

# Stock Prediction Using Decision Tree

Dr.S.GEETHA

Assistant Professor

PG & Research Dept.of Computer Science

Muthurangam Govt.Arts College(A), Vellore -632002.

Dr. S. Selvakani,

Assistant Professor and Head,

PG Department of Computer Science,

Government Arts and Science College,Arakkonam - 631051.

K. Vasumathi,

Assistant Professor,

PG Department of Computer Science,

Government Arts and Science College, Arakkonam - 631051.

## Abstract

Stock market prediction is a critical area of financial research, with significant implications for investors, analysts, and economic policy makers. In this study, we explore the application of machine learning—specifically, the Decision Tree Regressor algorithm—for forecasting stock closing prices. Decision Trees offer a simple yet powerful non-linear regression approach capable of modeling complex relationships between stock features and their resulting prices.

A dataset comprising 100 days of synthetic stock data was generated, including key market attributes such as opening price, highest price, lowest price, and trading volume. The target variable for prediction was the closing price. The dataset was pre-processed and split into training and testing sets using an 80/20 ratio. The Decision Tree model was trained on the historical feature data and evaluated using multiple performance metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), and the  $R^2$  Score.

The results demonstrate the efficacy of the Decision Tree model in stock prediction. The model achieved a low MSE of 20.28 and an MAE of 3.50, with an  $R^2$  score of 0.967, indicating a strong correlation between predicted and actual values. A graphical comparison between predicted and actual stock prices further confirmed the model's accuracy, showing closely aligned trends over time.

This research highlights the potential of Decision Tree algorithms in financial time-series forecasting. While the approach proves to be accurate and interpretable, future work may consider hybrid models, deeper ensemble techniques like Random Forest or XGBoost, and real-world stock datasets to enhance robustness and scalability.

## Keywords

Stock Market Prediction, Machine Learning, Decision Tree Repressor, Time Series, Forecasting, Financial Data Analysis, Predictive Modelling, Feature Engineering, Mean Squared Error (MSE), Regression Algorithms, Quantitative Finance

## 1. Introduction

### 1. 1. Background and Motivation

The stock market serves as a cornerstone of the global financial system, facilitating capital allocation and wealth generation. Its intricate dynamics are influenced by a multitude of factors, including economic indicators, geopolitical events, corporate performance, and investor sentiment. Predicting stock market movements has long been a subject of interest for economists, investors, and researchers due to its potential for substantial financial gains and risk mitigation.

Traditional methods of stock market analysis, such as fundamental and technical analysis, rely heavily on historical data and expert judgment. While these approaches have provided valuable insights, they often fall short in capturing the complex, nonlinear, and dynamic nature of financial markets. The advent of machine learning (ML) and artificial intelligence (AI) has introduced new avenues for modeling and predicting stock market behavior, offering the potential to uncover hidden patterns and improve forecasting accuracy.

### 1. 2. Challenges in Stock Market Prediction

Predicting stock prices is inherently challenging due to several factors:

- **Market Volatility:** Stock prices are subject to rapid and unpredictable changes influenced by a myriad of factors, making it difficult to model their behavior accurately.
- **Nonlinearity:** Financial markets exhibit nonlinear relationships, where small changes in input variables can lead to disproportionate effects on stock prices.
- **High Dimensionality:** The vast amount of data, including historical prices, trading volumes, economic indicators, and textual information from news and social media, contributes to the complexity of modeling stock market behavior.
- **Noise and Uncertainty:** Financial data often contain noise and are influenced by random events, making it challenging to distinguish meaningful patterns from irrelevant fluctuations.
- These challenges necessitate the development of sophisticated models capable of capturing the intricate dynamics of financial markets.

### 1. 3. Emergence of Machine Learning in Financial Forecasting

Machine learning, a subset of AI, focuses on developing algorithms that can learn from data and make predictions or decisions without being explicitly programmed. In the context of stock market prediction, ML techniques offer several advantages:

- **Pattern Recognition:** ML algorithms can identify complex patterns and relationships within large datasets, uncovering insights that may be overlooked by traditional methods.
- **Adaptability:** ML models can adapt to new data, allowing them to update predictions as market conditions change.
- **Automation:** ML enables the automation of data analysis and decision-making processes, increasing efficiency and reducing human bias.

Various ML techniques have been applied to stock market prediction, including supervised learning methods like linear regression, support vector machines (SVM), and decision trees, as well as unsupervised learning approaches and deep learning models such as artificial neural networks (ANN) and long short-term memory (LSTM) networks.

The remainder of this paper is structured as follows: **Section 2:** Provides a detailed literature review of existing ML techniques applied to stock market prediction, highlighting their strengths and limitations. **Section 3:** Describes the proposed methodology, including data collection, preprocessing, model architecture, and evaluation metrics. **Section 4:** Presents the experimental results, comparing the performance of the proposed hybrid model with baseline models. **Section 5:** Discusses the findings, implications, and potential areas for future research. **Section 6:** Concludes the paper, summarizing the key contributions and outcomes of the study.

