RESEARCH ARTICLE

Machine Learning for Electric Vehicle Stock Price Prediction: Analyzing Artificial Neural Network and Random Forest Performance

Montira Jarudecha[1] and **Sanya Khruahong**[1,*]

[1]Department of Computer Science and Information Technology, Faculty of Science, Naresuan University, Phitsanulok 65000, Thailand

*Corresponding author: Sanya Khruahong, sanyak@nu.ac.th

Article Information

Article History:
Received: 2 February 2025
Revised: 8 August 2025
Accepted: 24 August 2025
Published: 29 October 2025

Keywords:
Artificial Neural Network
Electric Vehicles
Machine Learning
Random Forest
Stock Market

Abstract

Forecasting the stock prices of electric vehicle (EV) companies presents a complex challenge due to market volatility and constantly changing external factors. This study aims to address a research gap in the literature, where comparative analyses of multiple machine learning models across several EV companies remain limited. Specifically, the study evaluates and compares the predictive performance of Artificial Neural Networks (ANN) and Random Forest (RF) in forecasting the stock prices of Tesla, BYD, Volkswagen, Geely, and GM using data from January 2018 to June 2023. The dataset comprises key stock market indicators—opening price, highest price, lowest price, volume, and closing price—augmented with COVID-19 pandemic data to reflect external influences on market behavior. Prior to analysis, missing values were handled using mean imputation, and data were normalized using Min-Max scaling to optimize model training. Performance was assessed using Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Mean Bias Error (MBE). The results indicate that RF generally outperforms ANN in forecasting stock prices across most companies, particularly GM (RMSE = 0.3760, MAPE = 0.8238, MBE = 0.0537) and Volkswagen (RMSE = 1.0437, MAPE = 0.6868, MBE = 0.0584). In contrast, ANN performed best for Geely (RMSE = 0.2240, MAPE = 1.4160, MBE = -0.0271), suggesting that ANN may be better suited for datasets with more consistent or specific characteristics, while RF delivered more stable performance across companies. A t-test revealed statistically significant differences in performance between RF and ANN for Volkswagen ($p = 0.0050$) and GM ($p < 0.001$), while no significant differences were found for Tesla, BYD, and Geely ($p > 0.05$), indicating that model selection should consider the specific data characteristics. This research contributes a novel approach by conducting cross-company ML model comparisons in the EV sector while incorporating external variables such as COVID-19, which are rarely addressed in prior work. The findings offer practical insights for investors, analysts, and market intelligence systems, emphasizing the importance of tailoring model selection to the characteristics of individual stock data and supporting the use of AI for more accurate investment decisions.

1. Introduction

Stock market prediction has gained significant attention in recent years, as more individuals seek profitable investment opportunities. Beyond serving as a secondary income source, stock market investments provide an alternative to traditional employment, with two main strategies: long-term investments, where stocks are held for extended periods to generate annual returns, and short-term trading (day trading), where stocks are fre-

ISSN 3088-1803 |

quently bought and sold based on price fluctuations. However, stock market investment requires experience and expertise in analyzing price trends, as uninformed decisions may lead to significant financial losses.

Traditional stock market forecasting methods (Li and Bastos, 2020; Sutheebanjard and Premchaiswadi, 2009) primarily rely on Technical Analysis (TA) and Fundamental Analysis (FA). TA examines historical price patterns and market behaviors using statistical techniques to predict future trends. In contrast, FA evaluates economic and financial indicators to assess the intrinsic value of stocks. Despite their effectiveness, both approaches demand extensive data collection and in-depth analysis, which can be time-consuming and prone to human error.

With advancements in Artificial Intelligence (AI) (Say et al., 2025), machine learning (ML) (Jordan and Mitchell, 2015) has emerged as a promising alternative for stock market prediction. AI-powered models can process vast amounts of data in real time, uncover hidden patterns, and adapt dynamically to new market conditions. ML algorithms, such as Artificial Neural Networks (ANN) (Agatonovic-Kustrin and Beresford, 2000) and Random Forest (RF) (Biau and Scornet, 2016), have demonstrated strong predictive capabilities in financial forecasting.

One of the most rapidly growing sectors in the stock market is the Electric Vehicle (EV) industry (Larminie and Lowry, 2012). EVs have gained global popularity due to their cost efficiency, environmental benefits, and reduced dependence on fossil fuels. According to recent reports, global EV sales reached over 10 million units in 2022 and are projected to grow to 14 million units by the end of 2023, representing a 35% year-over-year increase (Thompson, 2024). The top five EV manufacturers in Q1 2023 were BYD (China), Tesla (USA), Volkswagen (Germany), Geely (China), and GM (USA). As EV companies expand production and sales, their stock market performance has become an important area of study.

