**KANTIPUR ENGINEERING COLLEGE**

**DHAPAKHEL, LALITPUR**



**[Subject Code: CT654]**

**A MINOR PROJECT PROPOSAL ON**

**Stock Price Prediction**

**Submitted By:**

**Ankit Bhandari (KAN077BCT016)**

**Ankit Kunwar (KAN077BCT018)**

**Anup Bhattarai (KAN077BCT022)**

**Dibyanshu Niraula (KAN077BCT031)**

**IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF BACHELOR IN COMPUTER ENGINEERING**

**Submitted To:**

**Department of Computer and Electronics Engineering**

**Aug, 2023**

**Stock Price Prediction**

**Submitted By:**

**Ankit Bhandari (KAN077BCT016)**

**Ankit Kunwar (KAN077BCT018)**

**Anup Bhattarai (KAN077BCT022)**

**Dibyanshu Niraula (KAN077BCT031)**

**Submitted To:**

**Department of Computer and Electronics Engineering**

**Kantipur Engineering College**

**Dhapakhel, Lalitpur**

**Aug, 2023**

**LIST OF FIGURES:**

[Figure 0‑1:Gantt Chart 9](#_Toc140767030)

[Figure 0‑2:Steps involved in Developing Model 10](file:///C:\Users\bhatt\Downloads\lol.docx#_Toc140767032)

[Figure 0‑3:LSTM Model 12](#_Toc140767033)

[Figure 0‑4:Sub system of LSTM 13](#_Toc140767034)

[Figure 0‑5:Sub system of LSTM 13](#_Toc140767035)

[Figure 0‑6:Sub system of LSTM 14](#_Toc140767036)

[Figure 0‑7:Sub system of LSTM 14](#_Toc140767037)

[Figure 0‑8:Use Case Diagram 16](#_Toc140767038)

[Figure 0‑9:Data flow Diagram 16](#_Toc140767039)

[Figure 0‑10:Incremental Model 17](#_Toc140767040)

**LIST OF ABBREVIATIONS**

**ANN** Artificial Neural Network

**ADX** Average Directional Index

**LSTM** Long Short Term Memory

**RNN** Recurrent Neural Network

**RSI** Relative Strength Index

**TABLE OF CONTENTS**

[INTRODUCTION 6](#_Toc141871632)

[1.6.1 Hardware Requirement 7](#_Toc141871633)

[1.6.2 Software Requirement 7](#_Toc141871634)

[1.6.3 Feasibility Study: 8](#_Toc141871635)

[1.6.4 Gantt chart 9](#_Toc141871636)

[LITERATURE REVIEW 9](#_Toc141871637)

[METHODOLOGY 10](#_Toc141871638)

[3.1 Working of LSTM Model 12](#_Toc141871639)

[3.1.1 Step-by-Step LSTM Walk Through: 13](#_Toc141871640)

[3.2 Indicators for Buy, Sell, Hold Recommendation 15](#_Toc141871641)

[3.2 System Diagrams 16](#_Toc141871642)

[3.2.1 Use Case Diagram 16](#_Toc141871643)

[3.2.2 Data Flow Diagram 16](#_Toc141871644)

[3.3 SOFTWARE MODEL USED 17](#_Toc141871645)

[EPILOGUE 18](#_Toc141871646)

[4.1 EXPECTED OUTPUT: 18](#_Toc141871647)

[REFERENCES: 19](#_Toc141871648)

**CHAPTER 1**

# INTRODUCTION

**1.1BACKGROUND**

The Stock market is a constellation of marketplaces where securities like stocks and bond are bought and sold. It provides with easy, transparent access to investment assests.Stock market prediction is the act of trying to determine the future value of a company stock. The more accurate the prediction, the higher the profit potential. The prediction methodologies fall into two distinct categories i.e. fundamental analysis, technical analysis. Fundamental analysis is concerned with evaluating company’s past performances ,financial condition and market capitalization whereas technical analysis seeks to determine the possibilities of future stock price movement largely based on trends of the past price performance.

