STOCK MARKET PREDICTION USING DEEP LEARNING TECHNIQUES.

A research proposal submitted in partial fulfillment of the requirements of the degree of Bachelor of Science in Financial Engineering at the Jomo Kenyatta University of Agriculture and Technology.

2023

DECLARATION

This research is our original work and has not been presented for a degree award in any other university.

Name:	Reg.no	Signature
PARPETUAL WAWIRA GITARI	SCM222-0378/2019	……………...
LINDA MUTHONI MUTWIRI	SCM222-0714/2019	……………...
MUMBUA KYAI	SCM222-0389/2019	……………...
IVY WAMBUI NJOROGE	SCM222-0372/2019	……………...
BARTLET MUTISO KING'ONDU	SCM222-0391/2019	……………...

……………...

The research proposal has been submitted for examination with my approval as a university supervisor.

Signature……………………………… Date…………………………………….

DR. HERBERT IMBOGA

Department of Statistics and Actuarial Sciences, JKUAT.

ACKNOWLEDGEMENT

We primarily thank the Almighty God for His grace throughout our research work and enabling us to reach this stage in our career.

It would not have been possible to do this project proposal without the support of our families. Their care, understanding, prayers and continued support that has enabled us to reach this far and we are forever grateful.

It has also been a great honor and privilege to undergo training in this prestigious university.

We also are highly indebted to Dr. Susan Mwelu for her guidance and constant supervision throughout the research project proposal process.

LIST OF ABBREVIATIONS

A/D Accumulation/ Distribution

AI Artificial Intelligence

ANN Artificial Neural Network

AR AutoRegressive

ARIMA Autoregressive Integrated Moving Average

ARMA Autoregressive Moving Average

BAT British American Tobacco

CBK Central Bank of Kenya

CCI Commodity Channel Index

CNN Convolutional Neural Networks

EABL East African Breweries Limited

EMA Exponential Moving Average

EMH Efficient Market Hypothesis

GARCH Generalized Autoregressive Conditional Heteroskedasticity

GRU Gated Recurrent Unit

ICT Information Communication Technology

IQR Interquartile Range

KCB Kenya Commercial Bank

KNBS Kenya National Bureau of Statistics

LASSO Least Absolute Shrinkage and Selection Operator.

LSTM Long Short Term Memory

MACD Moving Average Convergence Divergence

MAE Mean Absolute Error

MAPE Mean Absolute Percentage Error

MLP Multilayer Perceptron

MSE Mean Squared Error

NSE Nairobi Security Exchange

PCA Principal Component Analysis

RMSE Root Mean Squared Error

RNN Recurrent Neural Network

RSI Relative Strength Index

RWT Random Walk Theory

SGD Stochastic Gradient Descent

WMA Weighted Moving Average

CONTENTS

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND OF THE STUDY

According to Investopedia (2021),a stock also known as a share or equity, represents ownership in a company. When a company goes public and offers its shares to the public, it divides its ownership into smaller units called stocks. These stocks are traded on stock exchanges. A stock exchange is a regulated marketplace where buyers and sellers trade stocks, bond and other securities, Investopedia (2021). Nairobi Security Exchange provides a platform for companies to list their shares and facilitates the buying and selling of those shares among investors.

Fundamentally, a stock market is associated with buyers and sellers dealing with stocks or shares. The stock market is known to be volatile, random and unpredictable due to continuously changing stream of data. The stock data is a sequence of the price of a given stock equally distributed by time intervals, i.e., time-series data. The stock traders mainly focus on accurate prediction of the shares to maximize profit. According to Cao, Y., Tay, F. E. H., & Yao, L. (2019),deep learning for stock prediction refers to the application of deep neural networks, a subset of machine learning algorithms, for forecasting stock prices or making investment decisions in the financial markets.

Globally stock markets are increasingly adopting deep learning techniques for stock prediction. With digital technologies and vast financial data, researchers and practitioners are exploring deep learning's potential to uncover patterns and improve prediction accuracy. From Wall Street to financial hubs in Europe and Asia, deep learning is enhancing stock market forecasting, enabling informed investment decisions and risk management strategies.

Africa is a diverse continent with different countries and economies, each with its unique perspective on stock markets. Predicting stock markets in Africa, like in any region, presents challenges and uncertainties. Key factors to consider when discussing the outlook of African stock markets include economic growth, commodity prices, the political and regulatory environment, investor sentiment, capital market development, and regional integration.In recent years, Africa's financial landscape has experienced rapid evolution, with stock exchanges emerging as crucial drivers of economic growth and investment opportunities. Deep learning for stock prediction is gradually gaining attention within the African context. As African stock exchanges embrace digitization and technological advancements, researchers and analysts are exploring the application of deep learning algorithms to predict stock prices, identify market trends, and enhance investment strategies.

