STOCK MARKET PREDICTION USING DEEP LEARNING

Presentation by

SAYAN GHOSH
Roll No: VU/PG/22/26/02-IVS-0018

SOURAV BERA
Roll No: VU/PG/22/26/02-IVS-0005

AHANA GHOSH
Roll No: VU/PG/22/26/02-IVS-0031

Under the guidance of

Mr. Bachchu Paul

Assistant Professor

Department of Computer Science

Vidyasagar University

Department of Computer Science, Vidyasagar University
Midnapore, West Bengal, 721102

Date of Submission- 22/07/2024

CONTENTS

INTRODUCTION

- The stock market is a structured marketplace where participants buy and sell securities like stocks.
- Stock market prediction uses analytical methods, including machine learning and technical analysis, to forecast future price movements and optimize investment strategies.
- Our project develops an LSTM-based model to predict Tesla's stock prices, utilizing historical data such as opening, closing, highest, lowest prices, and trading volume. Performance is evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE) and R-squared (R2) comparing results with traditional forecasting methods.

PROBLEM STATEMENT

Our objective is to develop an LSTM neural network system to accurately predict Tesla's next-day closing stock prices. By leveraging LSTM's ability to capture long-term dependencies in sequential data, we aim to improve prediction accuracy for Tesla's highly volatile stock. Our model is trained on historical data, with extensive preprocessing to uncover market patterns. This project seeks to provide actionable insights for investors, enhancing decision-making and risk management while contributing to the field of stock market prediction.

LITERATURE REVIEW

1) <u>Study on the prediction of stock price based on Machine Learning:</u>

- Mehtab and Sen[1] combined text mining and natural language processing with machine learning models like regression and classification to predict NIFTY index values, achieving high accuracy using a self-organizing fuzzy neural network (SOFNN).
- Kara et al.[2] employed support vector machines (SVM) and artificial neural networks to predict the Istanbul Stock Exchange National 100 Index, finding that artificial neural networks outperformed SVMs significantly.
- P. Mondal et al.[3] utilized an ARIMA model to predict share prices from data provided by the National Stock Exchange (NSE) of India, achieving over 85% accuracy across multiple industries.

LITERATURE REVIEW

2) Study on the prediction of stock price based on Deep Learning:

- Kilimci et al.[4] developed a model combining word embeddings (Word2Vec, FastText, GloVe) with deep learning techniques (CNN, RNN, LSTM) for predicting BIST 100 index direction. The Word2Vec and LSTM combination achieved the highest accuracy using Twitter data.
- Samarawickrama and Fernando used FFNN, SRNN, LSTM, and GRU to forecast Colombo Stock Exchange prices. FFNN achieved approximately 99% accuracy, while SRNN and LSTM had lower error rates but occasionally high errors, and GRU had consistently high error rates[5].
- Gao et al.[6] compared Multilayer Perceptron, LSTM, CNN, and Uncertainty Aware Attention for predicting next-day stock prices of the S&P 500, CSI 300, and Nikkei 225 indices. Uncertainty Aware Attention slightly outperformed the other models.

LITERATURE REVIEW

3) <u>Study on the prediction of stock price based on Convolutional Neural Networks :</u>

- Kusuma RMI et al.[7] applied CNN for predicting stock market movements, achieving over 90% accuracy using historical data from Yahoo Finance API, with sensitivity and accuracy as performance measures.

- Hoseinzade E et al. [8] CNNPred, a CNN-based framework, was used to predict movements in S&P 500, NASDAQ, NYSE, Dow Jones, and Russell indices, outperforming existing algorithms with F-measure as the evaluation metric.

- Another study on CNN for stock market movement prediction also reported over 90% accuracy, using historical data from Yahoo Finance API and measuring performance with sensitivity and accuracy[9].

LITERATURE REVIEW

4) <u>Study on the prediction of stock price based on Long Short-Term Memory networks:</u>

- B. Paul et al.[10] has conducted Voice-Based Railway Station Identification using Long Short Term Memory approach.
- Chen et al.[11] used the LSTM model to predict China stock returns, improving accuracy from 14.3% to 27.2% compared to random predictions by transforming historical data into 30-day sequences with ten features and 3-day labeling.
- Kim et al.[12] developed a feature fusion LSTM-CNN model, combining CNN for candlestick chart analysis and LSTM for historical price data, which outperformed individual models with lower prediction error on S&P 500 ETF data.