**1.2 PROBLEM STATEMENT**

The stock market is a complex and dynamic environment heavily influenced by a countless factors. The task of forecasting stock market movements is extremely difficult due to the non-linear and volatile nature of the market. Traditional statistical methods often fall short in capturing the intricate patterns and subtle dependencies in the data, leading to suboptimal predictions. As a result, there is an urgent need to explore advanced techniques to enhance the accuracy of stock market predictions. Despite the well-known investment adage, "buy low, sell high," investors face challenges in making informed investment decisions without a comprehensive understanding of stock market behavior. Investing in the right stock at the wrong time can lead to disastrous results, while investing in a seemingly mediocre stock at the opportune moment can yield significant profits. The ability to predict stock prices accurately can greatly enhance the profitability and risk management of financial investments. The objective of this research is to develop a robust and reliable stock market prediction model that surpasses conventional methods by leveraging the power of machine learning algorithms, particularly Long Short-Term Memory (LSTM) networks. LSTM, a type of recurrent neural network (RNN), excels at capturing sequential patterns and is well-suited for time series data like historical stock market prices.

**1.3 Objectives:**

1. To forecast stock prices accurately.
2. To develop a functional prototype.

**1.4 Application Scope**

As we know “The more accurate the prediction, the higher the profit potential”. Thus accurate prediction of stock price movement will enhance profit making potential.

The predicted data enables to gain insights into the financial status of companies and perform comparisons. Moreover, its utility extends to researchers, stock brokers, market makers, government entities, and the general public, offering valuable information and aiding in informed decision-making.

**1.6 Development Requirements**

### 1.6.1 Hardware Requirement

* Computer with Minimum 2 GB RAM
* Internet Connection

### 1.6.2 Software Requirement

* Python
* Anaconda Navigator
* Jupyter Notebook

### 1.6.3 Feasibility Study:

**a. Economic Feasibility:** The project demonstrates strong economic feasibility as it can be developed on a standard, cost-effective PC readily available in the market. Since the system is built using an open-source programming language, there are no associated development costs. All necessary research materials and resources are easily accessible online, further reducing financial burdens.

**b. Technical Feasibility:** The application benefits from favorable technical feasibility as it can be implemented on existing network and operating system services without significant challenges. The development relies on widely-used technologies such as Python programming, LSTM algorithm, and Scikit-learning pipelines, ensuring seamless integration and efficient execution. Thus, our project is deemed technically feasible.

**c. Operational Feasibility**

The system's operation will be straightforward, requiring only a computer and a user-friendly interface. The project demonstrates operational workability, ensuring ease of use for end users.

### 1.6.4 Gantt chart

Figure 0‑1: Gantt chart

**CHAPTER 2**

# LITERATURE REVIEW

Prediction of stock prices involves a significant understanding of market trends and values. Traditionally, gaining this expertise required considerable time and effort. However, with the advancements in technology and the availability of powerful machine learning algorithms, predicting stock trends has become more feasible.

Numerous methodologies and approaches have been proposed for analysts to predict future stock market values using machine learning algorithms. Some notable studies include 'Stock Market Forecasting using Machine Learning Algorithm' [1], and 'Prediction of Bombay Stock Exchange Market Returns using Artificial Neural Network (ANN) and Genetic Algorithm' [2]. These studies aim to provide valuable insights into predicting future stock market values, offering potential benefits for researchers and businesspeople seeking suitable methods for their purposes.

One of the methods used is Long Short-Term Memory (LSTM), a type of recurrent neural network, which has shown superior performance compared to other machine learning models like random forests and multilayer perceptron in predicting stock price patterns [4]. Researchers are also exploring more realistic trading strategies that go beyond simple buying and selling after a fixed amount of time. They consider additional factors such as timing, execution booking, and transaction costs to enhance the accuracy of predictions [5].Similarly several technical indicators have been used by analysts to predict the buy, sell or hold condition. Some noteworthy studies are 'Trend Identification with the Relative Strength Index (RSI) Technical Indicator – A Conceptual Study'[6], and 'The profitability of moving average trading rules in South Asian stock markets'[7].

By utilizing these prediction methodologies, investors can improve their investment strategies. Accurate stock price predictions can lead to significant profits and empower investors to make informed financial decisions, including buying, holding, or selling stocks.

**CHAPTER 3**

# METHODOLOGY

In this study, we will be demonstrating the application of the Long Short-Term Memory (LSTM) technique for predicting stock market indices.

FEATURE EXTRACTION

DATA

COLLECTION

DATA PRE-PROCESSING

0‑1

OUTPUT

GENERATION

TRAIN THE NEURAL NETWORK

Figure 0‑2: Steps involved in Developing Model

Our system is organized into several stages, outlined as follows:

Stage 1: Data Collection:

In this stage, the data related to the specific company's stock prices is collected from a reliable source. For instance, it could be obtained from financial APIs, financial databases, or specialized websites.

