

LSTM WITH XGBOOST ALGORITHM FOR INCREASED STOCK MARKET PREDICTION

Vijay Bhanu S

Research Scholar, Department of Business Administration, Annamalai University, India

Email ID: svbhanu22@gmail.com

Dr. S. Arulkumar

Research Supervisor, Department of Business Administration, Annamalai University, India

Email ID: arulmba123@gmail.com

Abstract

Stock market prediction is a challenging task due to the nature of financial markets which are volatile and non-linear. Traditional machine learning models often struggle to capture the complex temporal dependencies and underlying patterns in stock price movements. This paper presents a hybrid model that combines Long Short-Term Memory (LSTM) networks with eXtreme Gradient Boosting (XGBoost) to enhance stock market prediction accuracy. LSTM, a type of recurrent neural network (RNN), is well-suited for handling time series data, as it effectively models sequential dependencies. Financial time-series data are known to be volatile, unpredictable, non-linear and temporally dependent making stock market prediction a tough problem. A hybrid model of Long Short-Term Memory (LSTM)-XGBoost is proposed to improve the stock closing price prediction by using the sequential learning ability of LSTM and the non-linear modelling ability of XGBoost. Historical stock-market data for Alibaba (BABA) from 1 January 2016 to 1 October 2021 was used for the analysis. The data collection comprises important market characteristics such as Open, High, Low, Close, Adjusted Close, and Volume. The LSTM model learned the temporal relationships and patterns from the historical stock-price sequences. Then the residual errors and other non-linear relationships were modelled to improve the LSTM predictions with the XGBoost. Model Evaluation The model performance was evaluated using Mean Squared Error (MSE), Mean Absolute Error (MAE) and R2. The findings reveal that the standalone LSTM model achieved an MSE of 0.1219, MAE of 0.3163, and R2 of -4.1150. The hybrid LSTM-XGBoost model achieved an MSE of 0.0001, MAE of 0.0050, and R2 of 0.9972. The results show that the prediction metrics improved significantly by adding XGBoost to the LSTM framework. The study shows the potential of hybrid deep learning and boosting techniques in stock market prediction.

Keywords: RNN, LSTM, XGBoost, stock market, prediction, neural networks, accuracy, precision

I. INTRODUCTION

Understanding stock price movement is one of the oldest areas of study in finance and more recently in the discipline of machine learning and there are good reasons for that. Still predictions on stock prices hold true to be one onerous task due to the large degree of randomness, non-linearity and variations embedded in financial markets. Stock market data is complicated and is often a periodical series; ARIMA, as a classical statistical model, usually becomes ineffective in this context. At the same time, in the last few years, development of

LSTM networks for time series forecasting has attracted a quick interest. These models known as Long Short-Term Memory (LSTM) networks have been used to great effect for sequential data predictions, modelling reservoir computing. These models focusing on stock price prediction proved to be efficient for systematic survey of time dependent data, however they face a problem of overfitting or simply model collapse as well as hard time obtaining unsmoothed interrelationships in those data.

In order for further enhancement of the performance of the stock market prediction models, this paper suggests the implementation of a model which proposes an LSTM – XGBoost framework. XGBoost is a powerful ensemble technique that employs boosting, enabling it to learn non-linear interactions efficiently while also addressing the problem of overfitting. The goal of this hybrid model is to combine the strengths of the LSTM network in capturing temporal relationships and the strength of XGBoost in achieving good accuracy in time series prediction.

Stock market prediction has become an important research subject in finance, economics, data analytics and machine learning with the rising availability of financial data and the need to comprehend movement in stock prices. Investors, financial institutions, portfolio managers and researchers are always seeking trends in past market data that can provide them a clue about where prices are headed. However, it is still a challenging task to predict stock values due to the existence of many economic, technological, social, and market elements driving the financial markets. Stock prices are generally volatile, random, non-linear and non-stationary, and so it is difficult to apply the traditional forecasting methods effectively. In this paper, we propose a hybrid prediction model based on Long Short-Term Memory (LSTM) and eXtreme Gradient Boosting (XGBoost) algorithm to enhance the prediction of stock market.