TOOLS & TECHNOLOGY

Jupyter Lab

Python

Microsoft Excel

Kaggle

NumPy

Pandas

Matplotlib

Keras

TensorFlow

Streamlit

METHODOLOGY

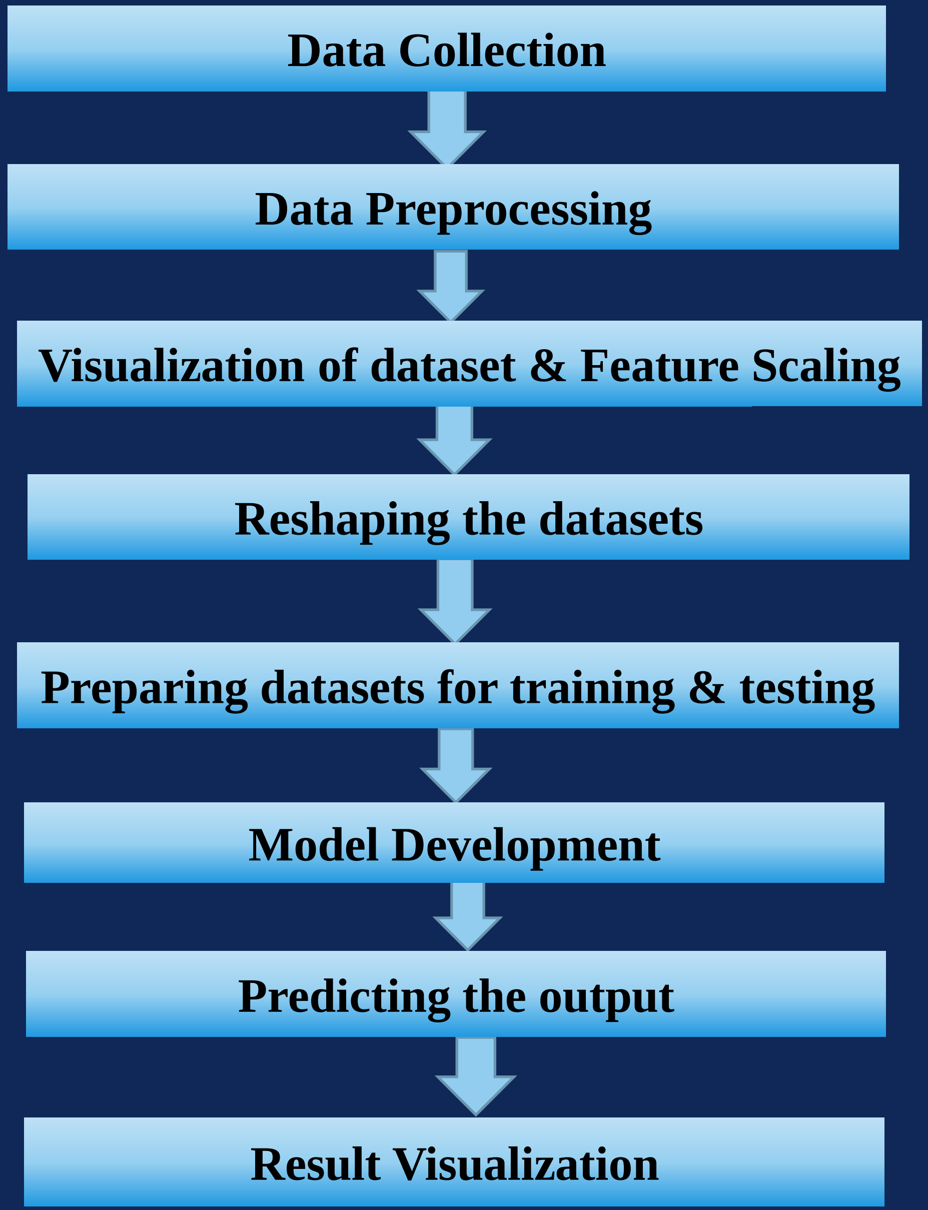


Fig: System Flow Diagram

DATA COLLECTION

Date	Open	High	Low	Close	Adj Close	Volume
2018-06-22	23.436001	23.483334	22.133333	22.242001	22.242001	**153991500**
2018-06-25	22.007999	22.564667	21.833332	22.200666	22.200666	**103969500**
2018-06-26	22.403334	22.903334	21.719999	22.799999	22.799999	**111787500**
2018-06-27	23.000000	23.386000	22.633333	22.966667	22.966667	**125005500**
2018-06-28	23.243999	23.801332	23.073999	23.328667	23.328667	**125970000**
2018-06-29	23.555332	23.590668	22.827333	22.863333	22.863333	**97386000**

- **Stock name: Tesla**
- **Data source: www.kaggle.com**
- **Data coverage: 5 years [2018-06-22 to 2023-06-21]**
- **Data details: 1257 rows and 7 columns**