Previous research (Behera et al., 2020; Daori et al., 2022; Wanjawa and Muchemi, 2014) has shown that ML models such as ANN and RF are highly effective for stock price prediction. Given the increasing significance of the EV market, this study aims to evaluate and compare the performance of ANN and RF in predicting EV stock prices. The dataset includes historical stock prices from Yahoo Finance (Lawrence et al., 2017) and COVID-19 pandemic data from API Ninjas (Ninjas, 2024), considering a one-year forecasting period. The results of this research will provide insights into the most suitable ML model for long-term EV stock market prediction, helping investors make informed decisions based on AI-driven analysis.

Stock market prediction has evolved from traditional methods like technical and fundamental analysis to machine learning models such as ANN and RF (Li and Bastos, 2020; Jordan and Mitchell, 2015). Prior studies (Behera et al., 2020; Wanjawa and Muchemi, 2014) have shown that ANN and RF can effectively identify patterns in financial data and improve forecasting accuracy. Given the rapid growth of the EV industry, research on EV stock price prediction is gaining attention, with machine learning proving more reliable than conventional approaches (Larminie and Lowry, 2012).

Recent studies (Daori et al., 2022; Zhang et al., 2022) have integrated external factors such as macroeconomic conditions and global events (e.g., COVID-19) into machine learning models to enhance prediction accuracy. While ANN and RF have been widely applied in financial forecasting, their comparative performance across multiple EV companies remains underexplored. Zheng et al. (2024) proposed a hybrid approach combining Long Short-Term Memory (LSTM) and RF to improve forecasting performance; however, their study was limited to Tesla's stock. This research aims to overcome those limitations by comparing ANN and RF across multiple EV companies while integrating external data, such as the COVID-19 pandemic, over a period of five and a half years (January 2018–June 2023), to better reflect real-world market dynamics.

ANN and RF were selected over other ML models due to their respective strengths: ANN can capture nonlinear data patterns, while RF effectively handles large datasets with diverse features and reduces the risk of overfitting. Both models also run faster than LSTM and require lower computational power, making them suitable for practical applications—particularly in environments that demand fast and accurate decision-making. Therefore, this study aims to address existing research gaps by comparing the performance of ANN and RF in forecasting stock prices of multiple EV companies. It integrates historical stock price data with external factors such as COVID-19 and evaluates model performance using RMSE, MAPE, and MBE, along with t-tests to assess statistical significance between the models. The findings provide practical guidance for selecting suitable models to support more accurate and reliable investment decisions in the EV stock market.

2. Machine Learning Models

Machine Learning (ML) refers to the ability of a computer system to learn and improve its performance autonomously based on available data, without requiring explicit programming of rules or predefined processes (El Naqa and Murphy, 2015; Hinton, 2011; Pomboomee et al., 2023; Pramote et al., 2023). The core concept of ML is to enable computers to analyze data independently, identify patterns, and develop models that can be used for prediction or decision-making. ML algorithms learn from historical data and experiences, continuously refining their models as new data is introduced. This adaptive learning capability allows ML systems to adjust their operations in response to changing conditions, resulting in smarter and more efficient performance over time.

In this study involves two machine learning models: RF (see Section 2.1) and ANN (see Section 2.2), detailed as follows.

2.1 Random Forest Model (RF)

The RF model (Biau, 2012; Speiser et al., 2019) is an ensemble learning method used for both classification and regression tasks. It is a versatile technique that can be applied to various types of problems, including categorizing data and predicting numerical values. Due to its high efficiency and capability to handle large and complex datasets, Random Forest has become widely adopted in research institutions and business applications.

The model consists of multiple decision trees, which are generated from randomly selected subsets of training data. Each decision tree is trained using randomly chosen features from the dataset. When making a prediction, the model aggregates results from all decision trees through ensemble learning and applies a voting mechanism to determine the most accurate prediction (Breiman, 2001).

In this study, the `random_state` parameter was set to 42 to ensure the reproducibility of results (Zhao et al., 2024), and the `n_estimators` parameter in the `RandomForestRegressor` class was set to 200, meaning the model generates 200 unique decision trees for prediction. Additionally, `GridSearchCV` was employed for automatic hyperparameter tuning to determine the optimal combination of key parameters such as `max_depth`, `min_samples_split`, and `min_samples_leaf` based on the specific characteristics of each company's dataset. This automated tuning approach improves forecasting accuracy while minimizing reliance on manual parameter selection, which can introduce bias.

2.2 Artificial Neural Network Model (ANN)

The ANN model (Agatonovic-Kustrin and Beresford, 2000) is a computational framework designed to mimic the structure of biological neural networks in the human brain. It is widely utilized in ML and AI applications to solve complex computational problems by simulating the way human neurons process information and learn from data.

