Stock Sense: Deep Learning based stock-market prediction tool

Meet Manoj Agarwal [1], Badrul Hasan TS [2], Shourya Gupta [3], Anmol Kaul [4]

[1] *Student, Computer Science and Engineering, Alliance University, Karnataka, India*
[2] *Student, Computer Science and Engineering, Alliance University, Karnataka, India*
[3] *Student, Computer Science and Engineering, Alliance University, Karnataka, India*
[4] *Student, Computer Science and Engineering, Alliance University, Karnataka, India*

ABSTRACT

In the past few years, many financial forecasting methods have employed deep learning techniques to learn more complex, non-linear temporal relationships in time-series data. This project involves predicting stock prices utilizing two sophisticated Recurrent Neural Network (RNN) structures: Long Short-Term Memory (LSTMs) and Bidirectional Long Short-Term Memory (Bi-LSTMs). This study uses historical stock market data provided by Yahoo Finance to create models that analyze stock market prices to project forward in time stock prices as well as determine the comparison between the two models.

The four major modules are: data collection, data preprocessing, model training and evaluation, and results visualization. The data collection module collects and extracts structured stock data including the following Open, High, Low, Close, and Volume values. The data preprocessing module takes the raw data and prepares it to be cleaned and normalized and assigns agglomeration of new processed data with several technical indicators (derived from financial data), which may capture the underlying market trends. The financial time-series structure must stay intact and having sliding windows allows for creating datasets a model can accept.

The core modelling phase involves creating and training LSTM and Bi-LSTM networks with TensorFlow/Keras. The measure of performance of LSTM and Bi-LSTM networks includes measuring Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R^2 Score. The Bi-LSTM exhibited superior accuracy (or perhaps better generalisation) than the uni-directional LSTM networks since the Bi LSTM requires the data to be shaped in both the forward and backward direction.

The final module is the visualization phase in which actual vs predicted price comparisons, training-validation loss curves, and technical indicators have been overlaid graphically to provide intuitive insights into the performance of the model. The findings suggest that Bi-LSTM consistently outperformed LSTM in terms of temporal dependencies associated with financial data; in this context, bidirectional architecture was useful for improving predictive accuracy. This project shows the practical feasibility of applying deep learning models for predicting financial data and sets the stage for even more advanced hybrid models that include additional market indicators and sentiment analysis in future refinements.

Keyword : - *Recurrent Neural Network, Long Short-Term Memory, Bidirectional Long Short-Term Memory, Root Mean Squared Error.*

---

1. Introduction

Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work[1]. At its essence, the stock market is a complex and constantly changing environment in which an individual buys and sells ownership shares in publicly-traded companies [2]. The stock market is a critical mechanism to help companies to raise capital to attract new investors for their growth, research, capital expenditures, and operating cash-flow requirements[3]. The institution of the stock market creates marketable ownership interests, or shares, that companies sell in the public marketplace. Along with the medium of exchange, the stock market offers the company, and investors, the opportunity to realize stock appreciation in value over time, and in some cases, receive dividends from profits the company distributes to shareholders.

Envision the stock market as a marketplace: instead of trading physical goods, shares of businesses (fractions of ownership) are being traded. The marketplace consists of exchanges, such as the Bombay Stock Exchange (BSE) and the National Stock Exchange of India (NSE) here in India. Exchanges are organized platforms that facilitate the buy and sell of stocks; they provide an orderly, regulated, and transparent environment for buyers and sellers to interact.

1.1 Introduction to LSTM

Recurrent Neural Networks (RNNs) follow a sequential data processing model, maintaining an internal state that recalls previous element inputs. The capability to inherently recall previous elements makes RNNs ideal for order and time-sensitive applications, specifically in natural language and time series analysis. However, RNNs face a fundamental problem: "vanishing gradients." The longer the input sequence, continuing to the end, the exponentially smaller the gradients will be on the way to optimizing a set of neural network weights contributory to supervised training to the extent the network itself learning to remember longer dependencies is eliminated.