DATA PREPROCESSING

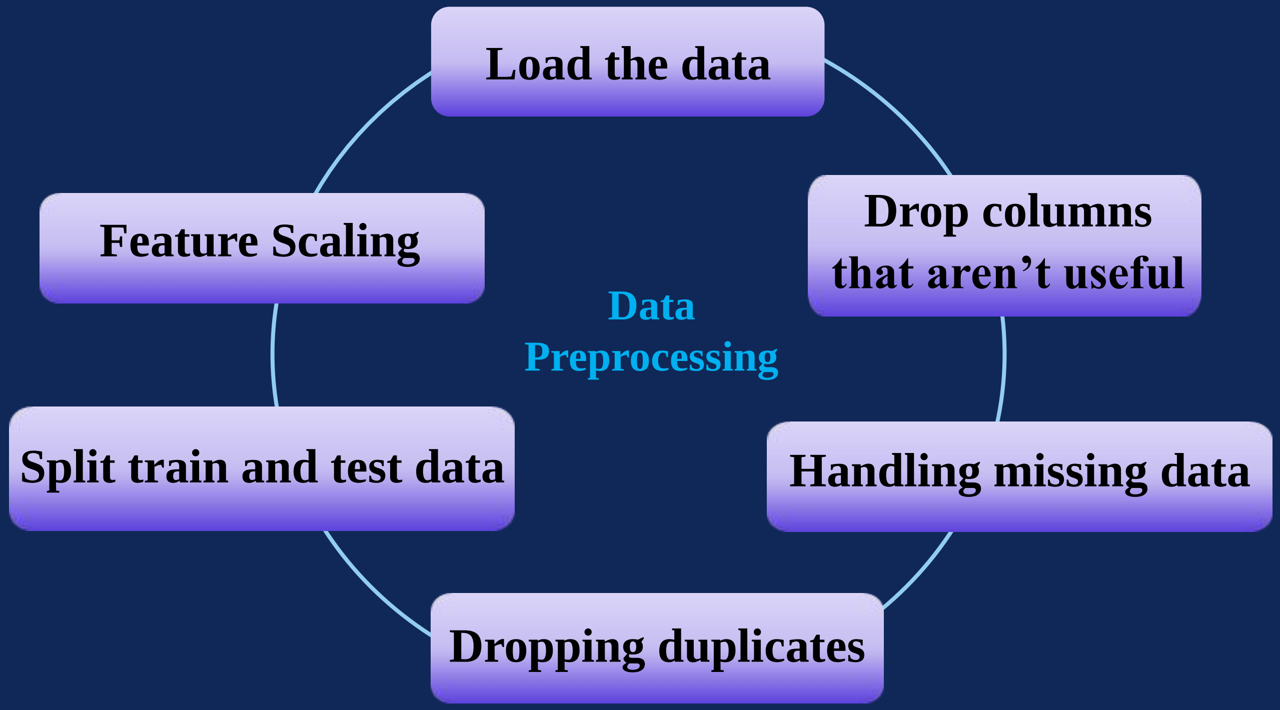


Fig: Data Preprocessing

MODEL DEVELOPMENT

- **Deep learning**, a powerful subset of machine learning, uses multi-layered neural networks to learn from large datasets for complex tasks.
- **Long Short-Term Memory (LSTM)** networks, an advanced RNN type, effective for time series forecasting and language modeling.

