Evaluating Transparency: A Cross-Model Exploration of Explainable AI in Financial Forecasting

1st Dylan Valensky
Computer Science and Software Engineering
Auckland University of Technology
Auckland, New Zealand
cpc5502@autuni.ac.nz

2nd Mahsa Mohaghegh
Computer Science and Software Engineering
Auckland University of Technology
Auckland, New Zealand
mahsa.mohaghegh@aut.ac.nz

***Abstract*—Advances in machine learning and deep learning have assisted progress in stock market predictions, which have presented unique opportunities for investors and traders to benefit from predictions. However, understanding how these models work is crucial to trusting their predictions. This research applies three prominent Explainable AI (XAI) techniques - SHAP, LIME, and Permutation Importance - to three distinct forecasting models: Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN), and Decision Tree. This work fills a gap in the literature by comparing XAI methods across a combination of ”black-box” and “white-box” models. Two key research questions guide this research: the performance of these XAI methods across the chosen models and the consistency between global and local explanations across these models. Findings reveal interesting model-specific behaviors, like CNN's emphasis on slightly older data and LSTM's focus on immediate past data. In the S&P 500 dataset context, features such as 'Close' and 'Adj Close' prices emerged as significant across models, while 'Volume' remained insignificant. This study offers a broader perspective on applying XAI in financial time series forecasting and helps enhance trust and comprehension among stakeholders relying on these model predictions.**

***Index Terms*—Explainable AI, Financial forecasting**

I. Introduction

Financial time series forecasting is an important field that seeks to predict future financial trends based on historical data. It is applied to areas such as stock markets so that investors can profit from the correct choices. Thus, the predictability of market movements is essential for traders, investors, and stakeholders alike. However, the fluctuations in stock prices are complex, and predicting them consistently can be challenging. Thanks to machine learning and deep learning advancements, there has been significant improvement in predicting stock market movements. Recently, advanced machine learning models, such as LSTM (Long Short-Term Memory) and CNN (Convolutional Neural Network), have become popular models in financial forecasting due to their accuracy. However, their inner workings remain largely unknown, leading to a lack of transparency and interpretability. On the other hand, models like Decision Trees are more transparent, termed ”white-box” models, and are more easily interpretable but do not provide the same levels of accuracy. Thus, the emerging dilemma is the trade-off between accuracy and interpretability. In light of these considerations, this study seeks to compare these models. It applied three prominent Explainable Artificial Intelligence (XAI) methods—SHAP, LIME, and Permutation Importance—to three distinct yet commonly used forecasting models: LSTM, CNN, and Decision Tree. These models were selected to ensure that each type of model is represented, with two of them (LSTM and CNN) being different black-box models that provide competitive performance, and one (Decision Tree) white-box model that provides transparency. The intention was to compare and contrast each model's forecasting behavior with the explanations provided by the respective XAI methods, mainly when applied to the financial time series data of the S&P 500. This approach is significant due to its potential to enhance trust in these models. It fills a gap in the existing literature by uniquely combining these three models with XAI methods in forecasting financial time series.

II. Literature Review

A. XAI Introduction

The decision-making processes of AI models, such as deep neural networks, have remained unknown, potentially holding their adoption back as the lack of interpretability can lead to mistrust, reluctance to adopt AI technologies, and ethical concerns. The idea of 'explainability' involves making sure that people can understand how a technology or system works and how easy it is for them to understand the results it gives. XAI is defined as the following: “Given an audience, an explainable Artificial Intelligence produces details or reasons to make its functioning clear or easy to understand.” [2]. XAI can be classified into three groups: understanding, diagnosing, and refining [16]. XAI holds significance for many different stakeholders such as developers, regulators, data scientists, users, consumers, and businesses [7].

B. XAI Taxonomies

XAI methods can be categorized into various taxonomies such as black-box models versus white-box models [7], model-specific versus model-agnostic [19], global explanations versus local explanations [5], and post-hoc versus ante-hoc [26].

1) White-Box versus black-box models: White-box model techniques involve AI models that are straightforward and self-explanatory. In contrast, black-box models have complexity in terms of decision-making processes. These models, such as deep neural networks, produce challenging predictions to decipher or explain.