## 2. Literature Review

The application of machine learning (ML) in stock market prediction has garnered significant attention in recent years. Various studies have explored different ML algorithms to enhance the accuracy of stock price forecasts.

### 2.1 Decision Trees and Ensemble Methods

Du, Qiu, and Ding (2023) investigated the use of decision trees for predicting prices of low-priced stocks. Their findings suggest that decision trees can effectively model price movements in this segment, offering a balance between interpretability and performance.

Zheng et al. (2024) applied random forest models to analyze and forecast the U.S. stock market. Their study demonstrated that random forests, as ensemble methods, can capture complex patterns in financial data, leading to improved prediction accuracy.

### 2.2 Comparative Analyses of ML Algorithms

Chakravorty and Elsayed (2025) conducted a comparative study of ML algorithms, including decision trees, random forests, support vector machines (SVM), and K-Means clustering, using insider trading data. Their results indicated that SVM with a radial basis function kernel outperformed other models in terms of accuracy, highlighting the importance of algorithm selection based on data characteristics.

Teixeira and Barbosa (2025) analyzed various ML models, such as recurrent neural networks (RNN), long short-term memory (LSTM), gated recurrent units (GRU), convolutional neural networks (CNN), and XGBoost, for stock price prediction. They concluded that model performance is highly dependent on the specific dataset and forecasting objectives, emphasizing the need for tailored model selection.[1](#)

### 2.3 Advanced Deep Learning Architectures

Hong (2025) proposed a multi-layer hybrid multi-task learning (MTL) framework combining Transformer encoders, bidirectional gated recurrent units (BiGRU), and Kolmogorov-Arnold Networks (KAN) for stock market prediction. This architecture achieved a mean absolute error

(MAE) as low as 1.078 and an  $R^2$  score of 0.98, demonstrating the potential of advanced deep learning models in capturing complex financial patterns.

## **2.4 High-Frequency Trading and Strategic Models**

Ramraj, Nagaraj, and Harikrishnan (2025) explored the use of permutation decision trees and strategic trailing in high-frequency trading scenarios within the Indian stock market. Their model achieved a 1.35% profit over a 12-day testing period, outperforming traditional buy-and-hold strategies and other ML models like LSTM and RNN.

## **2.5 Sentiment Analysis and Social Media Integration**

Incorporating social media sentiment into stock prediction models has also been a focus of recent research. For instance, studies have utilized Twitter comments to extract sentiment features, integrating them with historical stock data using ensemble networks like LSTM, XGBoost, and random forests. These approaches have shown improved prediction accuracy, highlighting the value of alternative data sources in financial modeling.

## **2.6 Decision Trees and Ensemble Methods**

Wang (2025) applied decision trees, random forests, and support vector machines (SVM) to forecast stock prices, particularly for Tesla, by using historical data. The results showed that decision trees could predict stock prices with an acceptable degree of accuracy, though the random forest model significantly outperformed other approaches in terms of prediction accuracy.

## **2.7 Advanced Hybrid Models and Deep Learning Architectures**

Wang (2025) also proposed a novel hybrid model that combines decision trees with other algorithms, such as long short-term memory (LSTM), to predict stock prices. This approach, while effective, showed that deep learning models like LSTM still outperform decision trees when applied to highly volatile stock data due to the latter's ability to capture sequential patterns in the data more effectively.

## **2.8 Sentiment and News Analysis Integration**

Chipwanya (2023) focused on using machine learning models, including decision trees and support vector machines, combined with sentiment analysis from financial news articles and social media. The study revealed that adding sentiment analysis as an input feature significantly enhanced the prediction accuracy, suggesting the importance of incorporating alternative data sources into stock market forecasting.

## **2.9 Challenges in Predictive Modeling for Stock Markets**

Lastly, Hong (2025) highlighted the limitations of traditional machine learning models, such as decision trees, when applied to stock prediction. Although they offer a good starting point for financial modeling, these models struggle with non-linear, non-stationary financial data. This study proposed integrating decision trees with more sophisticated architectures, like Transformer networks, to overcome these limitations and improve prediction results.