The structure of ANN consists of nodes (neurons) that are interconnected through weighted connections (Dongare et al., 2012), where each weight can be adjusted during the learning process. ANNs are composed of one or more layers, with each layer playing a specific role in processing data. In this study, the ANN model is structured as follows:

- **Input Layer:** Consists of 5 nodes, representing the variables COVID Cases, Open, High, Low, and Volume.
- **Hidden Layers:** Includes 2 hidden layers, with 100 and 50 nodes respectively, defining the model's complexity.
- **Output Layer:** Consists of 1 output node, representing the closing stock price.
- **Random State:** Set to 42 to ensure that results are reproducible each time the model is executed.
- **Hyperparameter Tuning:** Automated tuning was performed using `GridSearchCV` to select the optimal values for key parameters such as learning rate, batch size, and activation function, customized for each company's dataset to achieve the highest predictive performance.

ANNs are effective in solving classification and regression problems, as well as data transformation and forecasting. The learning process, known as training, involves adjusting the connection weights to minimize errors and enhance predictive accuracy. By continuously learning from new data, ANN can adapt and refine its predictive capabilities, similar to human cognitive learning, making it a powerful tool for stock market prediction (Wanjawa and Muchemi, 2014).

3. Research Methodology

This study employs historical stock market data from Yahoo Finance, covering the period from January 1, 2018, to June 30, 2023, for five major EV companies: BYD, Tesla, Volkswagen, Geely, and GM. These companies were selected due to their significant roles in the global EV industry in terms of sales volume, production capacity, and consistent visibility in the stock market.

3.1 Data Collection and Selection

The dataset used in this study consists of financial stock indicators and external economic factors that influence stock price movements. The data sources and selection criteria are detailed below.

1) Stock Market Data

Stock market data were retrieved from Yahoo Finance, which provides publicly available historical trading data. The dataset includes the following key stock indicators, which are commonly used in financial forecasting and stock market analysis:

- **Opening Price**: The price at which a stock begins trading each day.
- **Highest Price**: The maximum price reached during a trading session.
- **Lowest Price**: The minimum price recorded within the trading period.

- **Closing Price**: The final price at which a stock is traded before the market closes.
- **Trading Volume**: The number of shares exchanged during a given time frame.

These indicators were selected because they directly influence investor decision-making and provide insights into stock price trends and market fluctuations.

2) External Factors (COVID-19 Data)

To improve prediction accuracy, this study incorporates external market influences, particularly COVID-19 pandemic statistics. The data were retrieved from API Ninjas, which provides real-time and historical pandemic data. The following variables were included:

- **New daily infection cases**: To measure short-term fluctuations in market sentiment and trading behavior.
- **Cumulative infection cases**: To assess the long-term economic impact of the pandemic on the stock market.

COVID-19 was selected as an external factor due to its significant impact on global financial markets, supply chains, and investor sentiment. By integrating non-market influences, the study aims to provide a more comprehensive stock market forecasting model, capturing the effects of external disruptions on EV stock trends.

3.2 Forecasting Period and Data Preprocessing

1) Forecasting Period

This study focuses on predicting EV stock prices over a period of five and a half years, from January 2018 to June 2023. This timeframe was selected to capture a diverse range of market conditions, including pre-pandemic normalcy, the COVID-19 crisis, and the subsequent recovery phase. The training phase utilized data from January 2018 to May 2022, while the forecasting phase was conducted for the period from June 1, 2022, to June 30, 2023. This one-year test window allows for the evaluation of model performance under real-world market fluctuations.

2) Data Preprocessing

Before training the machine learning models, data preprocessing was conducted to enhance model accuracy and efficiency. The following preprocessing steps were applied:

- **Handling Missing Values**: For stock price data, linear interpolation was used to impute missing values, while COVID-19 case data employed the forward fill method to maintain continuity and minimize negative impacts on model training.
- **Data Normalization**: Since stock prices and COVID-19 cases exist on different scales, normalization was applied to scale numerical values within a standard range. This prevents larger values from dominating the learning process, ensuring that all variables contribute equally to predictions.
- **Data Splitting**: The dataset was divided into a training set (80%) and a testing set (20%) to enable the model to learn from a substantial portion of the data and assess its predictive performance on previously unseen data.

Normalization plays a crucial role in improving model performance and stability, as machine learning models perform better when input data are standardized and free of inconsistencies.

3.3 Machine Learning Models

This study evaluates two widely used machine learning models for stock price prediction: ANN and RF. These models were chosen because they have demonstrated high performance in financial forecasting and are widely used in stock market prediction research.

