Stock Market Prediction Using LSTM With Explainable AI

1st Nithesh G

Department of Computer Science and Engineering
Dayananda Sagar University
Bengaluru,India

2nd Nityanand B M

Engineering
Dayananda Sagar University
Bengaluru,India

3rd Rakesh Rathod
Department of Computer Science and Engineering
Dayananda Sagar University
Bengaluru,India

4th Sujeeth M
Department of Computer Science and Engineering
Dayananda Sagar University
Bengaluru, India

Guide Signature:

Abstract:

Forecasting stock market trends is a complex and dynamic task due to the influence of numerous fluctuating factors. This research presents a deep learning approach utilizing Long Short-Term Memory (LSTM) networks for predicting the future stock prices of HDFC Bank. The model is trained on historical closing price data from 2020 to 2025, sourced from Yahoo Finance. The time series data is preprocessed using normalization techniques, and sequences of past observations are structured to train the LSTM architecture effectively. To improve the interpretability of the model's predictions, Explainable AI methods, particularly SHAP (SHapley Additive exPlanations), are employed. The proposed model demonstrates strong predictive performance, achieving a Root Mean Squared Error (RMSE) of 0.0246 and a coefficient of determination (R^2) score of 0.9489. These results highlight the model's capability to accurately capture trends in stock price movement while also providing valuable insights into feature contributions. Unlike traditional statistical models, the LSTM network is capable of learning long-term dependencies, making it well-suited for sequential data such as stock prices. The integration of explainability not only adds transparency to the predictions but also enhances the trustworthiness of the model in real-world applications. Furthermore, this approach can be adapted to other financial instruments or extended with external indicators for broader market analysis. The combination of LSTM and SHAP enables a robust, interpretable framework for stock market forecasting that supports data-driven investment strategies.

Department of Computer Science and

Introduction :

Stock market prediction has long been a challenging problem due to the market's inherently volatile and non-linear behavior. Accurate forecasting of stock prices can provide significant advantages to investors, portfolio managers, and financial analysts by aiding in risk management and investment decision-making. Traditional models, including statistical and econometric approaches, often fall short in capturing the intricate patterns and temporal dependencies present in stock price data, limiting their effectiveness in dynamic market environments.

Recent developments in deep learning have opened new avenues for tackling time-series prediction problems. Long Short-Term Memory (LSTM) networks, a specialized type of recurrent neural network, have gained popularity for their ability to model sequential data with long-term dependencies. Unlike conventional models, LSTMs can remember important information over extended periods, making them well-suited for stock market data where historical trends and seasonality often influence future prices. This capability enables the LSTM to outperform many traditional forecasting methods in terms of accuracy and robustness.

However, one critical issue with deep learning models is their lack of interpretability, often being considered "black boxes." In financial applications, where decisions must be justified and risks clearly understood, the need for transparency is paramount. Explainable AI (XAI) frameworks aim to bridge this gap by providing insights into model predictions. SHAP (SHapley Additive exPlanations) is a widely used technique that attributes the contribution of each input feature to the model's output, allowing practitioners to understand how specific data points affect stock price predictions.

This research combines the strengths of LSTM networks and SHAP explainability to develop a reliable and interpretable stock price prediction model for HDFC Bank. Using historical market data from 2020 to 2025, the LSTM model is trained to capture temporal dependencies, while SHAP is employed to interpret the model's decision-making process. The model's performance is rigorously evaluated through metrics such as

RMSE and R^2, demonstrating strong predictive capabilities. By integrating explainability, this approach not only enhances forecasting accuracy but also fosters trust and transparency, making it valuable for practical financial analysis and investment strategies.

Furthermore, the integration of explainable AI with deep learning models addresses growing concerns regarding model accountability and ethical use of AI in finance. As automated trading and algorithmic investment strategies become more prevalent, stakeholders require assurance that model predictions are not only accurate but also understandable and justifiable. This transparency is essential for regulatory compliance, risk assessment, and building confidence among users. Moreover, the techniques developed in this study have the potential to be extended beyond stock prices to other financial instruments such as commodities, currencies, and derivatives, broadening their practical application. The fusion of LSTM and SHAP presents a promising step toward more interpretable, trustworthy, and effective financial forecasting tools in an increasingly data-driven market environment.