The stock market in Kenya refers to the Nairobi Securities Exchange (NSE), which is the primary securities exchange in the country. The NSE plays a crucial role in facilitating the buying and selling of securities, including stocks, bonds and other financial instruments. Shareholders however do not directly execute the trade, nor is there any meeting between buyers and sellers for sale negotiations. They trade by giving instructions to their Stockbrokers, who in turn execute orders by automatically matching mutually agreeable prices through the NSE trading software. Stockbrokers are the only authorized agencies that can act on behalf of investors in the stock trade business. Stockbrokers do not just execute client orders, they are also charged with the responsibility of advising clients on NSE trades (Government of Kenya, 2009). In the stock trade business, investors, shareholders, traders and clients usually mean the same thing. These are the entities that invest in the stock market by buying and selling stocks through Stockbrokers. In their advisory role, some Stockbrokers base their advice on their observation of short term price movements (trend). Other Stockbrokers do basic research into the fundamentals of the various stocks or undertake technical analysis before they advise their clients on investment decisions. However, none of these methods have any assurance of profitability and they usually state a caveat as such. While talking about the stock market, Graham (2003) explains technical approaches as those that generally urge investors to buy stocks because they are appreciating in value and sell when their value declines. He says this is the ‘popular’ method. This approach is based on the immediate or short term focus. He further says that many investors acquire common stocks for the mere ‘excitement and temptation’ of the market. This state of investing for the excitement of the market is described as ‘temperament’. Despite the popularity of technical analysis, he states that this approach is unreliable and is akin to ‘simple tossing of a coin’ (Graham, 2003) Since none of the current methods of investment used by NSE Stockbrokers guarantee trade at a profit, investors can be said to be leaving important investment decisions to mere temperament of Stockbrokers. Despite this lack of investment advice that assures investors of profits, this sector controls about half the wealth compared to bank deposits. It is therefore important to have carefully considered investment advice. The likelihood of financial losses, due to inadequate or incorrect advice, is detrimental to investors, especially the low-income citizens who may be investing their livelihoods in the stock exchange, hoping for some profit. Without assurance of some profit, potential investors have been reluctant to consider the NSE as a serious investment sector. Lack of interest in this sector is therefore likely to stagnate the growth of the NSE and hence fail to help the country achieve the desired economic growth.Stockbrokers need to be empowered to enable them have some capability to provide the best advice to their clients. Such empowerment should not only improve their reputation amongst clients, but is also for their own benefit, since they are paid a commission of each trade. If Stockbrokers do not have tools that can generate good advice, then they are bound to fail in their duty, leading to apathy from the investors. The current popular methods of technical analysis (using trends for guessing future price movements) and fundamental analysis (buy-and-hold principle for any company that is considered good) have no pointers on the exact price for a future trade. A system that can guide on the most likely next day price is therefore missing and necessary. Such a predictive tool is therefore desirable. It is for this reason that there is need to formulate a deep learning model that can be developed into a tool that can be used by Stockbrokers to advise investors on exactly which stock to invest in, for guaranteed profitability. Such a tool should not only show the trend but also the most probable stock prices for next day trade. Additionally, with such a system having generated a shortlist of good stocks, it is possible to further pick the most profitable of these stocks at any given time. Such a system should therefore be able to predict the next day events. Prediction is the foretelling of a future event or outcome, before it occurs. In contrast, trends (technical analysis) would show many stocks moving in a particular direction, without pinpointing the most profitable amongst them. The deep learning system can be able to provide a longer term investment plan so that the investor knows which stocks to buy and which ones to sell at any particular time in the future (prediction). Deep learning techniques such as LSTM and GRU are computer algorithms formulated using specific AI rules to learn from data and then be used for tasks such as prediction. The emergence of deep learning, a subset of machine learning revolutionized stock prediction. The main goal is to study and apply deep learning techniques to the stock market in order to predict stock behavior and thus act on those predictions to avoid investment risk and generate profit. Deep learning models, particularly

LSTM and GRU have shown exceptional capabilities in capturing complex patterns and non-linear relationships.

1.2 PROBLEM STATEMENT.

Trading in stocks is a significant economic activity in various economies, including Kenya. The current methods employed by Stockbrokers in executing trades and providing advice rely on their experience, technical analysis, or fundamental analysis. However, these methods are subjective, short-sighted, and limited in their ability to process large amounts of data effectively. Improper investment decisions can lead to substantial losses for investors, discouraging potential participants from entering the market. Therefore, there is a need for a deep learning algorithm that can guide investment decisions by predicting the most likely future stock prices. According to the information available on the websites of Kenyan Stockbrokers, there is little indication of the presence of deep learning algorithm that assist in advising clients on suitable stocks for trading. The websites emphasize the use of fundamental and technical analysis methods to predict future stock price movements. However, these methods only indicate the direction of price movement and not the specific value of future stock prices. Therefore, it is desirable to have an algorithm that not only provides directional guidance but also accurately predicts the most likely future stock prices. Stockbrokers, who execute trades and provide investment advice, are the appropriate target users for such an algorithm. By analyzing historical stock prices and leveraging deep learning techniques, such a system can provide precise price predictions, enabling investors to make profitable investment decisions. The predictive model will advise Stockbrokers on the optimal timing for buying or selling stocks based on intelligent analysis. This will empower Stockbrokers to provide more strategic advice to their clients, allowing them to invest in stocks that are likely to appreciate in value. Simultaneously, the tool will enable investors to maximize their returns by selling stocks at the right time. The assurance of returns will reduce uncertainty in the billion-shilling investment industry and foster increased participation in the Nairobi Securities Exchange (NSE).

1.3 JUSTIFICATION OF THE STUDY

Currently, Kenyan Stockbrokers rely on non-AI tools that may not be effective and lack predictive abilities. These methods only provide trends and don't accurately predict next-day stock prices. However, Stockbrokers play a crucial role as they bridge the gap between clients and the trading platform, offering influential advice and executing trades. By developing deep learning tools to assist Stockbrokers in providing predictive advice to clients, investors are more likely to make informed trading decisions. Assured returns on investment will attract more investors to participate in the Nairobi Securities Exchange (NSE), contributing to its growth and the overall economic development of Kenya.

The NSE plays a significant role in the country's economic growth by facilitating the movement of funds across various sectors. Investment in listed firms supports their development and improves the savings culture among citizens. Stocks provide an opportunity for citizens to profit, enhance their purchasing power, and actively participate in economic activities. This project also bridges the gap between theoretical considerations in Information and Communications Technology (ICT) systems and their practical implementation in solving real world financial problems. It encourages the academic community to apply ICT to tackle financial problems.

1.4 OBJECTIVES

1.4.1 GENERAL OBJECTIVE

To predict stock market prices using deep learning techniques.

1.4.2 SPECIFIC OBJECTIVES

1. To build Long Short Term Memory and Gated Recurrent Unit models for stock market prediction.
2. To optimize the Long Short Term Memory and Gated Recurrent Unit models for accurate stock market prediction.
3. To evaluate the trained Long Short Term Memory and Gated Recurrent Unit models.

1.5 SCOPE OF THE STUDY

The study aims to investigate the potential of LSTM and GRU neural networks for predicting stock market movements. By leveraging these advanced deep learning techniques, the research seeks to capture temporal dependencies and patterns in historical stock market data, addressing challenges such as non-linearity and volatility. Comprehensive datasets will be collected for training and testing the models, and various configurations will be explored to optimize performance in different market conditions. The study's outcomes could provide valuable insights for investors and financial analysts, contributing to the growing field of deep learning applications in finance.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This literature review explores the theoretical and empirical framework of using deep learning techniques for stock market prediction. It examines various studies, methodologies, and models employed in the field of deep learning for predicting stock market trends. The review focuses on the theoretical and empirical underpinnings of deep learning algorithms, their application in stock market prediction, and the challenges associated with this approach. The findings contribute to a comprehensive understanding of the theoretical and empirical foundations of deep learning techniques for stock market prediction and identifying research gaps from past studies.

2.2 Theoretical review

According to Schmidhuber(2015), forecasting is the process of predicting future events by analyzing historical data, and it finds applications in various fields, including business, economics, environmental science, and finance. Forecasting can be categorized into short-term (seconds to months), medium-term (1 to 2 years), and long-term (beyond 2 years) predictions. In financial markets, forecasting stock prices is a crucial aspect of investment decision-making.