4. Model Evaluation

To evaluate the performance of the models, a comparison was conducted between two machine learning techniques: RF and ANN. The evaluation was based on stock market data from five different companies: BYD, Tesla, Volkswagen, Geely, and GM. The models' performance was assessed using three key metrics: Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Mean Bias Error (MBE) (Chicco et al., 2021; Zhang et al., 2022).

4.1 Root Mean Square Error (RMSE)

Measures the average deviation between predicted and actual stock prices, where lower values indicate better predictive accuracy. The RMSE value can be calculated using the following equation:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - F_i)^2} \tag{1}$$

4.2 Mean Absolute Percentage Error (MAPE)

Assesses percentage errors in predictions, providing an understanding of relative forecasting performance. The

MAPE value can be calculated using the following equation:

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{|O_i - F_i|}{O_i} \times 100\right) \tag{2}$$

4.3 Mean Bias Error (MBE)

Evaluates whether the model tends to overestimate or underestimate stock prices on average. The MBE value can be calculated using the following equation:

$$\text{MBE} = \frac{1}{n}\sum_{i=1}^{n}(O_i - F_i) \tag{3}$$

where O_i represents the actual closing stock price, F_i represents the predicted closing stock price, and n denotes the total number of prediction periods.

These metrics ensure a comprehensive performance evaluation, allowing for a clear comparison between ANN and RF models in predicting EV stock prices.

5. Results and Discussion

The research findings are presented in Fig. 1, which illustrate the comparison between the actual closing stock prices and the predicted closing stock prices using the ANN model. Similarly, Fig. 2 depict the comparison between the actual and predicted closing stock prices obtained from the RF model.

Fig. 1 provide a detailed analysis of the ANN model's predictive performance across the stock prices of five major EV companies: Tesla, BYD, Volkswagen, Geely, and GM. The results highlight the model's ability to adapt to varying stock price behaviors, demonstrating both accuracy and robustness. For Tesla, the ANN model effectively tracks the overall trends despite the challenges posed by the stock's high volatility, with minor deviations only occurring during extreme peaks and troughs. Similarly, BYD, which experiences moderate fluctuations, exhibits predictions that closely follow both upward and downward price movements, showcasing the model's strength in handling dynamic trends. For Volkswagen, characterized by a consistent downward trend during the testing phase, the model aligns its predictions with remarkable accuracy, capturing the steady declines with minimal errors.

In comparison, Geely and GM exhibit relatively stable price trends, allowing the ANN model to achieve near-perfect predictions. Geely's data, marked by a steep decline during the training phase and stable low prices in the testing phase, provided ideal conditions for the model to excel. Particularly in the case of Geely, the ANN model outperformed RF significantly, which could be attributed to ANN's architecture that excels in learning simple linear and non-discontinuous patterns. In such scenarios with low noise and volatility, RF tends to overfit during the training phase, leading to deviation in the test phase, while ANN can better align with stable input data through its deep learning structure. Similarly, for GM, which exhibits moderate volatility, ANN maintained high forecasting accuracy. However, in cases with higher volatility and complex trends, such as Tesla, BYD, and Volkswagen, the RF model demonstrated a clear advantage over ANN.

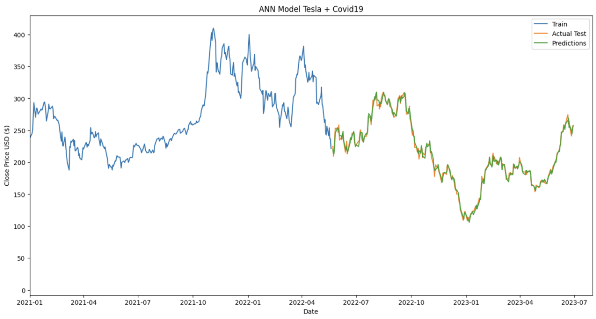
Fig. 2 showcase the performance of the RF model in predicting stock prices for Tesla, BYD, Volkswagen, Geely, and GM, with comparisons made against actual stock prices. For Tesla, the RF model effectively tracks the overall trends, demonstrating high accuracy despite Tesla's inherent price volatility. While the predicted prices align closely with the actual test prices, minor deviations occur during sharp fluctuations, particularly around price peaks and troughs. Similarly, BYD's stock prices, which show moderate variability, are predicted with notable precision, as the RF model successfully captures both upward and downward trends with minimal discrepancies, especially during volatile periods in 2022. For Volkswagen, characterized by a steady decline in stock prices, the RF model performs exceptionally well, aligning closely with actual prices throughout the testing period.