Time-series forecasting has been intensively studied using classical statistical approaches. Some temporal correlations can be captured using models such as ARIMA. However, their efficiency may be limited if the underlying data contains complicated non-linear relationships and fast changing patterns. Stock market observations are sequential in nature and the current price may be tied to information included in earlier observations. Therefore, a good prediction model should take into account the time dependence of the data and the possible non-linear relationship between different observations. The source paper lists these traits as significant obstacles in the stock price prediction and emphasises the limitations of traditional forecasting methods in addressing the complicated financial time-series data. The advent of deep learning has opened new horizons for a financial time-series forecast. One of the neural network architectures that are particularly ideal for sequential data are the LSTM networks, which are designed to remember relevant information over time. LSTMs are a unique type of recurrent neural network (RNN) that employ an internal memory to handle sequences. An LSTM differs from a simple recurrent network in that it is able to control the flow of information in its cell state and gates. The input gate determines which information is going to be added to the cell state, the forget gate determines which information is going to be removed from the cell state and the output gate determines which information is going to be outputted from the cell state. These techniques enable LSTM networks to learn dependencies over both short and large time scales in data sequences.

The use of LSTM in stock market prediction is relevant since the historical stock prices are time-dependent sequences. An LSTM model can use sequences of past stock observations as input and learn temporal patterns that may be related to future price changes. In the suggested approach, the historical data of stocks are fed as inputs in the LSTM architecture. The model consists of LSTM layers to learn the temporal patterns and dropout and batch-normalization components to limit the chances of overfitting. Then a dense layer is used to give the projected stock-price values. LSTM networks are capable of learning the temporal correlations well, but the prediction accuracy may still be influenced by the complicated and non-linear structure of financial data. In instance, the residual errors generated by an initial forecasting model may contain additional information that can be used to enhance forecasts. This gives us a chance to fuse deep learning with ensemble machine-learning methods. XGBoost is an efficient implementation of gradient-boosting technique, which produces a strong predictive model by merging numerous weak predictors, usually decision tree models. Each subsequent tree strives to minimise the mistakes caused by the preceding trees. This way the entire model can improve its predictions gradually. XGBoost additionally has built-in regularisation algorithms (L1 and L2 regularisation), tree pruning and computational optimisation. These properties make it a good candidate for describing complex non-linear connections in organised and tabular data.

The hybrid LSTM-XGBoost framework investigated in this study is built upon the combination of LSTM and XGBoost. The basic idea is to apply the LSTM to capture the time dependencies of stock-price sequences and to use the XGBoost to optimise the predictions by learning residual errors. The main time-series prediction is performed with the LSTM model, with XGBoost used to compensate for any non-linear effects that the LSTM model may have not caught. The combination is meant to harness the complimentary strengths of both techniques, rather than rely on a single prediction algorithm. The dataset utilised contains stock-market observations. The performance of the LSTM model and boosted LSTM model using XGBoost are analysed and compared. This comparison is performed by means of the prediction-error metrics which evaluate the gap between real and predicted stock values. Mean Squared Error (MSE) is the squared error of prediction. The lower the MSE, the closer the predictions are to the actual observations. MAE means Mean Absolute Error. It is the mean of the absolute errors between the true and predicted values. It is somewhat less affected by particular huge errors because the errors are not squared.

Hence, in this study, a hybrid LSTM-XGBoost algorithm is proposed for the improved performance of stock market prediction. The proposed approach is aimed to overcome the issues of volatility, non-linearity and temporal dependencies in financial time-series data by integrating the sequential learning capability with the gradient-boosting-based error optimisation. This research adds to the understanding of the integration of deep learning and ensemble machine-learning algorithms for stock-price forecasting and presents a comparative evaluation with historical Alibaba stock-market data. According to the original paper, further improvement can be accomplished by extending the observation period and increasing the amount of training data accessible to the model.