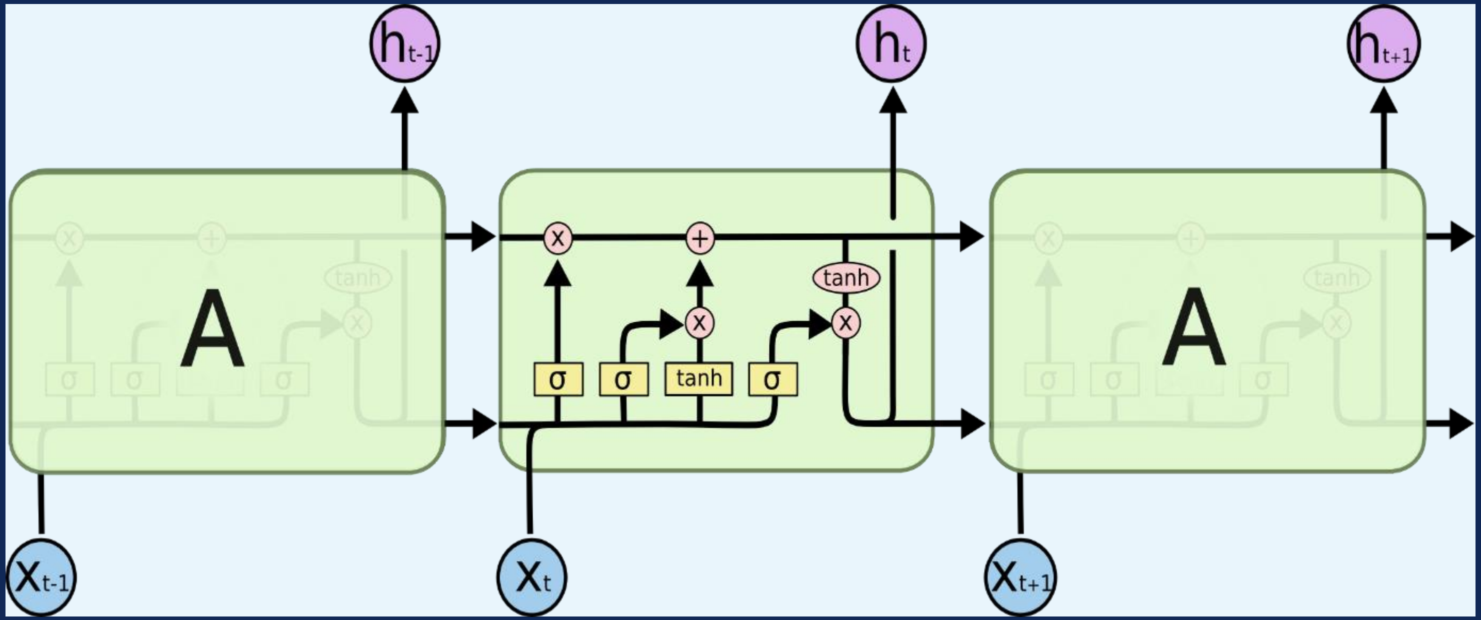
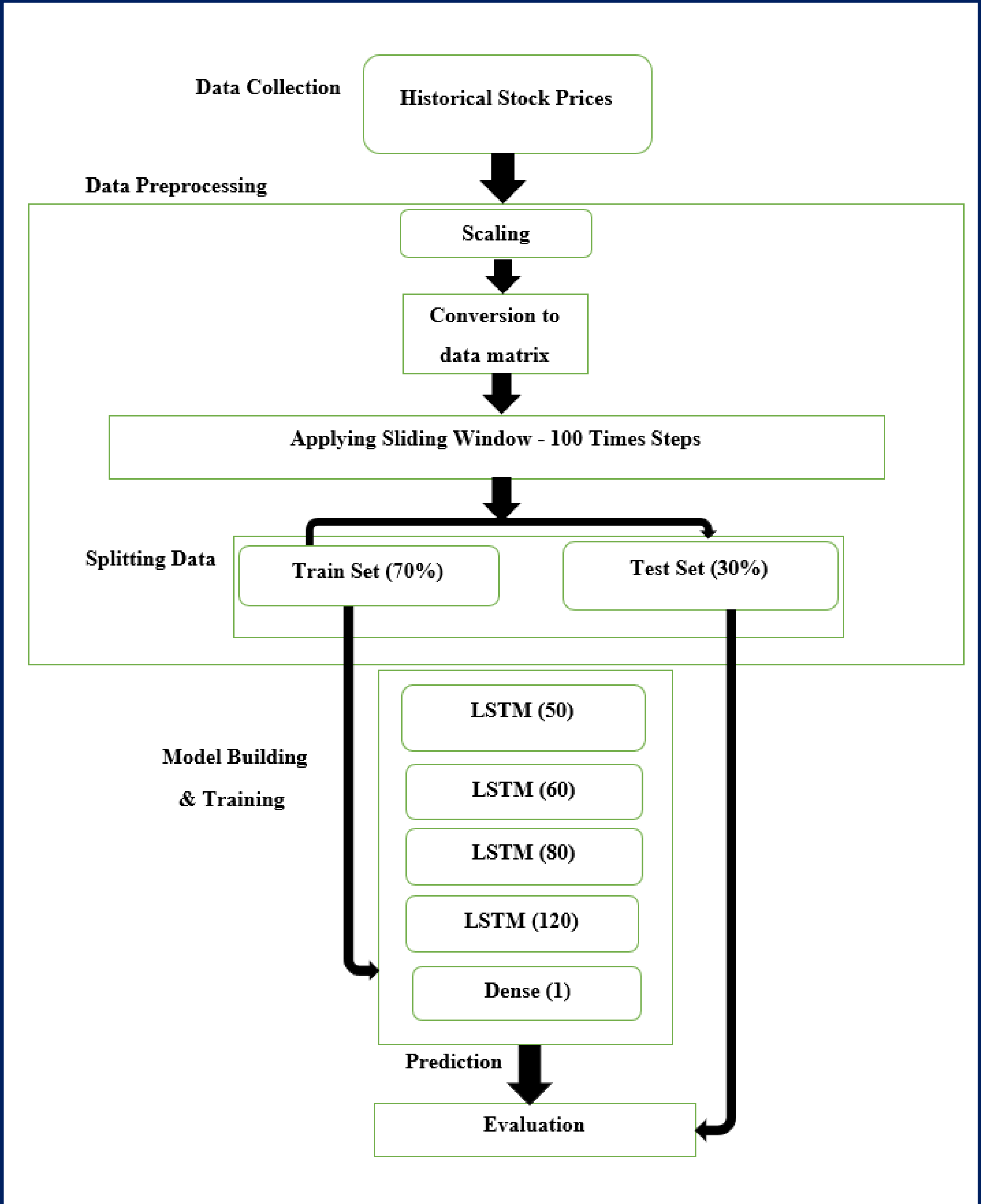


Fig: Long Short-Term Memory Architecture

MODEL DESIGN



MODEL SUMMARY

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 100, 50)	10,400
dropout (Dropout)	(None, 100, 50)	0
lstm_1 (LSTM)	(None, 100, 60)	26,640
dropout_1 (Dropout)	(None, 100, 60)	0
lstm_2 (LSTM)	(None, 100, 80)	45,120
dropout_2 (Dropout)	(None, 100, 80)	0
lstm_3 (LSTM)	(None, 120)	96,480
dropout_3 (Dropout)	(None, 120)	0
dense (Dense)	(None, 1)	121

