

Predicting Highly Volatile Stock Price with XGBoost: The Case of Tesla Inc. (2023–2025)

Rouaba Mohammed ^{*1}

¹ Ibn Khaldoun University Tiaret (Algeria), mohammed.rouaba@univ-tiaret.dz

Received: 11/06/2025

Accepted: 22/12/2025

Published: 28/12/2025

Abstract:

This study investigates the effectiveness of the XGBoost algorithm in forecasting the daily stock prices of Tesla Inc., a highly volatile equity in the U.S. stock market. Using historical data spanning from January 3, 2023 to June 9, 2025, collected from Yahoo Finance, the research constructs a predictive model based on machine learning and technical analysis principles. The methodology involves preprocessing and feature engineering using price and volume-based indicators, followed by hyperparameter tuning through GridSearchCV to enhance model performance. The results reveal that the XGBoost model achieves robust predictive accuracy, with a Mean Absolute Percentage Error (MAPE) of 7.6%, Mean Absolute Error (MAE) of 23.90, and Root Mean Squared Error (RMSE) of 33.33. These findings highlight the potential of gradient boosting methods in handling noisy and nonlinear stock market data.

Keywords: Stock price forecasting; Technical indicators; XGBoost; Tesla Inc.; Financial volatility.

Jel Classification Codes : C45, G11, G17

1. INTRODUCTION

Eugene Fama found that there is ample evidence supporting the Efficient Market Hypothesis, which suggests that security prices “fully reflect” relevant information at any given time: Therefore, neither technical nor fundamental analysis is effective in consistently finding abnormal returns under efficient market conditions (Malkiel, 2003).

But in reality, investors earn a certain amount of returns by applying various prediction methods, usually technical analysis, applied across

* *Corresponding author*

different asset markets (Stanković et al., 2015). Oriani & Coelho found that technical indicators when used as inputs can improve the accuracy of stock forecasts and the combination of various indicators as inputs can further improve forecasting performance (Oriani & Coelho, 2016).

Investment in the stock market is considered to have a high risk with high returns, attracting the attention of many investors and economists (Ajiga et al., 2024). However, information about stocks is usually incomplete, complex, uncertain and vague, making it challenging to predict future economic performance (Eggen et al., 2025). Time series data analysis helps in identifying patterns, trends, and periods or cycles present in data in the stock market (Hyndman & Athanasopoulos, 2018).

This makes time series analysis and forecasting an important area of research. The continuous evolution in the field of Machine Learning and Artificial Intelligence and the fact that financial market information is becoming more accessible to a large number of investors results in the emergence of advanced trading algorithms that consequently begin to have a significant influence on market behavior (Nobre & Neves, 2019).

Extreme Gradient Boosting, or XGBoost for short is an efficient open-source implementation of the gradient boosting algorithm developed by Tianqi Chen and Carlos Guestrin. The most important factor behind XGBoost's success is its scalability and speed in all scenarios (Rizkallah, 2025). The system runs ten times faster than popular existing solutions on a single machine and scales up to billions of examples in distributed or memory-constrained settings (Chen, 2016).

XGBoost dominates structured or tabular data sets on classification and regression predictive modeling problems as evidenced by the fact that the algorithm is always used by competition winners on the competitive data science platform Kaggle (Chen & Guestrin, 2016). Although XGBoost is designed for classification and regression on tabular data sets, it can also be used for time series forecasting (Chenchen, 2024).

1.1 Research Problem:

Despite the growing use of machine learning models in financial forecasting, accurately predicting stock prices in highly volatile markets

remains a difficult task due to uncertainty, rapid fluctuations, and complex nonlinear relationships in the data. Tesla Inc., known for its significant price volatility, provides an appropriate case study to evaluate the robustness of advanced prediction models.

Accordingly, this study seeks to address the following research problem:

How accurate is the XGBoost algorithm in predicting Tesla's stock price using technical indicators?

1.2 Research Hypotheses:

Based on the research problem and the reviewed literature, this study proposes the following hypotheses:

H1: The XGBoost model is capable of predicting Tesla's stock prices with a high level of accuracy.

H2: The inclusion of technical indicators (such as SMA, EMA, RSI, and MACD) significantly improves the predictive performance of the XGBoost model.

H3: The XGBoost model is effective in capturing nonlinear patterns in highly volatile stock price data.

