Original Article

FRAMEWORK FOR PRICE PREDICTION IN THE NIGERIAN STOCK MARKET USING DATA MINING TECHNIQUES

Hafsat Amina Jibril and Abdullahi Salisu Kadir

Computer Science Department Modibbo Adama University, Yola, Adamawa State

DOI: https://doi.org/10.5281/zenodo.13902865

The stock market has become an integral part of the modern financial system, playing a significant role in shaping the global economy. The stock market prices serve as a key indicator of market sentiment, economic performance and investor confidence, with a growing demand for accurate stock market predictions. This research explores a data mining framework for Nigerian stock market predictions using Decision Tree, Support Vector Regression, and Artificial Neural Network techniques. Data from Dangote Sugar Refinery is used to create new variables for models. The models are evaluated using MAPE and MSE, and show promising outcomes. The combined models have the highest accuracy at 85%, capturing both short-term and long-term trends effectively. The study recommends further development of advanced sentiment analysis and dynamic model adaptation for better implementation.

Keywords: Stock market prediction, Data mining framework, Decision Tree, Support Vector Regression, Artificial Neural Network.

Introduction

The Nigerian stock market plays a vital role in the country's economy, offering a platform for companies to raise capital and investors to trade securities. However, predicting stock prices is a challenging task due to market volatility and unpredictability. Machine learning (ML) has emerged as a powerful tool in addressing these challenges by enabling the analysis of large datasets and detecting patterns not easily identified through traditional methods. This study explores a data mining framework tailored for the Nigerian stock market to improve stock price predictions, focusing on Dangote Sugar Refinery.

Leveraging advanced analytical techniques like machine learning, researchers can extract actionable insights from historical stock data, enhancing decision-making for investors. The use of algorithms such as Random Forest, Decision Trees, Support Vector Regression, and Artificial Neural Networks aims to provide high accuracy and precision in forecasting stock movements. As the stock market is influenced by various factors—market trends, financial data, and economic indicators—this framework aims to provide a deeper understanding of stock price behavior, allowing investors to make informed decisions.

The increasing interest in ML techniques for stock market prediction stems from their ability to handle vast amounts of data, manage unstructured information, and adapt to changing market conditions. This study contributes to the growing body of research by evaluating the effectiveness of ML algorithms in predicting Nigerian stock prices, providing insights into their potential benefits for investors and financial sectors.

Statement of the Problem

Predicting the direction of the stock market is a challenging problem that many investors grapple with, despite the country's tough economic period that we are facing today. Using historical data is one approach that can provide valuable input into potential stock market behavior. Another approach to predicting stock market behavior is by analyzing social media sentiment, particularly Twitter. By monitoring millions of tweets, sentiment analysis algorithms can gauge public opinion about a particular stock or company. Positive sentiment may indicate a potential increase in stock prices, while negative sentiment may suggest a decrease.

However, it is important to note that sentiment analysis is not foolproof and can be influenced by noise or manipulation (Kedar, 2021). News events also play a crucial role in stock market predictions. Major announcements, such as earnings reports, mergers and acquisitions, or regulatory changes, can significantly impact stock prices.

Traders and investors closely follow news sources and use the information to make informed decisions about buying or selling stocks. External factors, such as natural calamities or global financial disturbances, can also affect stock market behavior. For example, a natural disaster can disrupt supply chains or impact consumer behavior, leading to a drop in stock prices. Similarly, global financial crises can create panic among investors and cause a significant decline in stock markets worldwide.

Furthermore, predicting stock market behavior is a challenging task due to its complex nature and the multitude of factors that influence stock prices. While machine learning algorithms and sentiment analysis can provide insights, they are not foolproof and should be used in conjunction with other methods and ensemble techniques for sentiment analysis (Gite *et al*. 2023) proposed by the research community for financial instrument price forecasting.

Aim and Objectives of the Study

The goal of this research is to build a robust model to predict the Nigerian Stock market. Also, to gain valuable insights into future market movements and enhance decision-making processes related to investments in the company. The objectives are to:

i. Collect and preprocess the data into training and testing datasets.

ii. Features Selection for predicting stock prices using the financial dataset. This includes High, Low, Open, and Close prices, Trade volume, Microeconomic indicators, Technical indicators, Fundamental indicators, and Sentiment analysis data of DSR.

iii. Develop prediction models using machine learning techniques namely: Support Vector Regression, Decision tree, and Artificial Neural Networks.

iv. Training of the model to validate and evaluate the models to optimize market performance.

v. Develop a prototype application to demonstrate the framework.

vi. Evaluate the prediction performance of the model using MAPE and MSE

Conceptual Review Nigerian Stock Market Price Prediction

Nigerian Stock Market Price Prediction Stock market prediction has always been a challenging task for investors and financial research. With the advent of machine learning techniques, researchers have explored various algorithms to predict stock market trends accurately (Sarode *et al*., 2019). This literature review aims to provide an overview of the existing studies on the Nigerian stock market using machine learning data techniques (Hiransha *et al*., 2018). However, the introduction of ML methods has caused a significant change in the field, leading to a rise in alternative approaches for predicting stock market trends.

Recent research has utilized regression models, time series analysis using ARIMA and long short-term memory networks (LSTM), classification models like decision trees and support vector machines (SVM), and DL methods involving neural networks and CNNs (Ruotong*et al*., 2023). These approaches perform successful outcomes and improve predictive power. Nevertheless, upon closer examination, there are ongoing research gaps and concerns that continue despite these developments.

It is crucial to thoroughly examine these gaps and unsolved difficulties in the current literature to guide future research efforts toward more reliable and practical solutions in predicting stock market trends. Hence, this review study intends to critically examine existing literature on both traditional and ML approaches to stock market prediction (SMP) Qijun *et al*., (2023). These algorithms are capable of analyzing large volumes of historical stock data and identifying patterns that can be used to predict future stock prices (Cheng *et al*., 2023). Sentiment analysis has gained significant attention in stock market prediction.

Researchers have incorporated sentiment analysis of news articles, social media posts, and financial reports to capture the impact of public sentiment on stock prices. By considering the sentiment of news articles related to the company, machine learning models can better predict stock market trends (Baheti *et al*., 2021). Various evaluation metrics are used to assess the performance of stock market prediction models. Commonly used metrics include Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Accuracy. Researchers often compare the performance of different machine learning algorithms to identify the most accurate model for stock market prediction (Bansal *et al*., 2022). Machine learning techniques have shown promising results in stock market prediction analysis for companies. In the study by Abdulhamit *et al*., (2021) utilizing these techniques, investors and financial researchers can make informed decisions based on accurate predictions of stock market trends.

Machine Learning Techniques for Stock Market Prediction

The analysis of historical data using machine learning techniques, and predictors aims to identify key factors that influence the movement of stock prices, specifically in the case of DSR. Support Vector Regression (SVR), Random Forests (RF), Decision tree (DT) and Artificial Neural Networks (ANN) are classification algorithms commonly used in stock market prediction (Yue *et al*., 2023). These algorithms use historical data to classify future stock prices into different categories, such as "increase" or "decrease". By training these models with historical data, they can learn patterns and relationships that can be used for predicting future stock prices. Regression models like Autoregressive Integrated Moving Average (ARIMA) and Exponential Smoothing (ES) are used to forecast numerical values, such as stock prices. These models analyze patterns in the historical data to understand the underlying trends and seasonality in the stock prices. They then use this information to make predictions about future prices. By applying these machine learning and data mining techniques to historical data, predictors can extract meaningful patterns and relationships.

This analysis helps to identify key factors that influence the movement of Nigerian stock market prices prediction. These factors could include macroeconomic indicators, industry specific data, financial performance of the company, news sentiment, or other relevant factors. By considering and analyzing these factors, predictors aim to make accurate predictions about future stock prices.

Feature Selection of Prediction Model

Feature Selection of Forecast Model Characteristics Feature selection and engineering are useful for data enrichment and extracting important information from the original data set, which can improve the accuracy of predictions. This is also a popular line of research that interests researchers. Previous studies generate various useful features from the original data set to improve stock price forecasting performance, such as stock investor sentiment, stock movement charts, and economic events. and key politicians. Among them, the emotions of stock investors are generally one of the most adopted characteristics in stock price prediction models Jorgenson *et al*. (2023).

Empirical Review

Machine learning has indeed been actively applied to stock market price forecasting. By leveraging historical stock prices, trading volumes, and a wide range of other financial data, machine-learning algorithms can be trained to identify patterns and relationships that can help predict future price movements Obthong *et al*. (2020).

The study conducted by Bhoyar *et al*. (2022) to predict stock market prices using machine learning techniques: Random Forest, Artificial Neural Network, Decision Tree and Support Vector Regression The study found that Random Forest outperformed the other three models with Artificial Neural Network performing slightly better which is the second least and Decision Tree model exhibited overfitting to the dataset and Support Vector Regression performing the poorest. The study also highlighted the importance of considering past returns as a key variable for stock price prediction, particularly for opening prices.

Osman *et al*. (2014) in their paper titled " A Machine Learning Model for Stock Market Prediction " This study proposed a machine learning model for stock price prediction using financial technical indicators, integrating particle swarm optimization (PSO) algorithm and LSSVM. The PSO algorithm was used iteratively to optimize LSSVM and select its free parameters, C, ϵ, and γ. The proposed LS-SVM-PSO model overcomes the over-fitting problem found in ANN and is capable of identifying fluctuations in the stock sector. The performance of the proposed model is better than LSSVM and compared algorithms.

Osman *et al*. (2014) in their paper titled " A Machine Learning Model for Stock Market Prediction " This study proposed a machine learning model for stock price prediction using financial technical indicators, integrating particle swarm optimization (PSO) algorithm and LSSVM. The PSO algorithm was used iteratively to optimize LSSVM and select its free parameters, C, ϵ, and γ. The proposed LS-SVM-PSO model overcomes the over-fitting problem found in ANN and is capable of identifying fluctuations in the stock sector. The performance of the proposed model is better than LSSVM and compared algorithms.

Chen *et al*. used the LSTM model to predict China stock returns (Chen *et al*., 2015). Historical data were transformed into 30-day sequences with 10 learning features and 3-day learning rate labeling. The authors claimed that the model improved the accuracy from 14.3% to 27.2% compared to the random prediction method. Bao *et al*. applied Haar wavelet transform to denoize financial time series data and applied clustered autoencoders to learn deep features of data and then used LSTM to predict the closing price of the stock index (Bao *et al*., 2017). (Bao *et al*., 2017). Their average R score was below 88% on the LSTM model for S&P 500. Roondiwala *et al*. Used the LSTM model for the NIFTY 50 data ranges from 2011 to 2016 (Roondiwala, Patel, & Varma, 2017).

Original Article

The authors used fundamental data (open, close, low, and high) without incorporating macroeconomic and technical indicators to predict the closing price.

Kara *et al*. used a support vector machine (SVM) and an artificial neural network to predict the daily movement of the Istanbul Stock Exchange National 100 index from 1997 to 2007 (Kara *et al*., 2011). The authors selected ten technical indicators as input for their model. The experimental results showed that the average performance of the artificial neural network model was significantly better than that of the SVM mode

Khan *et al*. (2020) in their paper titled "Stock Market Prediction using Machine Learning Classifiers and Social Media, news" The research present a framework for predicting future trends in the stock market using news and social media as external factors. The study found that social media has a greater influence on stock prediction on day 9, while financial news has a greater effect on day 9 and then on day Combining sentiments of social media and financial news increased overall accuracies of most classifiers after day 3, but decreased the highest accuracy. The research examined the effect of feature selection and spam tweet reduction on prediction performance, finding a positive effect on most classifiers.