To overcome this limitation, Hochreiter and Schmid Huber presented the LSTM (Long Short-Term Memory) network in 1997. LSTM is a specific architecture of RNN that was devised explicitly to remember and learn long-term information [10]. The LSTM's major advantage is the memory cell, which serves as an information highway where pertinent information travels through the network with little to no change, and an availability of forgetting any non-pertinent information.

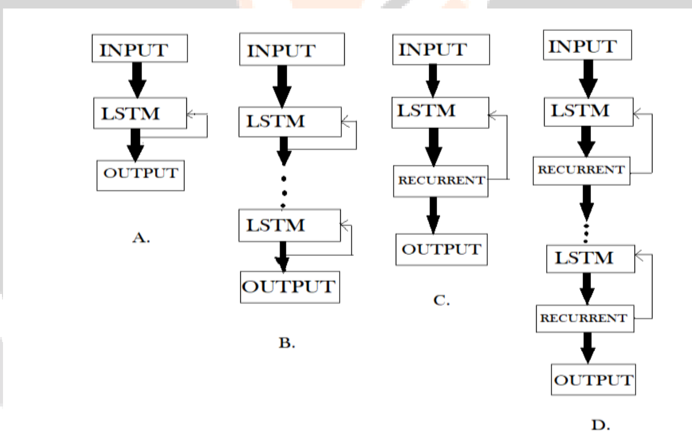


Fig-1: variations in LSTM networks

1.2 Introduction to Bi-LSTM

Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work Introduction related your research work [7]. Traditional Long Short-Term Memory (LSTM) networks process sequential data in a uni-directional form, meaning from past to future; yet, in many cases we find that the context of a certain point in the sequence, depends not just on what came before it but also on what comes after. The clearest example of this would be in natural language processing, where knowing the meaning of a word often requires knowing the words that come both before and after it.

In order to record such dependencies, the Bidirectional Long Short-Term Memory (BiLSTM) network encounters such dependencies, the Bidirectional Long Short-Term Memory (BiLSTM) network was created. A BiLSTM can be thought of simply as two LSTMs: the first processing the input sequence (in the forward direction of the sequence: beginning to end) and the other processing the sequence in the backward direction (end to beginning).

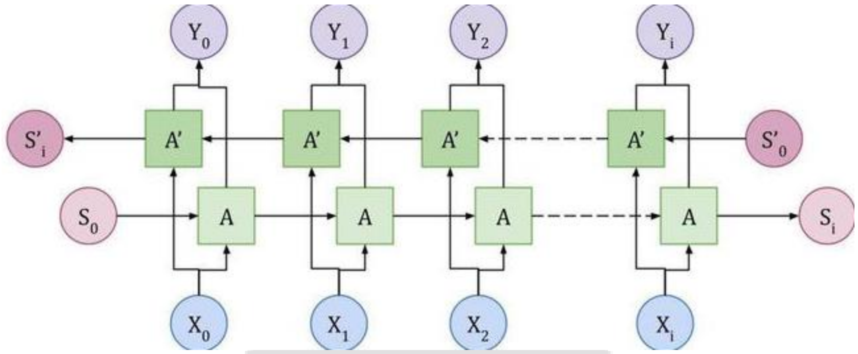


Fig-2: Architecture of BiLSTM

2. Methodology

The outlined system is divided into four distinct modules, all relevant to the construction of a reliable stock price prediction architecture using LSTM and Bi-LSTM models. The module structure allows developers to manage, test and improve upon each distinct component of the pipeline without disrupting the operational flow of the entire pipeline [8]. It also allows some variation from the base architecture, for example upgrades by building in the use of further indicator definitions, real-time APIs, or deployment capabilities.

2.1 Overview of the modules

- **Data Collection:**
  This module is the first component of the architecture because robust deep learning models rely on historical stock data that are accurate, relevant and fresh enough to be used in the training process. The module automates the task of pulling stock market data from Yahoo finance, by using the yfinance API, which cuts down on human error risk and ensures the data are clean & consistent as well as regularly updated.