2) Model-specific versus model-agnostic: Model-specific techniques suit the architecture of a particular machine-learning model. These methods use the model's internal mechanisms, extracting insights explaining how it arrives at its predictions. In contrast, model-agnostic techniques are independent and do not rely on a model's architecture to provide explanations.

3) Global explanations versus local explanations: Global explainability provides an overall understanding of a model's behavior across its entire input space. In contrast, local explainability provides insights into the rationale behind a specific prediction, offering information on individual instances.

4) Post-hoc versus ante-hoc: Post-hoc explainability approaches are separate from the models and provide insights into a model's learned behavior after training without altering its underlying structure. Whereas ante-hoc explainable methods are applied to models that possess inherent interpretability due to their simple and transparent structure by design.

C. XAI Methods

The most common model-agnostic methods are LIME [24], SHAP [18], Anchors [23] and LoRE [12].

1) Lime: LIME works by perturbing the features of an instance and creating a simple surrogate model to approximate the behavior of a model locally.

2) SHAP: SHAP provides local or global explanations by employing cooperative game theory to calculate Shapley values, representing each feature's average contribution to a prediction across all possible feature combinations.

3) Anchors: Anchors create simple, interpretable rules that explain the predictions of complex machine-learning models. The approach identifies conditions on the input features for a particular prediction.

4) LoRE: LoRE (Local Rule-based Explanations) generates local, interpretable rules that capture the decision boundaries around a specific instance.

5) Ablation: Ablation involves systematically removing one or more features from the input data or the model's architecture and observing how these removals affect the model's predictions [20]

6) Integrated Gradients: Integrated Gradients, proposed by [25], finds feature importance by integrating the gradients of the model's output concerning the input data over a path from a baseline input to the actual input. This technique measures the change in the model's output as each feature changes from the baseline to the actual input, providing feature attributions for a specific prediction.

7) Layer-wise Relevance Propagation: Layer-wise Relevance Propagation (LRP), propsed by [3], propagates the relevance of the model's output backward through its layers, attributing importance to individual neurons or features.

8) Tree-based: Tree-based methods, such as Tree SHAP, explain models like decision trees, random forests, and gradient-boosting machines. Reference [17] propose Tree SHAP, which extends SHAP values to tree-based models, providing feature importance for individual predictions. These methods are model-agnostic and offer local and global explanations for the importance of each feature [7].

9) Influence functions: Influence functions, proposed by [14], measure the influence of each training data point on the model's predictions. They help identify data points that significantly impact the model's decisions. Influence functions are model-agnostic and provide insights into the stability and robustness of the model.

10) Back-propagation based methods: These are post-hoc model-specific methods that calculate the gradient or variations of a specific output concerning the input through the back-propagation process [1]. Two such methods are Class Activation Mapping (CAM) and Gradient-weighted Class Activation Mapping (Grad-CAM), which are model-specific for CNNs.

D. Financial Forecasting Models

The most common regression task models are machine learning and deep learning models [11]. For machine learning, these include Decision Trees, Random Forest, Linear Regression, Naïve Bayes, and Support Vector Machines. For deep learning, these include Feedforward Neural Networks, Recurrent Neural Networks (RNN), and LSTMs. Reference [15] successfully predicted stock price fluctuations using LSTM architecture. Similarly, [13] shows the potential of several deep-learning neural network approaches for predicting stock prices with a focus on the S&P 500 index. Reference [8] also succeeded with deep learning models for financial forecasting. They perform a comparative analysis of different models, including RNN, LSTM, and CNN, for analyzing and predicting the trends of NIFTY 50 stock prices. The research concludes that the LSTM model outperforms the RNN and CNN models, offering lower Mean Squared Error (MSE) values. Meanwhile, [28]) focuses on employing LSTM, Multi-Layer Perceptron (MLP), and CNN to forecast the daily closing prices of stocks within the Casablanca Stock Market. Their findings show that CNN outperforms the other models based on MAE. These articles demonstrate the applicability and success of deep neural networks for predicting stock prices based on stock history data and show that different models can provide superior performance