LITERATURE SURVEY:

Predicting stock market prices has attracted extensive research attention due to its critical role in financial decision-making. Early approaches primarily relied on statistical models such as Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) to analyze time-series data. These models assume linear relationships and often fail to capture the complex, non-linear behaviors typical of stock market data. Consequently, researchers began exploring machine learning techniques, which offer greater flexibility and the ability to model non-linear patterns. Support Vector Machines (SVM), Random Forests, and Artificial Neural Networks (ANN) have been applied with varying degrees of success, although they often struggle to account for temporal dependencies inherent in sequential data.

The introduction of Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, marked a significant advancement in financial time-series forecasting. LSTMs are capable of learning long-term dependencies by maintaining a memory cell that regulates the flow of information, making them highly effective for sequential data such as stock prices. Many studies have demonstrated that LSTMs outperform traditional models by better capturing trends and patterns over time. For example, Fischer and Krauss (2018) showcased that LSTM networks could successfully forecast stock price movements by learning from historical data, outperforming conventional machine learning models. However, despite these improvements, LSTM models often lack transparency, limiting their interpretability and acceptance in critical financial applications.

To address the black-box nature of deep learning models, recent research has focused on integrating Explainable AI (XAI) techniques to shed light on model predictions. SHAP (SHapley Additive exPlanations), based on cooperative game theory, has gained prominence as a powerful tool to interpret complex models. It provides a unified measure of feature importance by quantifying each feature's contribution to the prediction. Lundberg and Lee (2017) introduced SHAP as a model-agnostic method that can explain outputs from any machine learning model, including deep neural networks. In stock market prediction, SHAP has been applied to understand how different input features such as past prices, volume, and technical indicators influence the model’s forecasts. This interpretability is crucial for gaining stakeholder trust and ensuring that models adhere to regulatory standards.

Several hybrid approaches have emerged that combine the strengths of LSTM models with explainability frameworks. For instance, recent studies have utilized LSTM models to predict stock prices while employing SHAP values to analyze the influence of each time-step and feature. These approaches enable practitioners to not only achieve high prediction accuracy but also gain insights into which temporal features drive market behavior. Furthermore, the application of such explainable models is not limited to stock prices but extends to other financial domains such as credit scoring, fraud detection, and portfolio optimization. This expanding scope highlights the growing importance of transparency in AI-driven financial systems.

Despite these advancements, challenges remain in effectively handling the noisy, volatile nature of stock market data and incorporating external factors such as news sentiment, macroeconomic indicators, and geopolitical events into predictive models. Future research is focusing on multimodal data fusion, combining textual news data and numerical time series, along with more advanced explainability techniques that can handle high-dimensional inputs. The integration of attention mechanisms within LSTM models is also being explored to enhance both prediction performance and interpretability. This evolving landscape indicates that combining deep learning with explainable AI holds great promise for developing robust, trustworthy, and actionable stock market forecasting tools.

RELATED WORK:

[1] Traditional statistical models such as ARIMA and GARCH have been widely used for stock market prediction due to their simplicity in modeling linear time series data. However, these models often fall short in capturing the complex nonlinear patterns and long-term dependencies present in financial markets.

[2] Machine learning algorithms like Support Vector Machines (SVM) and Random Forests have been applied to stock price forecasting with improved accuracy over classical methods. These models benefit from feature engineering but generally lack the ability to model temporal sequences effectively, which limits their predictive power in stock market applications.

[3] Recurrent Neural Networks, especially Long Short-Term Memory (LSTM) networks, have been demonstrated to excel in modeling sequential data by capturing long-term dependencies and nonlinear trends. Many studies report that LSTM-based models outperform conventional machine

learning techniques in financial forecasting tasks.

[4] Despite their performance advantages, LSTM models are often criticized for being "black-box" models that lack interpretability. To address this, explainable AI methods such as SHAP (SHapley Additive exPlanations) have been introduced, enabling detailed insights into feature contributions and improving the trustworthiness of deep learning models in finance.

[5] Recent hybrid approaches combining LSTM with explainability frameworks like SHAP have shown promise in providing accurate predictions while also offering interpretability. These methods help analysts understand which features or past time steps most significantly impact the model's predictions, facilitating more informed investment decisions.