1.3 Objectives of the Study:

The main objectives of this study are to:

- apply the XGBoost algorithm to forecast Tesla's daily stock prices;
- evaluate the predictive performance of the model using standard error metrics;
- examine the contribution of technical indicators to model accuracy;
- assess the suitability of XGBoost for highly volatile financial time series.

1.4 Previous Studies and Critical Review :

A growing body of literature has examined the application of machine learning techniques in stock price forecasting. Long et al. (2019) demonstrated that deep learning-based feature engineering can significantly improve the prediction of stock price movements by capturing complex nonlinear patterns. Nobre and Neves (2019) combined wavelet transforms

with XGBoost and reported enhanced trading performance in financial markets. Similarly, Nabipour et al. (2020) and Chong et al. (2017) found that ensemble and deep learning models often outperform traditional statistical approaches in modelling financial time series.

Several studies have emphasized the importance of technical indicators as input features. Oriani and Coelho (2016) showed that the use of indicators such as moving averages improves forecasting accuracy, while Yun et al. (2021) confirmed that integrating RSI and MACD within hybrid models leads to better predictive performance. Kara et al. (2011) and Patel et al. (2015) also reported that combining multiple indicators with machine learning techniques yields superior results compared to single-model approaches.

Despite these contributions, existing studies suffer from a number of limitations. First, many works focus on market indices or diversified portfolios rather than highly volatile individual stocks, which limits their applicability to assets such as Tesla. Second, several studies concentrate mainly on improving prediction accuracy without providing sufficient analysis of feature importance, reducing the interpretability and practical value of the results. Third, most prior research relies on historical periods that do not fully reflect recent market conditions characterized by heightened volatility and rapid structural changes. Furthermore, differences in datasets, evaluation metrics, and experimental settings make it difficult to directly compare results across studies.

In light of these gaps, the present study seeks to contribute to the literature by focusing on a highly volatile stock, Tesla Inc., using recent data and by assessing not only the predictive accuracy of the XGBoost model but also the role of technical indicators in enhancing model performance.

2. RESEARCH METHODS

2.1 XGBoost and Gradient Boosting Algorithm

In performing stock price prediction as a case study of this research, the XGBoost library is used which is an implementation of the Gradient Boosting algorithm designed for computational speed and model

performance (Chen & Guestrin, 2016). The Gradient Boosting algorithm combines weak learners into a single strong learner iteratively which can be explained as follows (Friedman, 2001):

a. Input: training set $\{(x_i, y_i)\}_{i=1}^n$, differential loss function $L(y, F(x))$, number of iterations M .

b. Process:

1. Initialize the model with a fixed value:

$$F_0(x) = \operatorname{argmin} \sum_{i=1}^n L(y_i, \gamma)$$

2. For $m = 1$ to M :

a. Calculate the pseudo-residual:

$$r_{im} = - \left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]$$

b. Adjust base learner (or weak learner, e.g. tree) closed under scaling $h_m(x)$ to the pseudoresidual, i.e. train it using the training set $\{(x_i, r_{im})\}_{i=1}^n$.

c. Calculate the multiplier γ_m by solving the following one-dimensional optimization problem:

$$\gamma_m = \operatorname{argmin} \sum_{i=1}^n L(y_i, F_{m-1}(x_i) + \gamma h(x_i))$$

d. Update the model

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x)$$

c. Output: $F_M(x)$.

2.2 Technical Indicators

Technical indicators are mathematical calculations based on historical prices, volume, or other information that aims to forecast the direction of prices in financial markets where those used in this study are (Taran-Morosan, 2009):

a. Simple Moving Average (SMA)

$$SMA_k = \frac{P_{n-k+1} + P_{n-k+2} + \dots + P_n}{k}$$

$$SMA_k = \frac{1}{k} \sum_{i=n-k+1}^n P_i$$

Where P_i : the stock price at time, k : the number of periods over which the average is calculated (i.e., the window size) and n the index of the most recent time period.

b. Exponential Moving Average (EMA)

$$EMA_k = \alpha x_k + (1 - \alpha)S_{k-1}$$

$$EMA_k = S_{k-1} + \alpha(x_k - S_{k-1})$$

With α as the smoothing factor and $0 \leq \alpha \leq 1$, S_k is the simple weighted average of the current observation x_k and prior statistics S_{k-1} .

c. Moving Average Convergence/Divergence (MACD)

$$MACD = EMA_{12} - EMA_{26}$$

d. Relative Strength Index (RSI)

$$RS = \frac{EMA(U, n)}{EMA(D, n)}$$

$$RSI = 100 - \frac{100}{1 + RS}$$

Where: U as the incremental period, D as the decline period, RS as an average ratio called relative strength factor and RSI as a conversion of RS into an index between 0 and 100.

