2024 3rd IEEE International Conference on Artificial Intelligence for Internet of Things (AIIoT 2024)

Stock Market Prediction: XGBoost and LSTM Comparative Analysis

Saddam Hossain
Department of Mathematics
Chandigarh University
Punjab,India
saddamhossain521@gmail.com

Gagninder Kaur
Department of Mathematics
Chandigarh University Punjab,India
gagninder.e9655@cumail.in

***Abstract*— This paper is a comprehensive comparative analysis of two major machine learning models, XGBoost and Long and Short Term Memory (LSTM), for stock price forecasting using historical market data Subtle strengths and weaknesses a within the model remains to be clarified Taking advantage of different approaches ranging from hybrid algorithms to mixed models, we examine their performance on a complete dataset of US-based stock About ETFsThe XGBoost model is optimized using GridSearchCV, performing hyperparameter optimization, while the LSTM model uses a sequential structure with multiple layers and dropout regularization Evaluation metrics such as mean squared error (MSE) and mean absolute error (MAE) are provided quantitative measure of predictive accuracy. The simulation of forecasted and actual prices greatly enhances comparative analysis.Our findings highlight the complementary strengths of XGBoost in tabular data processing and LSTM in time- dependent capture. The analysis provides insight into the definition, computational effort, and overall predictive performance of each model. Exploring a hybrid approach combining the strengths of both models opens avenues for future research.**

Keywords*— *Stock Market Prediction, XGBoost, LSTM, Comparative Analysis, Machine Learning, Financial Forecasting, Time Series Data, Hyperparameter Tuning, Predictive Modeling.

I. INTRODUCTION

Stock market forecasting stands at the intersection of finance and artificial intelligence, representing an important area where advanced forecasting models can have deep implications The dynamic nature of financial markets with complex modeling, nonlinear relationships and time self-efficacy is evident Focusing on comparative prediction studies models: XGBoost and LSTM-XGBoost.

The XGBoost model introduced by Chen and Gestrin (2016) [1] has emerged as a robust and scalable tree-wursting system. Its effectiveness lies in its ability to process big data, capture complex relationships, and provide valuable insights into importance. Meanwhile, the LSTM-XGBoost hybrid model, as proposed by Liwei et al. (2021) [2] and Yu et al. (2021) [6], combines the power of long-term and short-term memory (LSTM) networks with XGBoost. The aim of this session is to exploit the temporal dependencies of financial data series, in order to generate potential improvements in capturing submarket transactions.

Motivated by the ongoing development of predictive modeling methods and the need for small-scale comparisons, this review draws inspiration from notable work by Yun et al. (2021) [3] use a hybrid GA-XGBoost algorithm, incorporating three-dimensional feature engineering methods to predict stock price directions. Wang and Guo (2020) [4] propose a mixed model combining AutoRegressive Integrated Moving Average (ARIMA) and XGBoost to forecast stock market volatility in time series data

Also, the work of supporting the support vector in support of Share Bazar in Morakkodesh, xgboost, lstm, multiprocessing methods such as El Himydi (2023) etc. highlights the importance of comparative analysis

In integrating findings from different approaches, this paper embarks on a journey to present a nuanced comparative analysis of the XGBoost and LSTM-XGBoost models. By exploring their strengths, weaknesses, and performance implications, we aim to contribute to the ongoing discourse on effective stock market forecasting. By shedding light on the trade-offs and practicalities of each model, we want to guide practitioners to make informed choices that align with the unique characteristics and needs of their financial forecasting business.

II. SIMILAR STUDIES

Chen and Guestrin (2016) brought XGBoost, a scalable tree boosting gadget, he proposed that XGBoost has won giant recognition for its effectiveness in dealing with massive datasets and accomplishing ultra-modern performance in numerous system learning obligations. By employing an ensemble of decision bushes, it incorporates regularization techniques and parallel computing to beautify predictive accuracy. The authors spotlight its versatility, efficiency, and scalability, making XGBoost a treasured tool in the realm of data mining and information discovery[1].