The data used for stock price forecasting is typically time series data, LeCun, Bottou, Bengio, & Haffner (1998), which is a chronological sequence of observations for a specific variable, in this case, stock prices. Time series analysis helps identify patterns, trends, and cycles within the data. Early knowledge of bullish or bearish market trends can be advantageous for making wise investment choices, and recognizing patterns can aid in identifying the best-performing companies for specific periods.

According to Hochreiter and Schmidhuber(1997), Several methods are used for stock price forecasting, including fundamental analysis, technical analysis, and time series forecasting. Fundamental analysis estimates a company's share value by analyzing its financial indicators, making it suitable for long-term predictions. Technical analysis, on the other hand, relies on historical price data to identify current trends and is more suited for short-term forecasts.

Time series forecasting involves two main classes of algorithms: linear models and non-linear models. Linear models, such as AR, ARMA, and ARIMA, fit mathematical models to univariate time series data. However, these models have limitations, as they do not account for latent dynamics in the data and fail to identify interdependencies among various stocks.

Automated predictions in financial markets aim to incorporate traditional investment strategies into numerical algorithms, Cho,Merrienboer ,Bahdanau & Bengio (2014).This can lead to similarities in trading trends among companies using similar algorithms, potentially causing unusual market behavior and influencing each other's decisions.

According to Rosenblatt (1958), Efficient Market Hypothesis (EMH) and Random Walk Theory (RWT) are two significant theories used to understand stock market pricing movements. EMH suggests that all publicly available information is already reflected in stock prices, making profitable predictions impossible. RWT argues that past stock price movements cannot reliably predict future values, as stock prices change independently and randomly.

Trading strategies form the basis for making predictions in financial markets, with technical trading and fundamentals trading being the most popular approaches (Rumelhart, Hinton, & Williams 1986). Technical trading relies on analyzing past price movements to identify current trends, while fundamentals trading focuses on a company's intrinsic value based on financial data.

Deep learning, a subfield of machine learning, has gained immense popularity for its ability to automatically learn hierarchical representations from large datasets using neural networks with multiple layers (Schmidhuber, 2015). These neural networks, inspired by the human brain's structure and function, consist of interconnected nodes organized into layers. Each node processes inputs through a weighted sum and an activation function to produce an output, allowing information to flow through the network, eventually leading to the desired output.

The history of neural networks can be traced back to the work of (McCulloch and Pitts,1943) where they proposed an artificial neuron model based on logical calculus. In the following decades, Rosenblatt's perceptron model established the foundation for single-layer neural networks capable of linear separability.

The breakthrough that paved the way for modern deep learning was the development of the backpropagation algorithm (Rumelhart, Hinton, & Williams 1983). This algorithm enabled the efficient training of multilayer perceptrons (MLPs) with multiple interconnected layers of neurons, allowing the network to learn complex patterns.

The introduction of Convolutional Neural Networks (CNNs) by LeCun et al(1998), revolutionized the field of computer vision. CNNs utilize convolutional layers to extract spatial features from images, followed by pooling and fully connected layers for classification. Over time, CNN architectures have evolved, leading to significant improvements in tasks like image classification and object detection.

To handle sequential data with temporal dependencies, Recurrent Neural Networks (RNNs) were introduced. However, the standard RNN suffered from the vanishing gradient problem, hindering the learning of long-term dependencies. The solution came with Long Short-Term Memory (LSTM), proposed by (Hochreiter and Schmidhuber 1997), which introduced memory cells and gating mechanisms to capture and propagate information over long sequences, making it suitable for tasks like speech recognition and language modeling.

An alternative to LSTM, called Gated Recurrent Units (GRUs), was proposed by Cho et al (2014). GRUs simplify the gating mechanism while achieving comparable performance. They strike a balance between performance and computational efficiency and have found applications in various domains.

These advancements in neural network architectures, along with optimization algorithms and regularization techniques, have led to significant breakthroughs in diverse fields. CNNs excel in computer vision, while RNNs, LSTM, and GRUs have made significant contributions to natural language processing and speech recognition.

Deep learning has emerged as a powerful tool for solving complex problems across various domains. The continuous evolution of neural network architectures and techniques continues to drive the advancement of artificial intelligence and push the boundaries of what is possible in machine learning.

2.3 Critiques of the existing literature relevant to the study

The use of technical analysis, particularly trending, as a basis for predicting stock market prices has been explored in various markets. Ndiritu (2010) developed a support system for trending future prices in the NSE. In different studies, researchers applied technical analysis to predict stock prices in markets like the New York Stock Exchange (Deng, Mitsubichi, Shioda, Shimada & Sakurai 2011), the Tehran Stock Exchange (Aghababaeyan and Tamannasiddiqui 2011), the Bangladesh Stock Market (Khan, Alain & Hyussain 2011), and the Australian Stock Market (Pan, Tilakaratne &Yearwood 2005). While these studies achieved varying levels of accuracy, their tools were not commercialized or targeted for specific stockbrokers. Following the interest in stock price prediction, the focus shifted towards deep learning techniques.

Yixin (2022), explored the application of ARIMA, GARCH, and LSTM models for stock price prediction on the S&P 500 stock market. The LSTM model outperformed ARIMA and GARCH models, showcasing its ability to capture long-term dependencies in stock data. However, the limitations of the study included its focus on one stock market and the use of historical data. Further research is necessary to validate the findings across different stock markets and time periods.

Bhattacharjee and Bhattacharja (2019) conducted a comparative study between statistical and machine learning methods for stock price prediction using data from Tesla and Apple. The results indicated that machine learning methods, especially multi-layer perceptron and long short-term memory (LSTM), outperformed statistical methods in capturing non-linear relationships within stock data. Although the study only used historical data and not future predictions, it provided a clear explanation of the models and their evaluation metrics.

Hota, Chakravarty, Paikaray & Bhoya (2020) rexamined the application of four machine learning algorithms to predict the opening price of American Airlines stocks. The results demonstrated that the random forest algorithm outperformed other algorithms, including support vector regression, decision tree, and artificial neural network, in terms of predictive accuracy. The authors suggested using more recent data and exploring advanced evolutionary techniques for improving artificial neural networks. Despite the dataset's time limitation, the study offers valuable insights into the use of machine learning algorithms for stock price prediction.

Priya and Geetha (2022) present a machine learning model based on the random forest algorithm for predicting stock market prices. The model achieved an impressive 75% accuracy in its predictions. The authors acknowledged the limitations of using a single stock market dataset and the need for further exploration into other deep learning models and additional features. Nevertheless, the paper demonstrates the effectiveness of the random forest algorithm for stock price prediction and lays the groundwork for future improvements and research in the field.While the model's implementation and results are commendable, the small dataset and lack of comparison with other algorithms pose limitations that need to be addressed for further validation.