2.3 Hyperparameters in XGBoost

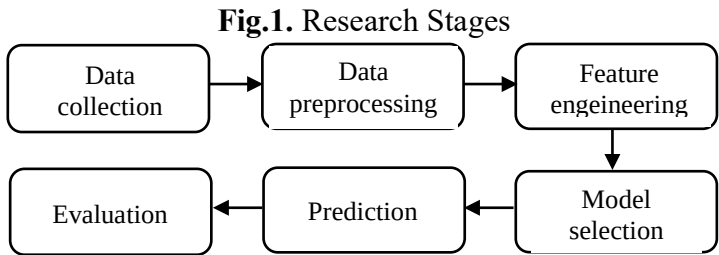
The following hyperparameters should be considered for tuning the Gradient Boosting ensemble and their effect on model performance (Chen & Guestrin, 2016):

- a. `gamma` is the minimum loss reduction required to create further partitions at the leaf nodes of the tree. The larger the `gamma`, the more conservative the algorithm.
- b. `learning_rate` is the step size shrinkage used in the update to prevent overfitting. After each boosting step, the new feature weights can

be obtained immediately, and `learning_rate` shrinks the feature weights to make the boosting process more conservative.

- c. `max_depth` is the maximum depth of the tree. Increasing this value will make the model more complex and more likely to overfit. 0 indicates no depth limit. Be careful as XGBoost aggressively consumes memory when training deep trees.
- d. `n_estimators` is the number of decision trees. Remember that decision trees are added to the model sequentially in an attempt to correct and improve the predictions made by previous trees. Thus, more trees are often better.

2.4 Research Stages



Source: Prepared by the researcher.

The XGBoost model to predict Tesla’s stock price is carried out in 5 stages as follows:

- a. Data collection, namely collecting Tesla’s stock price data from 2023 to 2025 downloaded from the site finance.yahoo.com and visualize it using `matplotlib` and `plotly` libraries.
- b. Data preprocessing is converting data into a format that is easier and more effective for further processing using the `pandas` library.
- c. Feature engineering, namely creating relevant features, namely technical indicators used in the data set using the `pandas` library and visualizing them using the `matplotlib` and `plotly` libraries.
- d. Model selection, namely entering data into the XGBoost model by splitting the data into training data (train) and test data (test) while also entering hyper parameters to determine the best results using the

XGBRegressor function in the XGBoost library.

- e. Prediction is predicting data using the predict function in the XGBoost library and visualizing it using the matplotlib and plotly libraries.
- f. Evaluation, namely determining the technical indicator features that play the most important role in modeling (feature importance) and measuring how accurate the prediction results made by the model (performance evaluation) using Mean Absolute Percentage Error (MAPE) with the equation:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t F_t}{A_t} \right|$$

where A_t is the actual value at time t while F_t is the forecast value at time t . The actual value is subtracted from the predicted value and then divided by the actual value and made into absolute values. The results are then visualized using the matplotlib and plotly libraries.

3. RESULTS AND DISCUSSION

Figure 2 shows the evolution of Tesla's closing prices during the period 03/01/2023 - 09/06/2025.

Fig.2. Tesla's Stock Closing Price Chart for the period 2023 to 2025



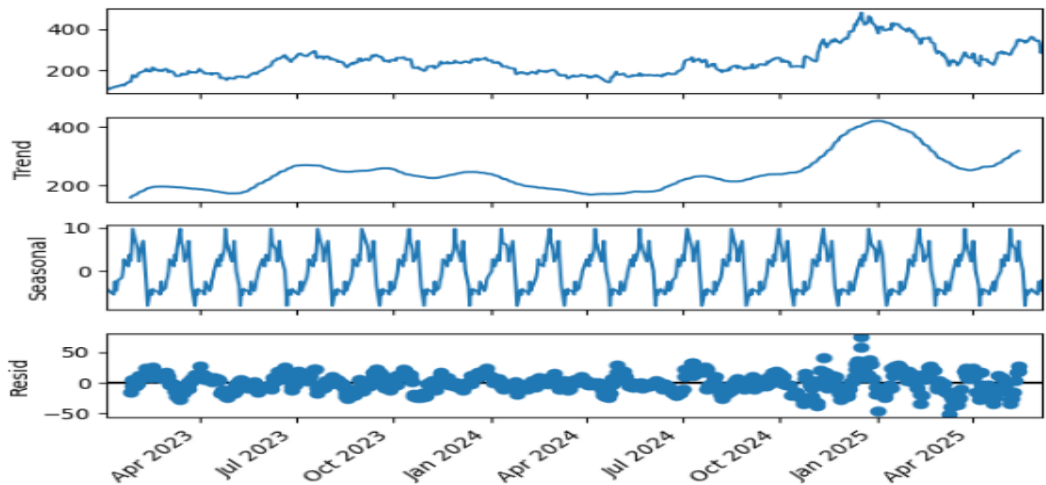
Source: Prepared by the researcher based on Python outputs.