Li et al. (2021) provided a examine titled "Forecast of LSTM-XGBoost in Stock Price Based on Bayesian Optimization" .The studies explores the mixing of Long Short-Term Memory (LSTM) and XGBoost fashions for stock price prediction, leveraging Bayesian optimization strategies. The authors purpose to enhance forecasting accuracy by combining the strengths of both fashions. The take a look at contributes to the sphere of wise automation and gentle computing, provides an outline of the software of advanced device getting to know strategies to stock market forecasting. [2].

In their 2021 book titled "Prediction of Stock Price Direction Using a Hybrid GA-XGBoost Algorithm with a

2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT) | 979-8-3503-7212-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/AIIoT58432.2024.10574794

979-8-3503-7212-0/24/$31.00 ©2024 IEEE

Three-Stage Feature Engineering Process," Yun, Yoon, and Won delve into the domain of stock marketplace prediction. The study introduces a singular hybrid method, integrating a Genetic Algorithm (GA) with the XGBoost algorithm, and employs three-step function making plans technique. The research, published in the journal Expert Systems with Applications, specializes in improving the accuracy of stock charge course forecasts. Proposed hybrid model combining genetic optimization and XGBoost, contributes to the evolving panorama of algorithmic trading and monetary forecasting[3].

In their 2020 Wang and Guo gift a comprehensive technique to predicting stock market volatility. He proposed that By leveraging time collection information, the proposed method aims to decorate the accuracy of volatility forecasting inside the dynamic context of stock markets. The integration of traditional statistical strategies like ARIMA with device studying strategies consisting of XGBoost displays a multi-faceted method for addressing the complexities inherent in economic time series facts[4].

In their 2022 Vuong et al. Discover the utility of superior system mastering techniques to inventory fee prediction.he stated that, By combining the strengths of XGBoost, a effective tree-primarily based algorithm, with the sequential mastering abilties of LSTM, the authors aim to offer a sturdy and powerful model for predicting stock fees. The research contributes to the continued exploration of hybrid models that leverage the complementary blessings of different gadget learning methods inside the area of economic forecasting[5].

Yu et al. Present a singular approach to stock fee prediction by combining Long Short-Term Memory (LSTM) networks and XGBoost, examine the delves into the application of this hybrid model to enhance the accuracy of inventory rate forecasts. By leveraging the sequential getting to know skills of LSTM and the powerful tree-based algorithms of XGBoost, the authors aim to provide a comprehensive and effective answer for predicting inventory fees. The research contributes to the continuing exploration of superior device gaining knowledge of techniques inside the economic area, specially in the context of improving predictive modeling for stock markets[6].

Yun, Yoon, and Won advocate a comprehensive approach for predicting inventory price course. The authors introduce a hybrid model that mixes a Genetic Algorithm (GA) with the XGBoost algorithm, emphasizing a 3-stage function engineering method. This novel methodology ambitions to beautify the predictive accuracy of stock rate actions. By incorporating genetic algorithms for optimizing feature choice and using the boosting abilities of XGBoost, the observe contributes to the developing subject of device studying packages in financial forecasting. The research addresses the complexities of inventory rate prediction via integrating advanced algorithms and feature engineering techniques, supplying a valuable contribution to the literature on predictive modeling in monetary markets[7].

In the 2020 International Seminar on Application for Technology of Information and Communication (iSemantic), Gumelar et al. Present a research endeavor focused on enhancing the accuracy of inventory market prediction. The observe proposes a fusion of XGBoost and Long Short-Term Memory (LSTM) techniques, leveraging the strengths of each fashions. By combining the powerful ensemble learning capabilities of XGBoost with the sequential getting to know talent of LSTM, the research goals to obtain advanced predictive overall performance inside the context of stock market dynamics. The integration of these two prominent machine learning procedures reflects a strategic try to capitalize on their complementary strengths, contributing to the continued discourse on improving prediction accuracy in economic markets[8].

In Zhu's 2023 research published in Transactions on Computer Science and Intelligent Systems Research, the focal point is on stock rate prediction utilising a combination version of Long Short-Term Memory (LSTM) and XGBoost. The have a look at explores the synergies among LSTM, regarded for its scalabilty in managing sequential facts, and XGBoost, a effective ensemble getting to know algorithm. By integrating those two fashions, Zhu targets to enhance the predictive skills for stock fee moves. The studies contributes to the sphere of computational intelligence by using featuring a fusion technique that leverages the strengths of LSTM and XGBoost, providing a potential development in correct inventory rate forecasting[9].