Gao, Wang & Zhou (2021)investigates the performance of LSTM and GRU models for stock market forecasting. The study reveals that both models can effectively predict stock prices, with LASSO dimension reduction generally producing better results than PCA. The authors provide valuable insights into the use of dimension reduction techniques and practical recommendations for stock market forecasting. However, the study's limitations include the evaluation on a single dataset, calling for further research to assess the generalizability of the results to other datasets and explore additional dimension reduction methods.

Shahi, Shrestha, Neupane, & Guo (2020) compares LSTM and GRU models for stock market forecasting. The authors employ a cooperative deep learning architecture, which results in LSTM outperforming GRU with an average accuracy of 52.3% compared to 50.7%. Incorporating financial news sentiment further improves accuracy to 53.2%. The paper's contributions include the comparative study, demonstrating the impact of sentiment analysis, and proposing the cooperative deep learning architecture. However, limitations include the focus on the S&P 500 index and the absence of comparison with other machine learning methods, calling for further research in different stock markets and algorithm comparisons.

2.4 Summary

This literature review delves into the theoretical and empirical aspects of using deep learning techniques for stock market prediction. It covers the challenges faced in stock market prediction, the different methods used for forecasting, and the importance of time series analysis for identifying patterns and trends. The review highlights the significance of deep learning, particularly LSTM and GRU models, for handling sequential data with temporal dependencies.

The review discusses the history of neural networks, from early artificial neuron models to the breakthrough of back propagation, CNNs, and RNNs. It emphasizes the importance of LSTM and

GRU in capturing long-term dependencies, making them suitable for tasks like speech recognition and language modeling.

Critiques of existing literature are addressed, with a focus on the use of technical and fundamental analysis for stock price prediction. The limitations of linear models and the advantages of deep learning models are discussed. The review also touches on the Efficient Market Hypothesis and Random Walk Theory, which influence stock market pricing movements.

Several papers are examined in the review. "Stock Price Prediction Using Machine Learning" showcases the effectiveness of LSTM over ARIMA and GARCH models, while "Stock Price Prediction: A Comparative Study between Traditional Statistical Approach and Machine Learning Approach" favors machine learning models, especially multi-layer perceptron and LSTM, over statistical methods. "Stock Market Prediction Using Machine Learning Techniques" highlights the superiority of the random forest algorithm for predicting opening prices of American Airlines stocks.

Overall, the literature review provides a comprehensive understanding of the theoretical and empirical foundations of deep learning techniques for stock market prediction. It emphasizes the potential and effectiveness of deep learning models, while acknowledging the limitations and areas for further research. The review contributes valuable insights to the field of stock market forecasting and highlights the continuous evolution of neural network architectures in driving advancements in artificial intelligence and machine learning.

2.5 Research gaps

The literature review on deep learning techniques for stock market prediction highlights several research gaps that warrant further investigation. Firstly, most studies focused on single stock or dataset, limiting the generalizability of their findings. Our research aims to validate the performance of deep learning models across multiple stocks and diverse datasets to establish their effectiveness in different financial contexts. Moreover, the reviewed studies predominantly relied on historical stock price data and, in some cases, sentiment analysis of financial news. To fill this gap, our research will explore the impact of incorporating additional relevant features, such as technical indicators, macroeconomic factors, or company-specific financial ratios, to enhance predictive capabilities. Additionally, examining the feasibility and performance of deep learning models in real-time or near-real-time stock market prediction will be valuable for assessing their practicality in live trading scenarios. By addressing these research gaps, the field of deep learning for stock market prediction will be enriched, making it more applicable and reliable in real-world financial contexts.

CHAPTER THREE

METHODOLOGY

3.1 Introduction

In this study, we will investigate the use of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) neural networks for stock market prediction. The methodology will involve data collection with relevant features, preprocessing, and partitioning for training, validation, and testing. LSTM and GRU architectures are designed to capture long-term dependencies, and hyper parameter tuning will be performed for optimal performance. The models will then be evaluated using performance metrics and compared to identify their strengths and weaknesses in predicting stock market prices. The study aims to provide valuable insights for investors, traders, and researchers seeking advanced deep learning techniques to enhance their decision-making.

3.2 Research Design

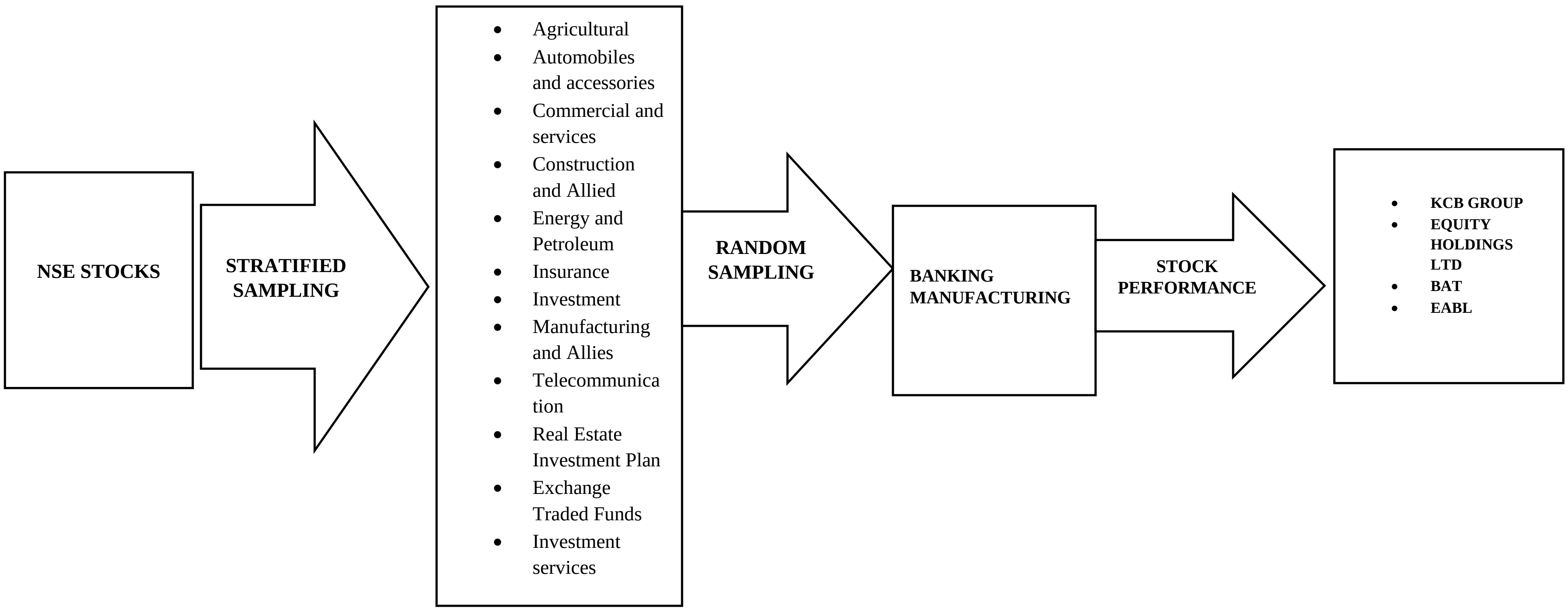
This research aims to explore the effectiveness of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) neural network architectures in predicting stock market trends. LSTM and GRU are powerful deep learning models known for their ability to capture long-term dependencies in time series data, making them potentially well-suited for stock market prediction. The research will collect historical stock price data, including open, high, low, and close prices, along with trading volumes, and additional macroeconomic indicators that may influence the stock market. The data will be preprocessed, and features will be engineered to improve the input representation for the models. The dataset will be split into training, validation, and test sets, and LSTM and GRU models will be implemented using a deep learning framework. The models will be trained and evaluated on the test set using various evaluation metrics, such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE). Statistical tests will be conducted to compare the predictive performance of LSTM and GRU models. Additionally, sensitivity analysis will identify factors influencing prediction accuracy, and the impact of macroeconomic indicators on the models' performance will be assessed. The expected outcomes include insights into the comparative performance of LSTM and GRU models in stock market prediction and their practical applicability in financial decision-making. By following this methodology, the research aims to contribute valuable knowledge to the use of deep learning techniques for forecasting financial markets.