ALGORITHMS USED:

1.Long Short-Term Memory (LSTM) Network

In this project, the Long Short-Term Memory (LSTM) network plays a central role in predicting future stock prices based on historical data. Stock market data is inherently sequential and time-dependent, where past prices influence future movements. Traditional neural networks struggle to capture such temporal relationships effectively, especially over long time periods. LSTM networks address this challenge by incorporating memory cells and gating mechanisms that can selectively retain or discard information over time.

By feeding the model with sequences of past closing prices, the LSTM learns patterns and trends that exist in the stock market data. The forget, input, and output gates allow the model to focus on important historical data while ignoring irrelevant fluctuations. This makes LSTM highly suitable for modeling the nonlinear and noisy behavior of stock prices. In this project, a stacked LSTM architecture is employed to capture complex temporal dependencies at multiple levels, improving prediction accuracy. The ability of LSTM to remember long-term dependencies helps the model anticipate price changes better than simpler methods, thus enhancing the quality of stock market forecasts.

LSTM Network Structure:

The LSTM model designed for this project is a **stacked LSTM network** consisting of three LSTM layers followed by a Dense output layer. This layered structure enables the model to learn complex patterns in sequential stock price data more effectively.

i. **First LSTM Layer:** This layer has 50 memory units (neurons) and is configured to return sequences, meaning it outputs a sequence of hidden states for each time step. This allows the next LSTM layer to receive a full sequence as input rather than just the final output.

ii. **Second LSTM Layer:** Similar to the first, it also contains 50 units and returns sequences, allowing for deeper temporal feature extraction.

iii. **Third LSTM Layer:** This layer has 50 units but returns only the final hidden state, summarizing the sequence information learned by the previous layers into a fixed-length vector.

iv. **Dense Layer:** A fully connected layer with a single neuron is added after the LSTM layers to produce the final predicted stock price.

The input shape for the model is (100, 1), where 100 represents the number of previous time steps (days) used to predict the next value, and 1 represents the single feature — the normalized closing price. This architecture enables the model to capture both short-term and long-term dependencies in the stock price data, improving forecasting performance.

2.Min-Max Scaling

To normalize the dataset within the range [0, 1], each feature value x is scaled as:

$$x' = x - x_{min} / x_{max} - x_{min}$$

x_{min} and x_{max} are the minimum and maximum values of the feature in the dataset,

x' is the scaled value.

This ensures that all features are on the same scale, which helps in faster and more stable training of neural networks.

3. Mean Squared Error (MSE) Loss Function

MSE measures the average squared difference between the predicted values y^i iy^i and the actual values yiy_iyi:

$$MSE = 1/N \sum (y_i - y\hat{}_i)\hat{}2$$

- n is the number of samples,
- y_i is the actual stock price,
- $y\hat{}_i$ is the predicted stock price.

4.Adam Optimizer

Adam (Adaptive Moment Estimation) is an advanced optimization algorithm widely used for training deep learning models. It combines the advantages of two other popular optimizers—**AdaGrad**, which works well with sparse gradients, and **RMSProp**, which works well in non-stationary settings. Adam maintains adaptive learning rates for each parameter by computing individual estimates of both the first-order moment (mean) and the second-order moment (uncentered variance) of the gradients.

5. SHAP (SHapley Additive exPlanations)

SHAP explains the output of a model fff by attributing the prediction to individual feature contributions based on Shapley values from cooperative game theory. The prediction f(x)f(x)f(x) for an input xxx can be represented as:

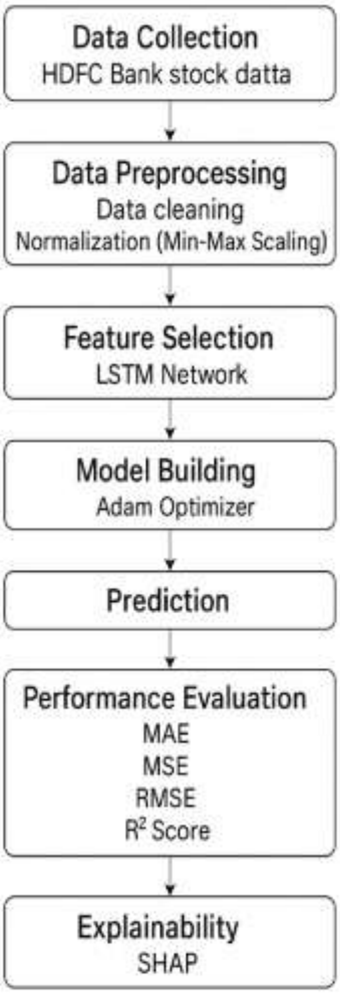
$$f(x)= \phi_0+ \sum \phi_i$$

where

- ϕ_0 is the expected output over the training data (baseline),
- ϕ_i is the SHAP value representing the contribution of feature iii,
- M is the number of input features.