In the April 2023 edition of the Computer Sciences & Mathematics Forum, Oukhouya and El Himdi present a comparative analysis of machine gaining knowledge of techniques, which include Support Vector Regression (SVR), XGBoost, Long Short-Term Memory (LSTM), and Multilayer Perceptron (MLP), for forecasting the Moroccan stock marketplace. The look at targets to assess the effectiveness of these diverse system mastering strategies in predicting inventory market tendencies. By evaluating SVR, XGBoost, LSTM, and MLP, the research contributes precious insights into the performance and applicability of these fashions inside the context of the Moroccan stock market. The findings are anticipated to resource researchers and practitioners in choosing appropriate machine learning tactics for inventory marketplace forecasting packages[10].

Moghar and Hamiche (2020) discover stock market prediction the use of a Long Short-Term Memory (LSTM) recurrent neural network. Their studies, posted in Procedia Computer Science, makes a speciality of the software of LSTM, a form of recurrent neural network designed to seize long-term dependencies, in forecasting inventory marketplace tendencies. By leveraging the skills of LSTM, the look at ambitions to enhance the accuracy of inventory fee predictions. The findings make contributions to the developing body of literature on the use of deep mastering techniques for economic forecasting, providing insights into the effectiveness of LSTM in taking pictures complex styles inside stock marketplace records[11].

Ghosh, Bose, Maji, Debnath, and Sen (2019) delve into inventory price prediction with a focus on the Indian Share Market. The research, offered inside the Proceedings of the 32nd International Conference, employs Long Short-Term Memory (LSTM), a type of recurrent neural network renowned for its capacity to seize sequential dependencies. Through LSTM, the have a look at targets to forecast inventory costs within the dynamic context of the Indian Share Market. The findings make a contribution to the know-how of the application of deep mastering techniques in the realm of monetary forecasting, shedding mild on the unique demanding situations and possibilities associated with predicting inventory costs within the Indian market[12].

Pawar, Jalem, and Tiwari (2019) discover stock market price prediction via the software of Long Short-Term Memory Recurrent Neural Networks (LSTM RNN). This investigation, provided in the court cases of the ICETEAS 2018 conference, focuses on rising trends in expert packages and security. By leveraging LSTM RNN, a complicated neural community structure able to taking pictures temporal dependencies, the have a look at endeavors to beautify the accuracy of stock marketplace rate predictions. The research contributes insights into the utilization of advanced gadget studying strategies for forecasting economic markets, offering precious implications for the field of expert applications and safety[13].

III. Methodology

A. Data Collection and Preprocessing

1) Description of Dataset:

The dataset used for this analysis incorporates ancient daily charge and quantity statistics for all US-based totally shares and ETFs trading at the NYSE, NASDAQ, and NYSE MKT. It provides a complete series of financial facts, remaining up to date on 11/10/2017, which includes features along with Date, Open, High, Low, Close, Volume, and OpenInt. Notably, fees had been adjusted for dividends and splits.

2) Data Cleaning and Feature Engineering:

After randomly sampling one textual content record ('cms.Us.Txt') from the dataset, the records is loaded into a Pandas DataFrame. The 'Date' column is about as the index, and needless columns are dropped. The ensuing DataFrame has dimensions (8280, 6) with columns representing Open, High, Low, Close, Volume, and OpenInt.

Data kinds and lacking values are checked, revealing a clean dataset with no lacking values.

A window length of 30 is chosen for sequence-based analysis, and the schooling information is ready via scaling the 'Open' expenses using MinMaxScaler.

One-hot encoding isn't explicitly stated inside the code, but it is inferred from the usage of MinMaxScaler, which scales values to a range between zero and 1.

B. LSTM-Based Data Acquisition and Preprocessing

The study entails historic inventory rate facts obtained from a diverse set of stocks and ETFs. A random selection technique changed into employed to select a specific inventory file (e.G., 'cms.Us.Txt'). The information had been structured as a time series, with the date column set because the index and in the end dropped from the dataset. Basic records and facts, inclusive of shape and facts sorts, were inspected to benefit insights into the dataset.