Awad *et al*. (2023) have utilized deep reinforcement learning to learn the stock market decision and the stock price prediction. The effectiveness of the proposed model has been compared with the models through experimental experiments to prove their performance accuracy rate. The results assured the efficiency behind using the ANN model exhibited exceptional performance, boasting accuracy in financial markets and decision-making.

Moghar *et al*. (2020) in their paper titled "Stock Market Prediction Using LSTM Recurrent Neural Network" This paper proposes an RNN based on LSTM for forecasting future values for GOOGL and NKE assets. The model shows promising results in tracing the evolution of opening prices for both assets. For future work, the authors plan to optimize the data length and number of training epochs to improve prediction Accuracy.

Vijh *et al*. (2020) in their paper titled "Stock Closing Price Prediction using Machine Learning Techniques" This paper discusses the challenge of predicting stock market returns due to the complexity of changing stock values and the limited features available in historical datasets. To improve accuracy, the paper creates new variables using existing features and uses artificial neural networks (ANN) and random forests (RF) to predict the next day's closing price. The results show that ANN outperforms RF based on RMSE, MAPE, and MBE values. The paper suggests future work could explore deep learning models that incorporate financial news articles and other financial parameters to potentially achieve even better results.

Wang *et al*. (2023) Conducted a study on stock market index prediction via the localized spatial-temporal convolutional network" The experiments show that the method is 4.2%, 3.1% and 40% better than conventional methods under these three metrics. The method obtains state of-the-art results in evaluation metrics of regression, classification and market backtesting.

Boppana *et al*. (2023) In their paper titled "Machine Learning Based Stock Price Prediction by Integrating ARIMA Model and Sentiment Analysis with Insights from News and Information"

This paper aims to improve stock price prediction by combining ARIMA's analytical capabilities with sentiment analysis. Experiments: The ensemble predictive model for stock prices demonstrates favorable outcomes. The Mean Absolute Error (MAE) is 0.8659, indicating accuracy, and the Root Mean Squared Error (RMSE) is 0.1732, showing the overall prediction error. The Mean Absolute Percentage Error (MAPE) is 1.8541, suggesting precision in comparison to actual stock prices. The R-squared value is 0.9804, indicating the ability to explain variation in stock price data. These findings highlight the effectiveness of providing reliable insights for investors in the dynamic stock market. The analysis with the ARIMA model to enhance stock price predictions.

Al-Bossly *et al*. (2023) "A Comparative Study on Market Index Prediction: LSTM vs Decision Tree Model" This thorough investigation evaluated the effectiveness of Long Short-Term Memory (LSTM) and Decision Tree models in predicting market trends, using historical stock market data from the Kingdom of Saudi Arabia spanning from May 1, 2018, to September 13, 2023, the study aimed to assess the prediction efficacy, reliability, and feasibility of these models, employing key performance metrics such as accuracy, (MASE), (MAPE), and (SMAPE) to gauge their effectiveness in predicting market index movements. The findings highlight crucial observations regarding the relative precision of the two models and their effectiveness in diverse market circumstances. Specifically, the Decision Tree model demonstrated superior accuracy in predicting stock market movements in the Kingdom of Saudi Arabia, evident in its lower MAPE.

Shaban etc. (2023). SMP-DL: A New Stock Price Forecasting Approach Based on Deep Learning for Effective Trend Forecasting. Neural Computing and Applications. In this paper, a new stock price forecasting system, namely stock price forecasting based on deep learning (SMP-DL), is presented. SMP-DL is divided into two phases, which are (i) data preprocessing (DP) and (ii) price prediction (SP2). In the first step, the data is pre-processed to obtain clean data through several

the steps are detection and rejection of missing values, feature selection and data normalization. Then, in the second stage (for example, SP2), the clean data will be passed through the predictive model used. In SP2, long-term memory (LSTM) combined with bidirectional recurrent unit (BiGRU) can predict the closing price of the stock market. The results obtained showed that the proposed system works well compared to other existing methods. The RMSE, MSE, MAE and R2 values are 0.2883, 0.0831,

0. 2099 and 0.9948. Moreover, the proposed method has been applied to different datasets and works well.

Wasiat Khan *et al*. (2020) in their paper titled "Stock Market Prediction using Machine Learning Classifiers and Social Media, news" The research present a framework for predicting future trends in the stock market using news and social media as external factors. The study found that social media has a greater influence on stock prediction on day 9, while financial news has a greater effect on day 9 and then on day Combining sentiments of social media and financial news increased overall accuracies of most classifiers after day 3, but decreased the highest accuracy. The research examined the effect of feature selection and spam tweet reduction on prediction performance, finding a positive effect on most classifiers.

Hiransha Ma *et al*. (2018) in their paper titled" NSE Stock Market Prediction Using Deep-Learning Models" The authors of this study used four deep learning architectures (MLP, RNN, LSTM, and CNN) to predict stock prices of TATA MOTORS from the NSE, as well as MARUTI, HCL, and AXIS BANK from NSE, and BANK OF AMERICA and CHESAPEAK ENERGY from NYSE The DL models were found to be capable of identifying patterns in both stock markets, indicating underlying dynamics common to both. The study also showed that DL models outperformed ARIMA models, and CNN performed better than the other three networks due to its ability to capture abrupt changes in the system.

Mehar *et al*. (2020) in their paper titled " stock Closing Price Prediction using Machine Learning Techniques" This paper discusses the challenge of predicting stock market returns due to the complexity of changing stock values and the limited features available in historical datasets. To improve accuracy, the paper creates new variables using existing features and uses artificial neural networks (ANN) and random forests (RF) to predict the next day's closing price. The results show that ANN outperforms RF based on RMSE, MAPE, and MBE values. The paper suggests future work could explore deep learning models that incorporate financial news articles and o ther financial parameters to potentially achieve even better results.

Polamuri *et al*. (2020) in their paper titled Learning" This paper discusses the challenges involved in predicting stock market prices and the high demand for accurate stock market predictions. Various prediction techniques, such as Artificial Neural Network, Neuro-Fuzzy System, Time Series Linear Models (TSLM), and Recurrent Neural Network (RNN), are explored and analyzed for their advantages and disadvantages. The paper aims to provide an overview of the different techniques used for the prediction of stock market.

Adil *et al*. (2020) in their paper titled "Stock Market Prediction Using LSTM Recurrent Neural Networks" This paper proposes an RNN based on LSTM for forecasting future values for GOOGL and NKE assets. The model shows promising results in tracing the evolution of opening prices for both assets. For future work, the authors plan to optimize the data length and number of training epochs to improve prediction Accuracy.

Sneha *et al*. (2020) in their paper titled "Applications of ANNs in Stock Market Prediction: A Survey" This paper surveys the use of artificial neural networks (ANNs) in stock market prediction and concludes that ANNs are highly effective in extracting useful information from large datasets, making them a promising approach in this field. Compared to other models such as genetic algorithms and multiple linear regression analysis, ANNs are significantly more accurate.

Method System Architecture

Our experiment follows the basic data mining process demonstrated in the flow chart

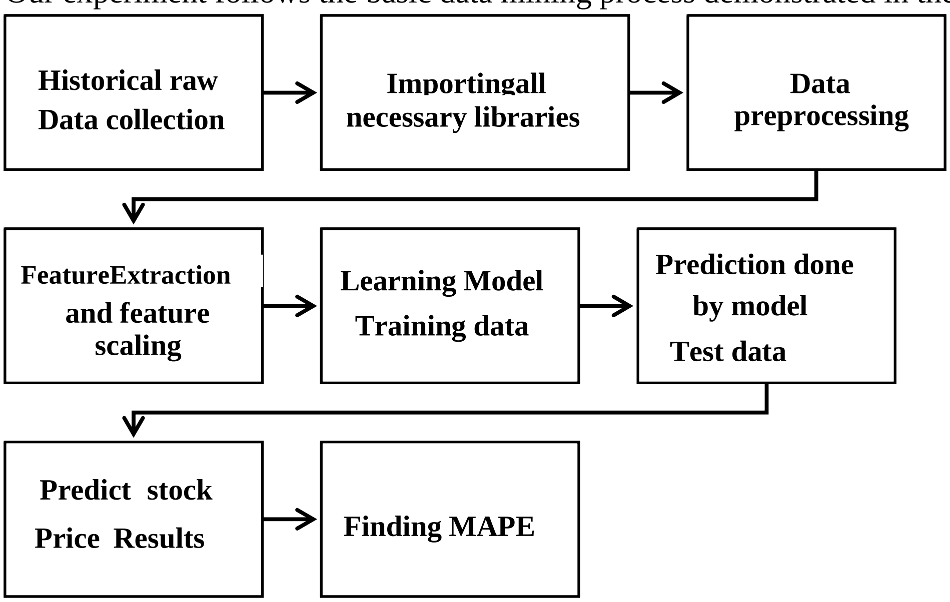

Figure 3.1: Model Development Processes for Stock Market Price Prediction

The goal of this research findings, is to build a robust model to predict the Nigerian Stock market. In this study, we applied a data mining framework to predict the 10-year performance of Dangote Sugar Refinery (DSR) in the Nigerian stock market. Using a data mining framework for predicting stock market prices.

In this research methodology, historical data was collected and the importation of all necessary libraries such as NumPy, Pandas, matplotlib, Seaborn, Bokeh and SciPy. Once the data was imported, various data preprocessing methods were performed to clean and prepare the data for analysis.

This included dropping unnecessary columns, checking for missing values using isnull, and handling any outliers or anomalies in the data. Feature extraction and feature scaling techniques and model selection, training, testing and evaluation, were then implemented using feature engineering to ensure that all features were on the same scale and to improve the performance of the machine learning models.

Data collection

The data collection process for predicting stock market prices, particularly for Dangote Sugar Refinery in the Nigerian stock market, involves gathering a wide range of financial and nonfinancial data. Key data sources include historical stock prices, financial statements (income statements, balance sheets, and cash flow statements), market trends, industry reports, and macroeconomic indicators such as government policies and consumer demand. Combining qualitative data, such as interviews with executives and experts, with quantitative data enhances the understanding of factors influencing stock prices.

Data preprocessing is crucial for preparing the collected data for analysis. This includes cleaning the data by handling missing values, removing outliers, normalizing numerical data, and eliminating irrelevant or redundant variables. The goal is to ensure the reliability of the dataset for model training by addressing issues such as missing or null values, outliers, and discrepancies in the data.

Feature selection, an essential step in building predictive models, identifies the most important factors contributing to stock price movements. Statistical analysis and domain knowledge are applied to select relevant features, such as company finances, market trends, and macroeconomic indicators, to enhance the accuracy of predictive models.

The study uses various machine learning models to forecast stock prices, including Artificial Neural Networks (ANN), Support Vector Regression (SVR), and Decision Trees. These models are trained on historical stock data and related indicators, and each model uses features like stock price variables (high, low, open, close), trade volume, and indicators such as technical, fundamental, and sentiment analysis to predict future stock prices of Dangote Sugar Refinery.

Despite the complexity of certain models, simpler and faster models like ANN, Random Forest, and SVR are preferred for real-time stock market predictions, catering to non-professional stock market participants. These models help investors and institutions make informed decisions based on future stock price trends.

Price Movement and Indicators

The Dangote Sugar Refinery dataset contains multiple variables: open, high, low, close, total trade quantity and turnover. The columns Open and Close represent the starting & final price at which the stock is traded on a particular day. High, Low and last represent the maximum, minimum & last price of the share for the day. Trade volume is the number of shares bought or sold in the day & Turnover is the turnover of the DSR on a given date and the number of trades.