- **Data Preprocessing:**
  The quality of the input data that impacts prediction performance. This module processes the source historical data from Yahoo finance, and generates clean and represented data forms suitable for time series forecasting. It forwards missing values, feature normalizes the variables other than stock price returns, and generates technical indicators (like moving averages of stock prices, for example) that capture trends in stock prices and investor behaviour. Technical indicators are treated as additional input features that improve model learning.

- **Model Creation, Training, Testing, Evaluation:**
  This is the most critical part of the system where all predictive modelling is being done [2]. This module implements both LSTM and Bi-LSTM neural networks together to capture and identify temporal patterns in stock prices. The module contains dataset slicing, hyperparameter selection, model training, and validation of model accuracy. It generates performance metrics for model comparison and allows for production outputs for presenting model information.

- **Visualization:**
  A major component of interpreting model information is visualization. This module will visualize all the various graphical forms of both the data and the outcome from the models. It will visualize stock price trends, versus model predictions versus actual values, and training loss graphs, along with LSTM performance versus Bi-LSTM performance comparisons [7]. This ultimately complements our understanding of the models as well as draw clear data driven conclusions about the effectiveness of these models.

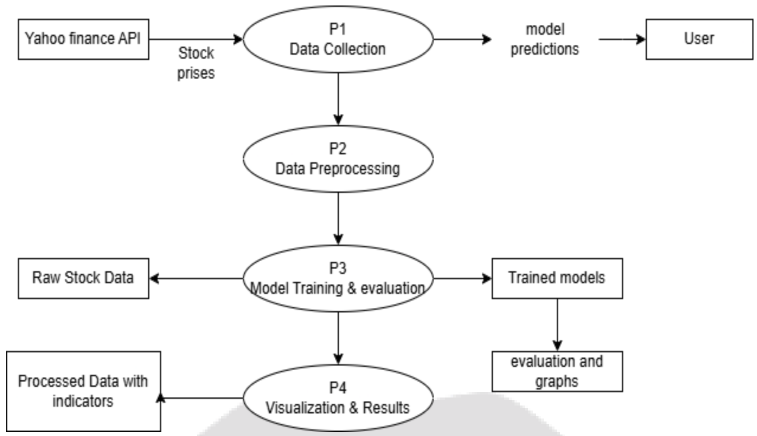


Fig-3: Data Flow Diagram (DFD) stock price prediction

2.2 Description of the Modules

2.2.1 Data Collection

- **Purpose:**
  This highlighted module is the basis to the system's implementation by gathering historical stock data. This module provides a means for the user to collect time intervals of reliable and structured financial data from Yahoo Finance.
- **Functionality:**
  - Retrieves the Open, High, Low, Close, Adjusted Close and Volume values associated with any selected stock symbols.
  - Automates the complete process of data extraction that reduces human error and the associated manual process of creating up to date records of stock data.
  - Stores data in a generic format for subsequent processing [8].
- **Outcome:**
  Clean, raw stock market data which forms the base for analysis and model training.
-

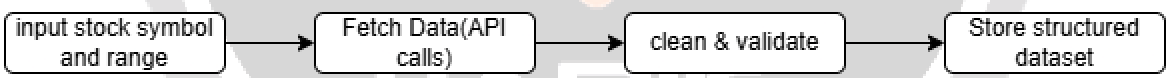


Fig-4: Data collection pipeline

2.2.2 Data Preprocessing Module

- **Purpose:**
  This module is responsible for transforming raw stock data into a form suitable for time-series modeling.
- **Functionality:**
  - Cleanses data (handling missing values if necessary or unknown inconsistencies);
  - Normalizes the numerical aspects of the open, high, low, close and volume values ensuring uniform scale of value across features;
  - Generates technical indicators such as Moving Averages (MA), Relative Strength Index (RSI), MACD which will be important features used for prediction [5].
  - Define ordering of data in tracked sequences (e.g., sliding windows) to be used to train models.
- **Outcome:**
  A refined and feature-rich dataset that captures both price trends and market behaviour, ready for model input.