In the case of Geely, where the stock prices stabilize after an initial sharp decline, the RF model achieves near-perfect predictions during the testing phase, reflecting its ability to adapt to stable trends. Similarly, GM's stock prices, which exhibit moderate fluctuations, are predicted with high accuracy, with the RF model effectively capturing both short-term variability and longer-term trends. Overall, the RF model demonstrates robust predictive capabilities across all five companies, particularly excelling with stocks that exhibit low to moderate volatility (e.g., Geely, GM), while remaining reliable for highly volatile stocks such as Tesla and BYD.

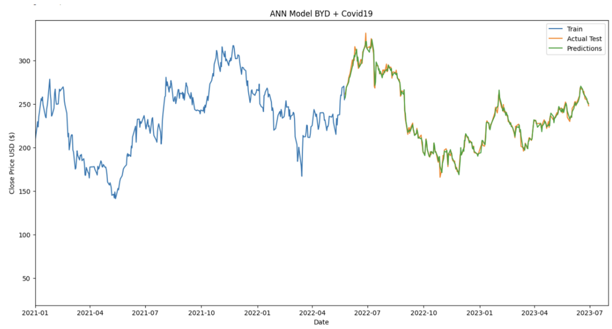
One major reason for RF's superior overall accuracy may stem from its ensemble learning nature, which aggregates results from multiple decision trees. This enhances its ability to manage data uncertainty and reduces the risk of overfitting. In contrast, while ANN is capable of deep learning, it requires careful tuning of architecture and parameters to perform optimally, and it may be more sensitive to outliers or noise in the data. Moreover, RF handles non-linear relationships well without requiring complex preprocessing, making it especially effective in diverse forecasting scenarios.

Table 1 provides a comparative analysis of the performance of ANN and RF models in forecasting the closing stock prices of five major electric vehicle (EV) companies: Tesla, BYD, Volkswagen, Geely, and GM. The evaluation metrics include RMSE, MAPE, MBE, and statistical significance testing using a paired t-test.

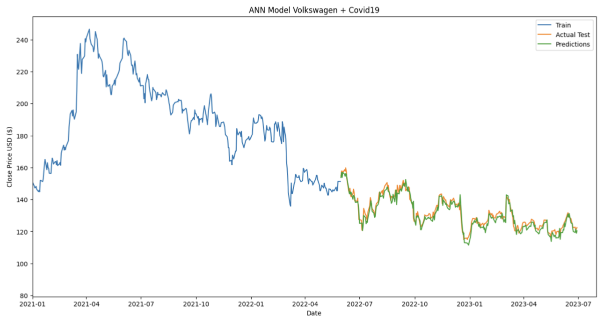
The experimental results show that in many cases, the RF model significantly outperforms ANN, particularly for stocks with moderate to high volatility such as



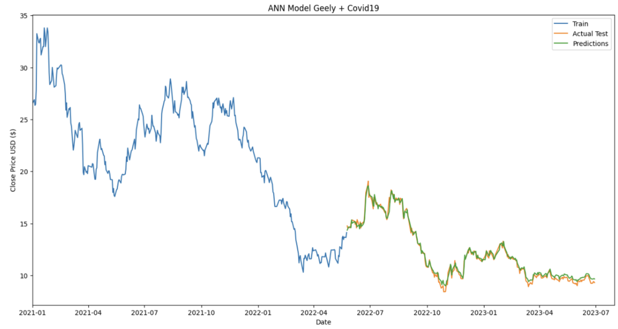
(a) Tesla



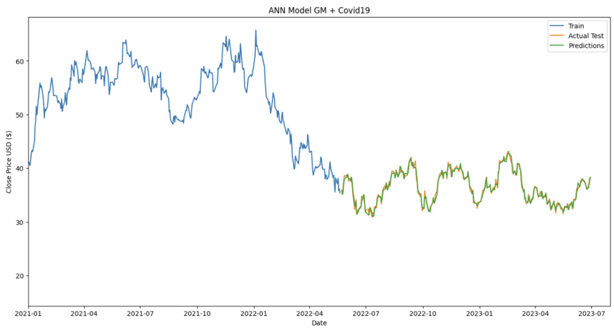
(b) BYD



(c) Volkswagen



(d) Geely



(e) GM

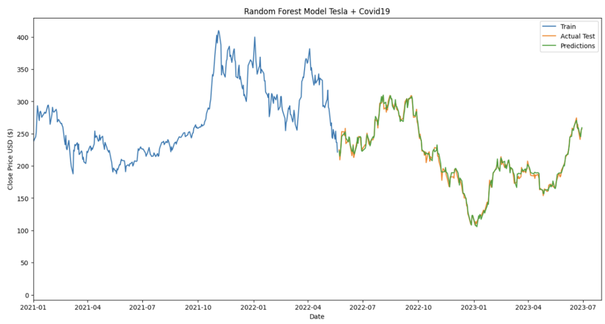
Figure 1. Comparison of predicted and actual stock prices for five electric vehicle companies using the ANN model.