Stage 2: Data Pre-Processing

a. Data cleaning: Missing values, outliers, or inaccuracies in the collected data are handled and cleaned to ensure high data quality.

b. Data normalization: The collected data is normalized to bring it to a common scale, typically between 0 and 1, to enhance model performance.

c. Data splitting: The dataset is divided into training and testing sets to evaluate the model's performance.

Stage 3: Feature Extraction

Moving average (200-day), Relative Strength Index (RSI), Accumulation/Distribution (A/D) Lines, and Average Directional Index (ADX) are calculated for the stock data. Based on the calculated features, a recommendation ('Buy', 'Sell', or 'Hold') is determined for the stock. After this, we carefully select the functions that will serve as input to the neural network. Specifically, we choose the features such as date, open price, big, medium price, near price, and trading volume. These chosen features will be utilized as input data for the neural network during the model training process.

Stage 4: Train the Neural Network

The stock's closing price data is normalized using Min-Max scaling to bring it within the range of 0 to 1. The data is split into training and test sets. In our LSTM model, we have designed a sequential architecture that comprises several layers. The initial layer is the input layer, which feeds the sequential data into the model. After that, we have integrated three LSTM layers, which are essential for capturing temporal dependencies and patterns in the data. LSTM (Long Short-Term Memory) layers are well-suited for time series data and can effectively handle long-term dependencies. Finally, we have the dense output layer, which is responsible for producing the final predictions. The dense layer is configured with a linear activation function, which is appropriate for regression tasks where we aim to predict numerical values. This choice of activation function allows the model to output continuous values, making it suitable for forecasting stock prices in this particular scenario.

Stage 5: Output Generalization

The actual and predicted stock prices are plotted for visualization. Predictions for the stock price for the next 30 days are made using the last 100 days' data and plotted along with the actual stock prices.

## 3.1 Working of LSTM Model

Long Short Term Memory networks – usually just called “LSTMs” – are a special kind of RNN, capable of learning long-term dependencies. LSTMs are explicitly designed to avoid the long-term dependency problem. Remembering information for long periods of time is practically their default behaviour, not something they struggle to learn!

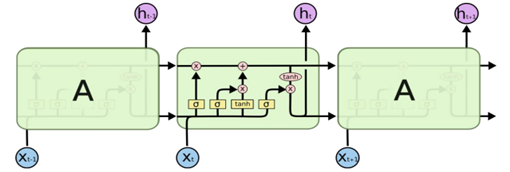


Figure 0‑3: LSTM Model

In the given illustration, every line conveys a complete vector, transferring information from one node's output to the inputs of other nodes. The pink circles depict pointwise operations, such as vector addition, whereas the yellow boxes represent neural network layers that are learned during the training process. Lines that converge indicate concatenation of vectors, while lines that fork signify the replication of content, with the copies being sent to distinct locations.

## 3.1.1 Step-by-Step LSTM Walk Through:

In the initial step of our LSTM, we utilize a sigmoid layer referred to as the "forget gate layer" to determine the information we want to discard from the cell state Ct−1. This gate layer examines ht−1​and xt, producing a value between 0 and 1 for each element in the cell state. A value of 1 signifies "retain completely," while a value of 0 indicates "discard completely."

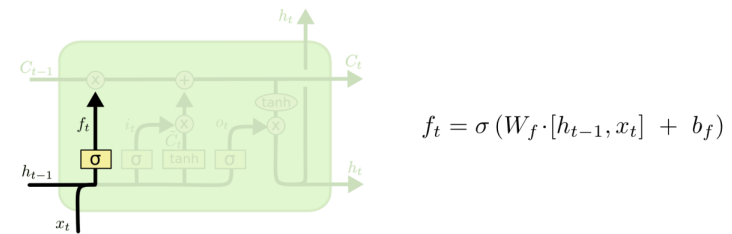


Figure 0‑4: Sub system of LSTM

The subsequent stage involves determining the additional data that will be stored in the cell state, comprising two key components. Initially, the "input gate layer," implemented as a sigmoid layer, determines the information to be updated. Subsequently, a "tanh layer" generates a vector of fresh potential values, 𝐶t, which could be incorporated into the state. Finally, these two elements are combined in the following step to create an update for the cell state.