3.3 The target population

The target population for the research methodology includes publicly traded stocks of specific companies, namely Kenya Commercial Bank (KCB), Equity Group Holdings (Equity), East African Breweries Limited (EABL), and British American Tobacco Kenya (BAT). These stocks are listed on the Nairobi Securities Exchange (NSE) and are among prominent players in the Kenyan financial market. The study aims to utilize deep learning models, such as LSTM and GRU, to predict future stock prices and trends for these four companies. By focusing on these specific stocks, the research seeks to provide valuable insights into their price movements and contribute to the understanding of the application of deep learning in stock market forecasting in the Kenyan context.

3.4 Sampling techniques and illustration.

In this study, we utilized both stratified and random sampling methodologies. Specifically, we focused on four stocks listed on the NSE: Equity Holdings LTD, KCB Group, BAT, and EABL. To ensure a representative sample, we divided the stocks into distinct strata based on different sectors of the economy. Subsequently, we randomly selected two strata from the entire pool of strata. Within each of the chosen strata, we further selected companies based on their stock's performance relative to the broader market.


NSE STOCKS
STRATIFIED SAMPLING
Agricultural
Automobiles and accessories
Commercial and services
Construction and Allied
Energy and Petroleum
Insurance
Investment
Manufacturing and Allies
Telecommunication
Real Estate Investment Plan
Exchange Traded Funds
Investment services
RANDOM SAMPLING
BANKING
MANUFACTURING
STOCK PERFORMANCE
KCB GROUP
EQUITY HOLDINGS LTD
BAT
EABL

3.5 The Instruments

MODELS

3.5.1 Long Short Term Memory

LSTM is a specific kind of RNN with a wide range of applications similar to time series analysis, document classification, speech, and voice recognition. In contrast with feedforward ANNs, the predictions made by RNNs are dependent on previous estimations. In real, RNNs are not employed extensively because they have a few deficiencies which cause impractical evaluations. The difference between LSTM and RNN is that every neuron in LSTM is a memory cell. The LSTM links the prior information to the current neuron.By the internal gate, the LSTM is able to solve the long-term dependence problem of the data.

The Long Short-Term Memory (LSTM) model comprises multiple LSTM cells, each equipped with three vital gates: an input gate, an output gate, and a forget gate. These gates control the flow of information within the model, with the input gate determining the relevant information passed to the next cell, and Equations (1) and (2) show its related formulas where h_{t-1} is output at the prior time (t − 1), and X_t is input at the current time (t) into Sigmoid function (S(t)). All W and b are the weight matrices and bias vectors that require to be learned during the training process. f_t defines how much information will be remembered or forgotten. The input gate defines which new information remember in cell state by Equations (3)–(4). The value of i_t is generated to determine how much new information cell state need to be remembered. A tanh function gains an election message to be added to the cell state by inputting the output h_{t-1} at the prior time (t − 1) and adding the current time t input information X_t. C_t gets the updated information that must be added to the cell state (Equation (5)). The output gate regulating information transfer to the subsequent layer, and the forget gate handling the retention or dismissal of previous time step information. The value of O_t is between 0 and 1; which is employed to indicate how many cells state information that need to output (Equation (6)). The result of h_t is the LSTM block’s output information at time t (Equation (7)) The cells are stacked in a feed-forward fashion, with the final output derived from the last LSTM cell. During training, optimization algorithms like stochastic gradient descent (SGD) or Adam are used to adjust the model's weights, minimizing the error between predicted and actual outputs. Once trained, the LSTM model is capable of making predictions on new data, utilizing its ability to capture long-term dependencies effectively in tasks such as time series forecasting.

$$f_t = \sigma(W_f.[h_{t-1}, X_t] + b_f) \quad (1)$$

$$S(t) = 1/\ 1 + e^{-t} \quad (2)$$

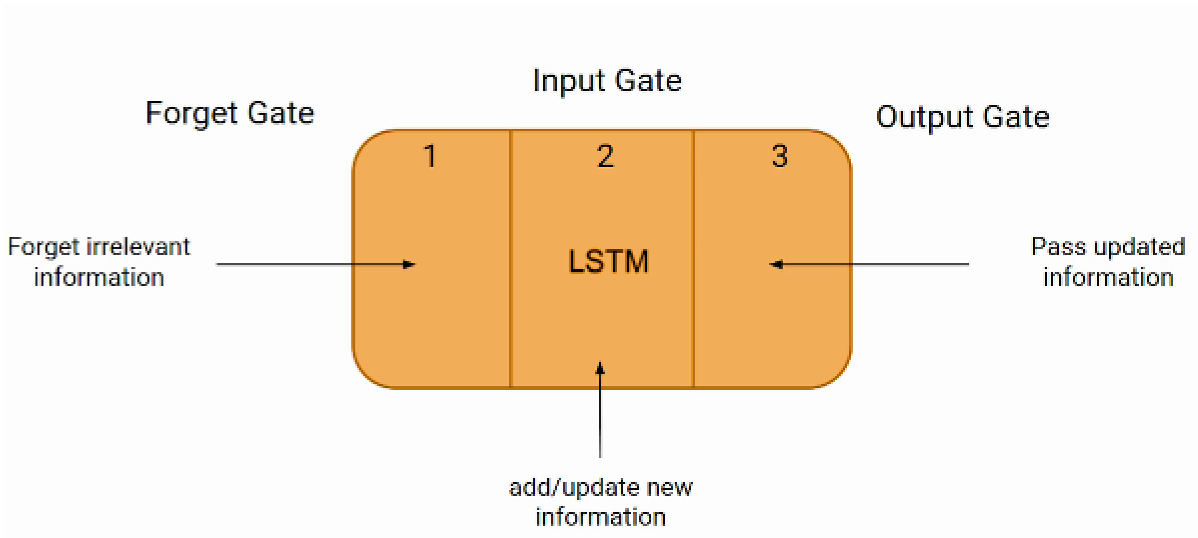
$$I_t = \sigma(W_i\ .[h_{t-1}, X_t\ ] + b_i) \quad (3)$$