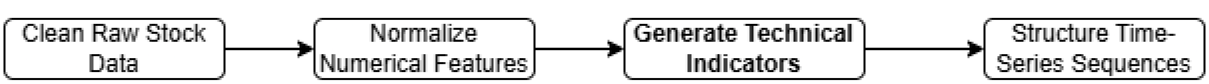


Fig-5: Data Preprocessing pipeline

2.2.3 Model Creation, Training, Testing, and Evaluation Module

- **Purpose:**
  This module represents the core machine learning component of the project, where the prediction models are built, trained, and evaluated.
- **Functionality:**
  - Uses two deep learning architectures (Long Short-Term Memory [LSTM] and Bi-directional LSTM [Bi-LSTM]);
  - Evaluates models trained on the processed stock data sequences based on performance metrics that include Mean Squared Error [MSE], Root Mean Squared Error [RMSE], and R^2 Score;
  - Compares LSTM vs. Bi-LSTM for prediction accuracy [6].
- **Outcome:**
  Trained prediction models capable of forecasting future stock prices, with results showing the comparative efficiency of Bi-LSTM over standard LSTM.

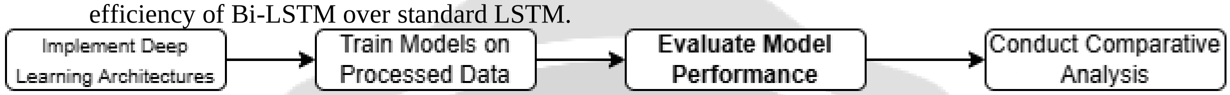


Fig-6: Model Training and Evaluation pipeline

2.2.4 Visualization Module

- **Purpose:**
  This module aids in interpreting and communicating the results effectively through visual means.
- **Functionality:**
  - Graphs actual vs. predicted stock prices to understand what models show close deployment to the actual market response.
  - Graph of the training loss and validation loss to inform understanding of model performance through the epochs.
  - Built a visual comparison of the LSTM and Bi-LSTM outputs.
  - Graph the trends of technical indicators against the stock getting insights into predictability..
- **Outcome:**
  Clear and insightful visual representations that support data-driven conclusions and enhance understanding of model behaviour.

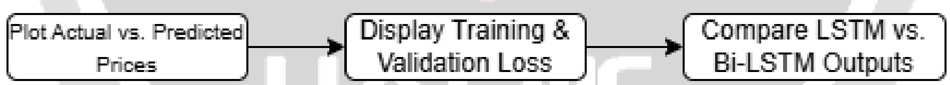


Fig-7: Deep learning Models Comparison pipeline

Apart from building the modules and everything as mentioned above, the implementation protocol also made sure that the modules would be seamless in connection with each other, to support passing of data from one stage in the prediction process to the next. The project was also developed in a Model-Driven Architecture (MDA) style, where the main function logic and dependencies of the platform could be decoupled [4].

3. Result and Discussions

This section summarizes the findings from the experiments with Long Short-Term Memory (LSTM) and Bidirectional Long Short-Term Memory (Bi-LSTM) models which were trained on stock price data. The aim of the study was to determine the efficacy of both models in predicting future stock prices based on historical trending, prices, and technical indicators.

3.1 Description

In this section, the experimental results of training both LSTM and Bi-LSTM models on historical stock market data are examined. The models were trained over various numbers of epochs—30, 60, and 100 epochs—as a manner to evaluate the effects of the duration in training on the performance stability, convergence rate, and predictability.

3.1.2 Model Architecture and Training Behaviour

Scientifically, both LSTM and Bi-LSTM models were built to capture temporal dependencies in sequential financial data. An LSTM network processes input sequences along a single direction (past-today-future) and a filtered signal is fed to the next node. A Bi-LSTM processes the same sequences in both directions (forward and backwards), thereby utilizing the contextual information from both past and future time steps. Bi-LSTM networks are strong for time-series predictions when a period's directional temporal context matters [2].