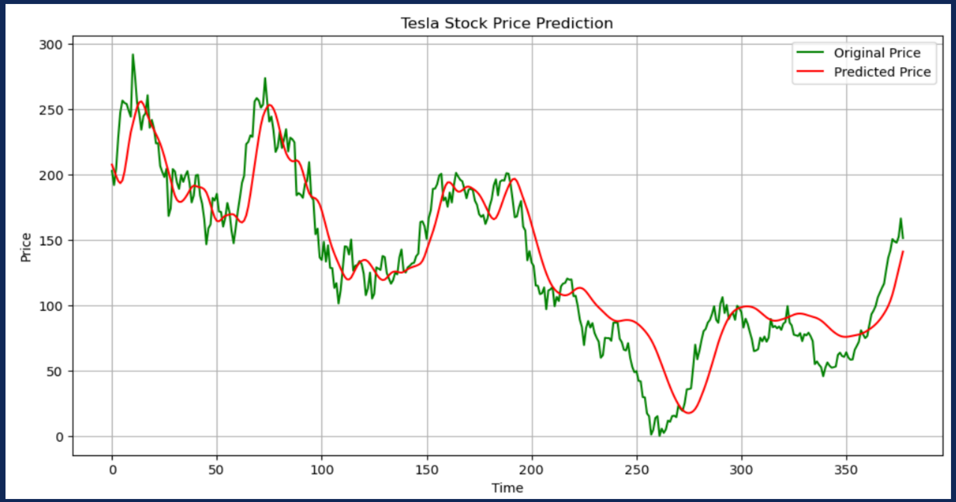
Total params: 178,761 (698.29 KB)

Trainable params: 178,761 (698.29 KB)

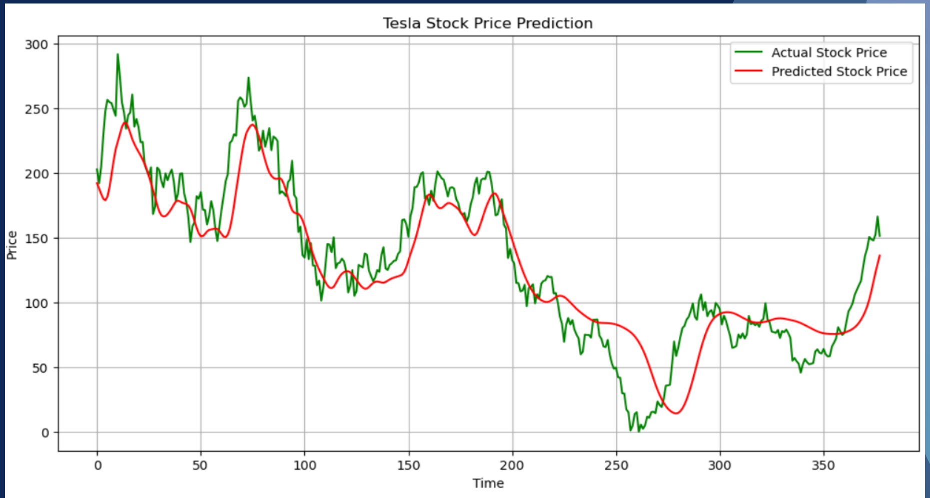
Non-trainable params: 0 (0.00 B)

Fig: Model Summary

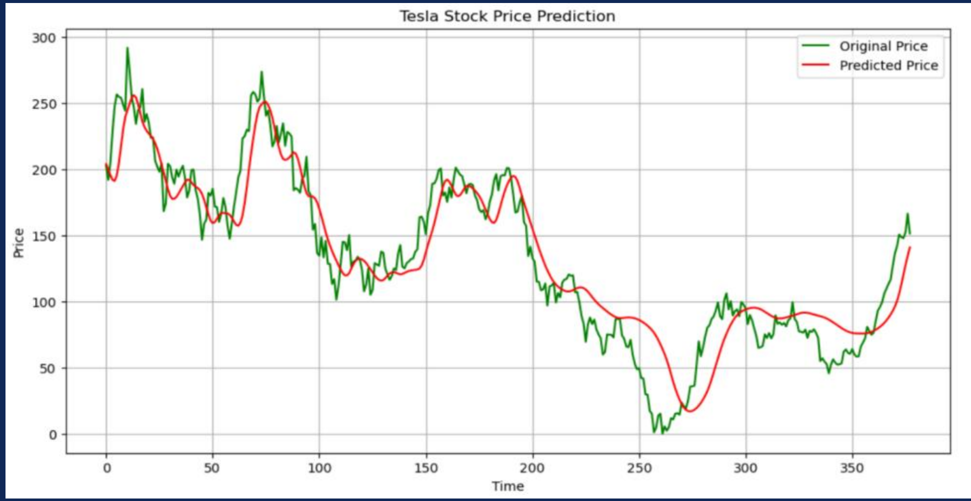
RESULT & DISCUSSION



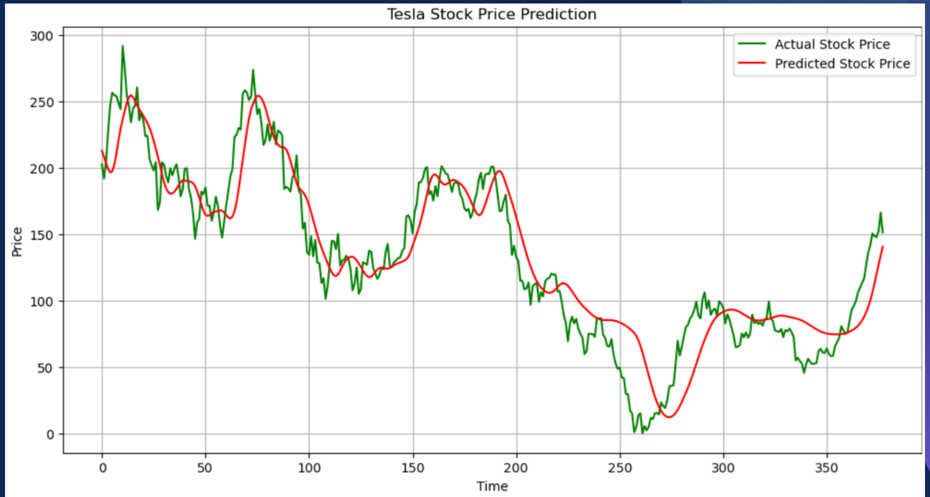
Original vs Predicted (40)