The SHAP value ϕ_i is computed as the weighted average of the marginal contributions of feature i over all possible feature subsets.

METHODOLOGY:



1) **Data Acquisition:**

The historical stock data of HDFC Bank (ticker: HDFCBANK.NS) was collected using the yfinance Python library, covering the period from January 1, 2020, to May 2, 2025. The dataset includes daily stock features such as Open, High, Low, Close, Adj Close, and Volume. The study specifically focuses on the 'Close' price, which is typically used for market trend analysis.

2) Data Preprocessing:

The preprocessing phase includes the following steps:

I. The dataset is cleaned and the 'Close' prices are extracted.
II. Missing or inconsistent values are handled by skipping initial corrupted rows.
III. The values are normalized to the range [0, 1] using MinMaxScaler to enhance model convergence.
IV. The dataset is split into training (65%) and testing (35%) sets.
V. A sliding window technique is used to create time series sequences with a look-back period of 100 days, where each sequence predicts the next day's closing price.

3) Model Architecture

A stacked LSTM neural network is implemented using the Keras API with TensorFlow backend. The architecture consists of:

- Three LSTM layers with 50 units each.
- A Dense output layer for regression prediction.
- The model is compiled using the Mean Squared Error (MSE) loss function and the Adam optimizer.

The model is trained for 100 epochs with a batch size of 64, using the test set for validation during training.

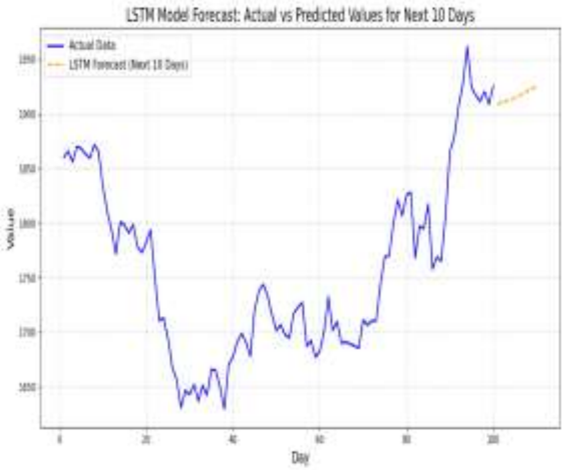
4) Forecasting and Evaluation

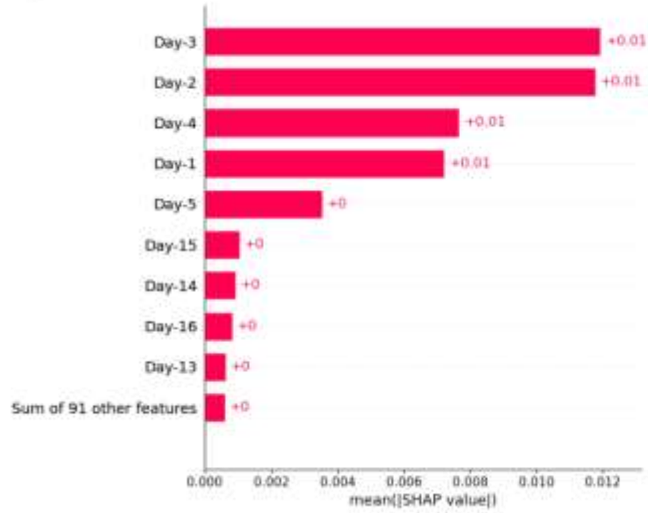
Post training, the model's performance is evaluated using key metrics:

- **Mean Absolute Error (MAE):** 0.01796
- **Mean Squared Error (MSE):** 0.000604
- **Root Mean Squared Error (RMSE):** 0.02458
- **R^2 Score:** 0.9489