II. LITERATURE REVIEW

Interest in the use of machine learning and deep learning techniques to stock-market prediction is growing as a consequence of the presence of complex trends within financial time-series data that cannot be easily represented using traditional forecasting approaches. Tiwari et al. (2017) studied the stock price prediction using data analytics. They stressed the need of analytical tools to derive relevant information from past market observations. Traditional methods often rely on statistical relationships in historical observations, whereas machine learning methods allow modelling of more complicated interactions in financial datasets. The field on stock-market prediction has therefore increasingly migrated to computer approaches that can handle vast amounts of historical information. In particular, recurrent neural networks have been attracting attention as stock prices are sequential observations in which previous values may include information for understanding forthcoming movements. The examined paper stresses the fact that stock-market data are characterised by unpredictability, non-linearity, volatility and temporal variation, which defy standard statistical forecasting methods. Such properties motivate us to study deep learning systems capable of learning sequential relations. This makes the LSTM networks particularly interesting, since the memory mechanism allows the network to keep information from prior observations and update it selectively. The current research therefore sets a basis for the implementation of LSTM for financial forecasting, but also shows that prediction performance can be influenced by the complexity of market activity and the risk of model overfitting.

Long Short-Term Memory networks have been studied as an important deep-learning technique for sequential prediction. They studied a deep-learning model to forecast buy and sell recommendations in the Stock Exchange of Thailand using LSTM. Their work shows the relevance of LSTM to stock-market applications because the architecture is specifically built to analyse sequential input. LSTM is an extension of the recurrent neural-network technology that includes tools to control the flow of input through the network. The main parts are the cell state, the hidden state, the input gate, the forget gate and the output gate. These components are used by the network to decide which information is kept, which is updated and which is sent on to later phases. This feature is especially helpful when dealing with financial time-series data, where observations are ordered in time and patterns may extend across numerous time periods. The research so implies that LSTM may provide an appropriate framework to model temporal dependencies that are not well modelled by simpler prediction models. However, the examined source also points out the drawbacks of LSTM-based stock prediction such as overfitting and challenges in capturing all the underlying linkages in complicated financial data. Thus, although LSTM provides a significant basis for stock-price prediction, other modelling methodologies could be needed to enhance the precision and resilience of its projections.

Another notable stream of work covers gradient-boosting algorithms, especially XGBoost, which has shown beneficial in simulating complex non-linear interactions. Mustapha and Saeed (2016) used extreme gradient boosting in the prediction of bioactive molecules which demonstrate the wider use of XGBoost as a machine learning technique for prediction. Bentejac and Csorgo also carried out a comparison study of XGBoost, which is indicative of the

continuous research interest in analysing the characteristics and performance of this algorithm. XGBoost is based on gradient boosting which combines many weak prediction models to make a stronger one. Successive decision tree models try to compensate the faults of earlier trees and so improve their forecast performance step by step. It also has regularisation with L1 and L2 penalties. It uses tree pruning and parallel processing. These properties make XGBoost suitable for structured datasets and issues with non-linear relationships. These features are relevant for financial forecasting, since the relationships in the stock-market might not follow simple linear patterns. The literature suggests that XGBoost might thus supplement deep-learning methods by giving an alternative way of learning non-linear patterns. This is especially true when the residual errors of an initial forecasting model may contain systematic information. Instead of considering these residuals as mere prediction errors that we cannot explain, we can train a second boosting model to understand what these residuals depend on and apply this knowledge to correct the original predictions Amthul Thawab et al., (2022).

More recent work has also examined hybrid techniques that combine LSTM with XGBoost, giving a platform for combining temporal learning with ensemble based prediction. Mekruksavanich, Jantawong, and Jitpattanakul (2022) combined LSTM with XGBoost to propose an LSTM-XGB deep learning model for human activity recognition. The study is not in the same domain as financial forecasting, but it shows the possibility of combining these two algorithms in the same predictive framework. The importance of such hybridisation to stock-market prediction is that the two methodologies have complementary properties. LSTM is good at learning temporal dependencies from sequential observations while XGBoost can learn nonlinear correlations and iteratively reduce predicted mistakes. The present source applies this idea especially to stock-market prediction by employing LSTM to generate forecasts, and XGBoost to optimise those forecasts using residual-error modelling. Historical stock-market data of Alibaba from 1 January 2016 to 1 October 2021 is utilised to compare the LSTM and enhanced LSTM techniques. The predicted metrics of the provided findings reveal a significant difference: the LSTM model has MSE, MAE, and R2 of 0.1219, 0.3163, and -4.1150, whereas the LSTM-XGBoost has MSE, MAE, and R2 of 0.0001, 0.0050, and 0.9972, respectively. The results offer an immediate research framework for investigating hybrid LSTM-XGBoost modelling as an approach to stock-market prediction. Overall, the literature suggests that the combination of sequential deep learning with gradient boosting provides a framework for tackling both temporal dependence and non-linear prediction errors.