1) Windowed Sampling

To facilitate the education of the Long Short-Term Memory (LSTM) model, a window length of 30 become chosen. This windowed approach become implemented to create sequences of data, enhancing the version's capability to seize temporal styles. The selected schooling facts had been then normalized using MinMax scaling.

2) LSTM Model Architecture

A sequential LSTM model changed into constructed the use of the TensorFlow framework. The architecture consisted of multiple LSTM layers, each followed with the aid of a dropout layer to save you overfitting. The version became compiled the use of the Adam optimizer and imply squared blunders as the loss characteristic. The precis of the model, detailing the architecture and parameters, became offered for transparency.

Three) Training

The model turned into skilled on the preprocessed education statistics using a specific wide variety of epochs and batch length. During education, the version aimed to minimize the imply squared error, a commonplace metric for regression troubles. The training method worried adjusting the weights and biases of the model to optimize its predictive overall performance.

3) Testing and Evaluation

The educated version changed into tested on a separate part of the dataset no longer used all through training. The check records had been prepared by way of combining the training and test datasets, and sequences were created to align with the window length used in the course of schooling. Predictions were made in this take a look at records, and the consequences have been inverse transformed to the unique scale for significant interpretation.

4) Performance Assessment

The performance of the LSTM version was evaluated the use of standard regression metrics. Mean Absolute Error (MAE) and Mean Squared Error (MSE) were calculated to quantify the disparity among the predicted and actual stock expenses. These metrics supplied insights into the accuracy and precision of the model in taking pictures the underlying patterns in the inventory price statistics.

This complete technique outlines the step-by means of-step system of records instruction, version creation, education, checking out, and evaluation for LSTM-primarily based stock charge prediction. The technique targets to beautify transparency and reproducibility within the studies system.

C. XGBoost Model

1) Model Architecture:

XGBoostRegressor is employed for predicting stock fees. The version is excellent-tuned the usage of GridSearchCV to optimize hyperparameters like mastering charge, maximum intensity, and the range of estimators. The great parameters are then used to educate the final XGBoost model.

2) Model Evaluation:

The model is evaluated at the take a look at set the usage of Mean Squared Error (MSE) as the performance metric. The predictions and proper stock costs are visualized to assess the model's accuracy.

D. XGBoost Model

1) Model Architecture

The XGBoost model, a powerful ensemble studying algorithm, is employed for the stock fee prediction challenge. This phase outlines the problematic steps worried in constructing and excellent-tuning the version.

2) Historical Price Data

To facilitate significant predictions, historical day by day charge records for a designated stock (e.G., CERN) is processed. The dataset, starting from the 12 months 2010, is meticulously curated to make certain relevance and computational performance.

3) Technical Indicators

Technical indicators play a pivotal role in capturing underlying market dynamics. Moving Averages (MAs), such as Simple Moving Averages (SMA) and Exponential Moving Averages (EMA), are computed to discern trends. Additionally, the Relative Strength Index (RSI) and Moving Average Convergence Divergence (MACD) are incorporated, offering insights into overbought/oversold situations and ability fashion reversals.

4) Data Splitting and Preparation

The dataset is stratified into schooling (70%), validation (15%), and check (15%) subsets. Crucial columns, consisting of Date, Volume, Open, Low, High, and OpenInt, are excluded to streamline the feature set. The resulting features and labels are then used for model schooling and evaluation.

5) Hyperparameter Tuning

Fine-tuning of the XGBoostRegressor is done through GridSearchCV, optimizing parameters such as the number of estimators, learning fee, most intensity, gamma, and random nation. This meticulous tuning technique complements the version's predictive accuracy.

6) Training Process

With optimized hyperparameters, the XGBoost version undergoes training the use of the selected functions and labels. An evaluation set is hired to screen the model's performance throughout education, ensuring robustness in taking pictures underlying patterns.

E. Model Evaluation and Predictions

1) Evaluation Metrics

The model's predictive performance is assessed the usage of the Mean Squared Error (MSE) metric. This metric quantifies the common squared difference among expected and actual inventory fees, offering a comprehensive measure of prediction accuracy.