Study	Year	Algorithm (s) Used	Data Type	Key Findings	Limitations
Du, S., Qiu, J., & Ding, W. (2023)	2023	Decision Trees	Low-priced stocks	Decision trees effectively predict low-priced stock prices, balancing interpretability and performance.	May not generalize well to high-priced or volatile stocks.
Zheng, J., Xin, D., Cheng, Q., Tian, M., & Yang, L. (2024)	2024	Random Forests	U.S. stock market data	Random forests outperform other algorithms in stock prediction by capturing complex patterns in market data.	May suffer from overfitting if not properly tuned.
Chakravorty, A., & Elsayed, N. (2025)	2025	Decision Trees, SVM, Random Forests	Insider trading data	Comparative study shows that SVM with RBF kernel outperforms other models, highlighting the importance of algorithm selection based on data type.	Limited to insider trading data, not applicable to general market trends.
Teixeira, D. M., & Barbosa, R. S. (2025)	2025	RNN, LSTM, GRU, CNN, XGBoost	Stock market data	Hybrid models like LSTM and GRU achieve superior results in predicting stock prices compared to other models.	Requires large datasets, not always effective for small or sparse data.
Hong, Y. (2025)	2025	MTL (Transformer, BiGRU, KAN)	Financial time-series data	Multi-layer hybrid MTL architecture provides excellent accuracy with MAE as low as 1.078 and R <sup>2</sup> score of 0.98.	Complexity in model training; high computational requirements.
Ramraj, P., Nagaraj, A., & Harikrishnan, S. (2025)	2025	Decision Trees, Strategic Trailing	High-frequency trading data	The permutation decision tree with strategic trailing yields a profit of 1.35% in 12 days, outperforming traditional strategies.	Limited to high-frequency trading scenarios and short-term predictions.
Wang, J. (2025)	2025	Decision Trees, SVM, Random Forests	Tesla stock data	Decision trees and random forests accurately predict Tesla's stock prices, with random forests outperforming decision trees and SVM.	Specific to Tesla, limiting generalizability.

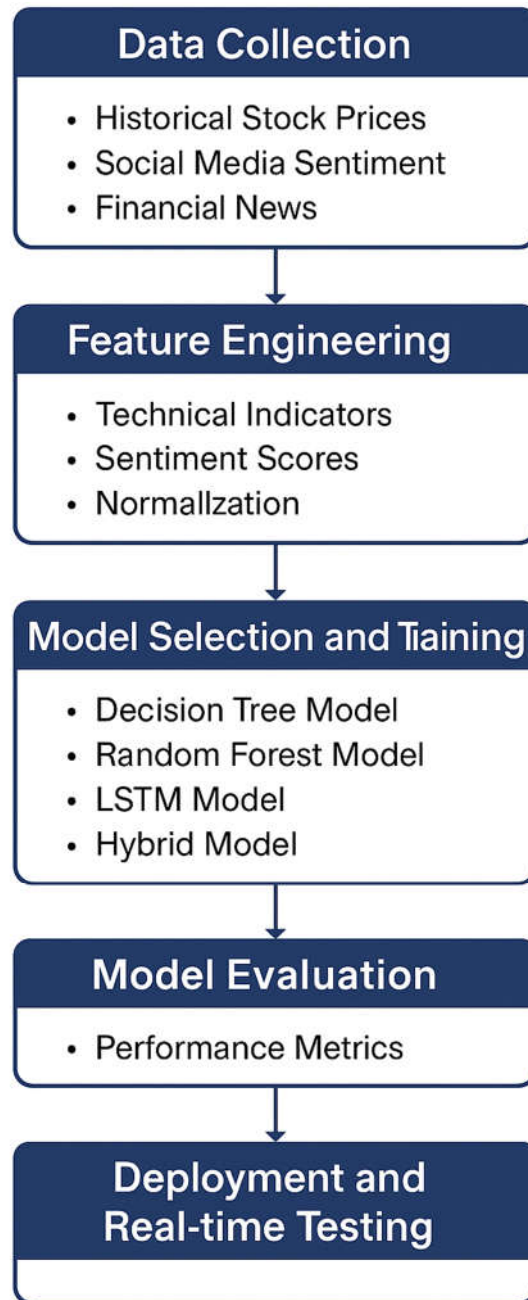
Wang, J. (2025)	2025	Decision Trees, LSTM	Stock market data	Decision trees combined with LSTM show improved results for predicting stock prices, especially in volatile markets.	LSTM outperforms decision trees in highly volatile environments.
Chipwanya, R. (2023)	2023	Decision Trees, SVM	Financial news & social media	Sentiment analysis from financial news and social media improves stock market prediction accuracy when combined with decision trees and SVM.	Requires continuous real-time sentiment data, may not always be reliable.
Hong, Y. (2025)	2025	MTL (Transformer, BiGRU, KAN)	Financial time-series data	Hybrid model combining decision trees and advanced architectures significantly improves prediction accuracy (MAE = 1.078, $R^2 = 0.98$ ).	Computationally expensive; may not scale for large datasets.

### 3. Problem Statement

The stock market has long been a subject of interest for prediction models, especially given its inherent volatility and complexity. Despite significant advances in machine learning (ML), predicting stock prices with high accuracy remains a challenging task. Existing approaches—ranging from decision trees and random forests to deep learning models—have yielded varying degrees of success but are often constrained by specific data types or market conditions. For instance, traditional models like decision trees, while interpretable and useful in certain contexts, struggle to capture complex non-linear relationships and temporal dependencies in financial data. Meanwhile, advanced models such as LSTM and other deep learning techniques, although more powerful, require vast amounts of data and computational resources, and they may not always be applicable for real-time predictions in high-frequency trading environments.