Original vs Predicted (60)



Original vs Predicted (80)



Original vs Predicted (100)

PERFORMANCE MATRICS

SEQUENCE LENGTH	MAE	MSE	R2 SCORE
40	6.10%	605.68	86%
60	8.94%	622.70	85%
80	10.50%	470.31	89%
100	14.60%	597.37	86%

Optimal Sequence Length: 80

- Mean Squared Error (MSE): 470.31 (lowest among tested sequence lengths)
- R2 Score: 0.89 (highest among tested sequence lengths)
- Mean Absolute Error (MAE): 10.50%
- Conclusion: Despite the MAE, the lowest MSE and highest R2 Score indicate that a sequence length of 80 offers the best balance of accuracy and model fit, providing consistent and accurate stock market predictions.

RESULT & DISCUSSION

MAE	7.89%
MSE	31.96
RMSE	81.66
R2 SCORE	94%

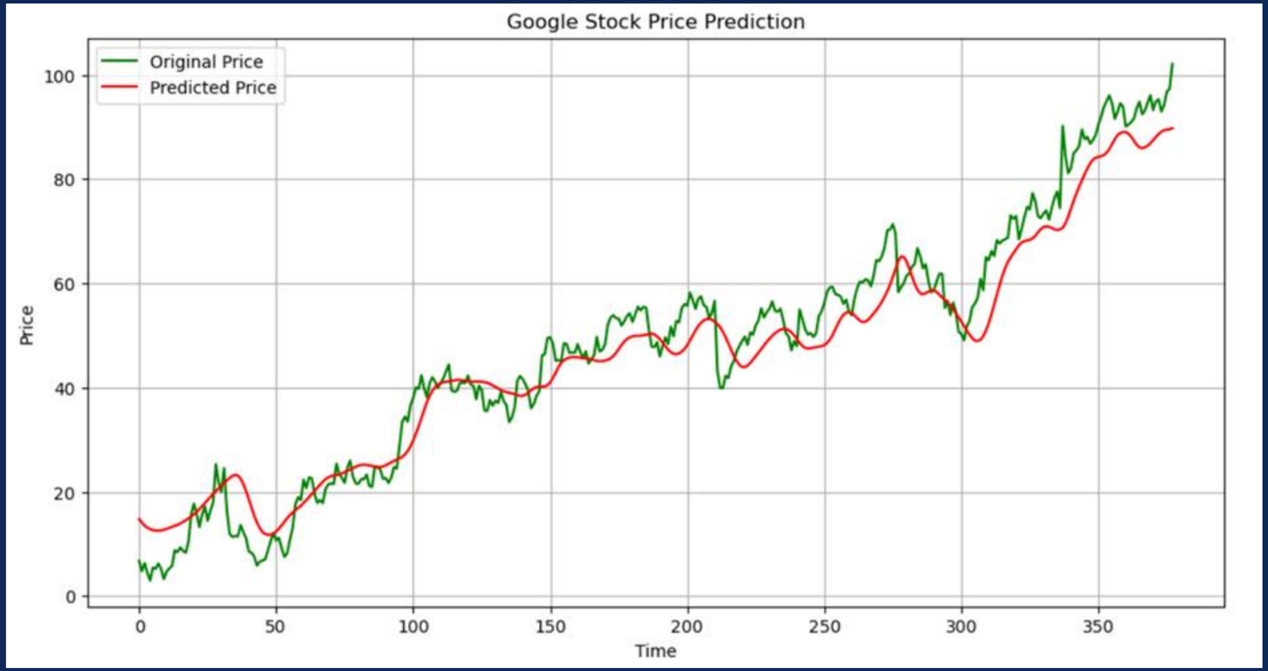


Fig: Prediction on Google Dataset

- Dataset: Google dataset
- Training: 100 epochs, 32 batch size
- Metrics: MAE, MSE, RMSE & R2 Score
- Conclusion: Demonstrated reliability and accuracy, effective for stock market forecasting

RESULT & DISCUSSION

MAE	129.93
MSE	9.43
R2 SCORE	92%

- Model Type: CNN-LSTM
- Dataset: Tesla
- Look-back Period: 60 days
- Layers: Conv1D, MaxPooling1D, LSTM, Dropout, Dense
- Training Duration: 50 epochs