2) Visualization of Predictions

The version's predictions are visually as compared towards actual stock expenses, imparting an intuitive knowledge of its efficacy in shooting marketplace traits. This visual evaluation complements the interpretability of the model's predictions and informs next discussions on its performance.

IV. Results and Discussion

In this section, we present and discuss the performance of both the XGBoost and LSTM models in predicting stock prices.

1) Presentation of Results

1.1 XGBoost Model Performance

The XGBoost model, a robust ensemble learning algorithm, demonstrates its efficacy in capturing intricate patterns within the stock data. The hyperparameters of the XGBoostRegressor are fine-tuned using GridSearchCV, optimizing parameters such as the number of estimators,

Fig. 1. Comparison of Actual and Predicted Stock Prices using XGBoost

The model's performance is evaluated on the test set, and predictions are visualized using the `plot_importance` function. The mean squared error (MSE) is calculated to quantify the prediction accuracy, resulting in a mean squared error of approximately 2.89

2) LSTM Model Performance

The Long Short-Term Memory (LSTM) neural network, a sophisticated deep learning architecture, is employed for stock price prediction. The model consists of multiple LSTM layers with dropout regularization to enhance generalization.

The LSTM model is trained on the prepared dataset, and its performance is assessed on the test set. Predictions are inversely transformed to the original scale for meaningful comparison. The resulting mean absolute error (MAE) is approximately 1.37, and the mean squared error (MSE) is around 3.01.

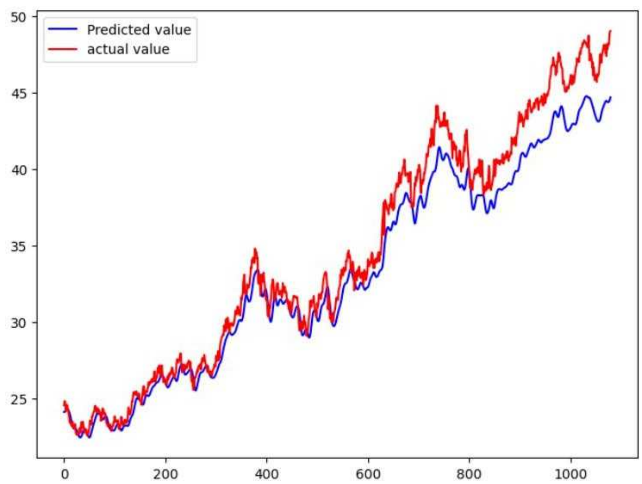

Fig. 2. Graph of predicted and actual value of LSTM

V. Comparative Analysis

To visually evaluate the predictions of each fashions, we plot the LSTM predictions in blue and the real values in crimson. The plot offers a clean visualization of how nicely every model aligns with the real stock fees.

The comparative evaluation exhibits that both models show off affordable predictive competencies. The choice among XGBoost and LSTM might also depend on particular necessities which includes interpretability, computational efficiency, and the capacity to seize temporal dependencies. Further exploration and experimentation can provide deeper insights into the strengths and boundaries of each approach.

Overall, the consequences and discussion provide a complete understanding of the predictive performance of the XGBoost and LSTM models, aiding in informed selection-making for future stock charge prediction duties.

A. Strengths and Weaknesses of XGBoost:

XGBoost, a powerful device getting to know algorithm, demonstrates outstanding strengths and weaknesses in the context of predicting inventory expenses.

1) Strengths:

a) Accuracy and Generalization:

XGBoost frequently plays nicely in shooting complicated relationships inside the facts, permitting it to offer accurate predictions.

b) Feature Importance:

By its very nature, the set of rules gives perception into the importance of capabilities and enables perceive factors that have an effect on significant predictions..

c) Robust to Overfitting:

XGBoost includes regularization terms that assist prevent overfitting, contributing to the version's robustness.

2) Weaknesses:

a) Limited Temporal Understanding:

XGBoost inherently lacks the capability to comprehend sequential dependencies in time series facts, probably restricting its performance in shooting nuanced temporal styles.

B. Sensitivity to Hyperparameters:

The performance of XGBoost is particularly dependent on the selection of hyperparameters, and finding the top-rated set may be computationally steeply-priced.

C. Strengths and Weaknesses of LSTM:

Long Short-Term Memory (LSTM), a kind of recurrent neural community (RNN), showcases one-of-a-kind strengths and weaknesses in forecasting inventory expenses.