Volkswagen and GM. For these companies, RF yields statistically significantly lower RMSE and MAPE values (p-value $= 0.0050$ and 1.03×10^{-11}, respectively), reflecting its strong ability to handle non-linearities and fluctuations in the market. This advantage may stem from RF's ensemble structure, which aggregates results from multiple decision trees, helping to mitigate overfitting and enhance result stability—leading to consistently accurate outcomes across various data characteristics.

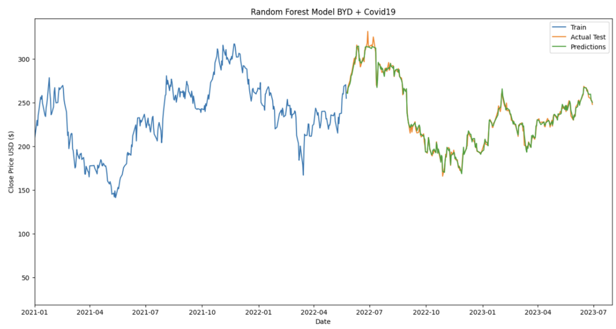
In contrast, a notable exception is observed in the case of Geely, where ANN outperforms RF across all metrics. Although the differences are not statistically significant (p-value $= 0.0839$), ANN achieves a lower RMSE (0.2240 vs. 0.2555), a lower MAPE (1.4160 vs. 1.7215), and an MBE closer to zero. This could be attributed to Geely's relatively stable stock prices following a clear initial downward trend, which aligns well with ANN's strength in learning simpler linear or sequential patterns. Such conditions are less prone to noise and volatility, allowing ANN to perform more effectively, while RF may overfit during training and produce more biased results on the test set.

For high-volatility stocks such as Tesla and BYD, RF still achieves lower RMSE values than ANN, although the differences are not statistically significant ($p > 0.05$). Nonetheless, these results indicate RF's superior capability in handling complex and unstable price behaviors, while ANN may require more refined architecture and parameter tuning to achieve comparable performance.

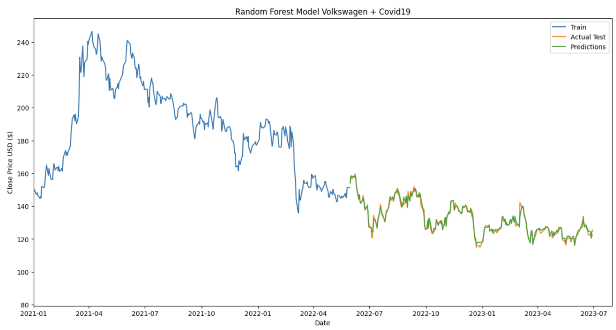
In conclusion, the RF model demonstrates consis-



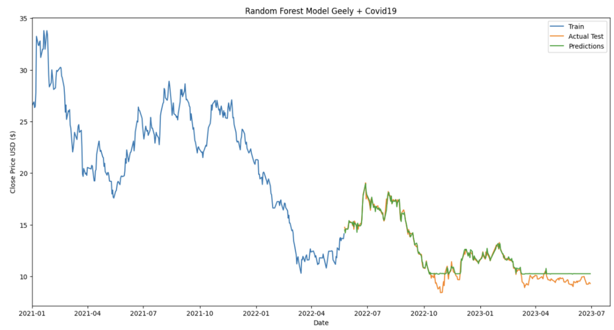
(a) Tesla



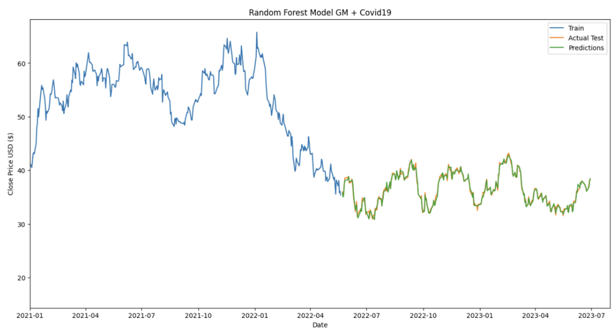
(b) BYD



(c) Volkswagen



(d) Geely



(e) GM

Figure 2. Comparison of predicted and actual stock prices for five electric vehicle companies using the RF model.

tent and accurate forecasting performance across various stock characteristics, particularly for stocks with high volatility or complex patterns. Meanwhile, ANN may yield better results in cases involving stable and less volatile price trends. Therefore, model selection should be carefully aligned with the specific nature of the dataset, as the compatibility between the data features and the model's learning mechanism plays a crucial role in determining predictive effectiveness.