In both the LSTM and Bi-LSTM models, the learning component was considered during training by using technical indicators (MA: Moving Averages, Bollinger Bands, MACD: Moving Average convergence divergence, RSI: Relative Strength Index) as enhanced feature vectors. Technical indicators are a form of domain-specific notation.

The training underwent an evaluation during the optimization cycle using loss functions (Mean Squared Error) as well as against the RMSE and R^2 Score metrics [10]. First at 30 epochs, both models begin their optimization cycle where the training and validation loss started to reduce albeit moderate levels, and no significant generalization to patterns was established. The LSTM learning was slow due to the structural sequential nature and often an underfitting early on, the Bi-LSTM needing to learn on both dimensions, began to learn the dependencies better at early epochs.

When reaching 60 epochs both models began to converge and their loss landscapes had started to flatten out which means they would have somewhat stable learning. At this epoch the Bi-LSTM has continued to outperform the LSTM for predictive consistency with lower residual variance and closer representation to short-range variations in stock price.

Upon completing testing at 100 epochs, the two models had transitioned to optimized predictions. The LSTM began to stabilize, but still exhibited erring behaviour relative to disposition in volatile market segments, but the Bi-LSTM model had maintained an overall generalization with low prediction error, and better performance in tracking the spikes/dips and overall trends with more correlation to actual market event outcomes. We can see that both models are able to hold on to their long-range dependencies, but the Bi-LSTM is better capable of understanding temporal abstraction.

Quantitative Performance

- **RMSE:** Bi-LSTM always maintains lower RMSE values in all epochs, reflecting smaller prediction intervals and lower variation from real values**.**
- **R^2 Score:** The coefficient of determination for Bi-LSTM increases with increasing epoch size, reflecting better explanation of variance in target values**.**
- **Loss Convergence:** Bi-LSTM training and validation loss converges more steadily, while LSTM occasionally has mild overfitting beyond 60 epochs.

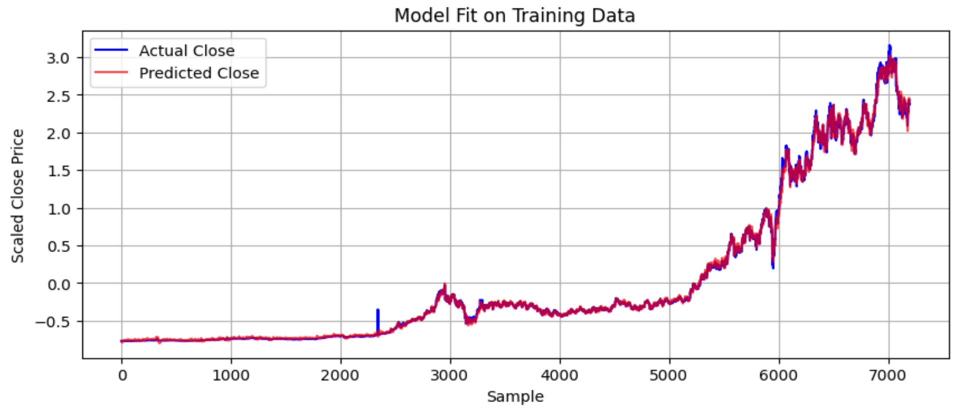
3.2 Graphs

Plots are key to guaranteeing model correctness and numerical output interpretability, otherwise mysterious. In the project, plots were not only made for looks, but for understanding model behavior. From loss across training epochs to actual vs. predicted price comparison, all plots are visual milestones for tracking the models' responses to data over time. Trends, deviations, and convergence patterns found in these plots often speak a thousand words more than numbers, helping in underfitting, overfitting, or learning instability diagnosis. These plots, in a way, translate abstract mathematical performance into reasonable real-world sense.