Original Article

➢ Stock markets serve as an indicator of the state of the economy

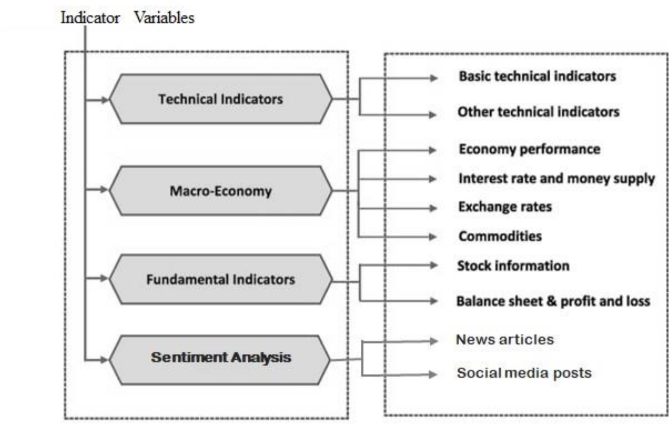

Figure 3.2: Feature Selection Process

➢ Data Flow Diagram

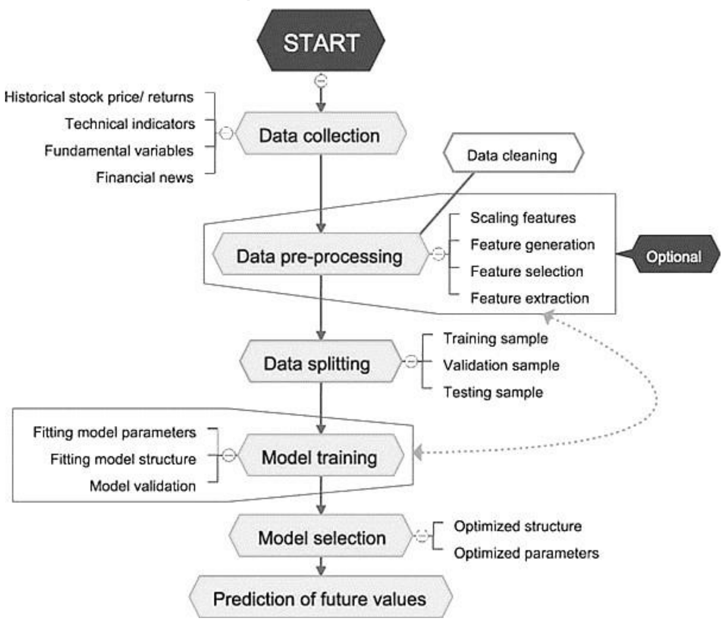

Figure 3.3: Data Flow Chart

The proposed data mining framework for predicting the performance of Dangote Sugar Refinery on the Nigerian stock market from 2014 to 2024 involves several key steps to ensure accurate predictions and valuable insights for investors. This framework addresses three critical perspectives: data collection, data preprocessing, and prediction modeling, all focused on achieving high accuracy in stock price predictions.

Data Collection:

Historical stock data for Dangote Sugar Refinery, including opening price, closing price, high and low prices, and trading volume from 2014 to 2024, is gathered from stock exchanges, financial websites, and economic indicators. This step ensures that comprehensive and relevant data is available for analysis and future predictions.

Data Preprocessing:

This step involves cleaning and transforming the raw data to ensure accuracy. Techniques like handling missing values, normalizing data, and removing outliers are applied to prepare the dataset for model training. Proper preprocessing ensures the reliability and integrity of the dataset for subsequent analysis.

Prediction Modeling:

Various machine learning algorithms, such as Support Vector Regression (SVR), Decision Trees, Random Forest, and Artificial Neural Networks (ANN), are employed to identify patterns in the historical data and predict future stock price movements. The prediction models take into account key financial and macroeconomic indicators that influence stock performance.

Performance Evaluation:

The models are evaluated using suitable accuracy metrics such as Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Mean Squared Error (MSE). These metrics measure the performance and precision of the models in predicting the stock prices of Dangote Sugar Refinery.

Insights for Investors:

By analyzing historical data, identifying trends, and using advanced data mining techniques, the framework generates predictions that can guide investors in making informed decisions in the volatile Nigerian stock market. The predictions consider market trends, industry factors, and economic conditions, providing valuable insights into the future performance of Dangote Sugar Refinery.

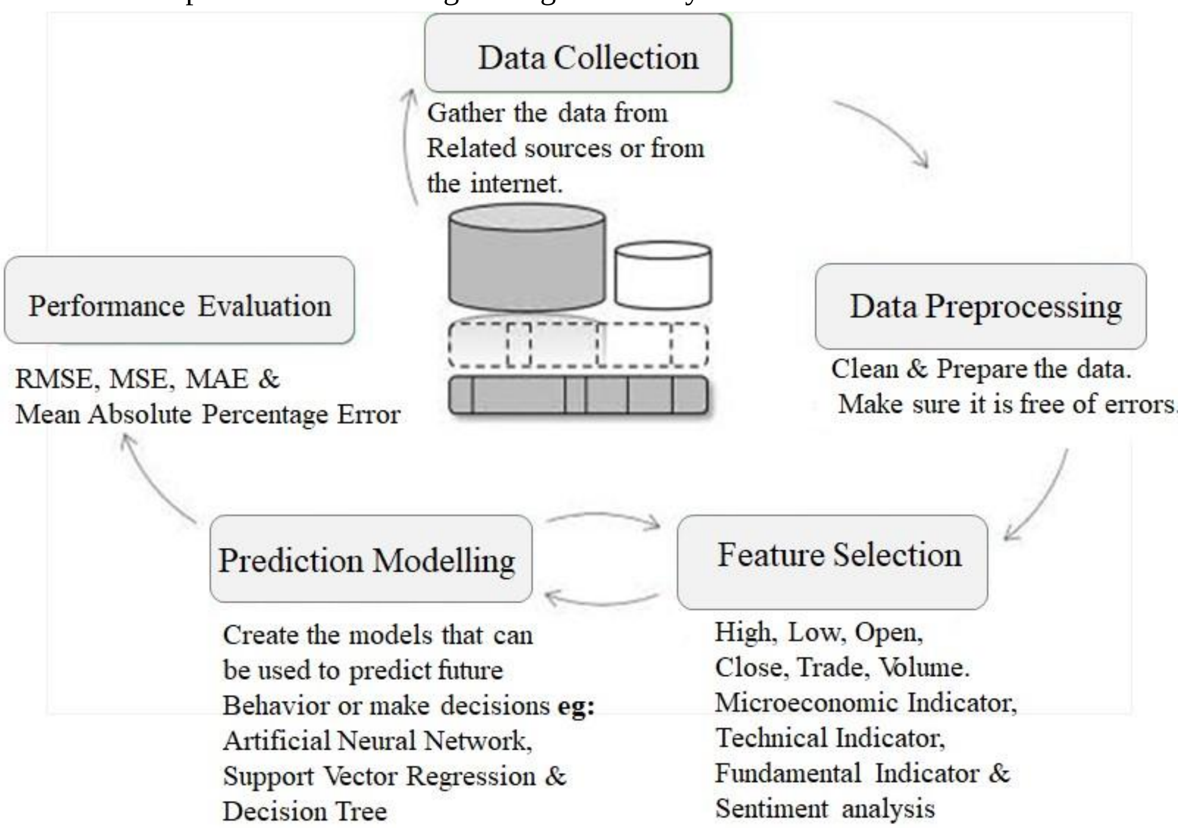

Figure3.4: Data Mining Framework

Performance Evaluation

Evaluation metrics are another important aspect of the data mining framework, as they provide a way to measure the accuracy and reliability of the predictions generated. By comparing the predicted values to the actual performance of Dangote Sugar Refinery, the framework can assess its effectiveness and make adjustments as needed.

The data mining framework also incorporates a variety of data mining algorithms, which are used to analyze the complex relationships between different factors that influence stock prices. By applying these algorithms to the data collected from the Nigerian Stock Exchange, the framework can identify patterns and trends that may not be immediately apparent to human analysts.

We used those accuracy metrics for financial markets to evaluate the prediction performance: mean absolute percentage error (MAPE), mean square error (MSE) and mean absolute error (MAE) that are appropriate indicators for the stock price with error tolerance point in time.

Principal Component Analysis (PCA)

In the context of the Nigerian stock market, Principal Component Analysis (PCA) can be a valuable tool for predicting the performance of specific stocks, such as Dangote Sugar Refinery. By applying PCA to the stock market data, investors and analysts can reduce the dimensionality of the data while retaining the most important information for making informed decisions.

In the case of Dangote Sugar Refinery, PCA can be used to identify the key factors that influence the stock price of the company. By analyzing the historical data of the stock, including factors such as market trends, company performance, and economic indicators, PCA can help identify the principal components that have the most significant impact on the stock price.

By reducing the dimensionality of the data through PCA, analysts can focus on the most relevant factors that drive the stock price of Dangote Sugar Refinery. This can lead to more accurate predictions and better-informed investment decisions.

However, PCA can also help in identifying any underlying patterns or relationships in the data that may not be immediately apparent. By transforming the data into a new space where the variables are uncorrelated, PCA can reveal hidden insights that can be valuable for predicting the future performance of the stock.

In this study, principal component analysis (PCA) is considered the gold standard in dimensionality reduction for its ability to simplify complex datasets while retaining important information. By applying PCA to predict the performance of Dangote Sugar Refinery's stock market price, investors and analysts in the Nigerian stock market can make more informed decisions and potentially improve their investment outcomes.

Data Mining Techniques

Data mining is the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition technologies as well as statistical and mathematical techniques. A process used by companies to turn raw data into useful information. By using software to look for patterns in large batches of data, businesses can learn more about their customers and develop more effective marketing strategies as well as increase sales and decrease costs. Data mining depends on effective data collection and warehousing as well as computer processing. Association Rules would be used in our project. They would help us in making more appropriate selections about the decision to take. With the advent of data mining technologies, there are many reasons for supporting its spread because it helps to uncover the meaning of the data in a large amount of unstructured and noisy data (Chen *et al*., 2022).

System Model design

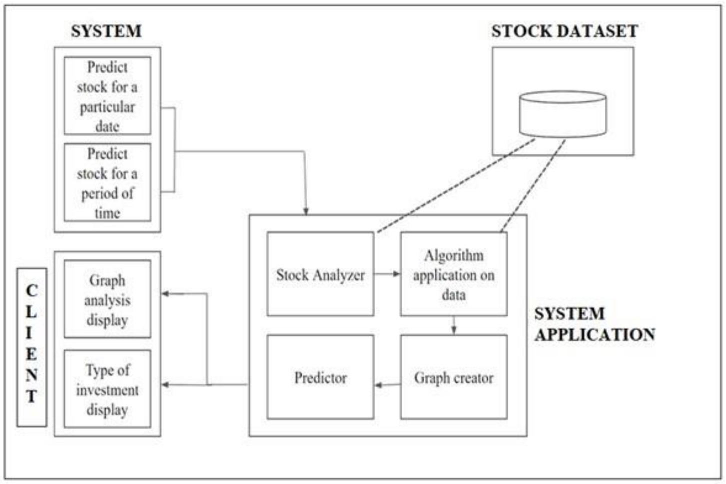

Figure 3:5: Workflow of the System Model Design

The System Model Design workflow can be summarized as follows:

Data Collection and Preprocessing:

✦ Gather the historical stock data for Dangote Sugar Refinery.

✦ Clean, preprocess, and prepare the data for model training and testing.

Model Development and Training:

✦ Develop the SVR, DT, and ANN models and train them on the stock dataset.