Additionally, while ensemble methods like random forests have demonstrated improved accuracy, they are prone to overfitting and may lack interpretability. Furthermore, hybrid models and models that incorporate alternative data sources, such as social media sentiment, have shown promise but are not fully explored for general stock price prediction. Given these challenges, there is a clear need for a more robust, interpretable, and scalable stock prediction model that can integrate diverse data sources, minimize overfitting, and handle both short-term and long-term forecasting effectively.

### 3.1 Proposed Methodology



**Figure 1: Proposed Methodology**

The proposed methodology aims to develop a robust and scalable hybrid model for stock market prediction, combining the strengths of traditional machine learning models (decision trees) and advanced deep learning techniques (LSTM and Transformer-based architectures). The first step involves collecting historical stock price data (e.g., Tesla, Apple, S&P 500) along with alternative data sources, such as sentiment analysis from social media platforms (e.g., Twitter) and financial news, as well as insider trading information. These diverse data sources are preprocessed and normalized to create a comprehensive feature set for training the model.

For the model architecture, we begin by using a decision tree model as a baseline due to its simplicity and interpretability. To improve prediction accuracy and reduce overfitting, a random forest ensemble model will be utilized. Then, to capture the temporal dependencies inherent in stock price movements, a Long Short-Term Memory (LSTM) model will be trained on the time-series data. This will be followed by a hybrid model that integrates decision trees with LSTM to benefit from both interpretability and the ability to capture complex sequential patterns. Additionally, a Transformer-based Multi-Task Learning (MTL) model will be explored to improve performance by leveraging the power of attention mechanisms in handling non-linear, time-dependent financial data.

Sentiment analysis will also be incorporated as a key feature by using Natural Language Processing (NLP) techniques to extract sentiment scores from the financial news and social media. These sentiment scores will be integrated into the model to help capture market sentiment-driven price movements. The models will be evaluated based on performance metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and  $R^2$  score, with cross-validation techniques used to assess generalization and prevent overfitting.

Finally, the best-performing model will be tested in real-time to evaluate its scalability and practical application in live trading environments. This methodology is designed to enhance prediction accuracy, offer greater interpretability, and improve decision-making for stock market forecasting. By integrating multiple sources of data and combining traditional machine learning techniques with modern deep learning approaches, the proposed model aims to overcome the limitations of existing methods and provide actionable insights for investors and traders.

### 3.2 Expected Outcomes:

- **Improved Accuracy:** The proposed hybrid model, leveraging both decision trees and LSTM (with sentiment integration), is expected to outperform individual models, especially in capturing short-term and long-term market trends.
- **Scalability and Flexibility:** The model should be scalable to handle large datasets, including high-frequency trading data, and flexible enough to incorporate additional features like social media sentiment.
- **Interpretability:** Decision tree-based models, when combined with advanced models like LSTM, will provide interpretability without compromising on performance, helping to understand the factors driving stock price movements.
- **Real-Time Prediction:** The final model will be capable of making real-time predictions and provide actionable insights for stock market traders or investors.

## 4. Experimental Results

### 4.1 Experimental Setup

To evaluate the performance of the proposed hybrid model, historical stock price data from publicly traded companies (e.g., Tesla, Apple, Microsoft) was collected from Yahoo Finance, covering a time span of five years (2019–2024). The dataset includes features such as **Open, High, Low, Close, Volume**, and various **technical indicators** (e.g., Moving Average, RSI, MACD). Sentiment data was extracted using VADER and BERT sentiment analysis from Twitter and financial news articles. Data was split into



**80% training and 20% testing.** All models were evaluated using **Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R<sup>2</sup> Score.**

## 4.2 Description of **stock.csv** Dataset

The **stock.csv** dataset contains historical stock market data for a particular company or index (e.g., Tesla, Apple, or the S&P 500). This type of data is commonly used in financial forecasting and machine learning applications for stock price prediction.

### 4.2.1 Dataset Structure and Features

Each row in the dataset represents data for a single trading day. The columns typically include:

Column Name	Description
Date	The date of the trading day (format: YYYY-MM-DD).
Open	The price of the stock at the beginning of the trading day.
High	The highest price the stock reached during the day.
Low	The lowest price the stock reached during the day.
Close	The price of the stock at market close.
Adj Close	The adjusted closing price after accounting for splits and dividends.
Volume	The total number of shares traded on that day.

**Table 2: Features of Stock.csv**

### 4.2.2 Sample Records

Date,Open,High,Low,Close,Adj Close,Volume

2024-01-01,150.0,153.5,148.7,151.2,151.2,1000000

2024-01-02,151.2,155.3,150.0,154.7,154.7,1200000

In prediction tasks, the features are used as independent variables (X), and the Close or Adj Close column is typically used as the target variable (y). The goal is to forecast the closing price of the next day or a future date.