$$\hat{C}\ _t = \tanh\ (W_c.[h_{t-1}, x_t] + b_c) \quad (4)$$

$$C_t = f_t \times C_{t-1} + i_t \times \hat{C} \quad (5)$$

$$O_t = \sigma(W_0.[h_{t-1}, X_t] + b_0) \quad (6)$$

$$h_t = 0_t \times \tanh\ (C_t) \quad (7)$$



3.5.2 Gated Recurrent Unit

The Gated Recurrent Unit (GRU) is a simplified version of the LSTM model, designed to capture long-term dependencies in sequential data. It also uses gates to control the flow of information, but it has only two gates: an update gate and a reset gate. The Gated Recurrent Unit (GRU) model is a type of recurrent neural network comprising multiple GRU cells. Each GRU cell is equipped with two fundamental gates: an update gate (z_t) and a reset gate (r_t). These gates allow the model to control the flow of information and handle long-term dependencies in sequential data. The update gate determines how much of the previous cell's hidden state should be retained and how much of the new candidate state should be integrated. On the other hand, the reset gate controls which portions of the past information to forget. Below are the equations for these gates respectively:

$$z_t = \sigma(w_z.[h_{t-1}, x_t]) \quad (8)$$

$$r_t = \sigma(w_r.[h_{t-1}, x_t]) \quad (9)$$

The candidate hidden state represents the new information that could be added to the hidden state.

$$h_t\sim = \tanh(w_h.[r_t \times h_{t-1}, x_t]) \quad (10)$$

The hidden state at the current time step is a combination of the previous hidden state and the candidate hidden state, controlled by the update gate.

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times h_t\sim \quad (11)$$

where:

z_t is the update gate at time step t.

h_{t-1} is the hidden state of the previous time step (t-1).

x_t is the input at the current time step t.

w_z is the weight matrix for the update gate.

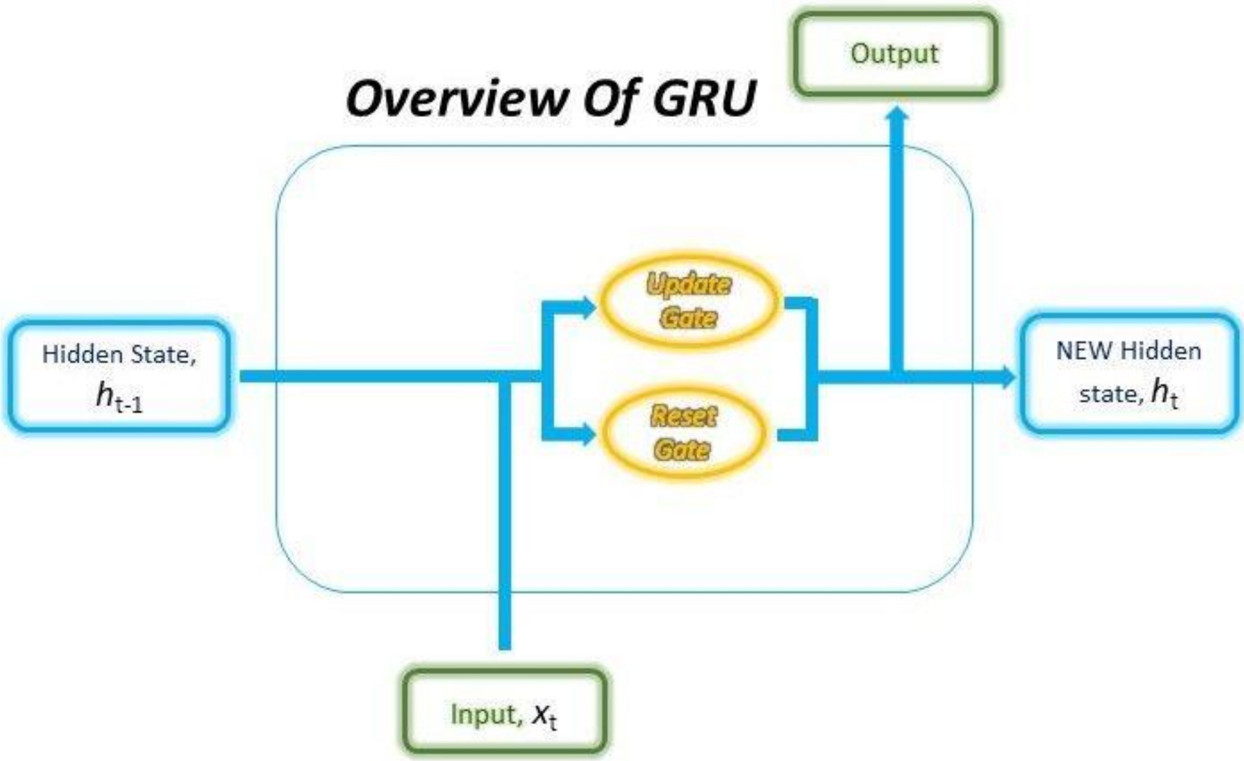
r_t is the reset gate at time step t.

$h_t\sim$ is The candidate hidden state at time step t.

w_h is the weight matrix for the candidate hidden state.

h_t is The hidden state at time step t.

Like LSTM, the GRU cells are stacked in a feed-forward manner, with the final output derived from the last GRU cell. During training, optimization algorithms such as stochastic gradient descent (SGD) or Adam are used to adjust the model's weights, minimizing the error between predicted and actual outputs. Once trained, the GRU model is capable of making predictions on new data, leveraging its ability to effectively capture long-term dependencies in tasks such as time series forecasting.



3.6 Data collection

The study will use secondary data from companies listed in the Nairobi Securities Exchange and licensed vendors from year 2018 to 2023. This data will typically include daily price information such as opening price, closing price, low price and trading volume. Macroeconomic variables data will be obtained from reputable economic databases such as; The Central Bank of Kenya and Kenya National Beareu of Statistics.