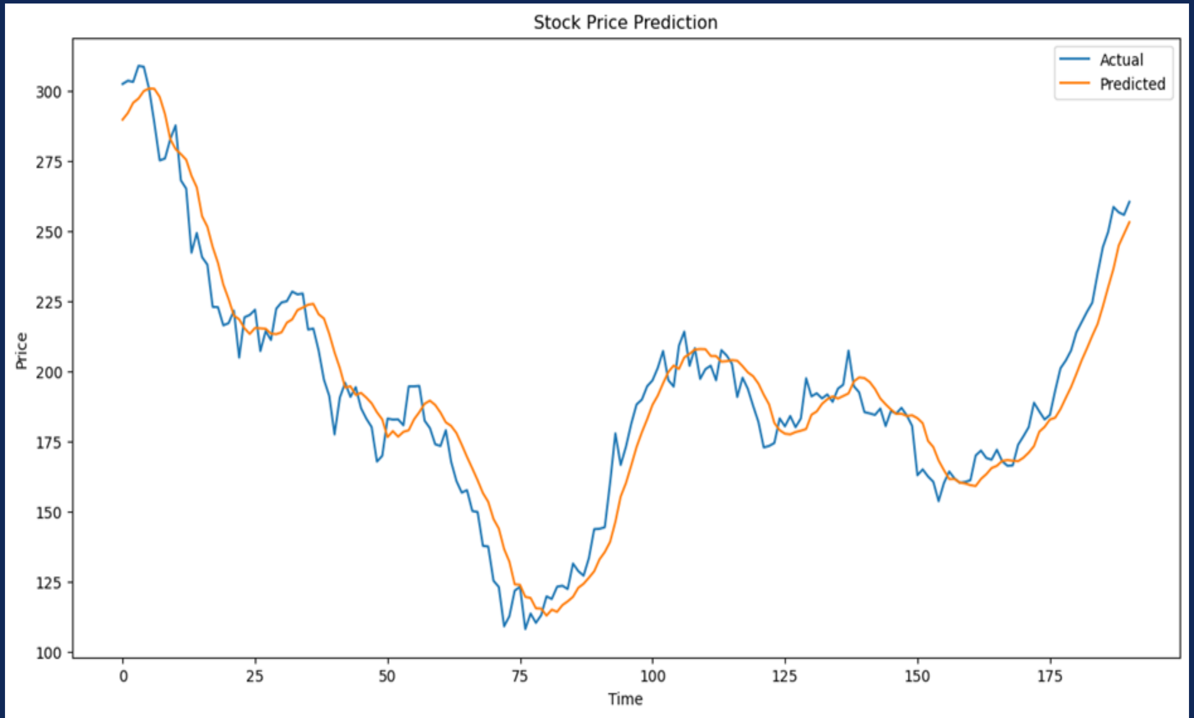
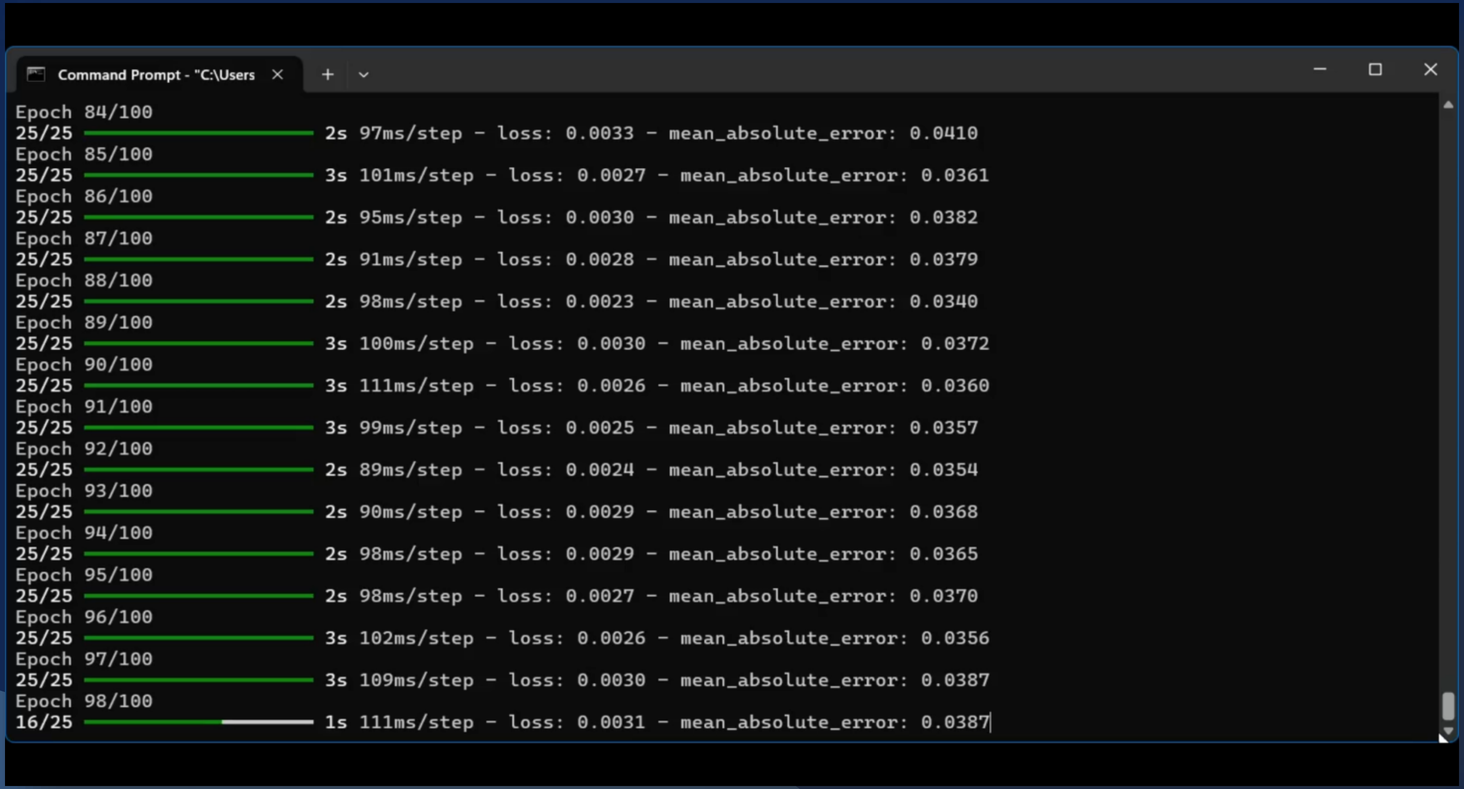