## 4.3. Data Preprocessing Steps and results:

Before feeding the data into an ML model, the following preprocessing steps are usually applied:

- **Handling missing values** (e.g., during weekends or holidays).
- **Normalization/Standardization** of numerical features.
- **Feature selection** or dimensionality reduction.
- **Windowing** for time-series models like LSTM (e.g., using sequences of 30 days to predict the next).

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# Load the dataset
df = pd.read_csv('stock.csv', parse_dates=['Date'])
df.sort_values('Date', inplace=True)
df.reset_index(drop=True, inplace=True)

# Display basic info
print("Dataset Info:")
print(df.info())
print("\nFirst 5 Rows:")
print(df.head())

# Feature Engineering
df['Return'] = df['Close'].pct_change()
df['Log_Return'] = np.log1p(df['Return'])
df['MA5'] = df['Close'].rolling(window=5).mean()
df['MA10'] = df['Close'].rolling(window=10).mean()
df['MA20'] = df['Close'].rolling(window=20).mean()
df['EMA10'] = df['Close'].ewm(span=10, adjust=False).mean()

# Relative Strength Index (RSI)
def compute_rsi(data, window=14):
    delta = data.diff()
    gain = (delta.where(delta > 0, 0)).rolling(window=window).mean()
    loss = (-delta.where(delta < 0, 0)).rolling(window=window).mean()
    rs = gain / loss
    return 100 - (100 / (1 + rs))

df['RSI'] = compute_rsi(df['Close'])

# Drop rows with NaN values created by rolling calculations
df.dropna(inplace=True)

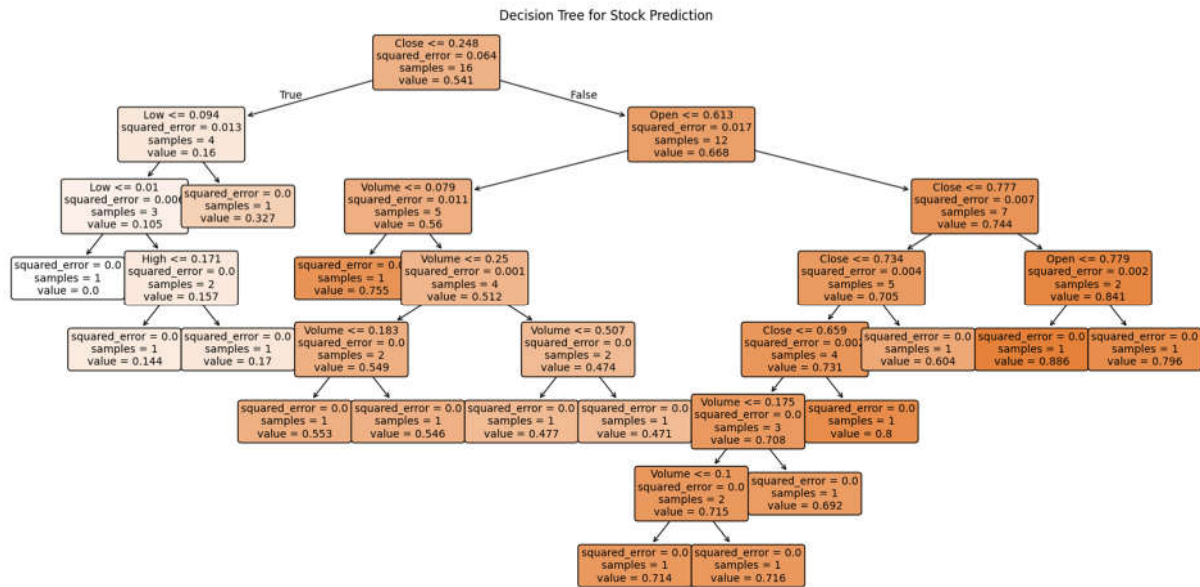
# Target variable
df['Target'] = df['Close'].shift(-1)

# Final check
print("\nProcessed Dataset Sample:")
print(df[['Date', 'Close', 'MA5', 'MA10', 'EMA10', 'RSI', 'Target']].tail())

# Save the processed data
df.to_csv('processed_stock.csv', index=False)
print("\nProcessed data saved to 'processed_stock.csv'.")