Figure 3 illustrates the daily closing prices of a set of stocks and the decomposition of the daily closing stocks by taking the decomposition function, resulting in the trend, seasonality and residual plots of the stock prices. It can be seen that the trend of this group of stock prices first rises and then tends to fall, which is consistent with the time series rule, and the

*Predicting Highly Volatile Stock Price with XGBoost : The Case of Tesla Inc.
(2023–2025)*

stock prices have obvious seasonality and the residuals tend to be smooth. Therefore, this set of stock data is suitable for price prediction model training.

Fig.3. Decomposition Chart (Observed, Trend, Seasonal, Residual) Tesla’s Stock Price for the Period 2023 to 2025



Source: Prepared by the researcher based on Python outputs.

Before performing model prediction, it is necessary to perform feature engineering using technical indicators as shown in table 1 (Zhang, 2022).

Table 1. Technical Indicators for Feature Engineering

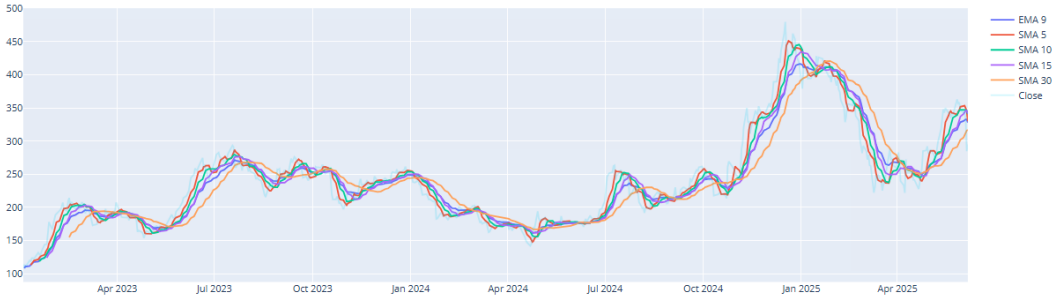
Technical Indicator	Description
EMA_9	9-day data weighted moving average stock price
SMA_5	5-day data of short-term average stock price movement
SMA_15	15-day data of short-term average stock price movement
SMA_30	30-day data of short-term average stock price movement
RSI	$\text{SMA}(\text{MAX}(\text{Close}-\text{LastClose},0),N,1)/\text{SMA}(\text{ABS}(\text{Close}-\text{LastClose}),N,1)*100$
MACD	EMA12-EMA26

Source: Prepared by the researcher.

The following is a visualization of engineering features using EMA (Exponential Moving Average) and SMA (Simple Moving Average)

technical indicators on Tesla's stock prices for the period 2023 to 2025 as shown in Figure 4.

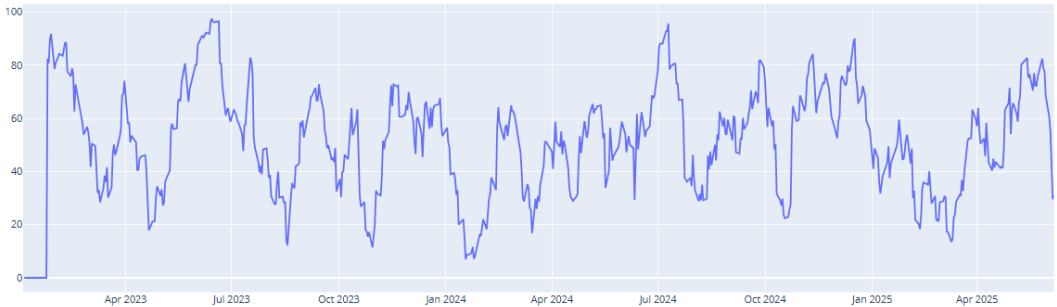
Fig.4. Chart of Technical Indicators (EMA and SMA) Tesla's Stock Price for the period 2023 to 2025



Source: Prepared by the researcher based on Python outputs.