III. METHODOLOGY

The aim is to conclude that using LSTM along with XGBoost algorithm is better for stock market predictions. To tackle the problem of stock market predictions in the most efficient and accurate way, the authors of this paper suggest a novel approach of using LSTM networks and XGBoost together. First, data on past stock prices, stock market indicators like prices and charts, and other necessary data for the analysis are collected from records such as Yahoo Finance. The construct of the LSTM consists of an input layer receiving sequences of historical stock data, several LSTM layers for temporal patterns learning, and dropout as well as batch normalization layers to avert overfitting. A dense layer is then used to output forecasted stock price values. XGBoost model is used to optimize the forecasts that occur due to the residuals

by the LSTM model. With the help of the XGBoost approach, it is finally possible to adjust the predictions by accounting for non-linear effects due to residuals minimization enhancing accuracy of model. To support the usage of the combined LSTM- XGBoost model, its performance is analyzed in relation to the traditional forecasting models using stock prices as a case study. The accuracy of the models will be a combined comparison of the following methods: MSE, RMSE, and MAE.

A. LSTM Model

LSTM stands for Long Short-Term Memory network, is it as RNN. It is used to learn, process and classify sequential data. Sentiment Analysis, language modeling, and speech recognition are some of the applications of the LSTM model. An important feature of the RNN is the hidden state. Some information about a sequence is stored in the hidden state. The hidden state is also known as the memory state because it remembers previous inputs.

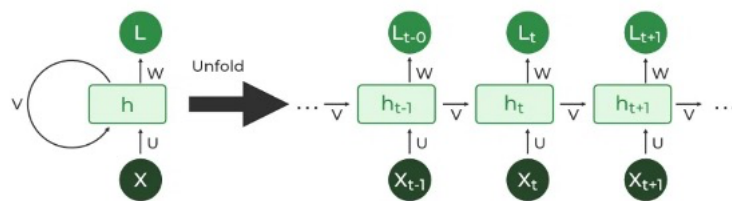


Figure 1. Long Short-Term Memory (LSTM) Network Architecture and Unfolded Recurrent Structure

An LSTM module (or cell) has 5 essential components which allow it to model both long-term and short-term data.

- Cell State (ct) – represents the internal memory of the cell
- Hidden State (ht) – Considering the current input, previous hidden state, and current cell state input, the output state information is calculated
- Input gate (it) – controls information flow from current input to cell state
- Forget gate (ft) – controls information flow from current input, previous cell state to current cell state
- Output gate (ot) – controls information flow from the current cell state to the hidden state

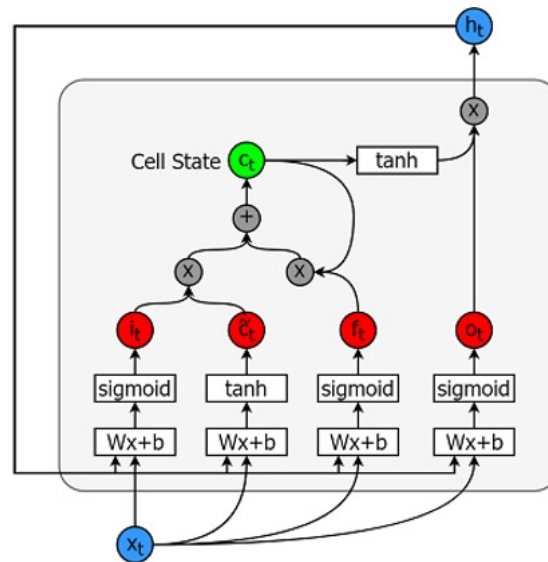


Figure 2. Internal Architecture of an LSTM Cell Showing Input, Forget, Cell State, and Output Gates