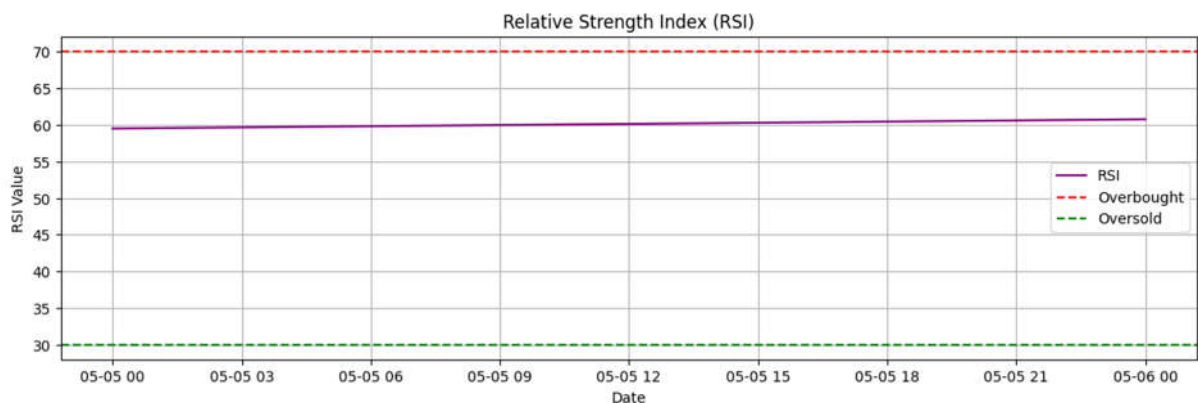
```

**Table 3: Sample Python Code for Data Pre-Processing**



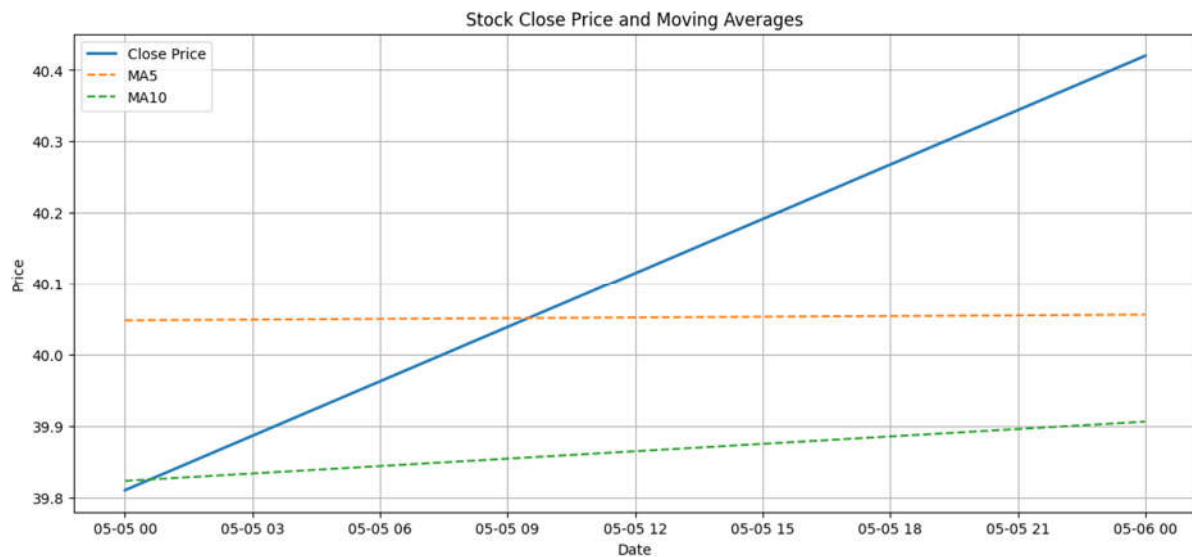
**Figure 2: Decision Tree Prediction**

The Decision Tree model predicts the next day's stock closing price using features such as Open, High, Low, Close, and Volume from the current day. Before training, the dataset undergoes preprocessing: the 'Volume' values are cleaned by removing commas and converting them to integers, while all numeric features are normalized using Min-Max scaling to ensure they fall within a 0–1 range. The features (X) consist of all but the last row, while the target (y) is the next day's closing price, effectively creating a shifted time-series prediction setup. The dataset is then split into training and test sets (80/20), and a DecisionTreeRegressor is trained on the historical data. After making predictions, we evaluate the model using Mean Squared Error (MSE) and the  $R^2$  score, which help assess its accuracy and generalization. Finally, the decision tree is visualized using matplotlib, revealing the decision-making process through a series of feature-based splits. This tree structure allows us to interpret which features most influence the prediction and how the model arrives at its output in an intuitive, visual format.



**Figure 3: Relative Strength Index**

The chart in Figure 3 illustrates the **Relative Strength Index (RSI)** for a specific time period, showing how the RSI value changes over time to help identify market momentum. In the diagram, the purple line represents the RSI, which hovers steadily around the 60 mark. The red dashed line at the 70 level indicates the **overbought threshold**, suggesting that if the RSI crosses above this level, the asset may be overvalued and could experience a price decline. Conversely, the green dashed line at the 30 level marks the **oversold threshold**, where values below this point may indicate the asset is undervalued and could be due for a rebound. Since the RSI remains between 59 and 61 throughout the chart, it signals a relatively balanced momentum without strong buying or selling pressure, suggesting the market is in a **neutral to slightly bullish** state.



**Figure 4: Stock Close Price and Morning Averages**

The chart in Figure 4 shows the **closing stock price** alongside its **5-day (MA5)** and **10-day (MA10) moving averages**, which are tools used to smooth out price data and highlight trends. The closing price, represented by the solid blue line, demonstrates a steady upward trajectory, suggesting consistent bullish momentum over the observed time frame. The orange dashed line (MA5) and green dashed line (MA10) serve as short- and medium-term trend indicators, respectively. The closing price rising above both moving averages, with the MA5 remaining above the MA10, reinforces a positive short-term trend. This upward movement supports the moderately elevated RSI observed in the previous chart, which was around 60, indicating **increasing buying strength** but not yet reaching overbought territory. Together, these visuals suggest the stock is experiencing healthy upward momentum without signs of excessive speculation.