3.6.1 Variables description

Historical prices data

Historical price data refers to the time series of past stock prices for a given security. It provides valuable information about the historical trends, patterns, and volatility in the stock market. By incorporating historical price data as an input variable in LSTM and GRU models, deep learning algorithms can learn from the historical price patterns and potentially capture complex relationships between past price movements and future stock performance.

Macroeconomic variables

Gross Domestic Product (GDP) reflects the overall economic health, and its growth impacts various sectors and market optimism. Inflation rate affects consumer spending and corporate profits, while interest rates influence borrowing costs and economic activity. By including these macroeconomic variables in deep learning models, we will be able identify complex relationships between economic factors and stock market movements, leading to more informed and accurate stock price predictions, despite the inherent challenges in financial markets. Changes in interest rates can influence investment decisions. Lower interest rates can encourage borrowing and investment, potentially boosting stock prices. Major stock market indices can act as general proxies for market sentiment and overall market performance. LSTM and GRU models can learn patterns and relationships between these indices and individual stocks to make more informed predictions.

Technical indicators

Technical indicators are commonly used in stock prediction models, including those based on LSTM and GRU architectures. These indicators are mathematical calculations based on historical stock price and volume data that provide insights into the market's momentum, trends, and potential reversal points. Below is a brief descriptions of some of these indicators:

Weighted 14-day moving average

Similar to the simple moving average, the weighted moving average assigns different weights to each closing price within the period. This can be useful in capturing short-term price movements with higher weights on recent prices, and long-term trends with lower weights on older prices.

Weighted 14-day moving average=$\frac{n*C_t+(n-1)*C_{t-1}+\cdots+C_{t-n+1}}{n+(n-1)+\cdots+1}$ (12)

Momentum

Momentum measures the difference between the current closing price and the closing price "n" periods ago. Including momentum as an input variable can provide information about the strength and direction of recent price movements, which can help the deep learning model identify potential continuation or reversal patterns.

Momentum=$C_t - C_{t-n+1}$ (13)

Stochastic K%

The Stochastic Oscillator, represented by %K, compares the current closing price to the highest and lowest prices over the last "n" periods. %K represents the relative position of the current price within this range. By including Stochastic %K as an input variable, the model can identify overbought and oversold conditions, which may signal potential reversals in stock price trends.

StochasticK%=$(\frac{C_t-LL_{t_{t-n+1}}}{HH_{t_{t-n+1}}-LL_{t_{t-n+1}}})*100$ (14)

Stochastic D%:

Stochastic %D is a smoothed version of Stochastic %K, providing a more stable signal for the deep learning model. By using both %K and %D, the model can capture short-term and longer-term trends in the stock's price movements.

StochasticD%=$(\frac{K_t+K_{t-1}+\cdots+K_{t-n+1}}{n})*100$ (15)

Relative Strength Index (RSI)

RSI measures the magnitude of recent price gains and losses over the last "n-1" periods. It quantifies the stock's momentum and can help the deep learning model identify potential price reversal points and overbought/oversold conditions.

Relative Strength index(RSI)=100- $\frac{100}{1+\frac{\sum_{i=1}^{n-1} UP_{t-i}}{\sum_{i=1}^{n-1} DW_{t-1}}}$ (16)

Signal (n)

Signal $(n)_t$ is used in the calculation of the Moving Average Convergence Divergence (MACD). By including the signal line as an input variable, the model can better analyze MACD crossovers and potential trend changes.

$\text{Signal}(n)_t = \text{MACD}_t * \frac{2}{n+1} + \text{Signal}(n)_{t-1} * (1 - \frac{2}{n+1})$ (17)

Larry William’s R %

Larry William's R %, similar to Stochastic %K, indicates the position of the current price relative to the highest and lowest prices over the last "n" periods. Including this indicator can provide additional insights into overbought and oversold conditions.

Larry William’s R%= $(\frac{HH_{t_{t-n+1}} - C_t}{HH_{t_{t-n+1}} - LL_{t_{t-n+1}}})*100$ (18)

Accumulation/Distribution (A/D) oscillator

The A/D oscillator considers the relationship between the high, low, and closing prices, providing information about the accumulation or distribution of the asset. This can be helpful for the deep learning model in identifying potential buying or selling pressure in the market.

Accumulation/ Distribution (A/D) oscillator = $\frac{H_t - C_t}{H_t - L_t}$ (19)

CCI (Commodity Channel Index)

CCI measures the current price relative to its average over "n" periods, taking into account the mean absolute deviation. By including CCI as an input variable, the deep learning model can gain insights into potential trend reversals and extreme price movements.

CCI (commodity channel Index) = $\frac{M_t - SM_t}{0.015 D_t}$ (20)

Where

n is the number of days

C_t is the closing price at time t

L_t and H_t is the low price and high price at time t, respectively

$LL_{t_t - n+1}$ and $HH_{t_t - n+1}$ is the lowest low and highest high prices in the last n days, respectively

UP_t and DW_t means upward price change and downward price change at time t, respectively

$$\text{EMA}(K)_t = \text{EMA}(K)_{t-1} * \left(1 - \frac{2}{k+1}\right) + C_t * \frac{2}{k+1} \quad (21)$$

Moving Average Convergence Divergence$(MACD)_t$= $\text{EMA}(12)_t$-$\text{EMA}(26)_t$ (22)

$$M_t = \frac{H_t + L_t + C_t}{3} \quad (23)$$

$$SM_t = \frac{\sum_{i=0}^{n-1} M_{t-1}}{n} \quad (24)$$

$$D_t = \frac{\sum_{i=0}^{n-1} |M_{t-i} - SM_t|}{n} \quad (25)$$

3.7 Process and evaluation

3.7.1 Process

In this research, we aim to explore the application of deep learning models, specifically Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), for stock prediction. Before using information for the training process, it is vital to take a preprocessing step. We employ data cleaning, which is the process of detecting and correcting inaccurate records from a dataset and refers to identifying inaccurate or irrelevant parts of the data and then replacing, modifying, or deleting the dirty data. The interquartile range (IQR score) is a measure of statistical dispersion and is robust against outliers, and this method is used to detect outliers and modify the dataset.

