Study of Stock Price Trends for Indian Share Market vs American Share Market Using LSTM

Sharma Ji
Department of CSE
Ajay Kumar Garg Engineering College.
Ghaziabad, India.
saurabh12d@gmail.com

Vishal Jayaswal
Department of CSE
Ajay Kumar Garg Engineering College.
Ghaziabad, India.
vishaljayaswal026@gmail.com

Rishabh Yadav
Department of CSE
Ajay Kumar Garg Engineering College.
Ghaziabad, India.
personalrsh11@gmail.com

Shreyansh Mohan
Department of CSE
Ajay Kumar Garg Engineering College.
Ghaziabad, India.
shreyanshmohan35@gmail.com

Vivek Kumar
Department of CSE
Ajay Kumar Garg Engineering College.
Ghaziabad, India.
vivekmishra28mgs@gmail.com

Shantanu Sirohi
Department of CSE
Ajay Kumar Garg Engineering College.
Ghaziabad, India.
shantanu4371@gmail.com

Satvik Maheshwari
Department of CSE
Ajay Kumar Garg Engineering College.
Ghaziabad, India.
maheshwarisatvik92@gmail.com

2024 1st International Conference on Advanced Computing and Emerging Technologies (ACET) | 979-8-3503-6772-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ACET61898.2024.10730548

***Abstract*— Stock market trading is a pivotal financial activity, fraught with uncertainty and volatility. Investors constantly seek methods to predict future trends, aiming to minimize losses and maximize profits. Achieving complete accuracy in forecasting remains a challenge, yet progress within Deep Learning (DL) and Machine Learning (ML) has prompted researchers to investigate several methods for predicting stock prices.This employs Long Short-Term Memory (LSTM) models to forecast stock prices of five Indian and five American companies across varying time horizons. In this underscores the need for customized predictive models to navigate the unique complexities of forecasting within different market contexts.**

***Keywords*— Deep learning, Recurrent neural networks, Long-Short term memory, Feature engineering, Stock Market**

I. Introduction

Stock markets serve as the quintessential arena for financial transactions. The fluctuating prices of stocks reflect a complex amalgamation of factors ranging from political shifts to global phenomena, underscoring the intricacies of market trends. Accurately predicting market trajectories remains a formidable challenge due to the enduring nature of global transformations and uncertainties[1]. Nevertheless, the quest for predictive accuracy persists, driving the exploration of predictive models leveraging advancements in Machine Learning algorithms.

LSTM networks have emerged as a potent tool for tackling various time-series prediction tasks, particularly in the realm of stock price forecasting. Being a variant of recurrent neural networks (RNNs)[2], LSTMs possess the capability to selectively retain or discard information over time, rendering them adept at modelling sequential data like stock prices. Over recent years, a plethora of studies have delved into leveraging LSTMs for stock price prediction, yielding promising outcomes. The proposed methodology harnesses historical stock price data as input, training an LSTM model to anticipate future stock prices[3]. The architecture of the LSTM network comprises multiple layers of LSTM cells, facilitating the capture of long-term dependencies and intricate patterns within the input data..

II. Idea About Our Problem and Solutions

One of the main goals of employing an LSTM model for stock price prediction is to forecast how stock prices will vary over time [4]. Using historical stock price data, the LSTM model can be trained to find patterns and trends in the data, which can subsequently be utilized to forecast how stock prices will fluctuate over time. Different performance metrics, such as mean squared error (MSE), root mean square error (RMSE), mean absolute error (MAE), or correlation coefficient (R) can be used to assess how well the LSTM model predicts changes in stock prices over time[5]. In technical analysis, a moving average (MA) is a stock indicator that is commonly employed in the financial industry.The moving average of a stock is calculated to help smooth out the price data by creating an average price that is updated continually. Financial analysts frequently utilize moving average as a technical indicator to spot possible buy or sell opportunities, in addition to using it as an input to the LSTM model[6]. For instance, a positive trend may be indicated if the current stock price is above its 50-day moving average, while a negative trend may be indicated if the current stock price is below its 50-day moving average. In general, moving averages are helpful tools for stock price prediction utilizing long short-term memory (LSTM) models because they can offer important insights into the general trend and momentum of a given stock over time.

III. Literature Review

Time series analysis is a key strategy in stock market prediction since time-dependent volatility are ubiquitous in this environment. Although Auto Regressive Integrated Moving Average (ARIMA) models are widely used, their linear structure frequently makes it difficult for them to fully represent the complexities of extremely turbulent markets.

In contrast, the advent of ML and data science heralds a new era, characterized by the development of robust predictive algorithms tailored to the complexities of financial markets.

In recent times, the fusion of statistical methods with learning models has refined numerous machine learning algorithms, showcasing enhanced effectiveness and reduced multicollinearity compared to traditional linear regression approaches. A plethora of ongoing research delves into the techniques of machine learning in finance, with a notable focus on employing tree-based models for predicting portfolio returns.