E. XAI Methods in Finance

XAI in finance is a relatively small but growing research topic. For instance, over 55% of the XAI in finance articles were published in 2020 and 2021, respectively [27]. Various finance tasks such as stock forecasting, credit risk evaluation, fraud detection, and investment strategy use black-box and white-box models respectively. The choice of AI model comes with a trade-off, with black-box models providing higher accuracy while being harder to interpret. In contrast, white-box models are easier to interpret while providing reduced accuracy. XAI methods have been used to explain black-box models in finance. Furthermore, these models have been used to make financial predictions on technical or textual data. For instance, [21] use SHAP to explain predictions made by the gradient boosting decision trees (GBDT) model on large drops in S&P 500 prices. Similarly, [10] uses XAI methods (ablation, permutation, random noise, and integrated gradients) to explain RNN models such as standard RNN, GRU, and LSTM. Reference [4] use LIME to provide explain stock market predictions. XAI methods have also been used to explain predictions of white-box models in finance. For instance, [6] proposes a rule extraction-based approach to stock return prediction that uses multiple information sources from stock trading and the internet. Likewise, [22] proposes a neuro-fuzzy model for stock price prediction. They attempt to address the trade-off between accuracy and interpretability. Furthermore, [9] adapted a Type 2 Fuzzy Logic variant to explain the forecast time-varying stock indices model in a regression task. While white-box models are used in financial regression tasks, they often sacrifice accuracy for interpretability.

III. Research Significance

The literature on XAI for financial forecasting is limited as studies only focus on certain types of models, whether black-box or white-box models. While numerous studies focus on XAI methods in financial forecasting, there is a significant gap in the literature concerning evaluating and comparing these methods between various black-box and white-box models. This restricted focus has led to a lack of understanding of XAI's applicability and performance across diverse forecasting models. A comparative study examining the effectiveness of XAI methods across different model types - both ”black-box” and ”white-box” - will contribute to understanding. This broader approach not only has the potential to offer more insights into the mechanics of different models but also benefit different stakeholders.

IV. Research Questions

Understanding different models' behavior and underlying logic is essential in financial forecasting, where a lot is at stake. Explanatory methods provide insights into these models. However, their effectiveness and consistency across different model architectures have yet to be discovered. Thus, this study aims to address the following research questions:

1) How do SHAP, LIME, and Permutation Importance perform when applied to LSTM, CNN, and Decision Tree models in financial forecasting?
2) How consistent are the insights from global versus local explanatory methods when applied to LSTM, CNN, and Decision Tree forecasting models?

V. Methodology

A. Dataset

The dataset was constructed with S&P 500 daily data from 20th April 1982 to 31st August 2023, excluding weekends. The dataset has six columns: ' Open', 'High', 'Low', 'Close', 'Adj Close', and 'Volume', respectively. In total, the dataset has 10,430 entries. These entries were split into training, validation, and testing.

B. Models

This study uses three distinct models for financial forecasting: an LSTM model, a CNN model, and a Decision Tree model.

1) LSTM: The LSTM model begins with a layer of 128 units, designed to return sequences, ensuring that the subsequent LSTM layer receives a sequence as input. The next layer, also containing 128 units, does not return sequences. The model concludes with an output-dense layer containing a single neuron.

2) CNN: The model begins with a convolutional layer comprising 32 filters, each of kernel size 2×2, paired with a Rectified Linear Unit (ReLU) activation function. A max pooling layer with a pool size of 2×2 is applied to reduce the spatial dimensions. Another convolutional layer is added, with 64 filters of the same kernel size and ReLU activation.

3) Decision Tree: The Decision Tree Regressor model was used from scikit-learn and trained with a random state of 42. The dataset was transformed into a supervised learning format to allow the model to learn the relationship between sequential observations.

C. Methods

The XAI methods used in this study include SHAP, LIME, and Permutation Importance, respectively. Each of these methods aims to enhance the interpretability of machine learning models by providing insights into how input features influence predictions.

1) SHAP: The 'DeepExplainer' method was used for the LSTM and CNN models respectively, which is particularly suited for deep learning architectures. For the Decision Tree model, the SHAP 'TreeExplainer' was used, which is suited for tree-based algorithms.

2) LIME: For all models, the input data had to be reshaped to be fed into LIME, designed primarily for tabular data. The input data for LIME was flattened, turning the sequential data into tabular form.

3) Permutation Importance: A baseline performance was created by predicting the original test data and computing the Mean Squared Error (MSE). The feature's values were permuted individually with this baseline, and the model's performance was recalculated. The difference between the baseline and the permuted performance indicated the feature's importance.