Graph 1: LSTM vs Bi-LSTM Predictions

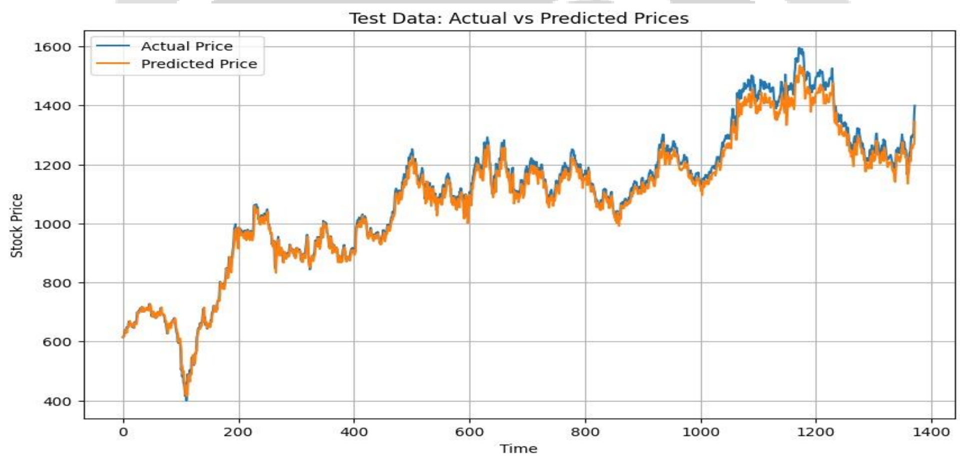
- **Observation:**
  At 100 epochs, The model begin learning the stock patterns.

BiLSTM:



Graph-1: Model prediction curve (BI-LSTM)

LSTM:

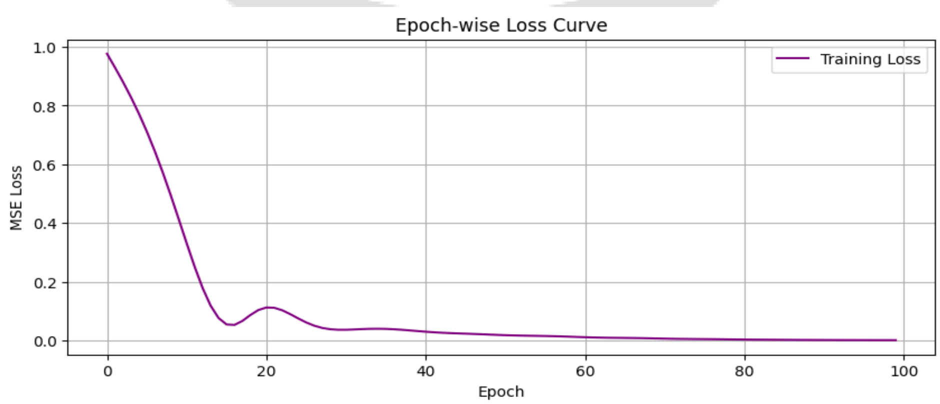


Graph-2: Model prediction curve (LSTM)