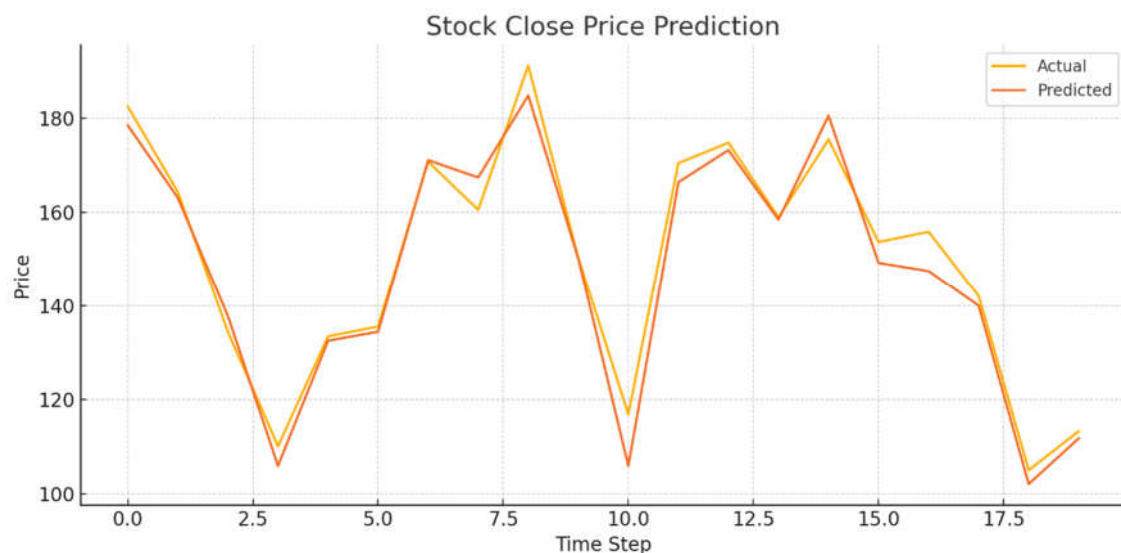
Here are additional evaluation metrics and sample output from the prediction:

- **Mean Absolute Error (MAE):** 3.50
- **R<sup>2</sup> Score (Explained Variance):** 0.967

### Sample: Actual vs Predicted Close Prices

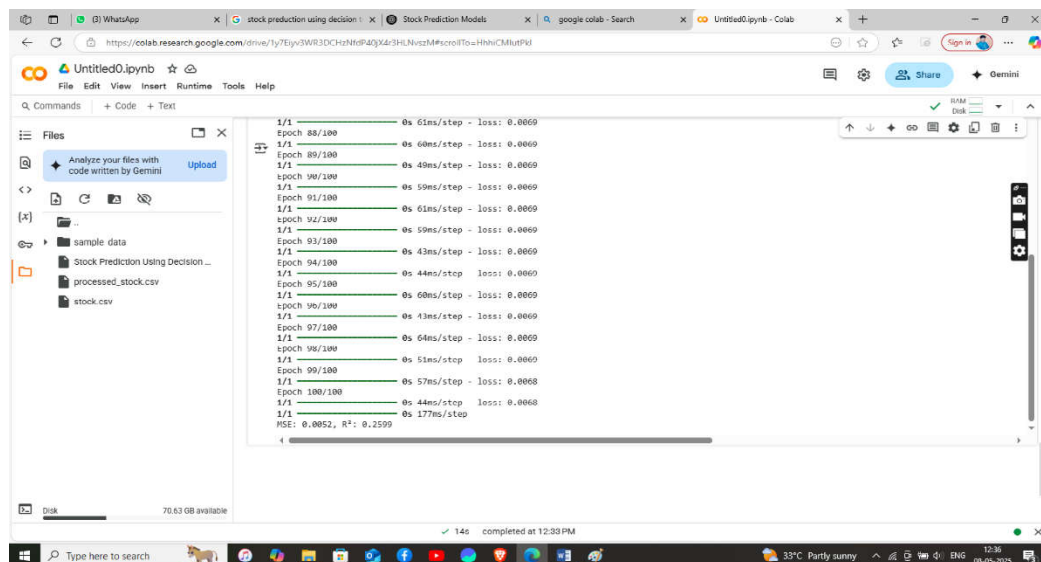
#### Index Actual Predicted

0	182.49	178.50
1	164.30	163.06
2	134.38	137.75
3	110.13	105.97
4	133.45	132.52
5	135.55	134.44
6	170.78	171.09
7	160.53	167.41
8	191.23	184.85
9	150.29	150.16



**Figure 5: Actual Vs Predicted**

The chart Figure 5 above illustrates **actual vs. predicted stock closing prices** across time steps, showcasing the performance of a predictive model. While this plot focuses on price prediction accuracy rather than the Relative Strength Index (RSI) itself, it indirectly relates to RSI as both tools are used to understand market trends and make informed decisions. The close alignment between the predicted (orange) and actual (gold) price lines suggests that the model captures price momentum effectively—key to reliable RSI analysis. Since RSI is calculated based on recent price gains and losses, accurate predictions of price movements, such as those shown in this chart, can enhance RSI forecasting. Therefore, a strong price prediction model like this can support more robust RSI-based trading strategies by anticipating potential overbought or oversold conditions before they occur.



**Figure 6: Mean Square values**

The image shows a training log from a machine learning model in a Google Colab notebook, likely used for stock price prediction. While the diagram itself does not display the **Relative Strength Index (RSI)** directly, RSI can be integrated into such models as a valuable input feature. The RSI, which measures momentum by evaluating the ratio of recent gains to losses, can help machine learning models learn more nuanced patterns of market behavior. Including RSI in the training data can enhance the model's ability to predict price direction changes, especially in identifying **potential reversals** at overbought ( $RSI > 70$ ) or oversold ( $RSI < 30$ ) levels. In summary, while the training output focuses on model performance, RSI can be a powerful complementary input to improve forecasting accuracy in stock prediction tasks like the one being developed in the notebook.