Fig: Prediction using CNN-LSTM model

FUTURE PREDICTION

- **Streamlit App Development**

We developed a Streamlit app for stock price forecasting with an intuitive user interface and interactive widgets for data input and model selection. Utilizing Plotly for data visualization, users can explore forecasted prices through interactive plots. After development, the app was deployed to a hosting platform, and rigorous testing ensured smooth functionality across devices and browsers. Stress testing was also conducted to evaluate performance under varying user traffic levels.



CONCLUSION

We applied LSTM networks to predict Tesla stock prices, demonstrating their effectiveness in capturing complex patterns and making accurate predictions. These findings are significant for both academia and the finance industry, suggesting further exploration in financial analysis. Our model's insights can help investors make informed decisions and improve strategies. Future work could enhance accuracy with additional features and hyperparameter tuning, highlighting LSTM's potential in stock market prediction.

FUTURE SCOPE

Future enhancements to our model will include:

- **Develop Hybrid Predictive Models**

  Combine LSTM with other neural network architectures

  Implement hybrid optimization algorithms to improve accuracy

- **Extend Sentiment Analysis**

  Use data from platforms like Facebook and Twitter

  Gauge market sentiment on stock price changes

ACKNOWLDEMENTS

We would like to thank our supervisor, **Mr. Bachchu Paul**, for his invaluable guidance, expertise, and continuous support. His insightful feedback, patience, and encouragement have been crucial in shaping this project and expanding our understanding of the subject matter. Also we would like to thank our family, friends and classmates for their unwavering support and encouragement throughout this project. Their support and motivation have made this journey more enjoyable and fulfilling for us.

REFERENCES

[1] Mehtab, S. and Sen, J.: A Robust Predictive Model for Stock Price Prediction Using Deep Learning and Natural Language Processing. In: Proc. of the 7th Int. Conf. on Business Analytics and Intelligence, Bangalore, India, December 5 – 7 (2019).

[2] Kara, Y., Acar Boyacioglu, M., & Ömer Kaan Baykan (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul stock exchange. Expert Systems with Applications, 38(5), 5311–5319.

[3] P. Mondal, L. Shit, S. Goswami, Study of effectiveness of time series modeling (ARIMA) in forecasting stock prices, Int. J. Comput. Sci. Eng. Appl. 4 (2) (2014) 13.

[4] Kilimci, Z. H., & Akyokus, S. The analysis of text categorization represented with word embeddings using homogeneous classifiers. 2019 IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA). IEEE. (2019).

[5] Samarawickrama, A.J.P.; Fernando, T.G.I. A Recurrent Neural Network Approach in Predicting Daily Stock Prices. In Proceedings of the 2017 IEEE International Conference on Industrial and Information Systems (ICIIS), Peradeniya, Sri Lanka, 15–16 December 2017; pp.1–6.

[6] Gao, P., Zhang, R., & Yang, X. (2020). The application of stock index price prediction with neural network. Mathematical and Computational Applications, 25(3), 53.

[7] Kusuma RMI, Ho TT, Kao WC, Ou YY and Hua KL, "Using deep learning neural networks and candlestick chart representation to predict stock market," arXiv preprint arXiv:1903.12258, 2019.

[8] Hoseinzade E, Haratizadeh S (2019) CNNpred: CNN-based stock market prediction using a diverse set of variables. Expert Syst Appl 129:273–285.

[9] Kusuma RMI, Ho TT, Kao WC, Ou YY and Hua KL, "Using deep learning neural networks and candlestick chart representation to predict stock market," arXiv preprint arXiv:1903.12258, 2019.

[10] Bachchu Paul, Somnath Bera, Tanushree Dey, and Santanu Phadikar. "Voice-Based Railway Station Identification using LSTM approach" In 2021, Proceedings of International Conference on Frontiers in Computing and Systems, Advances in Intelligent Systems and Computing 1255,

[11] Chen, K., Zhou, Y., & Dai, F. (2015). A LSTM-based method for stock returns prediction: A case study of China stock market. In 2015 IEEE international conference on big data (Big Data) (pp. 2823–2824).

[12] T. Kim and H. Y. Kim, "Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data," PloS one, vol. 14, no. 2, p. e0212320, April 2019.

THANK YOU