Graph 2: Training and Validation Loss Curves

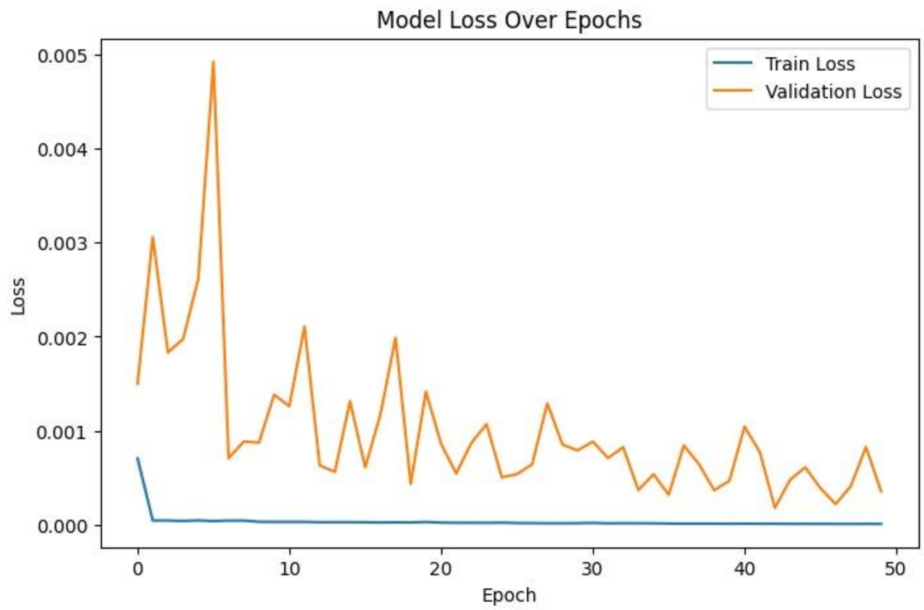
- **Observation:**
  Loss curves show decreasing trends for both models, indicating effective training. Bi-LSTM converges faster and with lower final loss, indicating more efficient learning.

BiLSTM:



Graph-3: Training and Validation Loss Curves (BI-LSTM)

LSTM:



Graph-4: Chart Training and Validation Loss Curves (LSTM)

Metric	LSTM	BiLSTM
Root Mean Squared Error (RMSE)	5.6123	0.0314

4. CONCLUSIONS

This project evaluated the performance of Bidirectional LSTM (Bi-LSTM) and Long Short-Term Memory (LSTM) neural networks for stock price forecasting using historical time-series data from Yahoo Finance. Through rigorous data preprocessing, technical indicator aggregation, and comparative modeling, we demonstrated that Bi-LSTM models are superior in modeling intricate sequential patterns in financial time-series data.

Modular architecture, right from data acquisition, preprocessing, model training, and visualization, facilitated scalability and adaptability. Various epochs (30, 60, 100) were evaluated, and the models' performance was verified using robust metrics such as RMSE and R^2 score. Empirical results evidently indicated that Bi-LSTM, with its bidirectional architecture, provided more precise predictions and generalized better, particularly when there was an availability of long-term trends from the past.

The visualization module further confirmed model robustness, enabling stakeholders to interpret and analyze predictions easily. Overall, the project successfully met its objectives and provided a strong foundation for future work in deep learning-based financial forecasting.

6. REFERENCES

[1]. S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, Vol. 9, No. 8, pp. 1735–1780, Nov. 1997.

[2] A. Graves and J. Schmidhuber, “Framewise phoneme classification with bidirectional LSTM and other neural network architectures,” *Neural Networks*, Vol. 18, No. 5–6, pp. 602–610, July 2005.

[3] G. E. Hinton, L. Deng, D. Yu et al., "Deep Neural Networks for Acoustic Modeling in Speech Recognition," *IEEE Signal Processing Magazine*, Vol. 29, No. 6, pp. 82–97, Nov. 2012.
[4] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, Vol. 14, No. 3, pp. 199–222, Aug. 2004.
[5] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Computation*, Vol. 12, No. 10, pp. 2451–2471, Oct. 2000.
[6] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*, 2nd ed., Springer, New York, 2002.
[7] M. Dixon, D. Klabjan, and J. Bang, "Classification-based Financial Markets Prediction using Deep Neural Networks," *Algorithmic Finance*, Vol. 6, No. 3–4, pp. 67–77, 2017.
[8] Y. Kim, H. Lee, and K. Oh, "Application of LSTM Neural Network for Predicting Stock Market Trends," *Sustainability*, Vol. 12, No. 1, Article 356, Jan. 2020.
[9] F. Chollet, *Deep Learning with Python*, 2nd ed., Manning Publications, USA, 2021.
[10] J. Brownlee, "How to Develop LSTM Models for Time Series Forecasting," *Machine Learning Mastery*, 2018. Available Online at: https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/