6. Conclusion

Predicting stock prices in the electric vehicle (EV) industry remains a challenging task due to high volatility and sensitivity to both internal market dynamics and external economic factors. This study focused on comparing the predictive capabilities of Artificial Neural Network (ANN) and Random Forest (RF) models in forecasting the closing stock prices of leading EV manufacturers: Tesla, BYD, Volkswagen, Geely, and GM. The input variables included key technical indicators such as opening price, high, low, close, and trading volume, contextualized within the post-COVID-19 economic recovery period, which has significantly influenced global investment behaviors.

The experimental results indicate that the RF model generally outperforms the ANN model, particularly for stocks with moderate to high volatility such as GM, Volkswagen, and BYD. In these cases, RF achieved significantly lower Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) values (e.g.,

Table 1. Comparison of ANN and RF model performance using RMSE, MAPE, MBE, and t-test.

Vehicle	ANN			RF			T-stat	P-value
Company	RMSE	MAPE	MBE	RMSE	MAPE	MBE		
Tesla	4.5878	1.7305	-0.0853	4.5177	1.7787	-0.7963	0.2125	0.8319
BYD	1.1574	1.1518	-0.0511	1.1019	1.0529	-0.0680	1.8737	0.0621
Volkswagen	1.1379	0.7635	0.1081	1.0437	0.6868	0.0584	2.8272	0.0050 **
Geely	0.2240	1.4160	-0.0271	0.2555	1.7215	-0.0776	-1.7348	0.0839
GM	0.4638	1.0853	0.0554	0.3760	0.8238	0.0537	7.1039	$1.03e^{-11}$ ***

Note: P-value < 0.05 indicates statistically significant differences.

p-value = 0.0050 and 1.03×10^{-11}, respectively), highlighting its ability to handle complex and non-linear data structures effectively. This advantage may stem from RF's ensemble nature, which aggregates predictions from multiple decision trees, thereby reducing overfitting and enhancing prediction stability across diverse datasets.

A notable exception was observed in the case of Geely, where the ANN model outperformed RF across all metrics. Although the differences were not statistically significant (p-value = 0.0839), ANN achieved a lower RMSE (0.2240 vs. 0.2555), lower MAPE (1.4160 vs. 1.7215), and a Mean Bias Error (MBE) closer to zero. This may be attributed to the relatively stable price trend of Geely following an initial sharp decline, making it more suitable for ANN's learning structure, which performs well with simpler, linear-like patterns and low-volatility data. In contrast, RF may have overfit to volatility during the training phase, resulting in greater deviations in the testing phase.

For highly volatile stocks such as Tesla and BYD, RF still delivered more accurate predictions in terms of RMSE, although the differences were not statistically significant ($p > 0.05$). These findings suggest that RF may better accommodate unstable and complex price behaviors, while ANN may require more careful tuning and structural adjustments to perform well under such conditions.

In summary, the RF model demonstrates consistent and robust forecasting performance across a variety of stock price patterns, especially for stocks with high volatility or complex trends. On the other hand, ANN may perform better when dealing with relatively stable price movements and lower volatility. Therefore, selecting an appropriate forecasting model should depend on the specific characteristics of the dataset, as alignment between data properties and model mechanisms directly affects prediction accuracy. Statistical validation through paired t-tests further reinforces the performance differences observed in companies such as GM and Volkswagen.

For future research, two primary directions are recommended:

1. Expanding the comparison to include other advanced machine learning models such as XGBoost, LightGBM, or LSTM, to evaluate their effectiveness under the same data context; and

2. Incorporating additional external variables into the forecasting model—such as crude oil prices, gold prices, global stock indices, or economic and political news—as time-series external indicators to improve dynamic forecasting capacity.

The findings from this study may serve as a foundational guideline for investors and financial analysts when selecting appropriate forecasting models for specific stock types, especially in fast-evolving industries like EVs. Moreover, they provide a practical framework for developing analytical tools to support informed investment decisions and mitigate risks associated with market uncertainty.

Declaration of AI Use

The authors declare that AI tools were used only to assist in language refinement and content organization. All outputs generated by AI were carefully reviewed and validated by the authors to ensure factual accuracy, originality, and compliance with ethical publication standards.

Declaration of Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Agatonovic-Kustrin, S. and Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis*, 22(5):717–727. DOI: 10.1016/s0731-7085(99)00272-1.

Behera, R., Das, S., Rath, S., Misra, S., and Damasevicius, R. (2020). Comparative study of real time machine learning models for stock prediction through streaming data. *JUCS - Journal of Universal Computer Science*, 26(9):1128–1147. DOI: 10.3897/jucs.2020.059.