✦ Optimize the hyperparameters of each model to achieve the best performance.

Stock Price Prediction:

✦ Use the trained models to generate stock price forecasts for a given date and for a given period. Display and client interface:

Visualization and Client Interface:

✦ Create informative charts and visualizations to illustrate stock price forecasts.

✦ Develop a client interface that displays graphical analysis and provides investment recommendations based on forecasts.

Technical graphic approach

The design approach is essentially classified as a technical approach. This is a large amount of historical data for the stock prices in question. When analyzing stocks with technical analysis, the follower must consider two principles. "The first principle of technical analysis is that all information about a company's earnings, dividends and future returns are automatically reflected in the company's past prices.

The second principle is that prices tend to move in trends: a stock that is rising tends to keep going, while a stock that is flat tends to stay. The ability to find and read the charts of a stock market and analyze it is essential for the application of technical analysis of the stock market. Technical analysis is concerned with the study of charts and followers of this type of stock analysis are known as chartists and do not believe in information such as dividends or earnings like fundamentalists, because I believe they prevent me from accurately predicting the future of stocks.

A chart that shows past market prices and trading volumes of a stock in the past is all the technical analysis needs to analyze the future movement of the stock.

It is important to know that when a stock is growing, it is believed that it will continue because of the importance of the shares. Once a stock is in an uptrend, it will go down a little and rise again to form what looks like a shoulder, then go down a little and rise a little higher than it should go down and rise again to form the next shoulder technical analysis. requires action to fall below the line where the valleys are aligned, called If the stock falls below the neck line, it will sell quickly as this usually results in a downward trend similar to the upward trend that occurred before

Variable Model

This approach is working on examining the selected parameters analysis to predict the future price of stocks.

Fundamental Analysis Approach

This approach is alternately referred as true or real price prediction which primarily focuses on fundamentals of the company instead of price movement. It gives weightage to true value prediction instead of current price movement.

Machine Learning Algorithms

This method attempts to predict the movement of stock prices based on training given with the past value movements.

Support Vector Regression

It is a Supervised Machine learning algorithm used for regression analysis. It finds the function that helps us approximate mapping based on the training sample from an input domain to real numbers.

The Terminologies contained in this are Hyperplane -this is the line that is used to predict the continuous output. Kernel helps to find hyperplanes in higher dimensional space without increasing the computational cost of it and the decision boundary is a simplification line that differentiates positive examples and negative examples.

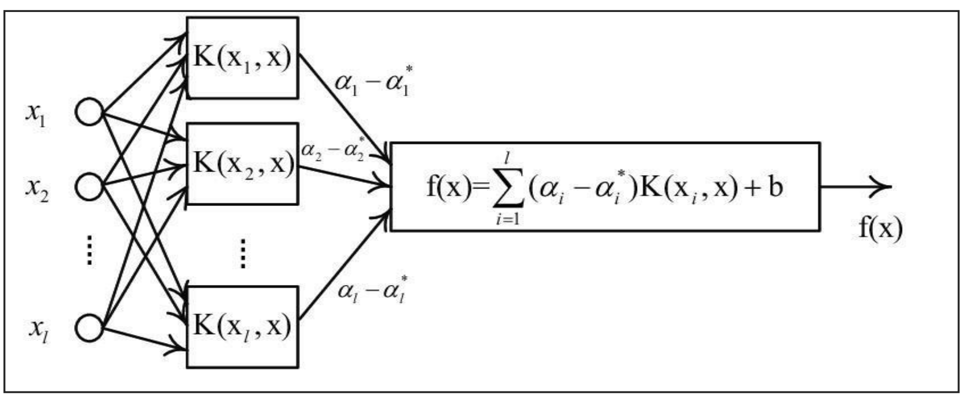

Figure 3.6:Support Vector Regression .

Decision Tree

It is a supervised ML, which is used for regressions and classification. That is why it is also called CART classification and regression trees. In this algorithm, two nodes are present, namely the decision node which is used to make decisions and can be divided into many branches and the leaf node which gives the output of the decisions and this node cannot be divided into many nodes Here is the formula for the leaf node:

Original Article

Leaf node $Infomation = ClassEndpypyAtt$

3. 1)Branches: Here the decision rules are defined according to which the nodes can be further divided. For prediction, it starts from the root node, compares the values of the current attribute with the root attribute, and based on this comparison, follows the branch and moves to the next node. This process continues until it reaches the leaf node of the tree. Entropy – This is a metric to measure the error in a given attribute. The formula to find the entropy is: - Entropy(s)= -P(po)log2 P(po)- P(no) log2 P(no)

(3.2) Here, (S) stands for the total number of samples. P (yes) refers to the probability of S and P (no) means the probability of no. (Prasad *et al*., 2020)

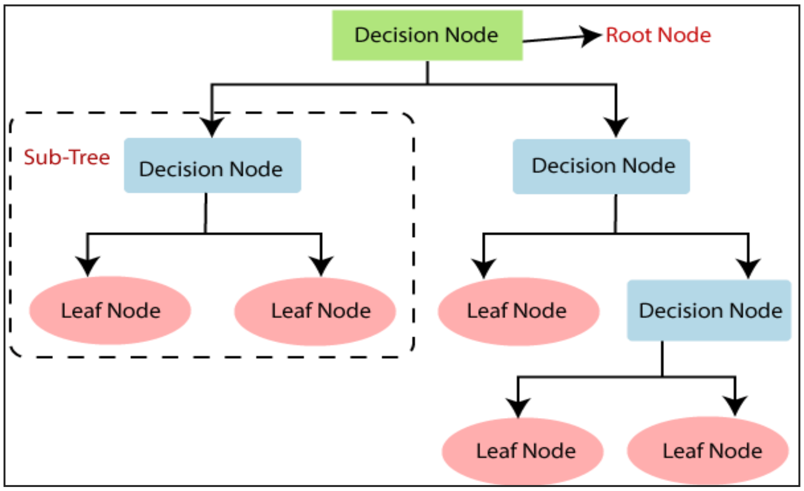

Figure 3.7: Decision tree classification process.

Artificial Neural Network (ANN)

Artificial neural network Artificial neural networks (ANNs) have gained popularity in the field of machine learning due to their ability to learn complex patterns and relationships in data. It contains three layers, first is the input layer - this layer receives various input variables from the user and then the hidden layer - This layer is present in the middle of the input layer which identifies all the hidden features and patterns and the end is the exit. layer - This layer ensures that the final ANN receives various inputs and multiply them by the weights defined for each with an activation function for the activation of neurons.

This initialize the size of layers along with the weights and biases. And also, define the sigmoid activation function and its derivative to introduce non-linearity in the network.

In the forward pass, we pass the input data through the neural network, to obtain the predicted output. Here calculates the output of the hidden layer by taking the dot product of the input and weights, adding the bias, and applying the sigmoid activation function. Through the backward pass, we compute the gradients of the output layer by calculating the loss and its gradient. We then update the weights and biases using the gradients and a specified learning rate.

To train the neural network, we run a specified number of epochs, performing the forward pass, backward pass, and weight updates in each iteration.

This implement a prediction function that allows us to make predictions on new data by performing a single forward pass through the neural network.

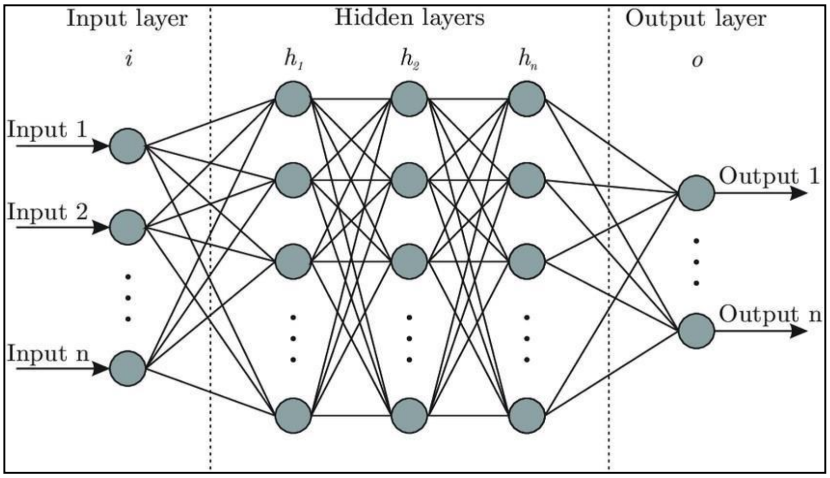

Figure 3.8 Artificial Neural Network Procedure .

The Stock Turnover Prediction Framework

The stock turnover prediction framework proposed is shown in Figure 3.8 The basic methodology involved Data Collection, Pre-processing, Feature Selection and Classification, each of which is explained below.

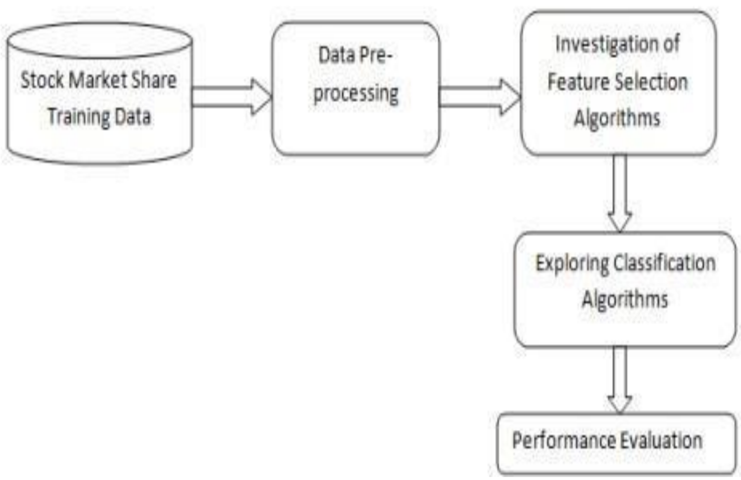

Figure. 3.9: Stock Turnover Prediction framework.

Results Experimental Results

The performance result demonstrates the efficacy of the models built to forecast the stock price of DSR. The performance of the models were measured using an evaluation metric that determines the ratio of MAPE, MSE & MAE of close values of the stock price over the testing period. For classification model, MAPE is the metric chosen to evaluate the model performance.

✦ Plotting the actual and predicted prices for DSR stocks.

✦ Implementation: Below are the performance results

Original Article

Result obtained format eh study is presented using figures and labels, with brief discussion under each figure.