The mathematical equations of each state/gate are as follows:

- Forget Gate (f_t) =
- Input Gate (i_t) =
- Cell State (c_t) =
- Output Gate (o_t) =
- Hidden State (h_t) =

B. XGBoost Algorithm

XGBoost is an efficient implementation of gradient boosting algorithm, which is parallel and optimized for speed. It operates by aggregating the outputs of numerous weak predictors – usually decision stump – into a single strong predictor model. Each new tree in XGBoost tries to fix the mistakes or errors of previous trees especially on the samples where the model was not very good. XGBoost includes method such as the L1 and L2 regularization to reduce overfitting, as well as optimizations such as parallelization and tree pruning to improve computational time and increase scalability of the model. It is good at dealing with missing data, scaling for large data, and modeling high-degree nonlinearity, making itself widely used in multiple machine learning tasks, especially in the structured and tabular datasets.

IV. ANALYSIS

A. Data set

The comparison study is performed using the historical stock-market data of Alibaba (BABA) and the dataset used is from 1 January 2016 to 1 October 2021. The dataset is in time-series format where the stock-market observations are ordered in chronological order so that the prediction models can learn patterns from earlier observations of the market. Historical variables such as Open, High, Low, Close, Adjusted Close and Volume give information on daily market behaviour and are the basis for the development of the forecasting models. This dataset is also applicable for comparing the performance of the traditional LSTM model with

the suggested LSTM model with XGBoost, as both methods can be evaluated based on the same historical observations. The source material clearly states that the comparison is made with the Alibaba stock-market dataset and the indicated observation period.

The analysis is based on a sequential prediction approach, where previous data are used to discover patterns correlated with the future values of the stock price. The LSTM model is capable of dealing with such consecutive observations, and learning temporal connections through its memory and gating mechanisms. In the improved model, the residual errors learned from the original LSTM predictions by the LSTM model are further modified by the use of XGBoost. This results in a two-stage prediction process: XGBoost handles the remaining non-linear relations in the prediction errors, while LSTM learns the temporal patterns.

	Open	High	Low	Close	Adj Close	Volume
Date						
2016-01-04	78.180000	78.309998	75.180000	76.690002	76.690002	23066300
2016-01-05	77.919998	78.680000	77.260002	78.629997	78.629997	14258900
2016-01-06	77.120003	78.485001	76.970001	77.330002	77.330002	11569300
2016-01-07	73.290001	75.500000	71.540001	72.720001	72.720001	27288100
2016-01-08	74.330002	74.660004	70.669998	70.800003	70.800003	20814600
...	...	...	...	...	...	...
2021-09-24	147.710007	148.000000	144.570007	145.080002	145.080002	32067400
2021-09-27	144.919998	150.990005	144.440002	150.179993	150.179993	24287400
2021-09-28	152.160004	153.729996	148.860001	152.389999	152.389999	26299600
2021-09-29	150.460007	152.100006	147.479996	147.580002	147.580002	17361300
2021-09-30	147.029999	149.580002	146.863998	148.050003	148.050003	14291500

Figure 3. Alibaba Stock Market Dataset Showing Historical Open, High, Low, Close, Adjusted Close, and Volume Values

B. Accuracy

The prediction models are evaluated based on ****Mean Squared Error (MSE)**, **Mean Absolute Error (MAE)** and **R² (coefficient of determination)****. These measures provide distinct viewpoints of the accuracy of prediction and enable the performance of the base LSTM model to be compared with the LSTM-XGBoost model.

1. Mean Squared Error (MSE)

MSE is mean squared error. It is average of squares of errors. Because the prediction errors are squared, higher errors are weighted more heavily in the metric. Hence, lower MSE means that predicted values are closer to real values. This statistic is very useful for detecting models that tend to create quite big prediction errors. MSE is utilised in this analysis to check the added XGBoost would minimise the prediction error made by the LSTM model or not.