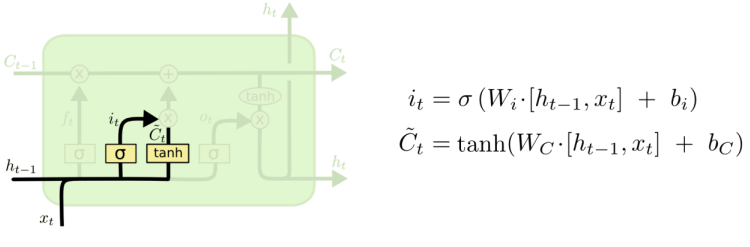


Figure 0‑5:Sub system of LSTM

Now, we proceed with updating the previous cell state,Ct−1, to the new cell state, Ct. Having already determined the necessary actions in the prior steps, our focus shifts to their actual implementation.

Firstly, we multiply the old state by ft, effectively discarding the information we had chosen to forget earlier. Then, we add 𝑖 \* 𝐶t. which represents the new candidate values scaled by the extent of update decided for each individual state value. In this way, the cell state is successfully updated.

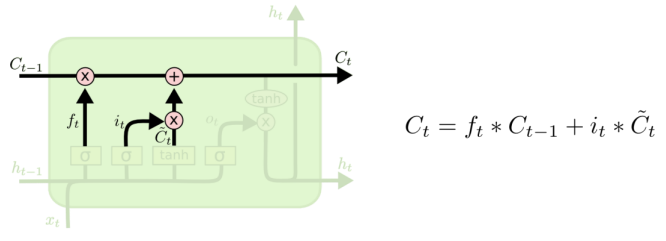


Figure 0‑6: Sub system of LSTM

Lastly, the determination of the output is required. The output will be derived from our cell state, but in a filtered manner. Initially, we pass the cell state through a sigmoid layer, responsible for deciding which aspects of the cell state will be included in the output. Subsequently, we subject the cell state to the tanh function to constrain the values within the range of -1 and 1, and then we multiply it by the output obtained from the sigmoid gate. This ensures that we only output the selected portions that were decided earlier, resulting in the final filtered output.

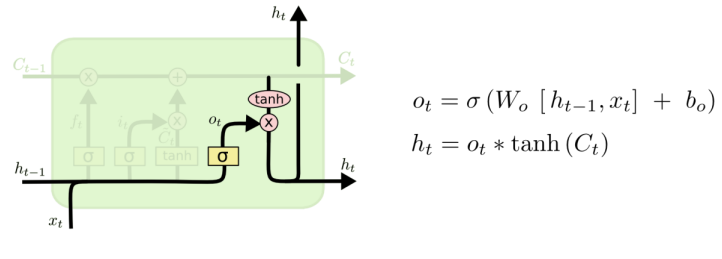


Figure 0‑7: Sub system of LSTM

## 3.2 Indicators for Buy, Sell, Hold Recommendation

i. Accumulation/Distribution Line

A/D lines represent the accumulation and distribution of a security based on its price and trading volume. The indicator considers whether the closing price is higher or lower than the previous day's price and adjusts the volume accordingly. Positive A/D values suggest accumulation (buying pressure), while negative values suggest distribution (selling pressure).

ii. Moving Average (MA):

A moving average is a trend-following indicator that smooths out price data over a specified period. It calculates the average price of an asset over that period and provides a clearer view of the underlying trend by reducing noise and short-term fluctuations.

iii. The Relative Strength Index (RSI):

It is another popular technical indicator used in financial markets to assess the strength and momentum of price movements. The most basic use of an RSI is as an [overbought](https://www.investopedia.com/terms/o/overbought.asp) and [oversold](https://www.investopedia.com/terms/o/oversold.asp) indicator. When RSI moves above 70, the asset is considered overbought and could decline. When the RSI is below 30, the asset is oversold and could rally.

iv. Average Directional Index (ADX):

ADX measures the strength of a trend, whether it is bullish or bearish. When the ADX is above 40, the trend is considered to have a lot of directional strength, either up or down, depending on the direction the price is moving.

When the ADX indicator is below 20, the trend is considered to be weak or non-trending. Based on these indicators, the system will generate trading recommendations for a specific stock as follows:

* If the closing price is above the 200-day moving average, RSI is above 70, A/D lines are rising, and ADX is above 25, the system will recommend a "Buy" action.
* Conversely, if the closing price is below the 200-day moving average, RSI is below 30, A/D lines are falling, and ADX is above 25, the system will suggest a "Sell" action.
* For all other scenarios where the specified conditions are not met, the system will provide a "Hold" recommendation.