VI. Results

1) SHAP: As shown in Fig. 1, the CNN model mainly considers features from the t-2 and t-3 time steps. Interestingly, it does not factor in t-1 features, suggesting that the most recent data could be more predictive, at least in the context of this model and data set. This occurrence might indicate that the underlying patterns the model detects are better captured with slightly older data. Furthermore, among all the features, the 'Close' price emerged as the most influential. This occurrence is particularly evident in the t-3 time step, where 'Close' registers the highest feature value. In the t-2 step, 'Close' also remains a dominant feature. The map shows that features significantly impact the model from around t-5. Notably, the 'High' price feature around this time frame exhibits a negative SHAP value, which implies that higher values of this feature reduce the model's output. 'Volume' has no significant impact across all steps. In contrast, The LSTM model heavily relies on t-1 features. To a lesser extent, it also considers t-2. Within the t-1 time step, 'Adj Close' is the most influential feature, followed by 'Close', 'Low', and 'High'. The 'Open' price also has a noticeable impact in t-1, albeit lesser than the above-mentioned features. As with the CNN model, 'Volume' has no significant impact across all steps. The t-2 step exhibits some influence for 'Adj Close', 'Close', 'Low', and 'High', although not as pronounced as in t-1. The darker discoloration is observable around t-6, signifying features that reduce the predictions. The Decision Tree model had similar behavior as shown in Fig. 2. Features from t-1 were the most significant with 'Close' being paramount followed by 'Low' and 'Adj Close'.

2) LIME: As shown in Fig. 3, the CNN model emphasized features from t-2 and t-3 time steps for the individual prediction while overlooking t-1 entirely. In the t-3 time step, the 'Close' feature is the most significant, followed by the 'Low' feature. 'Adj Close', 'High', and 'Open' prices also impact the prediction somewhat, while 'Volume' has no impact.

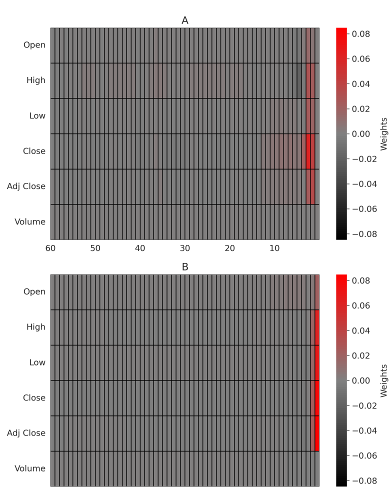

Fig. 1. SHAP feature importance for CNN (A) and LSTM (B).

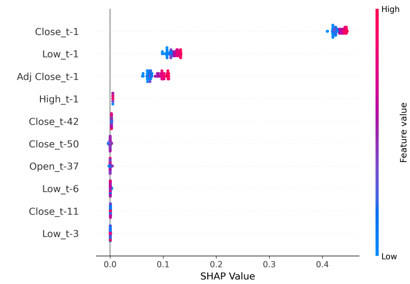

Fig. 2. SHAP feature importance for Decision Tree.

The influence pattern in t-2 is similar to that of t-3, with 'Close' having the most impact followed by 'Low'. 'Adj Close' and 'High' carry some weight, but 'Open' and 'Volume' are inconsequential. Of all the features across these steps, 'Close' has the most significant impact. The LSTM model's prediction is made predominantly on t-1, with t-2 being less significant. In the t-1 time step, 'Adj Close' is the most significant feature, followed by 'Close', 'Low', and 'High'. The 'Open' price makes a marginal contribution, whereas 'Volume' has no impact. The map shows that the t-2 time step features such as 'Adj Close', 'Close', 'Low', and 'High' have some influence, but their impact pales compared to t-1 features. This observation cements the dominance of the four major t-1 features, specifically underscoring their significance in the LSTM's prediction mechanism. The Decision Tree model's prediction is made mainly from t-1 with 'Close' being the most significant followed by 'Adj Close' and 'Low' as shown in Fig. 4. Interestingly, 'High' from t-22 and 'Adj Close' from t-45 had a marginal influence on the model's prediction.