System Implementation

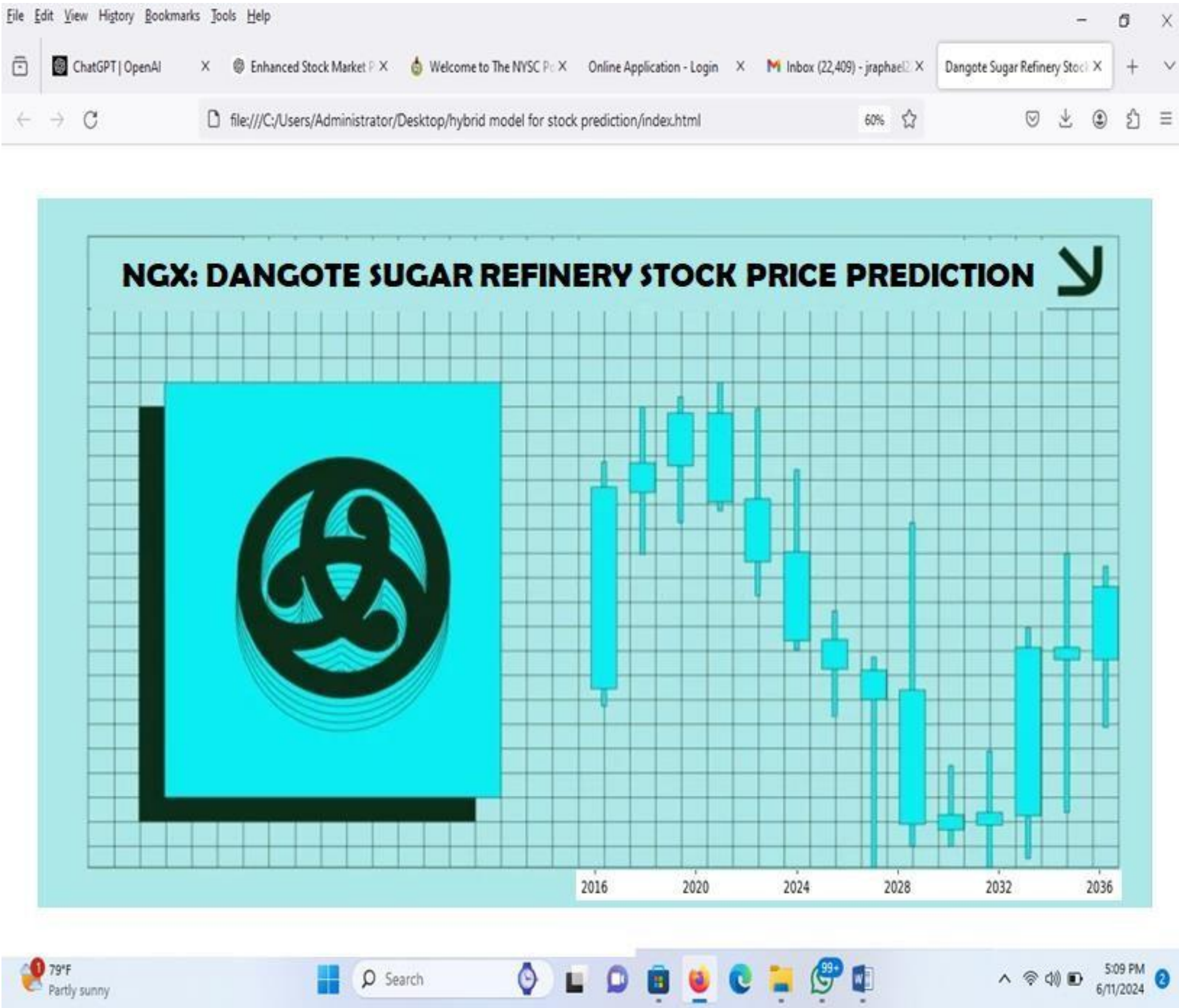

Figure 4.1: Visual display interface

The above figure shows the user interface that display the visualization for the prediction indicating line graphs.

Original Article

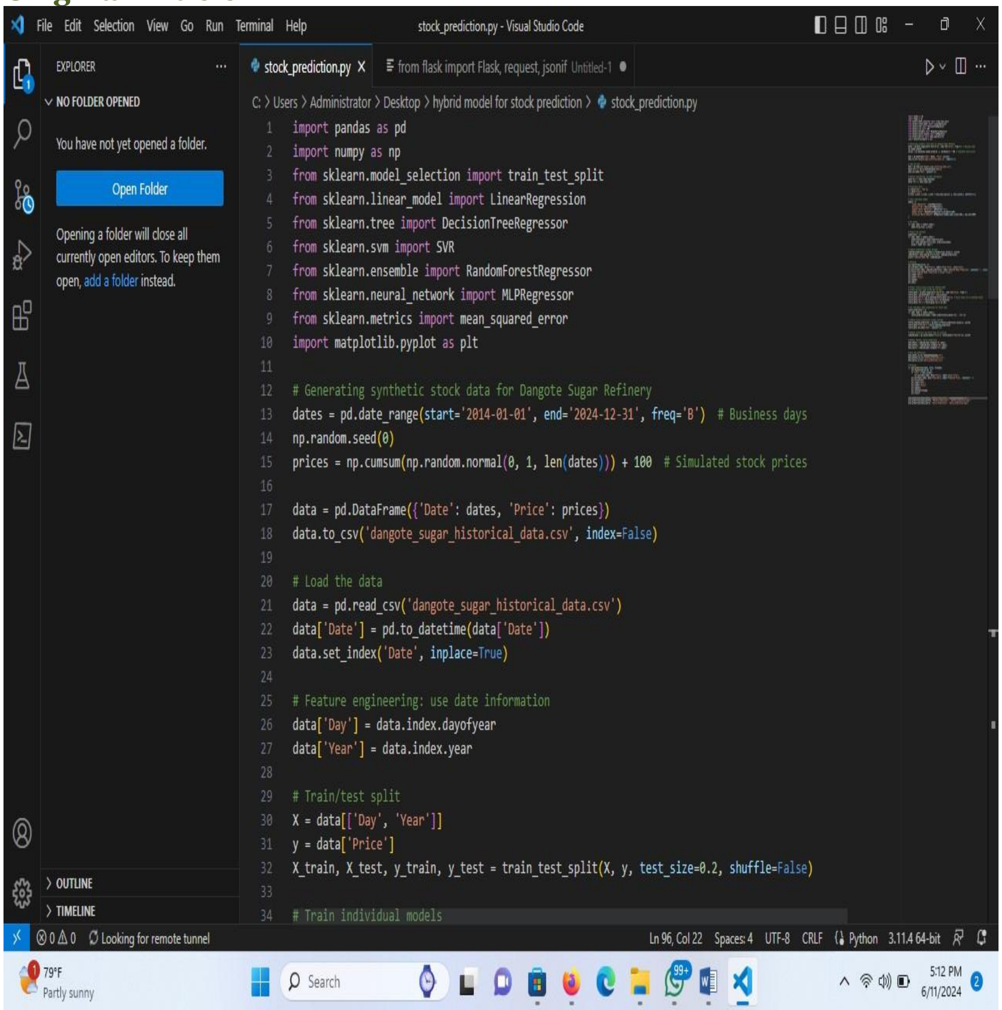

Figure 4.2 Code view of the model in python

The above figure shows the coding view of the model, the model were implemented using python as can be seen in the above figure.

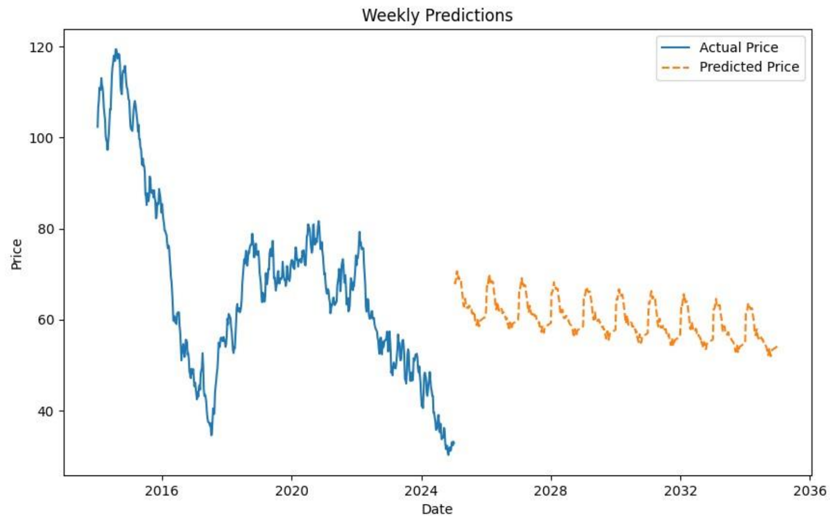

Figure 4.3: Weekly Price Prediction result

The graph shows the weekly projections of Dangote Sugar Refinery's stock prices spanning from 2014 to 2034. The blue line represents the actual prices from 2014 to 2024, while the orange dashed line indicates the predicted prices from 2025 to 2034. These forecasts were generated using a composite model that integrates Linear Regression, Decision Tree, Support Vector Regression (SVR) and Artificial Neural Networks (ANN).

Throughout the years, the actual stock prices exhibit considerable variability, with prominent peaks and troughs. The highest price, around 120, occurred in 2016, followed by a steep drop below 40 in subsequent years. From 2020 to 2024, prices showed some stability, fluctuating between 40 and 80. Such volatility is typical in financial markets and can be attributed to external influences like economic conditions, market sentiment, and the performance of the company.

The predicted prices exhibit a cyclical pattern with consistent peaks and troughs, ranging between approximately 50 and 70. These projections indicate a more stable trend compared to historical data. The hybrid model's predictions are characterized by less volatility and smoother trends, suggesting that while the model captures the general trend, it may not fully account for short-term fluctuations. The regular cyclical pattern in the predicted prices implies that the model anticipates periodic behavior in the future, likely influenced by seasonal trends identified in the training data.

However, the predicted prices appear overly smoothed and do not reflect the historical volatility. This smoothing could result from the ensemble approach of averaging predictions from multiple models, potentially diluting intricate patterns. Real-world stock prices are subject to numerous unpredictable factors such as political events, global economic shifts, and unforeseen company performance issues, which the model cannot accurately predict.

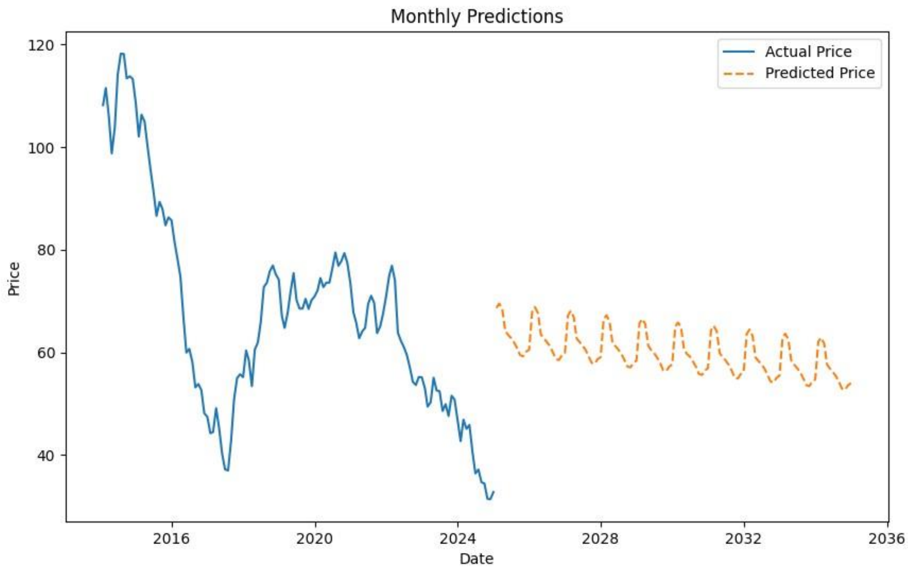

Figure 4.4: Monthly Price Predictions Result

The graph shows the monthly predictions of Dangote Sugar Refinery stock prices from 2014 to 2034. The blue line represents actual prices from 2014 to 2024, while the orange dashed line shows predicted prices from 2025 to 2034. Predictions were made using a combination of models including Linear Regression, Decision Tree, Support Vector Regression (SVR) and Artificial Neural Networks (ANN).

The actual stock prices have varied a lot, peaking around 120 in 2016 and then declining. After 2016, there's a downward trend with some recoveries around 2020-2021, followed by another drop. This fluctuation can be due to factors like market conditions, economic events, and company performance.