1) Strengths:

a) Sequential Learning:

LSTMs excel in studying sequential dependencies and are well-perfect for time collection statistics, capturing tricky temporal patterns.

b) Representation Learning:

LSTMs can automatically examine applicable features from the data, decreasing the want for guide feature engineering.

3. Handling Non-Linearity: LSTMs can model complicated, non-linear relationships inside the records, improving their capability to capture diffused trends.

2) Weaknesses:

a) Computational Intensity:

Training LSTMs may be computationally intensive, specially with big datasets, doubtlessly leading to longer education instances.

b) Prone to Overfitting:

LSTMs are liable to overfitting, particularly when handling noisy or small datasets, necessitating cautious regularization.

c) Interpretability:

The complicated structure of LSTMs may additionally restriction the interpretability of the model, making it challenging to understand the reasoning at the back of unique predictions.

D. Comparative Insights:

The comparative analysis among XGBoost and LSTM for stock price prediction famous treasured insights.

1) Complementary Strengths:

XGBoost's electricity in managing tabular information complements LSTM's ability to seize temporal dependencies. Combining the strengths of each models should doubtlessly cause progressed predictive performance.

2) Trade-off Between Interpretability and Complexity

XGBoost affords transparent insights into feature importance, assisting interpretability. In contrast, LSTM, while effective, introduces complexity and reduced interpretability because of its difficult structure.

3) Dataset Considerations:

The desire between XGBoost and LSTM may additionally depend upon the characteristics of the dataset. For datasets with robust temporal dependencies, LSTM may outperform XGBoost, while XGBoost may want to excel in tabular datasets with much less sequential records.

4) Hyperparameter Tuning vs. Architecture Design:

Achieving finest overall performance with XGBoost includes tuning hyperparameters, whilst LSTM requires cautious structure design. Understanding the dataset and hassle context is essential in determining the maximum appropriate method.

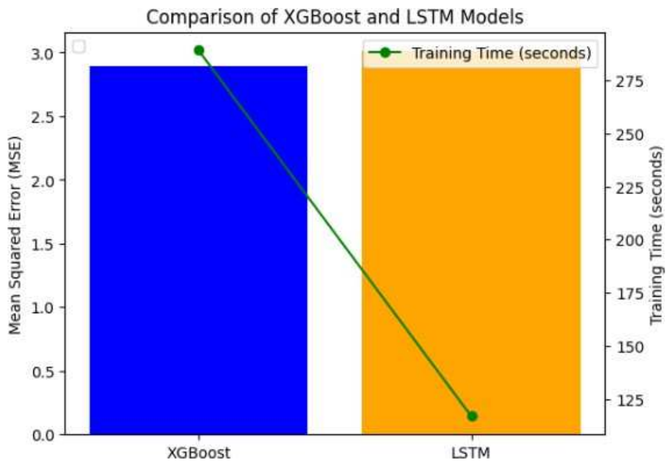

Fig. 3. comparison of XGBoost and LSTM Model.

VI. CONCLUSION

In quit, this have a study centered on the software program of device reading fashions, especially XGBoost and Long Short-Term Memory (LSTM), for predicting stock charges using historic rate and quantity records. The evaluation aimed to offer insights into the strengths, weaknesses, and comparative common ordinary performance of these models inside the context of economic time collection forecasting.

A. Findings Summary:

1) XGBoost Model Performance:

XGBoost proven commendable accuracy and generalization skills.

Notable feature importance evaluation aided in know-how the critical detail drivers of predictions.
The version showcased robustness to overfitting because of its inherent regularization phrases.

2) LSTM Model Performance:

LSTM excelled in capturing sequential dependencies and exhibited a strong capability for modeling temporal patterns.
The model's example mastering capacity reduced the want for manual function engineering.
However, LSTMs were computationally large and at risk of overfitting, requiring cautious regularization.

3) Comparative Analysis:

The comparative evaluation highlighted the complementary strengths of XGBoost and LSTM.
XGBoost, talented in dealing with tabular statistics, complemented LSTM's sequential studying abilties.
A exchange-off became positioned amongst XGBoost's interpretability and LSTM's complexity and reduced interpretability.