Biau, G. (2012). Analysis of a random forests model. *Journal of Machine Learning Research*, 13(1):1063–1095. DOI: 10.5555/2188385.2343682.

Biau, G. and Scornet, E. (2016). A random forest guided tour. *TEST*, 25(2):197–227. DOI: 10.1007/s11749-016-0481-7.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32. DOI: 10.1023/a:1010933404324.

Chicco, D., Warrens, M. J., and Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7:e623. DOI: 10.7717/peerj-cs.623.

Daori, H., Alharthi, M., Alanazi, A., Alzahrani, G., Aborokbah, M., and Aljehane, N. (2022). Predicting stock prices using the random forest classifier. *Preprint*. DOI: 10.21203/rs.3.rs-2266733/v1.

Dongare, A. D., Kharde, R. R., and Kachare, A. D. (2012). Introduction to artificial neural network. *International Journal of Engineering and Innovative Technology (IJEIT)*, 2(1):189–194.

El Naqa, I. and Murphy, M. J. (2015). *What Is Machine Learning?*, page 3–11. Springer International Publishing. DOI: 10.1007/978-3-319-18305-3_1.

Hinton, G. E. (2011). Machine learning for neuroscience. *Neural Systems & Circuits*, 1(1). DOI: 10.1186/2042-1001-1-12.

Jordan, M. I. and Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260. DOI: 10.1126/science.aaa8415.

Larminie, J. and Lowry, J. (2012). *Electric Vehicle Technology Explained*. Wiley.

Lawrence, A., Ryans, J. P., Sun, E., and Laptev, N. (2017). Earnings announcement promotions: A Yahoo finance field experiment. *SSRN Electronic Journal*. DOI: 10.2139/ssrn.2940223.

Li, A. W. and Bastos, G. S. (2020). Stock market forecasting using deep learning and technical analysis: A systematic review. *IEEE Access*, 8:185232–185242. DOI: 10.1109/access.2020.3030226.

Ninjas, A. (2024). Covid-19 api. urlhttps://api-ninjas.com/api/covid19. Retrieved December 2024, from `https://api-ninjas.com/api/covid19`.

Pomboomee, P., Lonlue, P., Praha, P., Mungmor, P., and Khruahong, S. (2023). Classification grading of Nam Dok Mai See-Thong mango by deep learning and transfer learning. In *2023 20th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, page 161–166. IEEE. DOI: 10.1109/jcsse58229.2023.10202009.

Pramote, O.-U., Khruahong, S., Jitanan, S., and Kong, X. (2023). Quality grading of crown flowers using convolutional neural network. *ICIC Express Letters*, 17(2):143–152. DOI: 10.24507/icicel.17.02.143.

Say, K., Srisawatsakul, C., and Boontarig, W. (2025). TalkTutorAI: Empowering English language proficiency through ChatGPT-assisted conversational practice – a case study at Ubon Ratchathani Rajabhat University, Thailand. *Journal of Applied Informatics and Technology*, 7(1):12–26. `https://ph01.tci-thaijo.org/index.php/jait/article/view/254818`.

Speiser, J. L., Miller, M. E., Tooze, J., and Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. *Expert Systems with Applications*, 134:93–101. DOI: 10.1016/j.eswa.2019.05.028.

Sutheebanjard, P. and Premchaiswadi, W. (2009). Factors analysis on stock exchange of Thailand (SET) index movement. In *2009 7th International Conference on ICT and Knowledge Engineering*, page 69–74. IEEE. DOI: 10.1109/ictke.2009.5397320.

Thompson, C. (2024). Analyzing the growth and challenges of the EV industry in the IEA's global EV outlook 2024. *eResearch*. Analyst Article, Retrieved January 2025.

Wanjawa, B. W. and Muchemi, L. (2014). ANN model to predict stock prices at stock exchange markets. `https://arxiv.org/abs/1502.06434`.

Zhang, R., Chen, S., Zhang, Z., and Zhu, W. (2022). Genetic algorithm in multimedia dynamic prediction of groundwater in Open-Pit Mine. *Computational Intelligence and Neuroscience*, 2022:1–6. DOI: 10.1155/2022/8556103.

Zhao, J., Lee, C.-D., Chen, G., and Zhang, J. (2024). Research on the prediction application of multiple classification datasets based on random forest model. In *2024 IEEE 6th International Conference on Power, Intelligent Computing and Systems (ICPICS)*, page 156–161. IEEE. DOI: 10.1109/icpics62053.2024.10795875.

Zheng, J., Xin, D., Cheng, Q., Tian, M., and Yang, L. (2024). *The Random Forest Model for analyzing and Forecasting the US Stock Market under the background of smart finance*, page 82–90. Atlantis Press International BV. DOI: 10.2991/978-94-6463-419-8_11.