3) Permutation Influence: Permutation Importance provides a global explanation. As shown in Fig. 5, the CNN model has 'Low' and 'Close' as the most significant features, contributing approximately 29.63% and 29.39% of the total feature importance, respectively. The next prominent feature is 'High', which contributes 18.31%. 'Open' and 'Adj Close'

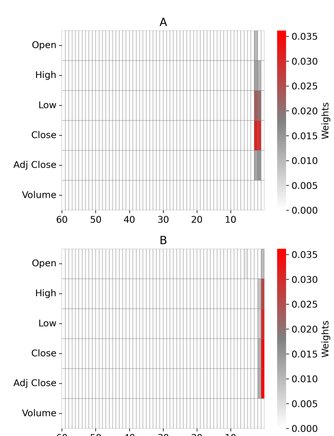

Fig. 3. LIME feature importance for CNN (A) and LSTM (B).

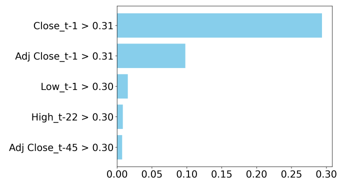

Fig. 4. LIME feature importance for Decision Tree.

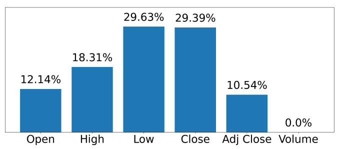

Fig. 5. Permutation Importance feature value (%) for CNN.

contribute 12.14% and 10.54%, respectively, while 'Volume' has no importance. The LSTM model has 'Adj Close' as the most influential, accounting for 24.75% of the total feature importance as shown in Fig. 6. This feature is followed closely by 'Close' and 'Low' with contributions of 22.1% and 21%, respectively. 'High' and 'Open' have an influence as well, with 19.52% and 12.63% of the total importance, respectively, while 'Volume' has a negligible influence. Similarly, the Decision Tree model also has 'Adj Close' as the most significant feature, however, it makes up 88.58% of the total influence as shown in Fig. 7.

VII. Discussion

This study used three XAI methods: SHAP, LIME, and Permutation Importance. Each technique provided unique insights that revealed patterns and behaviors embedded in the models. Regarding input influence as per the temporal aspects, the SHAP results, which offer global explanations, showed that the CNN model placed significant weight on the t-2 and t-3 time steps. The model skipped the most recent data from t-1, suggesting that the immediate past might only sometimes be the most predictive for the S&P 500 dataset.

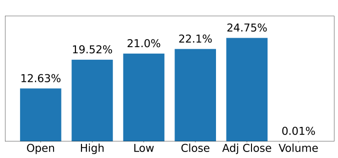

Fig. 6. Permutation Importance feature value (%) for LSTM.

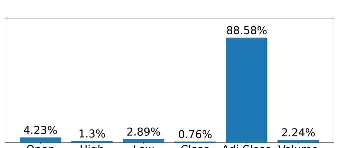

Fig. 7. Permutation Importance feature value (%) for Decision Tree.

This behaviour suggests that the CNN model detects underlying patterns better when considering slightly older data for time series forecasting. This observation is consistent with LIME's local explanations, reinforcing its validity. The consistency between SHAP's global and LIME's local explanations, especially for the CNN model's t-2 and t-3 time step preference, provides confidence in the model's behaviour. In contrast, the LSTM favored the immediate past, t-1, while ignoring the older time steps. This behavior is interesting as LSTM models are designed to remember long-term dependencies. However, this model showed a surprising focus almost entirely on the immediate past, t-1, while placing less significance on older time steps. This behavior goes against expectations as LSTMs can remember and use older data but predominantly use the most recent time step. This behavior suggests a couple of possibilities. It might indicate that the LSTM determined the older time steps were not particularly influential for the task at hand, reinforcing the notion that the sequential nature of the S&P 500 might render the most recent data most important. It could also reflect specific training dynamics where the model found no substantial advantage in considering older temporal inputs. The Decision Tree also showed favor towards the immediate past, t-1, with both LIME and SHAP showing similar behavior. This result is consistent with the inner workings of Decision Tree models as they make decisions based mainly on the input features provided to them at the current step. Regarding feature significance, all three models elevated the importance of 'Close' and 'Adj Close' prices. The CNN model's emphasis on the 'Close' feature remained consistent across global (SHAP) and local (LIME) explanatory models. Conversely, the LSTM placed greater importance on 'Adj Close'. Meanwhile, there were disparities with the Decision Tree model with 'Close' being the most important feature according to SHAP and LIME but permutation importance showed 'Adj Close' as the most important. Furthermore, the CNN model had contrasting behavior associated with 'High' and 'Low' prices. The SHAP results for the CNN indicated a negative association with 'High' prices for older time steps. Meanwhile, 'Low' prices emerged as significantly influential, especially in the permutation importance analysis. However, not all features received such prominence. A common observation across all models and all methods was the marginal contribution of 'Volume.' Its consistent lack of significance makes sense as volume does not tend to influence prices but rather indicates trading activity. A noteworthy observation is the broader feature consideration in the LSTM model compared to the CNN. It clearly prefers 'Adj Close' and integrates influences from 'Close', 'Low', and 'High'. The reason for this can be attributed to the LSTM's architectural design. By capturing long-term dependencies, LSTMs can integrate a more comprehensive set of features over multiple time steps.