The accuracy of the models was checked using four metrics:

1. **MSE: Mean Squared Error** – measures how close a regression line is to a set of data points. The lower the values of MSE the closer the predictions to actual values.
2. **MAE: measures the average absolute difference** between the actual values and the predicted values. It is less sensitive to outliers than MSE because it doesn't square the errors.
3. **R2: measures how well the model's predictions fit the actual data.** It ranges from $-\infty$ to 1:

$R^2 = 1$: Perfect fit
 $R^2 = 0$: Predictions are no better than a horizontal line (mean of actual values)
 $R^2 < 0$: Predictions are worse than just predicting the mean

Figure 4. Interpretation of R^2 (Coefficient of Determination) Values

2. Mean Absolute Error (MAE)

MAE calculates the average absolute deviation between the actual stock values and the anticipated values by the model. MAE is not squared and hence is less sensitive to single large deviations as opposed to MSE. Lower MAE means actual observations and predictions are nearer. Complementing the MAE with the MSE measure gives a complementary evaluation of the prediction performance and assists in establishing whether the hybrid model produces consistently fewer forecasting errors.

3. R^2 (Coefficient of Determination)

R^2 measures how well the actual observations are fitted by the model's predictions. A score closer to 1 suggests a better fit between the anticipated and observed values, whereas a negative value shows that the model performs worse than a simple forecast based on the mean of the observed data. In this investigation the overall predictive performance of the two models is evaluated using R^2 , MSE and MAE.

The obtained data reveal a clear difference between the two approaches. The LSTM model achieved MSE = 0.1219, MAE = 0.3163, $R^2 = -4.1150$. The LSTM-XGBoost model obtained MSE of 0.0001, MAE of 0.0050, and R^2 of 0.9972 for comparison. Thus, for the presented test results, the hybrid model provided significantly lower values of MSE and MAE and a significantly higher value of R^2 compared to the standard LSTM model.

C. Plots before and after using XGBoost Algorithm for LSTM Model

Graphical analysis gives a visual comparison of the actual observations of stock-prices and the forecasts made by the models. The graphic presented in the paper displays the historical movement of the stock-price along with the anticipated values, which allows visual examination of the consistency between the model predictions and actual observations. Such a comparison is important since the numerical metrics summarise the prediction errors while the graphs illustrate how close the predicted series matches the movement of the real stock-price series over time.

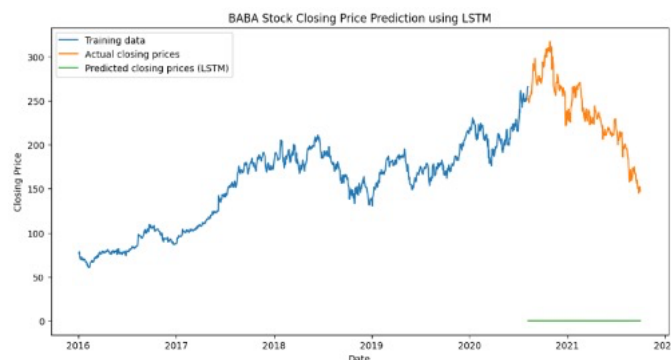


Figure 5. Alibaba Stock Closing Price Prediction Using the LSTM Model

The LSTM prediction plot depicts how the model attempts to mimic the underlying behaviour of the Alibaba stock price during the observation period. The given error metrics also show

that the base LSTM model has a significant prediction error, especially in terms of the MSE of 0.1219, MAE of 0.3163 and negative R^2 of -4.1150 . The prediction performance changes significantly after the use of XGBoost. The XGBoost part learns from the errors or residual information of the LSTM prediction to correct the forecast and capture extra non-linear effects.