The following is a visualization of engineering features using the RSI (Relative Strength Index) technical indicator on the stock price of Tesla for the period 2023 to 2025 as shown in Figure 5.

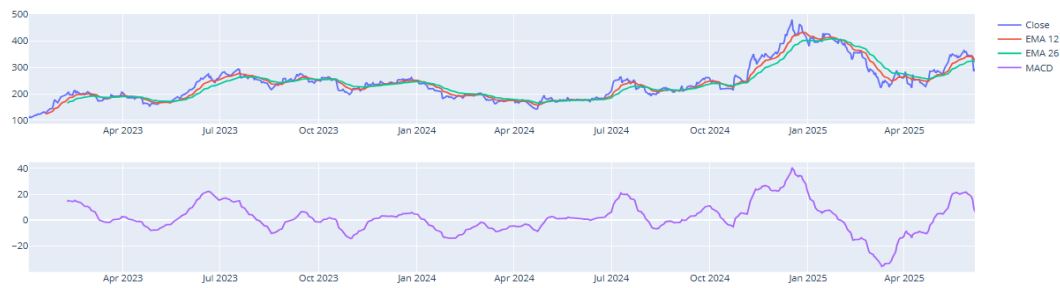
Fig.5. RSI Technical Indicator Chart of Tesla's Stock Price for the period 2023 to 2025



Source: Prepared by the researcher based on Python outputs.

The following is a visualization of feature engineering using the MACD (Moving Average Convergence Divergence) technical indicator on the stock price of Tesla's for the period 2023 to 2025 as shown in Figure 6.

Fig.6. MACD Technical Indicator Chart of Tesla’s Stock Price for the period 2023 to 2025



Source: Prepared by the researcher based on Python outputs.

In this study, stock price data from the period January 3, 2023 to December 10, 2024 (80% of total data set) was utilized as the training set to develop and calibrate the prediction model. Subsequently, the data from December 11, 2024 to June 9, 2025 was reserved as the testing set to evaluate the generalization performance of the final XGBoost model after completing the training phase. This division ensures a clear temporal separation between model development and evaluation, thereby reducing the risk of information leakage. The predictive performance on the test data is illustrated in Figure 7.

Fig.7. Data Split Graph (Train, Test) Tesla’s Stock Price for the period 2023 to 2025



Source: Prepared by the researcher based on Python outputs.

When using the XGBoost algorithm, it is necessary to search for hyper parameters according to the characteristics of the stock price. The task of hyper parameter tuning is left to GridSearchCV which will try every

possible parameter among all candidates, and the best performing parameter is the final result as shown in table 3. The main disadvantage of using GridSearchCV is that it is more time consuming, but it guarantees the accuracy and validity of the results (Zhang, 2022).

Table 2. Hyper Parameter Search Settings

Hyper Parameters	Value
n_estimators	[100, 200, 300, 400]
learning_rate	[0.001, 0.005, 0.01, 0.05]
max_depth	[2,4,6,8,10, 12, 14]
gamma	[0.001, 0.005, 0.01, 0.02]
random_state	[42]

Source: Prepared by the researcher.

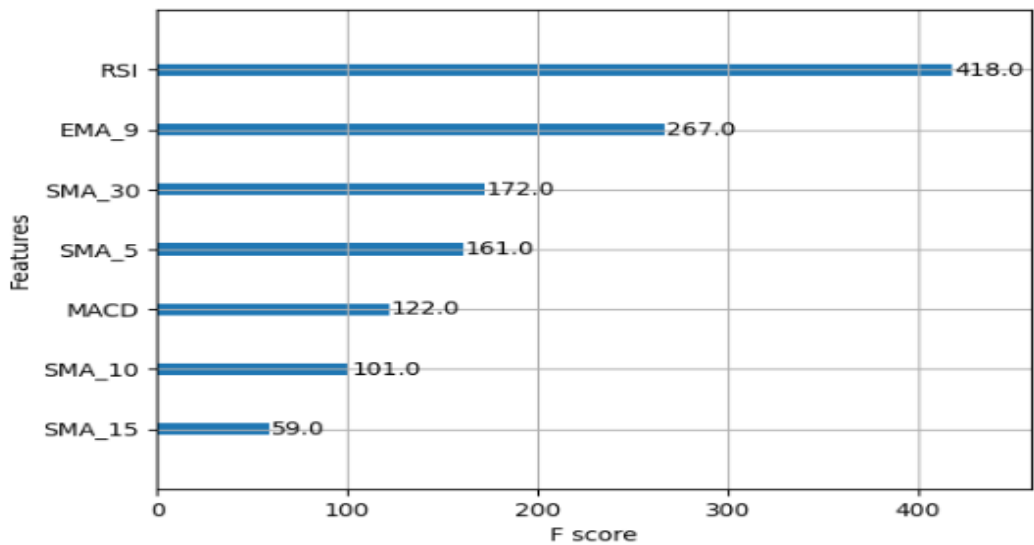
Table 3. Best Hyper Parameter Search Results