VIII. Conclusion

This research aimed to explore and understand the inner workings of three distinct popular forecasting models using different XAI methods. Specifically, an LSTM, a CNN, and a Decision Tree were used for this research. LSTM and CNN considered "black box" models are known to be challenging to interpret while the Decision Tree, a "white box" model, offers inherent transparency. Considering these diverse models, this study aimed to provide a broad perspective on applying XAI in financial time series forecasting. It is worth mentioning that this approach of combining these three particular models in XAI is unique in the current literature. Furthermore, this approach allowed the comparative analysis between models about how they approach financial time series forecasting and the adaptability of the XAI methods to these respective models. Thus, this research gained a deeper understanding of each model's inner workings. The implications of this research help build trust among stakeholders and investors who rely on these forecasts. However, this research does have its limitations. Firstly, the choice of LSTM and CNN models employed were not of the most complex variants. While this was done to achieve transparency and interpretation, it may not represent high performance in financial time series forecasting. More complex models might provide better predictive performance. A future study could adapt these methods to more complex models. Secondly, these models only consider technical data. Financial forecasting is a multidimensional problem, where both technical and textual data influence price behavior. This research focused solely on technical data, neglecting potentially valuable information in textual data sources like news articles and financial reports. Thus, a future study could adapt XAI methods to deep learning models that forecast using both technical and textual data. Finally, LIME treats time series data as tabular data, disregarding its temporal nature and sequential dependencies. Similarly, SHAP does not inherently account for temporal relationships and patterns in time series data, potentially overlooking critical recurrent spatiotemporal dependencies. This oversimplification ignores the temporal nature of financial time series, where the order of data points and sequential dependencies are significant. Thus, future research should explore how to adapt XAI techniques to account for temporal relationships and sequential patterns more effectively.

REFERENCES

[1] M. Ancona, E. Ceolini, C. Öztireli, and M. A. Gross,"Towards better understanding of gradient-based attribution methods for deep neural networks," in Proceedings of the 9th International Joint Conference on Computational Intelligence, 2018, pp. 169–191.

[2] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," Information Fusion, vol. 58, pp. 82-115, 2020.

[3] S. Bach, A. Binder, G. Montavon, F. Klauschen, K. Müller, and W. Samek, "On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation" PLOS ONE, vol. 10, no. 7, p. e0130140, 2015.

[4] T. B. Çelik, Ö. İcan, and E. Bulut, "Extending machine learning prediction capabilities by explainable AI in financial time series prediction," Applied Soft Computing, vol. 132, no. 1, pp. 1-14, 2023.

[5] M. Du, N. Liu, and X. Hu, "Techniques for interpretable machine learning," Communications of the ACM, vol. 63, no. 1, pp. 68-77, 2020.

[6] J. Duan and J. Zeng, "Forecasting stock returns using multiple information sources based on rules extraction," in 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2015.

[7] R. Dwivedi, D. Dave, H. Naik, S. Singhal, R. Omer, P. Patel, B. Qian, Z. Wen, T. Shah, G. Morgan, and R. Ranjan, "Explainable AI (XAI): Core ideas, techniques, and solutions," ACM Computing Surveys, vol. 55, no. 9, pp. 1-33, 2023.

