using a custom training algorithm. Users can make decisions to buy or sell stocks based on the model's predictions. The system will calculate profit or loss based on the closing price of a stock for the day, which we consider as the target variable.

Forecasting stock market behavior can seem complicated because there are many unknown factors that don't seem to follow a straightforward pattern. However, by using machine learning effectively, we can teach computers to recognize trends and make informed guesses by looking at past and current data.

4.2 System Flow Diagram

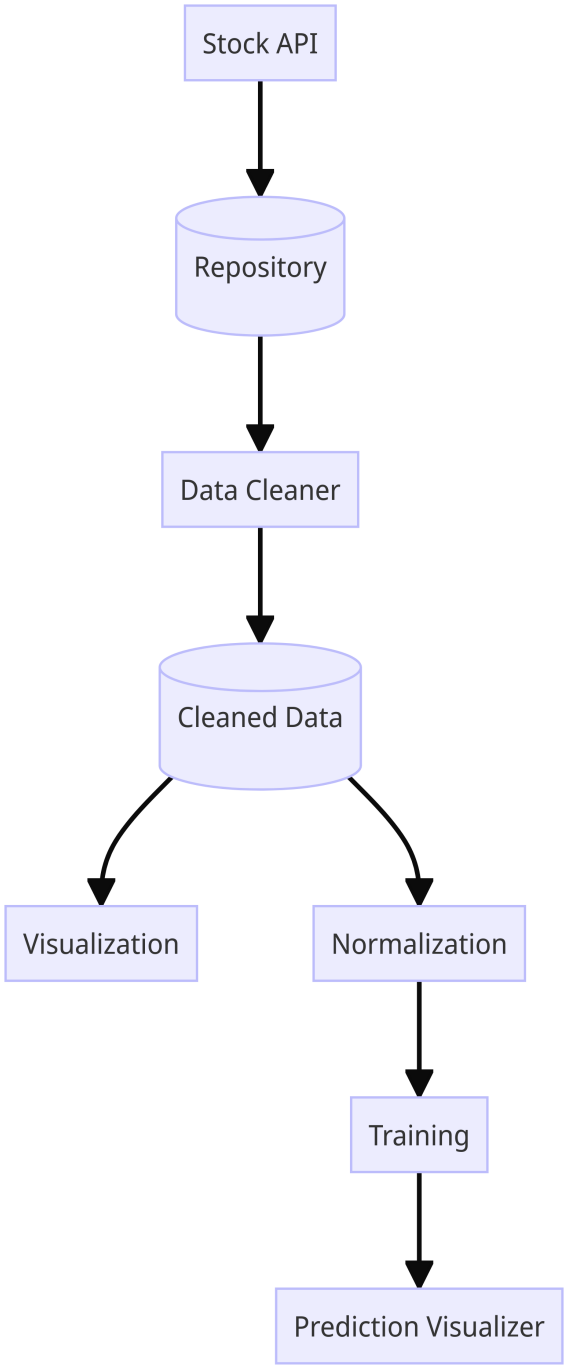

Figure 4.1: System Flow Diagram

4.3 Activity Diagram

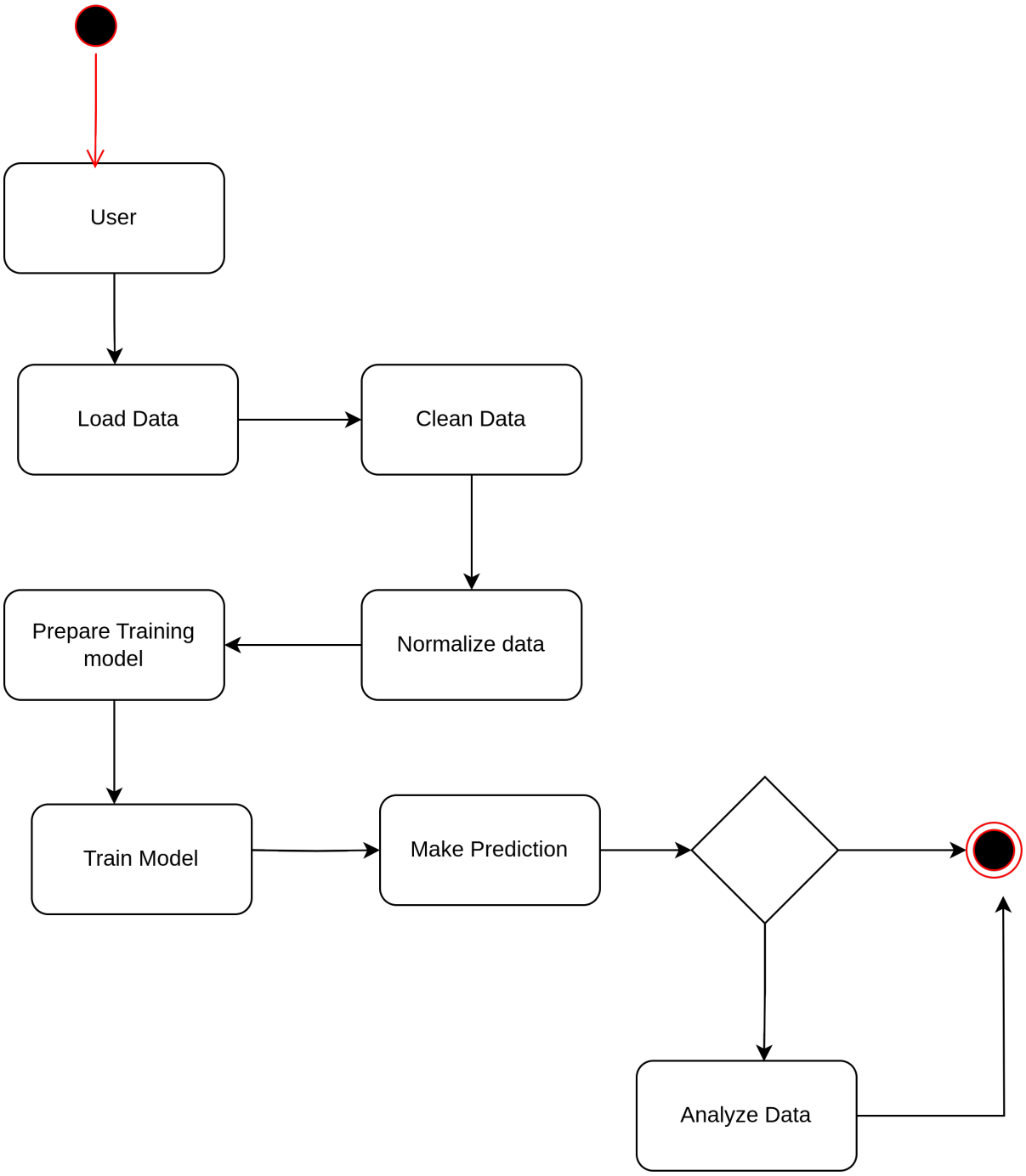

Figure 4.2: Activity Diagram

4.4 Class Diagram

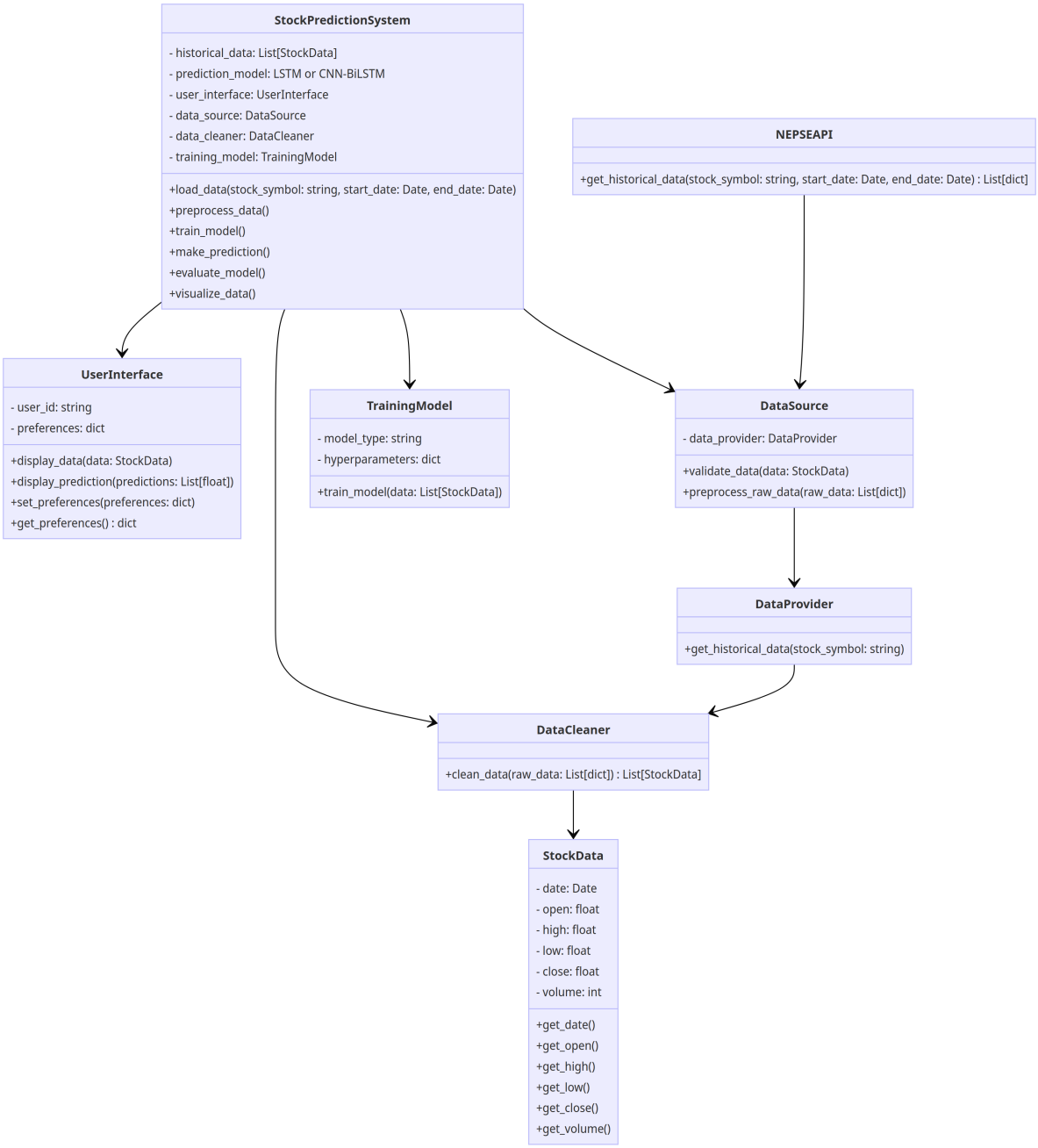

Figure 4.3: Class Diagram

CHAPTER 5
IMPLEMENTATION PLAN

In this chapter we discuss about the implementation of our system and tools to be used

5.1 Algorithm

5.1.1 LSTM Algortihm

LSTM stands for long short-term memory networks, used in the field of Deep Learning. It is a variety of recurrent neural networks (RNNs) that are capable of learning long-term dependencies, especially in sequence prediction problems. LSTM has feedback connections, i.e., it is capable of processing the entire sequence of data, apart from single data points such as images. This finds application in speech recognition, machine translation,Stock Price Prediction etc. LSTM is a special kind of RNN, which shows outstanding performance on a large variety of problems[19] .