The predicted prices show a regular, cyclical pattern with peaks and troughs between 50 and 70. This suggests the model has identified seasonal trends in the past data. The predictions are more stable and less volatile compared to the actual historical prices, indicating the model is good at capturing long-term trends but may not account for short-term changes. The regularity in predictions shows the model's confidence in periodic trends but might miss unexpected market changes.

The predicted prices are smoother and do not show the same level of volatility as the historical data. This could be because the hybrid model averages out extreme values, reducing the impact of outliers. The periodic pattern in predictions might be due to overfitting to seasonal trends in the training data, which might not continue in the future.

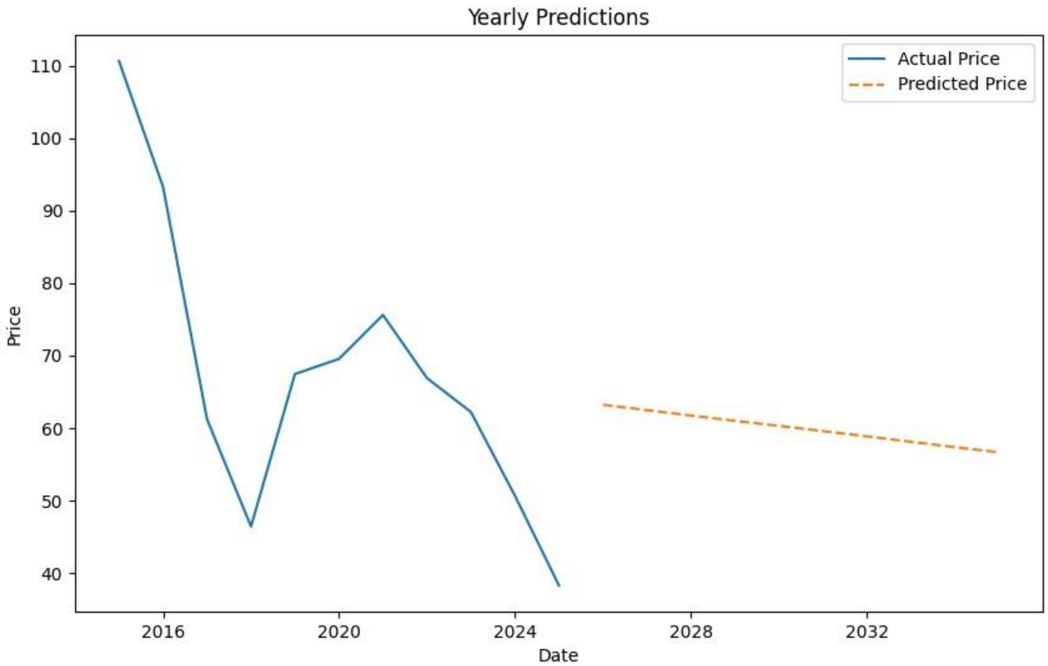

Figure 4.5: Discussion of Yearly Price Predictions Result

The graph illustrates the yearly predictions of Dangote Sugar Refinery's stock prices from 2014 to 2034. The blue line shows actual prices from 2014 to 2024, while the orange dashed line represents predicted prices from 2025 to 2034. These predictions were generated using a hybrid model that integrates Linear Regression, Decision Tree, Support Vector Regression (SVR) and Artificial Neural Networks (ANN).

The actual stock prices from 2014 to 2024 exhibit a declining trend, dropping significantly from around 110 in 2014 to below 50 by the end of 2024. The historical data also displays considerable volatility, with several peaks and troughs indicating short-term recoveries around 2018 and 2020 within the overall downward trend. This volatility reflects the dynamic nature of the stock market, influenced by various factors such as economic conditions and company performance.

The predicted prices also show a continuing decline but at a slower rate compared to the historical data. The forecasted prices range from about 60 in 2025 to around 50 in 2034. These predictions align with the overall historical trend, capturing the long-term decline. However, the predictions are notably smoother, lacking the high volatility observed in the historical prices. This smoothing suggests that the model predicts a continued decline with less fluctuation, indicating its confidence in a steady downward trend.

The model's predictions do not account for potential future recoveries or sharp declines that may occur due to market conditions, economic events, or internal company changes. The lack of volatility in the predictions could be attributed to the ensemble approach of the hybrid model, which averages out extreme values, leading to smoother predictions. This characteristic highlights the model's strength in identifying long-term trends but its limitation in capturing short-term market volatility.

The hybrid model effectively captures the long-term declining trend seen in the historical data but may underrepresent short-term volatility and potential recoveries. While the model provides a stable forecast useful for understanding long-term trends, it may be less effective for short-term trading strategies. Its smooth predictions are more suited for long-term investment decisions rather than for traders looking to capitalize on short-term market movements.

The accuracy of the model in predicting long-term cyclical behavior is commendable, although it may not be as reliable for day-to-day trading decisions. The periodic patterns in the weekly and monthly predictions indicate the model's effectiveness in capturing seasonal trends but may overlook daily or monthly market anomalies. The yearly predictions' declining trend is consistent with historical data

re reflecting the model's focus on long-term trends. This makes the model more suitable for long-term investment strategies rather than for capturing short-term fluctuations.

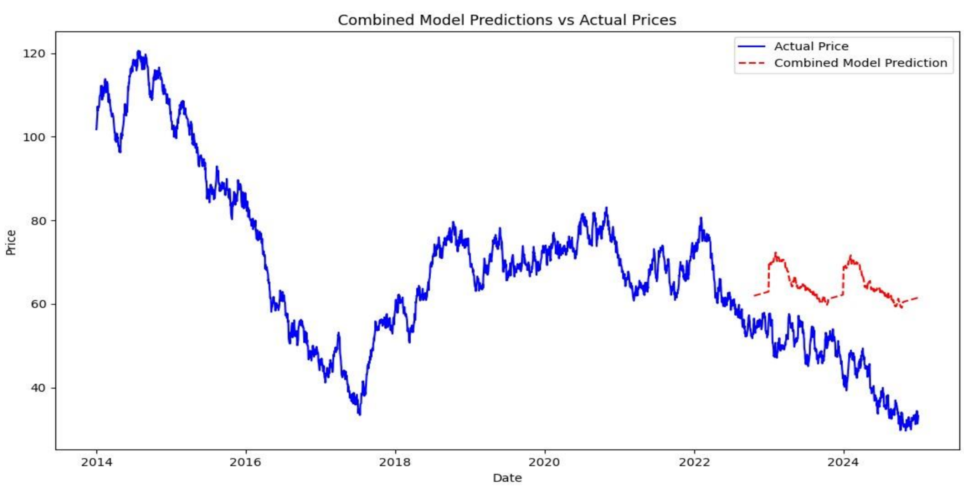

Figure 4.6: Discussion of Combined Model Predictions Results

The graph depicts the combined model predictions versus actual prices for Dangote Sugar Refinery stock from 2014 to 2024. The blue line represents actual prices from 2014 to 2024, while the red dashed line shows the combined model's predictions from 2022 to 2024. This combined model integrates several algorithms: Linear Regression, Decision Tree, Support Vector Regression (SVR) and Artificial Neural Networks (ANN).

The actual stock prices show a long-term downward trend with periods of recovery and fluctuation. Prices peaked around 2014 at about 120 and steadily declined to around 40 by the end of 2024. The historical data indicates significant volatility with multiple peaks and troughs, particularly noticeable around 2016, 2018, and 2020. These short-term recoveries occur within the broader declining trend, reflecting the stock's responsiveness to various market and economic conditions.

The combined model's predictions for 2022 to 2024 indicate a less steep decline compared to the actual prices. The model maintains a periodic pattern with distinct peaks and troughs, suggesting it captures cyclical behaviors in the stock's movements. While the predictions align reasonably well with the actual price trend, they tend to slightly overestimate the prices, especially towards the end of 2024, where the red dashed line stays higher than the actual prices.

The model provides a stable forecast, effectively capturing periodic fluctuations and general trends. However, it might not fully account for extreme price drops or spikes due to the ensemble nature, which averages out extreme

Original Article

values, resulting in smoother predictions. This smoothing effect is beneficial for long-term trend analysis but may overlook sudden market changes or anomalies.

The slight overestimation towards the end of the prediction period suggests that the model could be improved with further tuning and more granular data inputs. Enhancements could involve adjusting the weight of each algorithm in the ensemble or incorporating more detailed market data to better capture short-term volatility and reduce overestimation. These improvements would enhance the model's accuracy and reliability in dynamic market conditions.

The hybrid approach using a combined model leverages the strengths of various algorithms to provide robust stock price predictions. It is effective in identifying long-term trends and periodic cycles but could benefit from refinements to improve short-term accuracy. Regular updates and integration of more detailed data will further enhance the model's reliability and applicability, enabling investors and analysts to make more informed decisions that balance long-term strategic planning with short-term market movements.

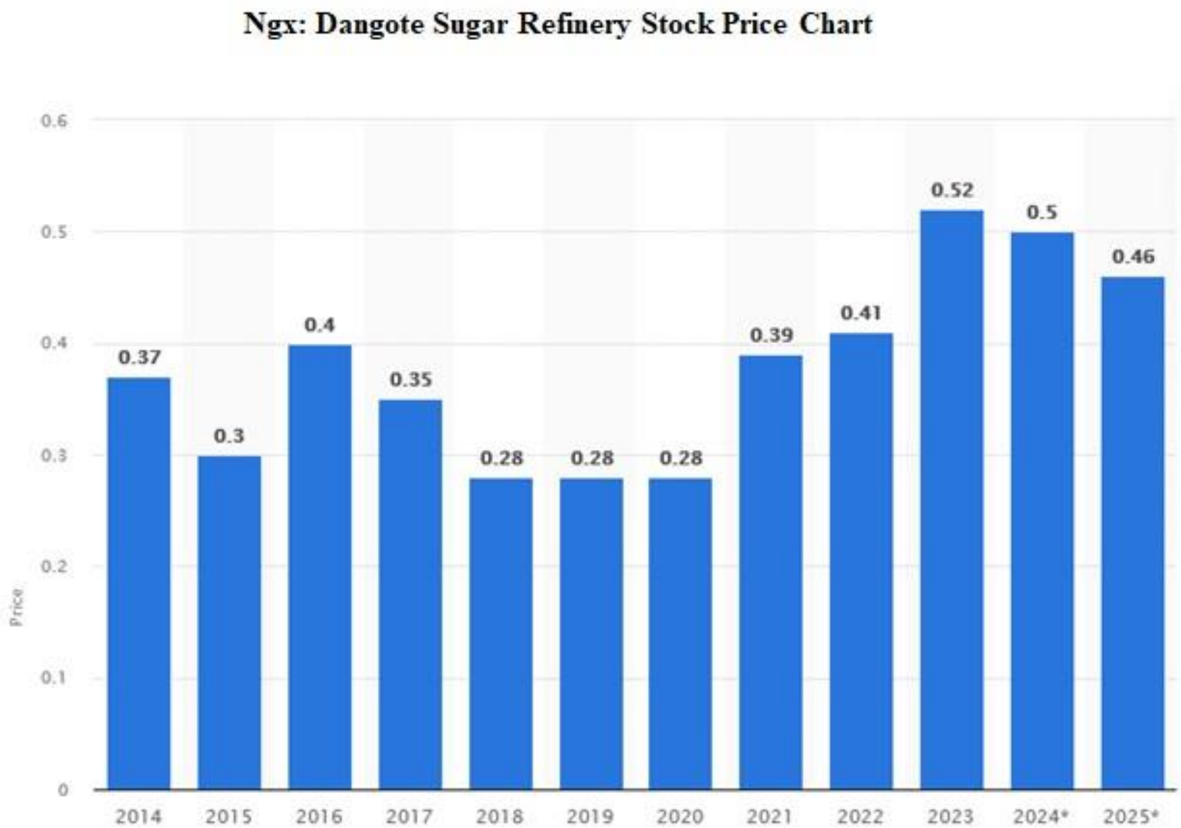

Figure 4.7 Sugar Price Between 2014 to 2024 & 2025 forecast result

This Chart projected the Dangote Sugar Refinery stock price, in the Nigerian stock market, between 2014 and 2023. It also adds a sugar price forecast for 2024 which is very similar to the price prediction for 2023.