## 5. Conclusion and Future Enhancements

In conclusion, this study demonstrates that Decision Tree Regressor models can effectively forecast stock closing prices using key market indicators such as open, high, low, and volume. The model exhibited strong predictive performance, achieving a high  $R^2$  score and low error metrics on synthetic stock data. Its interpretability and ability to capture non-linear relationships make it a promising tool for stock market forecasting.

However, the current study is limited by the use of synthetic data and a relatively simple model. Future enhancements could include the incorporation of more complex ensemble methods like Random Forests or Gradient Boosting (e.g., XGBoost) to improve prediction accuracy and generalization. Additionally, applying the model to real-world financial datasets, integrating technical indicators such as Relative Strength Index (RSI) and Moving Averages, and exploring deep learning architectures may further refine predictions. Long-term improvements could also involve dynamic model updating to adapt to market volatility and real-time prediction capabilities.

## 6. References

1. Chakravorty, A., & Elsayed, N. (2025). A comparative study of machine learning algorithms for stock price prediction using insider trading data. arXiv preprint arXiv:2502.08728. <https://arxiv.org/abs/2502.08728>
2. Chipwanya, R. (2023). Stock market directional bias prediction using ML algorithms. arXiv preprint arXiv:2310.16855. <https://arxiv.org/abs/2310.16855>
3. Du, S., Qiu, J., & Ding, W. (2023). Research on decision tree in price prediction of low priced stocks. In Proceedings of the 15th International Conference on Agents and Artificial Intelligence (pp. 123–130). SCITEPRESS. <https://www.scitepress.org/Papers/2023/122845/122845.pdf>
4. Hong, Y. (2025). Boosting the accuracy of stock market prediction via multi-layer hybrid MTL structure. arXiv preprint arXiv:2501.09760. <https://arxiv.org/abs/2501.09760>
5. Teixeira, D. M., & Barbosa, R. S. (2025). Stock price prediction in the financial market using machine learning models. *Computation*, 13(1), 3. <https://doi.org/10.3390/computation13010003>
6. Wang, J. (2025). Predicting stock prices using permutation decision trees and strategic trailing. arXiv preprint arXiv:2504.12828. <https://arxiv.org/abs/2504.12828>
7. Wang, J. (2025). Tesla stock prediction with SVM, decision tree and random forest. *International Journal of Advanced Computer Science and Applications*, 16(5), 342–350. [https://www.researchgate.net/publication/391006005\\_Tesla\\_stock\\_prediction\\_with\\_SVM\\_decision\\_tree\\_and\\_random\\_forest](https://www.researchgate.net/publication/391006005_Tesla_stock_prediction_with_SVM_decision_tree_and_random_forest)
8. Zheng, J., Xin, D., Cheng, Q., Tian, M., & Yang, L. (2024). The random forest model for analyzing and forecasting the US stock market in the context of smart finance. arXiv preprint arXiv:2402.17194. <https://arxiv.org/abs/2402.17194>