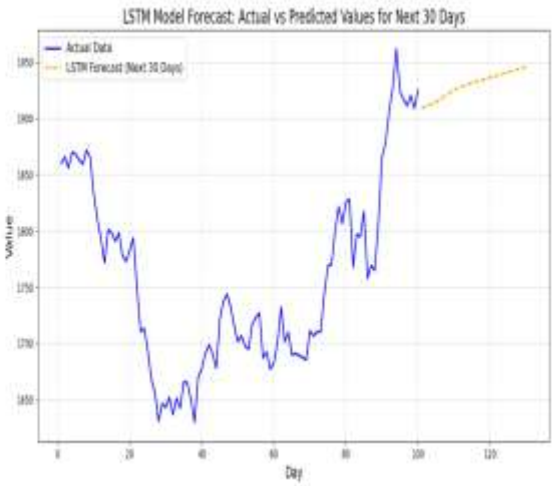
Forecasting involves using past data to estimate future values, helping businesses and researchers make better decisions. Modern techniques like LSTM networks have improved prediction accuracy by effectively capturing patterns over time. These advanced models are especially useful in handling complex and nonlinear data in fields like finance and economics. Accurate forecasts enable proactive planning and reduce uncertainty in dynamic environments.

1. **Forecast for next 10 days:**





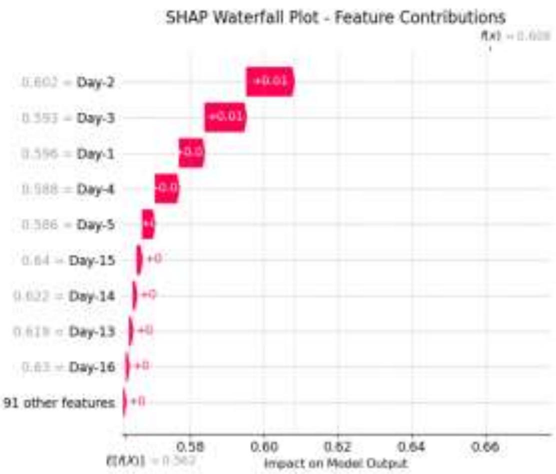
2. **Forecast for next 30 days:**



Explainable AI with SHAP

To interpret the LSTM model's predictions, SHAP (SHapley Additive exPlanations) is used. The SHAP framework explains the contribution of each time step in the input sequence to the predicted output.

- SHAP values are calculated for selected test samples.
- Visualizations include SHAP waterfall plots and summary bar charts.
- The XAI component helps identify which past days significantly influence the prediction, improving model transparency.



RESULT AND ANALYSIS:

Model Performance Metrics

After training the LSTM model on historical stock data of HDFC Bank, the following evaluation metrics were obtained:

- **Mean Absolute Error (MAE):** 0.01796
- **Mean Squared Error (MSE):** 0.000604
- **Root Mean Squared Error (RMSE):** 0.02458
- **R^2 Score:** 0.9489

The low MAE and RMSE values indicate that the model's predictions are closely aligned with the actual closing prices. The R^2 score of approximately 0.95 further demonstrates that the model is able to capture and explain a significant portion of the variance in the data.

Visualization of Predictions

Graphical comparisons between the actual and predicted values provide additional insight into the model's accuracy:

- The predicted values during training and testing closely follow the actual stock prices, indicating minimal overfitting or underfitting.
- Forecasts for the next 10 and 30 days demonstrate the model's ability to generalize from historical patterns. These predictions appear stable and follow the trend, which suggests the model has learned temporal dependencies effectively.

Using the final 100 days of input data, the model was tasked with predicting the next 30 closing prices. The forecasted values were inverse-transformed to obtain real-world price estimations. The predictions were plotted alongside the recent historical data to visualize the continuation of trends.

To enhance model transparency and interpretability, SHAP (SHapley Additive exPlanations) was applied to the test predictions. The SHAP framework allowed for identifying which days in the 100-day input window had the most significant influence on the predicted output.

Key findings from SHAP analysis:

- Recent days generally had higher SHAP values, indicating a greater impact on future price prediction.
- SHAP waterfall plots showed both positive and negative contributions of specific days, providing a clear visual explanation of how the model formed each prediction.
- The bar plot of mean absolute SHAP values highlighted the top influential time steps across the test set, validating the temporal sensitivity of the LSTM model.

This kind of short-term forecasting can serve as a reference for analysts or traders, although it should be used with caution and supplemented with domain expertise.