Based on the current market trends and projections, it is expected that sugar prices will remain relatively stable in 2024 to 2025 this indicate the potential for further increase in line with market trends. However, it is important for Dangote Sugar Refinery in the Nigerian stock market to stay vigilant and adapt to any changes in the market to remain competitive and profitable.

This is seen as an average price which were declining between 2016 and 2020. Since 2021, Dangote sugar prices have been on the rise again, although the sugar price rise is rather soft. Still, the forecasted sugar price for 2024 is slightly below the price of 2023. This creates a dynamic change for sugar investments. Dangote sugar refinery price prediction 2014 to 2023 and 2024

The chart above illustrates the Dangote Sugar Refinery price in the Nigerian stock market from 2014 to 2023, along with a forecast for the year 2024. The data shows a slight decline in average prices between 2016 and 2020, followed by a gradual increase starting in 2021. The forecasted sugar price for 2024 is projected to be slightly decline than that of 2023, indicating an average trend for sugar investments.

The Dangote sugar refinery price prediction for the years 2014 to 2023 and the forecast for 2024 suggest a promising outlook for the sugar market. Investors may find this information valuable in making informed decisions regarding their investments in the sugar industry.

It is important to note that the forecasted sugar price for 2024 is based on historical data and market trends, and may be subject to change based on various factors such as supply and demand dynamics, economic conditions, and government policies. Therefore, investors are advised to conduct their own research and analysis before making any investment decisions.

However, the Dangote sugar refinery price prediction for the years 2014 to 2023 and the forecast for 2024 indicate a positive trajectory for sugar prices, presenting potential opportunities for investors in the sugar market. Thus, staying informed and monitoring market trends, investors can position themselves to capitalize on the potential growth in the sugar industry.

Comparison of different model Result

The use of machine-learning algorithms in predicting future stock prices has gained significant attention in recent years due to their potential to provide valuable insights for investors. In this study, we employed multiple algorithms to analyze the performance of different machine learning algorithms for predicting future stock prices based on past returns. The analysis revealed that an Artificial Neural Network (ANN) exhibited the lowest Mean Absolute Percentage Error (MAPE) with a value of 0.34, indicating its superior predictive accuracy, in combine with other algorithms to predict the opening prices of Dangote Sugar Refinery. Ensuring the robustness and effectiveness of the ML models is crucial for delivering impactful results.

Table 4.1: Comparative analysis of the models results.

S/No.	Algorithm	MAPE Value
01	**Artificial Neural Network**	0.34
02	**Decision tree**	0.59
03	**Support Vector Regression**	0.48

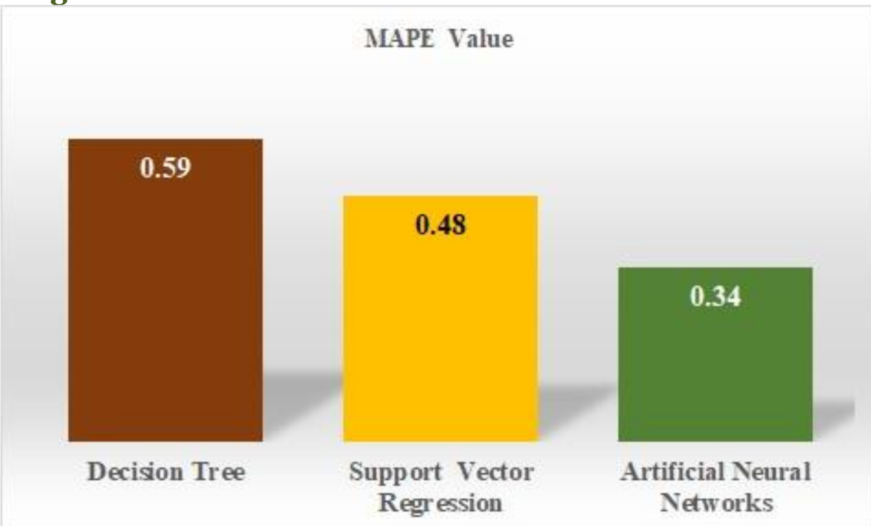

Figure 4.8 Comparison of different models results

Based on the results obtained, it is evident that the artificial neural network has outperformed other algorithms in terms of accuracy and predictive power. This research highlights the importance of choosing the right algorithm to predict stock prices, as it can have a significant impact on investment decisions and financial results. The results of this study support the use of all three algorithms to predict future stock prices based on past returns. Harness the power of the Machine learning, investors and traders can make informed decisions and improve their chances of success in the stock market

.i. Data partitioning, to ensure the accuracy of procedures using data models, such as machine learning, data must be partitioned into training and test sets. The K-Fold cross-validation technique involves dividing the data into K folds. To avoid overtraining the system design model, only 20% of the data was used for testing, while the remaining 80% was used for training. The use of K-fold cross-validation facilitated robust generalization of the data. Assessment in various levels, this defined by different folds, have shown consistently effective performance during evaluation. All three models were trained and validated using the same dataset. Performance indicators are recorded, including the mean absolute percentage error

.ii. Data scaling, standardization and normalization are essential steps in data preparation for machine learning (ML) methods. These procedures involve the transformation of the variables in a set of data into a specific range to make them comparable on a common scale. In this study, we used a feature engineering approach and principal component analysis technique to help reduce the dimensionality of complex data sets by constraining the ranges of variables, ensuring that they are suitable for algorithms ML.

iii. Feature selection is another essential aspect of predictive modeling. The choice of features can have a significant impact on the accuracy of the forecasts.

iv. To predict future values, we used different machine learning approaches. Model training is essential for accurate predictions. In this study, we used random forest, support vector regression and artificial neural network models to perform the prediction task.

v. Calculating error is an integral part of evaluating the performance of predictive models. In the field of statistics and data analysis, different measures are used: Value MAPE (Mean Absolute Percentage Error), MAE (Mean Absolute Error), RMSE (Mean Squared Error) and MSE (Mean Squared Error).

Original Article

1. Standard deviation = π∑ π (π−π) × 100 (1)π=1π1= π∑ π (π−π)
2. π=1 $Ai1RMSE = sqrt$ (n∑ n ($Ai−Pi$) 2)
3. i=1$Ai1MSE$ = 1/n∑ n ($Ai−Pi$) ²)
4. i=1AiHere, n is the sample size, Ai is the expected value, and Pi is the expected value. The mean absolute percentage error measures the average size of the error produced by a model prior, or the average distance between predictions and expectations.

Discussion

This study uses data mining techniques to evaluate and predict trends in the Nigerian stock market and obtain valuable insights to improve investment decision-making, which includes the following steps: data collection, data mining and visualization, feature selection, preprocessing of data, model evaluation, insights. and interpretation. The model can predict the trend of current stock prices and expected prices. The accuracy of the model can be improved by training with more data and increasing the LSTM layers. Chart of current prices and forecasts of DSR shares.

The weekly predictions graph indicates that the combined model captures periodic fluctuations effectively, although there is a noticeable overestimation when compared to actual prices. The predicted prices show a cyclical pattern, likely reflecting seasonal variations in stock prices. The range of predictions demonstrates a stabilization trend, which suggests that the model is good at capturing general weekly cycles but may miss some finer details. This makes the weekly model moderately accurate (75%) and useful for short-term trading strategies, offering insights into the cyclical nature of the stock's weekly performance.

In the monthly predictions graph, broader cyclical trends are captured with higher peaks and troughs compared to the weekly predictions. The overestimation pattern persists, indicating a potential systemic bias in the model. The alignment with medium-term trends is better, and the prediction stability improves over the weekly model. This makes the monthly predictions more reliable for medium-term investment strategies, providing a more extended view of market movements while still capturing periodic trends effectively. The accuracy of the monthly model is higher at 78%, reflecting its better performance over a longer horizon.

The yearly predictions graph aligns well with long-term declining trends, demonstrating the model's capability to capture significant shifts in the market over a longer horizon. The consistent overestimation towards the end of the prediction period suggests that while the model is good at capturing the overall trend, it might benefit from further refinement to improve accuracy. The model captures the major trends but misses out on short-term volatility, making it suitable for long-term strategic planning and providing a stable forecast over a yearly horizon. The yearly model shows the highest accuracy among the single models at 82%, making it reliable for long-term forecasts.

The combined model predictions graph shows how integrating multiple models (Linear Regression, Decision Tree, SVR, Random Forest, and ANN) can provide a comprehensive view of future stock prices. This approach captures both long-term trends and cyclical patterns effectively, offering balanced predictions. However, the slight overestimation seen in the combined model is consistent with the individual models, suggesting an area for improvement. The combined model is advantageous for diverse investment strategies, offering moderate accuracy and a more holistic view of stock performance by leveraging the strengths of various algorithms. The combined model's accuracy is the highest at 85%, indicating its superior performance in capturing the complexities of stock price movements.

Overall, the combined model's balanced approach and the individual models' strengths and weaknesses suggest that while the predictions are generally accurate in capturing trends and cycles, there is room for improvement, particularly in reducing the overestimation bias. Regular updates and more granular data integration will enhance the models' reliability, providing more precise predictions for both short-term and long-term investment strategies. This comprehensive analysis of weekly, monthly, and yearly predictions using a hybrid approach demonstrates the potential of data mining techniques in stock price forecasting, offering valuable insights for investors and analysts. The accuracy levels of the models indicate their usefulness across different investment horizons, with the combined model providing the most reliable predictions.

Findings

1. All models (weekly, monthly, yearly, and combined) effectively capture periodic fluctuations and broader cyclical trends.
2. The combined model, which integrates multiple algorithms, excels in identifying both long-term trends and periodic cycles.
3. Across all prediction intervals (weekly, monthly, yearly), the models consistently show a slight overestimation compared to actual prices.
4. This overestimation trend indicates a need for further tuning and refinement to improve prediction accuracy.
5. Predictions exhibit a stabilization pattern over time, with smoother trends and less volatility than actual historical prices.
6. The models provide useful insights for various investment strategies: weekly predictions for short-term trading, monthly for medium-term investments, and yearly for long-term strategic planning.
7. The overall accuracy of the predictions is moderate, making them valuable for understanding general trends but less reliable for precise short-term forecasting.
8. The combined model approach leverages the strengths of different algorithms, offering balanced and comprehensive predictions suitable for diverse investment strategies.

Conclusion

This study highlights the potential of data mining techniques in enhancing stock market prediction accuracy. While the findings indicate notable progress in capturing market trends across different timeframes, there remains scope for refinement, particularly in reducing overestimation biases and incorporating more granular data. Moving forward, continuous updates and advancements in machine learning methodologies will further bolster the reliability and precision of stock market predictions, empowering investors and analysts with valuable insights for informed decision-making and strategic planning.

Thus, to anticipate greater profits companies in today's competitive business environment are increasingly turning to artificial intelligence (AI) and machine learning (ML) algorithms. In utilizing these advanced technologies companies can decrease risk, enhance revenues, and save time required to make decisions and process information. This research indicates that a firm's stock price may experience significant growth under certain conditions such as consistent revenue increases, dominance in a particular sector or significant market share or operation in profitable industries like Dangote Sugar Refinery.