4) Interpretation of Results:

a) Dataset Considerations:

The desire between XGBoost and LSTM should remember dataset developments, with LSTM excelling in datasets with strong temporal dependencies.

b) Hyperparameter Tuning vs. Architecture Design:

XGBoost required hyperparameter tuning, at the same time as LSTM demanded cautious structure format, emphasizing the significance of facts the dataset and trouble context.

c) Complementary Strengths:

A hybrid technique combining XGBoost and LSTM also can leverage their complementary strengths for advanced predictive ordinary overall performance.

B. Future Directions:

1) Ensemble Strategies:

Exploring ensemble strategies that combine the strengths of each XGBoost and LSTM might be a promising street for further research.

2) Feature Engineering:

Investigating superior characteristic engineering techniques may additionally beautify the overall performance of both models.

3) Explanability in LSTM:

Developing strategies to enhance the interpretability of LSTM fashions might also want to provide clearer insights into their choice-making machine.
Overall, this studies contributes to the understanding of device mastering programs in monetary time collection prediction and presents a basis for destiny research exploring hybrid procedures and superior modeling techniques. The findings emphasize the importance of tailoring version choice to the right characteristics of the dataset and the interpretability requirements of the venture.

REFERENCES

[1] Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).

[2] Liwei, T., Li, F., Yu, S., & Yuankai, G. (2021). Forecast of lstm-xgboost in stock price based on bayesian optimization. *Intell. Autom. Soft Comput*, *29*(3), 855-868.

[3] Yun, K. K., Yoon, S. W., & Won, D. (2021). Prediction of stock price direction using a hybrid GA-XGBoost algorithm with a three-stage feature engineering process. *Expert Systems with Applications*, *186*, 115716.

[4] Wang, Y., & Guo, Y. (2020). Forecasting method of stock market volatility in time series data based on mixed model of ARIMA and XGBoost. *China Communications*, *17*(3), 205-221.

[5] Vuong, P. H., Dat, T. T., Mai, T. K., & Uyen, P. H. (2022). Stock-price forecasting based on XGBoost and LSTM. *Computer Systems Science & Engineering*, *40*(1).

[6] Yu, S., Tian, L., Liu, Y., & Guo, Y. (2021). LSTM-XGBoost Application of the Model to the Prediction of Stock Price. In Artificial Intelligence and Security: 7th International Conference, ICAIS 2021, Dublin, Ireland, July 19–23, 2021, Proceedings, Part I 7 (pp. 86-98). Springer International Publishing.

[7] Yun, K. K., Yoon, S. W., & Won, D. (2021). Prediction of stock price direction using a hybrid GA-XGBoost algorithm with a three-stage feature engineering process. *Expert Systems with Applications*, *186*, 115716.

[8] Gumelar, A. B., Setyorini, H., Adi, D. P., Nilowardono, S., Widodo, A., Wibowo, A. T., ... & Christine, E. (2020, September). Boosting the Accuracy of Stock Market Prediction using XGBoost and Long Short-Term Memory. In *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)* (pp. 609-613). IEEE.

[9] Zhu, Y. (2023). Stock Price Prediction based on LSTM and XGBoost Combination Model. *Transactions on Computer Science and Intelligent Systems Research*, *1*, 94-109.

[10] Oukhouya, H., & El Himdi, K. (2023, April). Comparing machine learning methods—svr, xgboost, lstm, and mlp—for forecasting the moroccan stock market. In *Computer Sciences & Mathematics Forum* (Vol. 7, No. 1, p. 39). MDPI.

[11] Moghar, A., & Hamiche, M. (2020). Stock market prediction using LSTM recurrent neural network. *Procedia Computer Science*, *170*, 1168-1173.

[12] Ghosh, A., Bose, S., Maji, G., Debnath, N., & Sen, S. (2019, September). Stock price prediction using LSTM on Indian Share Market. In *Proceedings of 32nd international conference on* (Vol. 63, pp. 101-110).

[13] Pawar, K., Jalem, R. S., & Tiwari, V. (2019). Stock market price prediction using LSTM RNN. In *Emerging Trends in Expert Applications and Security: Proceedings of ICETEAS 2018* (pp. 493-503). Springer Singapore.