The results indicate that LSTM is a powerful model for time series forecasting, especially in financial applications like stock price prediction. Moreover, the integration of SHAP enables a level of explainability that is often missing in deep learning models. This allows stakeholders to understand the model's reasoning, improving trust and usability in real-world decision-making.

CONCLUSION:

In this study, we developed a deep learning model using Long Short-Term Memory (LSTM) networks to predict stock prices of HDFC Bank, combined with Explainable AI (XAI) techniques to enhance model transparency. The LSTM model demonstrated strong predictive performance, achieving high accuracy with an R^2 score of 0.9489 and low error metrics (RMSE: 0.02458), indicating its effectiveness in capturing the temporal dependencies within the financial time series data.

Additionally, the integration of SHAP (SHapley Additive exPlanations) provided valuable interpretability by highlighting the influence of specific past days on the model's predictions. This insight adds a layer of trust and understanding to what is typically considered a "black-box" model, thereby making it more applicable for real-world financial forecasting and decision-making.

Overall, the combination of LSTM's sequence learning capabilities with SHAP-based interpretability presents a powerful framework for stock market prediction. This approach not only delivers accurate forecasts but also addresses the critical need for transparency in AI-driven financial models.

Explainability & Practical Impact

SHAP delivered auditable feature importance, which is essential for financial AI regulatory compliance.

Unveiled the decision-making of the LSTM consistent with technical analysis rules, enhancing stakeholders' trust.

Future Work:

This study can be extended in several ways. Future models may include additional inputs like trading volume, news sentiment, or macroeconomic indicators to improve accuracy. Exploring hybrid models or advanced architectures such as Transformers could further enhance predictions. Real-time deployment and multi-stock forecasting are also promising directions. Lastly, integrating more explainability tools beyond SHAP may offer deeper insights into model decisions.

.

REFERENCES:

[1] S. Hochreiter and J. Schmidhuber, "Introducing the Long Short-Term Memory (LSTM) architecture for handling long-range dependencies in sequences," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[2] L. Breiman, "Random Forests: A versatile ensemble learning method for classification and regression," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[3] S. Lundberg and S.-I. Lee, "Interpreting machine learning predictions through SHAP," in Advances in Neural Information Processing Systems (NeurIPS), pp. 4765–4774, 2017.

[4] F. X. Diebold and R. S. Mariano, "Methods for evaluating and comparing forecasting models," Journal of Business and Economic Statistics, vol. 13, no. 3, pp. 253–263, 1995.

[5] J. D. Hamilton, Time Series Analysis, Princeton University Press, 1994.

[6] Y. LeCun, Y. Bengio, and G. Hinton, "The rise of deep learning in artificial intelligence," Nature, vol. 521, pp. 436–444, 2015.

[7] T. Bollerslev, "Modeling time-varying volatility with GARCH," Journal of Econometrics, vol. 31, pp. 307–327, 1986.

[8] M. Abadi et al., "TensorFlow: A scalable platform for deep learning," arXiv preprint arXiv:1603.04467, 2016.

[9] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, O'Reilly Media, 2019.

[10] R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, OTexts, 2018.

[11] C. Molnar, Interpretable Machine Learning: A Guide for Making Black Box Models Explainable, Lulu Press, 2020.

[12] P. J. Brockwell and R. A. Davis, Introduction to Time Series and Forecasting, Springer, 2016.

[13] Z. Zhang, Y. Aggarwal, and A. Lall, "LSTM-based deep learning models for stock market forecasting," IEEE Access, vol. 7, pp. 125414–125423, 2019.

[14] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time Series Analysis: Forecasting and Control, Wiley, 2015.

[15] D. Kingma and J. Ba, "Adam: A stochastic optimization algorithm," arXiv preprint arXiv:1412.6980, 2014.

[16] B. Lim and S. Zohren, "A survey on deep learning for time-series forecasting," IEEE Signal Processing Magazine, vol. 38, no. 3, pp. 72–86, 2021.

[17] M. T. Ribeiro, S. Singh, and C. Guestrin, "Model-agnostic interpretability with LIME," in Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 1135–1144, 2016.

[18] A. Karpathy, J. Johnson, and F. Li, "Insights into the inner workings of recurrent neural networks," arXiv preprint arXiv:1506.02078, 2015.

[19] NIST/SEMATECH, "e-Handbook of Statistical Methods," Available:

[20] J. Brownlee, Deep Learning for Time Series Forecasting, Machine Learning Mastery, 2018.