## System Diagrams

## Use Case Diagram

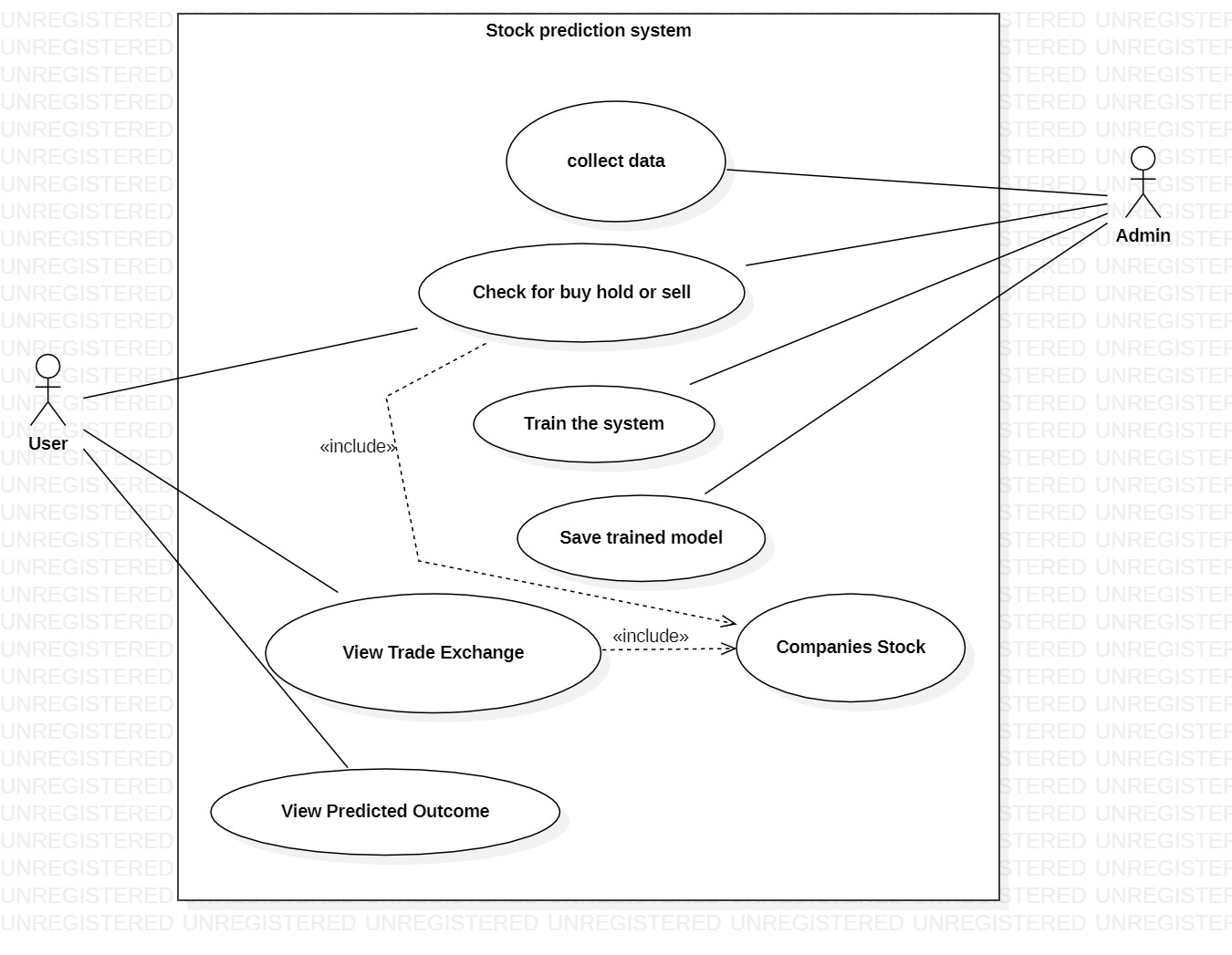


Figure 0‑8: Use Case Diagram

## Data Flow Diagram

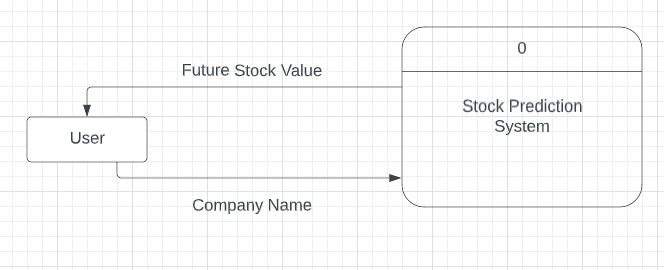


Figure 0‑9: Data flow Diagram

## 3.3 SOFTWARE MODEL USED

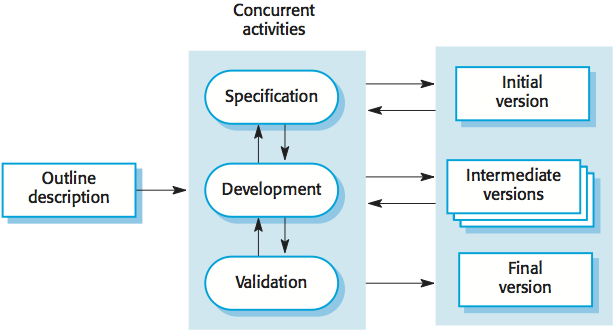


Figure 0‑10: Incremental Model

In the incremental model for developing the stock analysis system, the following steps will be taken:

Increment 1: Data Collection

* Collect historical stock price data for specific company stock using the API.

Increment 2: Analysis and Recommendations

* Calculate the 200-day moving average to identify trends and support/resistance levels.
* Calculate the Relative Strength Index (RSI) to measure momentum and overbought/oversold conditions.
* Calculate the Accumulation/Distribution (A/D) Lines to analyse the buying and selling pressure based on price and volume.
* Calculate the Average Directional Index (ADX) to assess trend strength.
* Generate trading recommendations for AAPL stock based on the analysed indicators. Recommendations include "Buy," "Sell," or "Hold."

Increment 3: LSTM Model Training

* Prepare the historical stock price data for training the LSTM model.
* Normalize the data using Min-Max scaling to bring it within the range of 0 to 1.
* Split the data into training and test sets for the LSTM model.
* Train the LSTM model on the training data.

Increment 4: LSTM Model Evaluation and Prediction

* Make predictions on the training and test data using the trained LSTM model.
* Plot the actual and predicted stock prices to assess the model's accuracy.

Increment 5: Future Price Prediction

* Predict the stock price for the next 30 days using the last 100 days' data with the LSTM model.
* Combine the actual stock prices with the predicted future prices for plotting.

**CHAPTER 4**

# EPILOGUE

## 4.1 EXPECTED OUTPUT:

* Closing Price History Visualization