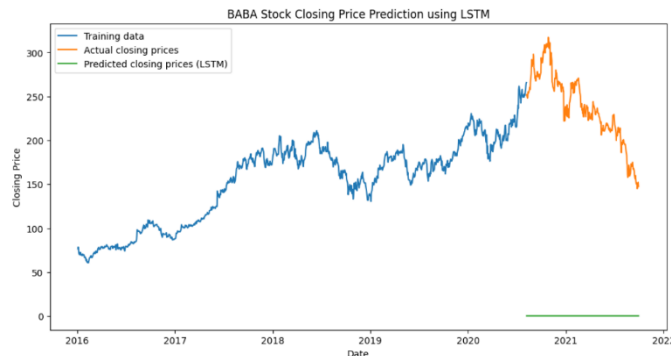


Figure 6. Alibaba Stock Closing Price Prediction Using the LSTM Model

Briefly, the investigation reveals that the LSTM-XGBoost hybrid framework yields much improved reported prediction metrics over the standalone LSTM model for the Alibaba dataset. This combination enables LSTM to cope with the temporal features of stock-price data and XGBoost to provide an extra learning phase for non-linear residual patterns. The given graphical comparison and assessment metrics jointly provide proof of the difference in the prediction performance of the two methodologies.

V. RESULTS

The difference in the predictions of both models is seen greatly after the predictions and the actual values were plotted.

1. LSTM MODEL Accuracy rates

```
LSTM Model Metrics:
MSE: 0.1219
MAE: 0.3163
R2: -4.1150
```

2. LSTM MODEL WITH XGBOOST ALGORITHM Accuracy rates

```
Boosted LSTM Model Metrics:
MSE: 0.0001
MAE: 0.0050
R2: 0.9972
```

VI. CONCLUSION

This work studied the efficiency of a hybrid LSTM-XGBoost model for stock-market prediction using historical Alibaba stock-market data from 1 January 2016 to 1 October 2021. The investigation examined the predictive capacity of a stand-alone LSTM model with an LSTM model using the XGBoost algorithm. We used LSTM to understand the temporal dependencies and sequential patterns in the historical stock-market observations and XGBoost

to refine the predictions of LSTM by learning from the residual errors and non-linear correlations.

The results reported show a significant improvement after combining XGBoost with LSTM. The MSE, MAE and R2 of the LSTM model alone were 0.1219, 0.3163, and -4.1150, respectively, while the boosted LSTM model attained an MSE of 0.0001, MAE of 0.0050, and R2 of 0.9972. These results show that given the dataset and evaluation reported, the hybrid strategy yielded significantly fewer prediction errors and a better fit to the observed values. Therefore, the combination of LSTM with XGBoost provides a suitable framework for increasing the stock-market prediction performance. Future studies can expand the period of the dataset and the training observations to test the robustness of the approach to diverse market circumstances.

VII. FUTURE ENHANCEMENTS

The accuracy of the LSTM models can be increased increasing the time period of the dataset and increasing the training sets for the model to get a better understanding of the dataset.

VIII. REFERENCES

References

1. Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2019). A comparative analysis of XGBoost. arXiv. <https://doi.org/10.48550/arXiv.1911.01914> (arXiv)
2. Mekruksavanich, S., Jantawong, P., & Jitpattanakul, A. (2022). LSTM-XGB: A new deep learning model for human activity recognition based on LSTM and XGBoost. In 2022 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON). IEEE. <https://doi.org/10.1109/ECTIDAMTNCON53731.2022.9720409> (ResearchGate)
3. Mustapha, I. B., & Saeed, F. (2016). Bioactive molecule prediction using extreme gradient boosting. *Molecules*, 21(8), 983. <https://doi.org/10.3390/molecules21080983> (PubMed Central (PMC))
4. Amthul Thawab, M., Zubaidunisa, S. M., & Shariff, I. Y. (2022). Impact of internationalization on brand building and business expansion. *AIP Conference Proceedings*, 2393(1), 020061.
5. Sanboon, T., Keatruangkamala, K., & Jaiyen, S. (2019). A deep learning model for predicting buy and sell recommendations in Stock Exchange of Thailand using long short-term memory. In 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS) (pp. 757–760). IEEE. <https://doi.org/10.1109/CCOMS.2019.8821776> (ResearchGate)
6. Tiwari, S., Bharadwaj, A., & Gupta, S. (2017). Stock price prediction using data analytics. In 2017 International Conference on Advances in Computing, Communication and Control (ICAC3) (pp. 1–5). IEEE. <https://doi.org/10.1109/ICAC3.2017.8318783> (ResearchGate)