Hyper Parameters	Value
n_estimators	400
learning_rate	0.05
max_depth	2
gamma	0.001
random_state	42

Source: Prepared by the researcher based on Python outputs.

Figure 8 displays the results of the feature importance analysis conducted to assess the relative contribution of technical indicators in predicting Tesla's stock prices for the year 2025. Among the selected features, the Relative Strength Index (RSI) emerged as the most influential indicator, recording the highest importance score compared to other indicators such as EMA, SMA and MACD. This finding highlights the predictive power of momentum-based indicators in capturing price dynamics, particularly in the context of highly volatile assets like Tesla.

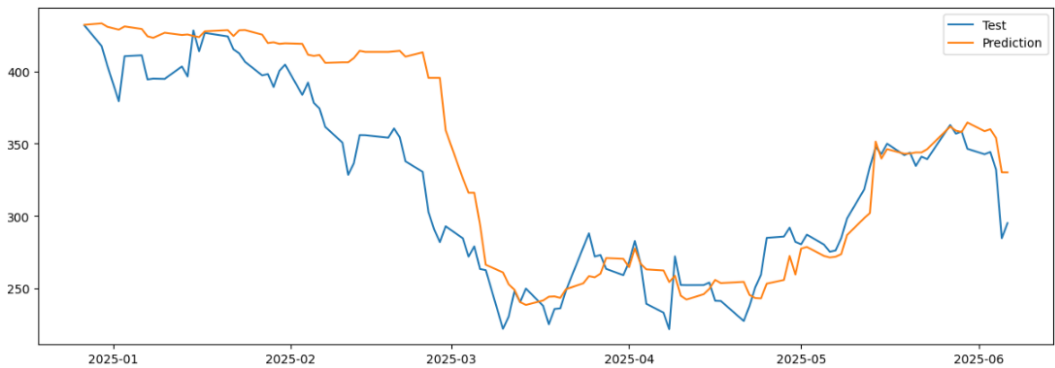
Fig.8. Feature Importance Graph of Tesla’s Stock Price Prediction in 2025



Source: Prepared by the researcher based on Python outputs.

The predictive performance of the XGBoost model for the test period is presented in Figure 9. The model was trained using optimal hyperparameters $\gamma = 0.001$, $\text{learning_rate} = 0.05$, $\text{max_depth} = 2$, and $n_estimators = 400$ identified through a comprehensive tuning process.

Fig.9. Tesla’s Stock Price Prediction Chart for the 2025 period



Source: Prepared by the researcher based on Python outputs.

The model produced a Mean Absolute Error (MAE) of 23.90, a Mean Squared Error (MSE) of 1111.17, a Root Mean Squared Error (RMSE) of 33.33, a Mean Absolute Scaled Error (MASE) of 2.16, and a Mean

Absolute Percentage Error (MAPE) of 7.6%. These metrics indicate a relatively high level of predictive accuracy, especially considering Tesla's well-known volatility.

The low MAPE suggests that the model is capable of capturing the underlying trends in the stock price with a reasonable degree of reliability. Furthermore, the results highlight the robustness of the XGBoost algorithm in modeling financial time series characterized by nonlinear patterns and noisy fluctuations.

Visual inspection of the predicted versus actual prices reveals that the model tracks short-term fluctuations and overall trend movements reasonably well, though it occasionally lags during periods of extreme volatility. This is consistent with the behavior of tree-based models, which are more sensitive to recent historical patterns than to abrupt structural changes.