[8] Z. Fathali, Z. Kodia, and L. Ben Said, "Stock market prediction of NIFTY 50 index applying machine learning techniques," Applied Artificial Intelligence, vol. 36, no. 1, pp. 3101-3123, 2022.

[9] M. M. Ferdaus, R. K. Chakrabortty, and M. J. Ryan, "Multiobjective automated type-2 parsimonious learning machine to forecast time-varying stock indices online," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 52, no. 5, pp. 2874-2887, 2022.

[10] W. Freeborough and T. Van Zyl, "Investigating explainability methods in recurrent neural network architectures for financial time series data," Applied Sciences, vol. 12, no. 3, p. 1427, 2022.

[11] L. Gasparėnienė, R. Remeikiene, A. Sosidko, and V. Vėbraitė, "Modelling of S&P 500 index price based on U.S. economic indicators: Machine learning approach," Engineering Economics, vol. 32, no. 4, pp. 362-375, 2021.

[12] R. Guidotti, A. Monreale, F. Giannotti, D. Pedreschi, S. Ruggieri, and F. Turini, "Factual and counterfactual explanations for black-box decision making," IEEE Intelligent Systems, vol. 34, no. 6, pp. 14-23, 2019.

[13] F. Kamalov, L. Smail, and I. Gurrib, "Stock price forecast with deep learning," in 2020 International Conference on Decision Aid Sciences and Application (DASA), 2020.

[14] P. W. Koh and P. Liang, "Understanding black-box predictions via influence functions," in Proceedings of the 34th International Conference on Machine Learning, 2017.

[15] B. Lin, W. Chu, and C. Wang, "Application of stock analysis using deep learning," in 2018 7th International Congress on Advanced Applied Informatics (IIAI-AAI), 2018.

[16] S. Liu, X. Wang, M. Liu, and J. Zhu, "Towards better analysis of machine learning models: A visual analytics perspective," Visual Informatics, vol. 1, no. 1, pp. 48-56, 2017.

[17] S. M. Lundberg, G. G. Erion, and S. I. Lee, "Consistent individualized feature attribution for tree ensembles," arXiv, 2019. [Online]. Available: https://arxiv.org/abs/1802.03888.

[18] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017, pp. 4768–4777.

[19] C. Meske, E. Bunde, J. Schneider, and M. Gersch, "Explainable artificial intelligence: Objectives, stakeholders, and future research opportunities," Information Systems Management, vol. 39, no. 1, pp. 53-63, 2020.

[20] R. Meyes, M. Lu, C. W. De Puiseau, and T. Meisen, "Ablation studies in artificial neural networks," 2019.

[21] J. J. Ohana, S. Ohana, E. Benhamou, D. Saltiel, and B. Guez, "Explainable AI (XAI) models applied to the multi-agent environment of financial markets," in Explainable and transparent AI and multi-agent systems, 2021, pp. 189–207.

[22] S. Rajab and V. Sharma, "An interpretable neuro-fuzzy approach to stock price forecasting," Soft Computing, 2017.

[23] M. Ribeiro, S. Singh, and C. Anchor, "Anchors: High-precision model-agnostic explanations," in Proceedings of the 32nd AAAI Conference on Artificial Intelligence, 2018, pp. 1527–1535.

[24] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you?: Explaining the predictions of any classifier," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1135–1144.

[25] M. Sundararajan, A. Taly, and Q. Yan, "Axiomatic attribution for deep networks," in Proceedings of the 34th International Conference on Machine Learning, 2017, pp. 3319–3328.

[26] A. Theissler, F. Spinnato, U. Schlegel, and R. Guidotti, "Explainable AI for time series classification: A review, taxonomy and research directions," IEEE Access, vol. 10, no. 1, pp. 100700-100724, 2022.

[27] P. Weber, K. V. Carl, and O. Hinz, "Applications of explainable artificial intelligence in finance—a systematic review of finance, information systems, and computer science literature," Management Review Quarterly, 2023.

[28] C. Yinka-Banjo, M. I. Akinyem, and B. Er-Rabbany, "Stock market prediction using a hybrid of deep learning models," International Journal of Financial Management and Economics, vol. 2, no. 2, pp. 1-16, 2023.