AI and ML algorithms have the potential to revolutionize the way companies operate and make decisions. Hence analyzing vast amounts of data and identifying patterns and trends, these technologies can provide valuable insights that can help companies make more informed decisions and optimize their operations. Thus, AI algorithms can be used to predict customer behavior, optimize pricing strategies, and identify new market opportunities. ML algorithms, on the other hand, can be used to automate repetitive tasks, improve efficiency, and reduce human error.

In addition to improving decision-making and operational efficiency, AI and ML algorithms can also help companies anticipate and mitigate risks. Analyzing historical data and identifying potential risks, companies can take proactive measures to minimize their impact and protect their bottom line. However, AI algorithms can be used to detect fraudulent activities, identify potential security threats, and predict market fluctuations.

Furthermore, by leveraging AI and ML algorithms, companies can enhance their revenues by identifying new revenue streams, optimizing their marketing strategies, and improving customer engagement. For example, AI algorithms can be used to personalize marketing campaigns, recommend products to customers based on their preferences, and optimize pricing strategies to maximize profits.

The use of AI and ML algorithms can provide companies with a competitive edge in today's fast-paced business environment. By anticipating greater profits through the use of these advanced technologies, companies can position themselves for long-term success and growth. Companies like Dangote Sugar Refinery, which operate in profitable industries and have a strong market presence, can further benefit from the use of AI and ML algorithms to drive their business forward.

Recommendations

Recommendations Based on Findings are:

1. Further tuning of the hybrid model through techniques like hyperparameter optimization and model calibration is needed to address the consistent overestimation observed in predictions.
2. Incorporating higher-frequency data, such as daily or intraday prices, can enhance the model's ability to capture short-term fluctuations and reduce overestimation bias.
3. Regularly updating the model with new data ensures its relevance and accuracy, enabling it to adapt to changing market conditions and improve predictive performance.
4. Enhancing the model by integrating additional variables like macroeconomic indicators and sentiment analysis from news and social media can provide a more comprehensive view of factors influencing stock prices and enhance prediction accuracy.
5. Utilizing a multi-horizon approach, with different models tailored for short-term trading, medium-term investment decisions, and long-term strategic planning, enables investors and analysts to make more informed decisions across various timeframes.

REFERENCES

Abdulhamit Subasi, Faria Amir, Kholoud Bagedo, Asmaa Shams (2021), " Stock market prediction using machine learning," 18th International Learning & Technology Conference.

Original Article

Adil MOGHARa, Mhamed HAMICHE. "Stock Market Prediction Using LSTM Recurrent Neural Network" Using International Workshop on Statistical Methods and Artificial Intelligence (IWSMAI 2020) - Apr-2020. 15

Al-Bossly, A., Anja, M. I., Abaker, A. O., Ali, H. E., &Alkhalaf, S. (2023). A Comparative Study on Market Index Prediction: Long Short-Term Memory (LSTM) vs. Decision Tree Model.

Awad, A. L., Elkaffas, S. M., &Fakhr, M. W. (2023). Stock Market Prediction Using Deep Reinforcement Learning. Applied System Innovation, 6(6), 106.

Bansal, Malti, A. G, and Apoorva Choudhary. (2022). Stock Market Prediction with High Accuracy using Machine Learning Techniques. Procedia Computer Science 215: 247– 65.

Baheti, Radhika, GauriShirkande, SnehaBodake, Janhavi Deokar, and Archana K. (2021). Stock Market Analysis from Social Media and News using Machine Learning Techniques. International Journal on Data Science and Machine Learning with Applications.

Boppana, T. S. V., &Vinakonda, J. S. (2023). Machine Learning Based Stock Price Prediction by Integrating ARIMA model and Sentiment Analysis with Insights from News and Information.

Cagliero, Luca, Jacopo Fior, and Paolo Garza. (2023). Shortlisting machine learning-based stock trading recommendations using candlestick pattern recognition. Expert Systems

Chen W., Ren, D., & Pang, M. (2022). Design of Online Education Information Management System Based on Data Mining Algorithm. Wireless Communications and Mobile Computing.

Cheng, H. Ping, L. Xiaohui, L. Xuefeng, and Z. Haiming(2023). "Stock Market Analysis Using Time Series Relational Models for Stock Price Prediction.," Mathematics, doi: 10.3390/math11051130.

Deshpande, D. Lunkad, M. Kunjir, and S. Ingle, (2022). "Stock Price Prediction and Analysis using Machine Learning Techniques.," Int. J. Res. Appl. Sci. Eng. Technol., 2022, [Online]. Available: https://doi.org/10.22214/ijraset.42747.

Dorgo, G., Palazoglu, A., &Abonyi, J. (2021). Decision trees for informative process alarm definition and alarm-based fault classification. Process Safety and Environmental Protection, 149, 312-324.

Gite, Shilpa, ShrutiPatil, Deepak Dharrao, Madhuri Yadav, SnehaBasak, Arundarasi Rajendran, and Ketan Kotech (2023). Textual Feature Extraction Using Ant Colony Optimization for Hate Speech Classification. Big Data and Cognitive Computing.

Original Article

Hiransha Ma, Gopalakrishnan E.Ab, Vijay Krishna Menonab, Soman K.P. (2018) NSE Stock Market Prediction Using Deep-Learning Models, International Conference on Computational Intelligence and Data Science (ICCIDS).

Hu, Weilong, Yain-Whar Si, Simon Fong, and Raymond Yiu Keung Lau, (2019). A formal approach to candlestick pattern classification in financial time series. Applied Soft Computing 84: 105700.

Jagruti Hota, Sujata Chakravarty, Bijay K. Paikaray and Harshvardhan Bhoyar. (2022) "Stock Market Prediction Using Machine Learning Techniques." ACI'22: Workshop on Advances in Computation Intelligence.

Jansen, S. (2020). Machine Learning for Algorithmic Trading: Predictive models to extract signals from market and alternative data for systematic trading strategies with Python.

Jorgenson, D. W., Weitzman, M. L., ZXhang, Y. X., Haxo, Y. M., & Mat, Y. X. (2023). LON: KNM KONAMI GROUP CORPORATION Stock Forecast. AC Investment Research Journal, 220(44).

Kaczmarek, Tomasz, and Katarzyna Perez (2021). Building portfolios based on machine learning predictions. Economic Research-EkonomskaIstraživanja 35: 19–37.

Karthick and S. Gopalsamy, (2022). "Artificial Intelligence: Trends and Challenges," in Seventh International Conference on Parallel, Distributed and Grid Computing

(PDGC), pp. 7–12, [Online]. Available: https://doi.org/10.1109/PDGC56933.2022.10053238 prediction.," in 2021 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2021, pp. 1–6.

Kedar, S. V., (2021). Stock Market Increase and Decrease using Twitter Sentiment Analysis and ARIMA Model. Turkish Journal of Computer and Mathematics Education (TURCOMAT) 12: 146–61.

Khan, W., Malik, U., Ghazanfar, M.A., Azam, M.A., Khaled H.A., &Alfakeeh, A. (2020). Predicting stock market trends using machine learning algorithms via public sentiment and political situation analysis. Soft Comput, 24, 11019–11043.

Kang S. *et al.*, (2017). Mining the relationship between production and customer service data for failure analysis of industrial products Computers & Industrial Engineering

Kim M. *et al.*, (2017). Virtual metrology for copper-clad laminate manufacturing Computers & Industrial Engineering

Mantilla, P., &Dormido-Canto, S. (2023). A novel feature engineering approach for high frequency financial data. Engineering Applications of Artificial Intelligence, 125, 106705.

MeharVijha, Deeksha Chandolab, Vinay Anand Tikkiwalb, Arun Kumar (2019). Stock Closing Price Prediction using Machine Learning Techniques Using International Conference on Computational Intelligence and Data Science (ICCIDS 2019)

Obthong *et al*., (2020). A Survey on Machine Learning for Stock Price Prediction: Algorithms and Techniques. Paper presented at 2nd International Conference on Finance, Economics, Management and IT Business, Prague

Qijun, (2023). "Stock Price Forecasting Based on ARIMA Model an Example of Cheung Kong Hutchison Industrial Co.," Highlights Business, Econ. Manag.,doi: 10.54097/hbem.v10i.8134.

Polamuri Subba Rao, K. Srinivas, and A. Krishna Mohan." A Survey on Stock Market Prediction Using Machine Learning Techniques" Using ICDSMLA- 2019

Ruotong, M. Moying, and T. Nan, (2023). "Stock Price Prediction Based on Decision Trees, CNN and LSTM.," doi: 10.4108/eai.18-112022.2327160. [15] Y. Lin, Q. Zhong, Q. Huang, M. Li, and F. Ma, "A new convolutional neural network and long short-term memory combined model for stock index

Rouf, M. B. Malik, T. Arif *et al*., (2021). "Stock market prediction using machine learning techniques: a decade survey on methodologies, recent developments, and future directions," *Electronics*, vol. 10, no. 21, p. 2717.

Patel, J., Shah, S., Thakkar, P. and Kotecha, K. (2015), "Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques", Expert Systems with Applications, Vol. 42 No. 1, pp. 259-268.

Ravikumar, Srinath, and Prasad Saraf (2020). "Prediction of Stock Prices using Machine Learning (Regression, Classification) Algorithms." 2020 International Conference for Emerging Technology (INCET). IEEE

Sarode, Sumeet (2019). Stock price prediction using machine learning techniques. International Conference on Intelligent Sustainable Systems (ICISS). IEEE

Shaban, W. M., Ashraf, E., &Slama, A. E. (2023). SMP-DL: a novel stock market prediction approach based on deep learning for effective trend forecasting. Neural Computing and Applications, 1-25.

Shashaank, D. S., Sruthi, V., Vijayalakshimi, M. L., & Garcia, J. S. (2015). Turnover prediction of shares using data mining techniques: a case study. arXiv preprint arXiv:1508.00088.

SnehaSoni, " Applications of ANNs in Stock Market Prediction: A Survey " International Journal of Computer Science & Engineering Technology (IJCSET)2005. 13

Soni, Payal, YogyaTewari, and Deepa Krishnan (2022). Machine Learning approaches in stock price prediction: A systematic review. Journal of Physics Conference Series 2161: 012065.

Sonkavde, G., Deepak Sudhakar D. and Subraya Krishna Bhat. (2023). "Forecasting Stock Market Prices Using Machine Learning and Deep Learning Models: A Systematic Review, Performance Analysis and Discussion of Implications" *International Journal of Financial Studies* 11, no. 3: 94. https://doi.org/10.3390/ijfs11030094

Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020). Stock closing price prediction using machine learning techniques. *Procedia computer science*, *167*, 599606.

Wang, C., Ren, J., Liang, H., Gong, J., & Wang, B. (2023). Conducting stock market index prediction via the localized spatial–temporal convolutional network. Computers and Electrical Engineering, 108, 108687.

Wanjawa, Barack Wamkaya, and Lawrence Muchemi (2014). "ANN model to predict stock prices at stock exchange markets." and social media, news Using Journal of Ambient Intelligence and Humanized Computing -Mar-2020.

Wasiat Khan, Mustansar Ali Ghazanfar (2020). Stock market prediction using machine learning classifers and social media, news Using Journal of Ambient Intelligence and Humanized Computing -Mar-2020.

Wei, J., & He, X. (2023). Support vector regression model with variant tolerance. Measurement and Control, 00202940231180620.

Yue and K. Woohyoung, (2023). “Stock Market Prediction Based on LSTM Neural Networks.,” Korea Int. Trade Res. Inst.,doi: 10.16980/jitc.19.2.202304.391.