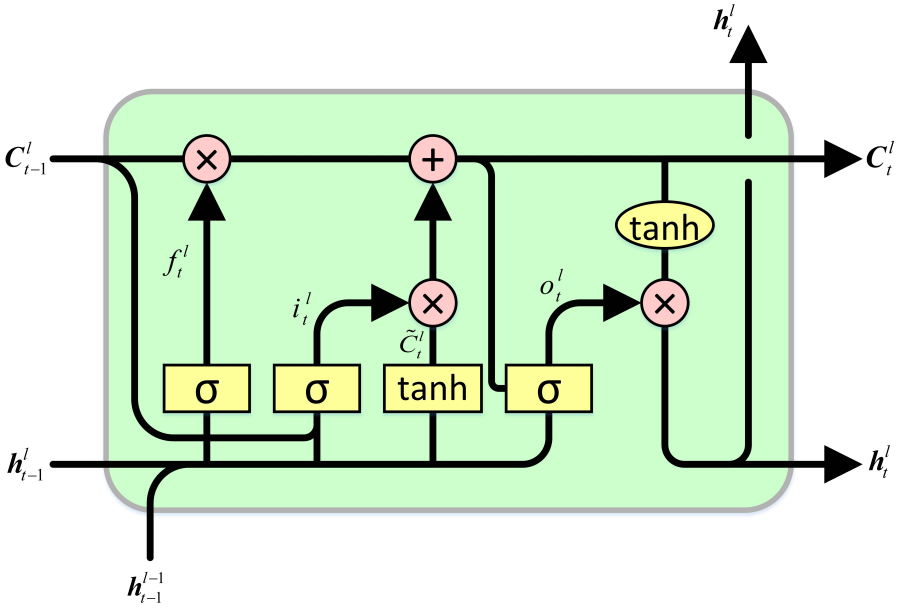

Figure 5.1: LSTM Architecture

5.2 Tools to be used

5.2.1 Languages and Libraries

1. **Python**: Python is a High level general purpose programming language which is widely used for Data Science , ML and AI related works.

2. **Pandas**: Pandas is an open-source data manipulation and analysis library for the Python programming language. It provides easy-to-use data structures and functions for working with structured data, such as tables, spreadsheets, and time series data.

3. **Numpy**: NumPy, which stands for Numerical Python, is a fundamental Python library for numerical and mathematical operations. It provides support for large, multi-dimensional arrays and matrices, along with a collection of mathematical functions to operate on these arrays efficiently.

4. **Django**: Django is a high-level open-source web framework for building robust, maintainable, and secure web applications using Python.

5.2.2 Tools and Softwares

1. **Jupyter Notebook**: Jupyter Notebook is an open-source web application that allows us to create and share documents that contain live code, equations, visualizations, and narrative text. It is widely used for interactive computing, data analysis, machine learning, and scientific research. Key features of Jupyter Notebook

2. **Vs Code**: Visual Studio Code (VS Code) is a popular, free, open-source code editor developed by Microsoft. It's widely used by developers for a variety of programming languages and technologies

3. **git**: Git is a popular open-source tool used for version controlling and project tracking.

5.2.3 Hardware and System Requirements

1. i5+ or equivalent processor laptop or desktop.

2. 4GB+ Ram and Dedicated Graphics Card(optional).

3. Linux Mac or Windows7+ operating system.

5.3 Expected Outcome

The expected outcome of this project is to develop a user-friendly Stock Forecasting system that equips investors with accurate stock predictions and risk management tools. Through extensive research and the integration of the latest trends in finance and stock prediction, this system aims to provide users with the means to make informed investment decisions. Ultimately, it will enable users to potentially enhance their profitability and reduce financial risks while navigating the dynamic world of stock trading.

CHAPTER 6
CONCLUSION

In conclusion, the development of a stock price prediction system utilizing a machine learning model, specifically LSTM, holds great promise in enhancing the financial decision-making process. With the expected outcome of this system, investors and analysts will have a robust tool at their disposal, offering accurate stock predictions and risk management capabilities. The extensive research and incorporation of the latest industry trends underscore the commitment to providing users with valuable insights into the stock market's complexities. By empowering users to make informed investment decisions, this system aims to not only improve profitability but also mitigate financial risks. In an ever-evolving financial landscape, such a system can be a valuable asset for those seeking to navigate the stock market with greater confidence and effectiveness.

BIBLIOGRAPHY

[1] F. Momin, S. Patel, K. Shide, and P. A. C. Taskar, "Stock market prediction system using machine learning approach," 2019, October 2019.