* Stock Analysis Results for buy, sell of hold.
* A chart comparing the actual closing prices of specific company stock with the predicted stock prices.
* A chart displaying the actual closing prices of specific company stock along with the predicted future stock prices for the next 30 days.

# REFERENCES:

[1] Y. Perwej and A. Perwej, “Prediction of the Bombay stock exchange (bse) market returns using artificial neural network and genetic algorithm,” 2012.

[2] N. H. Hung and Y. Zhaojun, “Profitability of applying simple moving average trading rules for the Vietnamese stock market,” Journal of Business Management, vol.2,no. 3, pp. 223–1, 2013.

[3] K. S. Kannan, P. S. Sekar, M. M. Sathik, and P. Arumugam, “Financial stock market forecast using data mining techniques,” vol. 1, p. 4, 2010.

[4] Sanket Chaudhari , K Rajeswari , Sushma Vispute, 2021, A Review on using Long-Short Term Memory for Prediction of Stock Price, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 10, Issue 11 (November 2021),

[5] David M. Q. Nelson, Adriano C. M. Pereira, and Renato A. de Oliveira, “Stock Market’s Price Movement Prediction With LSTM Neural Networks,” in International Joint Conference on Neural Networks (IJCNN), 2017, pp. 1419–1426. doi: 10.1109/IJCNN.2017.7966019.

[6] Panigrahi, A. K., Vachhani, K., Chaudhury, S. K. (2021). Trend Identification with the Relative Strength Index (RSI) Technical Indicator – A Conceptual Study. Journal of Management and Research Analysis, 8(4), 159-169.

[7] Gunasekarage, A., & Power, D. M. (2001). The profitability of moving average trading rules in South Asian stock markets. Emerging Markets Review, 2(1), 17-33. <https://doi.org/10.1016/S1566-0141(00)00017-0>

[8] "Investopedia."Top 7 Technical Analysis Tools." Investopedia, [www.investopedia.com/top-7-technical-analysis-tools-4773275](http://www.investopedia.com/top-7-technical-analysis-tools-4773275). Accessed 20 July 2023.