Indeed, as an important point, to prevent the effect of the larger value of an indicator on the smaller ones, the values of technical indicators for all groups are normalized independently. Data normalization refers to rescaling actual numeric features into a 0 to 1 range and is employed in machine learning to create a training model less sensitive to the scale of variables. The data processing and analysis will involve several key steps. Firstly, historical stock market data will be collected, including daily or intraday price and volume data, as well as relevant financial indicators. Next, the data will be preprocessed to handle missing values, normalize the features, and potentially perform feature engineering to extract relevant patterns. Subsequently, the dataset will be split into training, validation, and testing sets to train and evaluate the LSTM and GRU models. The deep learning models will be implemented using appropriate libraries such as TensorFlow or PyTorch. We will experiment with various hyperparameters and architecture configurations to optimize model performance. Additionally, backtesting and cross-validation techniques will be employed to assess the robustness of the predictive models. The research will conclude with a comprehensive analysis of the LSTM and GRU models' performance in predicting stock prices, evaluating their accuracy, precision. The results will be discussed to draw meaningful conclusions regarding the feasibility and effectiveness of deep learning techniques in stock market prediction.

3.7.2 Evaluation

Evaluation Measures

In this section four metrics used in the study are introduced.

Mean Absolute Percentage Error

Mean Absolute Percentage Error (MAPE) is often employed to assess the performance of the prediction methods. MAPE is also used as a measure of prediction accuracy for forecasting methods in the machine learning area, it commonly presents accuracy as a percentage. Its equation is shown below:

$$\text{MAPE}=\frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right| \tag{26}$$

Where y_i is the actual value and $\hat{y}_i$ is the forecast value. In the formula, the absolute value of the difference between those is divided by $\hat{y}_i$ .The absolute value is summed for every forecasted value and divided by the number of data. Finally, the percentage error is made by multiplying to 100.

Mean Absolute Error

Mean absolute error (MAE) is a measure of the difference between two values. MAE is an average of the difference between the prediction and the actual values. MAE is a usual measure of prediction error for regression analysis in the machine learning area. The formula is shown below:

$$\text{MAE}=\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{27}$$

Where y_i is the true value and $\hat{y}_i$ is the prediction value. In the formula, the absolute value of the difference between those is divided by *n* (number of samples) and summed for every forecasted value.

Relative Root Mean Square Error

Root Mean Square Error (RMSE) is the standard deviation of the prediction errors in regression work. Prediction errors or residuals show the distance between real values and a prediction model, and how they are spread out around the model. The metric indicates how data is concentrated near the best fitting model. RMSE is the square root of the average of squared differences between predictions and actual observations. Relative Root Mean Square Error (RRMSE) is similar to RMSE and this takes the total squared error and normalizes it by dividing by the total squared error of the predictor model. The formula is shown below:

$$\text{RMSE}= \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{28}$$

Where y_i is the observed value, $\hat{y}_i$ is the prediction value and n is the number of samples.

Mean Squared Error

The Mean Squared Error (MSE) measures the quality of a predictors and its value is always non-negative (values closer to zero are better). The MSE is the second moment of the error (about the origin), and incorporates both the variance of the prediction model (how widely spread the predictions are from one data sample to another) and its bias (how close the average predicted value is from the observation). The formula is shown below:

$$\text{MSE}=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \tag{29}$$

Where y_i is the observed value, $\hat{y}_i$ is the prediction value and n is the number of samples.

REFERENCES

Aghababaeyan, R., & TamannaSiddiqui, N. (2011). Forecasting the Tehran Stock Market by Artificial Neural Network. International Journal of Advanced Computer Science and Applications, Special Issue on Artificial Intelligence .

Bhattacharjee, I., & Bhattacharja, P. (2019). Stock Price Prediction: A Comparative Study between Traditional Statistical Approach and Machine Learning Approach. International Journal of Computer Applications, 975(8887), 1-6.

Cao, Y., Tay, F. E. H., & Yao, L. (2019). Deep LSTM with Attention for Large Vocabulary Language Modeling. Proceedings of the 2019 Conference of the Association for Computational Linguistics (ACL 2019).

Cho, K., van Merrienboer, B., Bahdanau, D., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.

Deng, S., Mitsubuchi, T., Shioda, K., Shimada, T., & Sakurai, A. (2011). Combining Technical Analysis with Sentiment Analysis for Stock Price Prediction. 2011 IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing, (pp. 800-807).

Gao, Y., Wang, R., & Zhou, E. (2021). Stock Prediction Based on Optimized LSTM and GRU Models. Journal of Computational and Theoretical Nanoscience, 18(8), 4912-4918.

Government of Kenya. (2009). The Capital Markets Act, Laws of Kenya, Cap 485A.

Graham, B. (2003). The Intelligent Investor (Revised ed.). New York: HarperCollins

Guo, Y. (2022). Stock Price Prediction Using Machine Learning. Journal of Finance and Investment Analysis, 1(1), 45-57.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

Hota, J., Chakravarty, S., Paikaray, B. K., & Bhoyar, H. (2020). Stock Market Prediction Using Machine Learning Techniques. International Journal of Advanced Research in Computer Science, 11(2), 35-42.

Investopedia. (2021, April 15). Stock Exchange. Retrieved September 9, 2021, from https://www.investopedia.com/terms/s/stockexchange.asp

Investopedia. (2021, April 15). Stock. Retrieved September 9, 2021, from https://www.investopedia.com/terms/s/stock.asp

Khan, Z. H., Alin, T. S., & Hussain, A. (2011). Price Prediction of Share Market using Artificial Neural Network (ANN). International Journal of Computer Applications (0975– 8887) , 22 (2).

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics, 5(4), 115-133.

Ndiritu, J. M. (2010). Technical Trading Support System (TTSS): A Stock Market Analyst Support System. MSc. Thesis, University of Nairobi, School of Computing and Informatics, Nairobi.

Pan, H., Tilakaratne, C., & Yearwood, J. (2005). Predicting Australian Stock Market Index Using Neural Networks Exploiting Dynamical Swings and Intermarket Influences. Journal of Research and Practice in Information Technology , 37 (1).

Publishers Inc.

Retrieved April 12, 2013, from Capital Markets Authority :http://cma.or.ke/index.php?option=com_docman&view=docman&Itemid=123 92

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and retrieval. Psychological review, 65(6), 386-408.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. Psychological review, 65(6), 386-408.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533-536.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533-536.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural Networks, 61, 85-117. Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.

Shahi, T. B., Shrestha, A., Neupane, A., & Guo, W. (2020). Stock Price Forecasting with Deep Learning: A Comparative Study. IEEE Transactions on Neural Networks and Learning Systems, 31(4), 1144-1155.

Viveka Priya, N., & Geetha, S. (2022). Stock Prediction Using Machine Learning Techniques. Journal of Finance and Economics, 10(2), 123-137.

WORK PLAN

Action Plan	May	June	July	Aug	Sep	Oct	Nov
Study the theoretical Background							
Literature review							
Methodology and proposal Presentation							
Data analysis and findings							
Prepare the project report							

Gantt chart