The findings of this study are consistent with a growing body of literature emphasizing the effectiveness of machine learning and ensemble methods in stock price forecasting. For instance, Nobre and Neves (2019) reported that combining XGBoost with advanced feature extraction techniques leads to superior trading and forecasting performance in financial markets. Similarly, Nabipour et al. (2020) and Chong et al. (2017) found that ensemble and deep learning models often outperform traditional statistical approaches when dealing with nonlinear and noisy financial time series.

The importance of technical indicators observed in this study also aligns with the results of Oriani and Coelho (2016), who demonstrated that moving averages significantly improve forecasting accuracy, and with Yun et al. (2021), who showed that integrating RSI and MACD enhances predictive performance in hybrid models. The dominance of RSI in the feature importance ranking supports these conclusions and confirms the role of momentum indicators in capturing short-term market behavior.

Compared to studies such as Zhang (2022), which applied XGBoost to stock prediction and reported competitive error rates, the MAPE of 7.6% obtained in this research falls within a comparable range, despite focusing

on a particularly volatile stock and a more recent period characterized by heightened market uncertainty. This suggests that XGBoost remains robust across different market conditions and datasets.

However, unlike many previous works that focus on market indices or diversified portfolios, this study concentrates on a single highly volatile stock, thereby extending prior research by demonstrating that XGBoost can maintain reliable performance even under extreme price fluctuations. In addition, the explicit analysis of feature importance addresses a common limitation in earlier studies, which often prioritize accuracy over interpretability.

Overall, the results corroborate earlier evidence that gradient boosting techniques are well suited for financial time series forecasting, while also contributing new insights by validating their applicability to recent, highly volatile market data and by highlighting the critical role of technical indicators in enhancing model performance.

4. CONCLUSION

This paper demonstrates the applicability and effectiveness of the XGBoost algorithm in predicting the daily stock prices of a highly volatile asset, namely Tesla Inc. The model's relatively low forecasting error metrics validate the strength of gradient boosting methods in dealing with complex, non-linear time series in financial markets.

By combining historical stock data with machine learning techniques and technical indicators, the study provides empirical evidence that supports the integration of AI tools into financial forecasting frameworks. The findings are particularly relevant for retail traders, portfolio managers, and data scientists seeking to enhance decision-making under market uncertainty.

Future research may expand this work by integrating sentiment analysis, macroeconomic variables, or by comparing the performance of XGBoost with other deep learning models such as LSTM or GRU. Additionally, further evaluation across different market conditions or industry sectors could help generalize the effectiveness of the approach.

5. Bibliography List:

- Ajiga, D. I., Adeleye, R. A., Tubokirifuruar, T. S., Bello, B. G., Ndubuisi, N. L., Asuzu, O. F., & Owolabi, O. R. (2024). Machine learning for stock market forecasting: a review of models and accuracy. *Finance & Accounting Research Journal*, 6(2), 112-124.
- Chen, T. (2016). XGBoost: A Scalable Tree Boosting System. *Cornell University*.
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*,
- Chenchen, L. (2024). Time-series water prediction based on feature clusteringfusion and XGboost. *Proceedings of the 6th International Conference on Computing and Data Science*,
- Eggen, S., Espe, T. J., Grude, K., Ristad, M., & Sandberg, R. (2025). Financial time series uncertainty: A review of probabilistic AI applications. *Journal of Economic Surveys*.
- Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.
- Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of economic perspectives*, 17(1), 59-82.
- Nobre, J., & Neves, R. F. (2019). Combining principal component analysis, discrete wavelet transform and XGBoost to trade in the financial markets. *Expert Systems with Applications*, 125, 181-194.
- Oriani, F. B., & Coelho, G. P. (2016). Evaluating the impact of technical indicators on stock forecasting. 2016 IEEE symposium series on Computational Intelligence (SSCI),
- Rizkallah, L. W. (2025). Enhancing the performance of gradient boosting trees on regression problems. *Journal of big data*, 12(1), 35.
- Stanković, J., Marković, I., & Stojanović, M. (2015). Investment strategy optimization using technical analysis and predictive modeling in emerging markets. *Procedia Economics and Finance*, 19, 51-62.
- Taran-Morosan, A. (2009). Some Technical Analysis Indicators. *Revista Economica*, 46(3), 116-121.
- Zhang, Y. (2022). Stock price prediction method based on XGboost algorithm. 2022 International Conference on Bigdata Blockchain and Economy Management (ICBBEM 2022),