[2] J. L. Ticknor, "A bayesian regularized artificial neural network for stock market forecasting," vol. 40, no. 14, 2013, 15 October 2013.

[3] W. A. Mohammed, "Challenges of stock prediction," in *Valuation Challenges and Solutions in Contemporary Businesses*, IGI Global, 2020, pp. 234–252.

[4] J. Chen, Y. Wen, Y. Nanehkaran, M. Suzauddola, W. Chen, and D. Zhang, "Machine learning techniques for stock price prediction and graphic signal recognition," *Engineering Applications of Artificial Intelligence*, vol. 121, p. 106 038, 2023, ISSN: 0952-1976. DOI: https://doi.org/10.1016/j.engappai.2023.106038. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0952197623002221.

[5] B. Mandelbrot, "The variation of certain speculative prices," *Journal of Business*, vol. 36, no. 4, pp. 394–419, 1963.

[6] D. Zhang and S. Lou, "The application research of neural network and bp algorithm in stock price pattern classification and prediction," *Future Generation Computer Systems*, vol. 115, pp. 872–879, 2021, ISSN: 0167-739X. DOI: https://doi.org/10.1016/j.future.2020.10.009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167739X20329861.

[7] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[8] M. Beniwal, A. Singh, and N. Kumar, "Forecasting long-term stock prices of global indices: A forward-validating genetic algorithm optimization approach for support vector regression," *Applied Soft Computing*, vol. 145, p. 110 566, 2023, ISSN: 1568-4946. DOI: https://doi.org/10.1016/j.asoc.2023.110566. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494623005847.

[9] W. Lu, J. Li, J. Wang, and et al., "A cnn-bilstm-am method for stock price prediction," *Neural Computing & Applications*, vol. 33, pp. 4741–4753, 2021. DOI: 10.1007/s00521-020-05532-z. [Online]. Available: https://doi.org/10.1007/s00521-020-05532-z.

[10] Z. Hu, Y. Zhao, and M. Khushi, "A survey of forex and stock price prediction using deep learning," *Applied System Innovation*, vol. 4, no. 1, 2021, ISSN: 2571-5577. DOI: 10.3390/asi4010009. [Online]. Available: https://www.mdpi.com/2571-5577/4/1/9.

[11] I. Klioutchnikov, M. Sigova, and N. Beizerov, "Chaos theory in finance," *Procedia Computer Science*, vol. 119, pp. 368–375, 2017, 6th International Young Scientist Conference on Computational Science, YSC 2017, 01-03 November 2017, Kotka, Finland, ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2017.11.196. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050917324067.

[12] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, 2011. DOI: 10.1016/j.jocs.2010.12.007.

[13] G. Zhang, W. Xie, and Y. Zhu, "Stock price prediction with time lags using a hybrid deep learning framework," *Applied Soft Computing*, vol. 70, pp. 525–533, 2018. DOI: 10.1016/j.asoc.2018.05.042.

[14] M. Hagenau, M. Liebmann, M. Hedwig, and D. Neumann, "Automated news reading: Stock price prediction based on financial news using context-specific features," in *2012 45th Hawaii International Conference on System Sciences*, 2012, pp. 1040–1049. DOI: 10.1109/HICSS.2012.129.

[15] R. Orús, S. Mugel, and E. Lizaso, "Quantum computing for finance: Overview and prospects," *Reviews in Physics*, vol. 4, p. 100 028, 2019.

[16] Pi Tech Private Limited, *Chukul nepse*, All rights reserved. [Online]. Available: %5Curl%7Bhttps://chukul.com/#/%7D.

[17] NEPSE, *Nepal share*, ALL rights reserved . [Online]. Available: %7Bhttps://www.nepalstock.com/%7D.

[18] B. Boehm, "A survey of agile development methodologies," *Laurie Williams*, vol. 45, p. 119, 2007.

[19] I. JAAMU, "What is lstm? introduction to long short term memory," 2023. [Online]. Available: https://intellipaat.com/blog/what-is-lstm/.