When evaluating predictive accuracy, metrics such as Root Mean Square Error (RMSE), Symmetric Mean Absolute Percentage Error (SMAPE), and R^2 value (R-squared) serve as essential benchmarks. These metrics allow for a detailed comparison between actual and predicted values, with Mean Squared Error serving as a key evaluation tool. Upon analysing the RMSE function, it becomes evident that the root mean square error loss values consistently fall within the [0,1] range, meeting expectations. This indicates that the LSTM model's parameters effectively enable stock prediction capabilities [11].

Through empirical analysis and critical evaluation, this seeks to illuminate the path towards predictive accuracy in the ever-evolving landscape of financial markets. Researches have been done using RNN and LSTM model on Indian stocks using high, low, close and open price of each day.

RNN and LSTM in Sequential Data Analysis

Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNN) are essential tools for digesting sequential data and predicting future trends.

LSTM emerged as a solution to the limitations of traditional RNNs, which struggled with vanishing or exploding gradients during backpropagation through time. Developed by Sepp Hochreiter and Jurgen Schmid Huber, LSTM introduced a novel architecture capable of learning long-term dependencies and handling noisy or incompressible sequences effectively. LSTM addresses the challenge of retaining information over extended time intervals, making it suitable for various applications such as image captioning, handwriting generation, and chatbots [12].

TABLE 1
Literature Review

S No.	Paper Name	Author Name	Year	Description
1	Predicting Stock Market using LSTM	Manan Shah, Druthi Seth	2023	AI helps anticipate stocks more accurately, resolving previous issues and demonstrating promise for increased accuracy.
2	Forecasting Stock Market Indices Using RNN hybrid Models	Hyunsun Song, Hyunjun choi	2023	Cutting edge hybrid models surpass conventional techniques in improving stock index prediction.
3	An efficient time series RNN for stock price prediction	Minrong Lu, Xuerong Lu	2024	The time series data in this work is processed using the sliding window approach.In this time series data processing is done using this method..
4	Prediction of High Variation in Stock price using LSTM	Gaurav Batha, Rinkle Rani	2022	In this article, the model is trained and tested using LSTM in conjunction with an optimizer and sigmoid function.
5	Optimizing LSTM for time-series prediction in Indian Stocks	Anita Yadav, C.K Jha, Aditi Sharan	2020	This work optimizes LSTM models for stock market prediction, thereby resolving the issue of hyperparameter selection in time series forecasting.

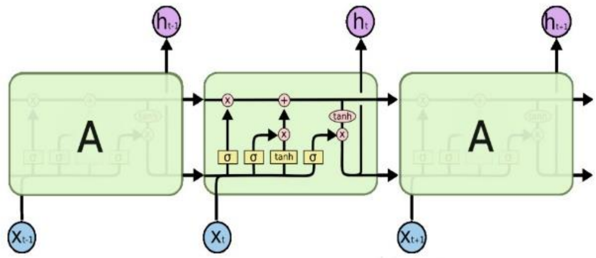

Fig. 1. The skeleton of LSTM model

The input gate, forget gate, memory gate, and output gate are the four interacting layers that make up the LSTM architecture. These gates, which control information flow inside the network by allowing it to selectively keep or discard previous knowledge, are made up of sigmoid neural network layers [3]. LSTM's effectiveness lies in its ability to maintain a cell state which runs through the entire network and stores relevant information over time. LSTM shows more accuracy than other sequential data machine learning models. LSTM also overcome the gradient weight loss problem in traditional recurrent neural network.

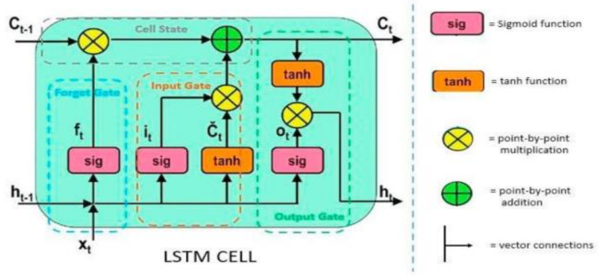

Fig. 2. LSTM architecture

Traditional RNNs lack the capability to retain long-term memory, leading to performance degradation when faced with distant dependencies. However, RNNs excel in capturing short-term dependencies and context-based reasoning, making them valuable for tasks where recent information is crucial for predictions [11]. Because RNNs are recurrent, they can handle sequential input by using knowledge from earlier time steps, which makes context-aware analysis possible.

IV. Methodology and Data Acquisition Procedures

Step 1: Use APIs or web scraping tools to collect historical stock price data. Libraries like yfinance can be used for this purpose.

Step 2: Pre-processing of Raw Data:
a) Data discretization
b) Data transformation
c) Data cleaning
d) Data integration

Step 3: Generally, 80% of the data is used for training and 20% is used for testing after the dataset is divided into training and testing sets.

Step 4: Feature Selection: Choosing relevant attributes for feeding into the neural network.

Step 5: Training the Neural Network (NN) Model

Step 6: Comparing the generated output with target values to calculate error. Backpropagation algorithm adjusts neural network biases and weights to minimize error.

Step 7: Update of Test Dataset: Repeat of Step 2 for the test dataset.

Step 8: Error Calculation and Net Growth of Companies

Step 9: Utilization of Keras and their function APIs for visualization of predictions.

Step 10: Investigation of Different Time Intervals

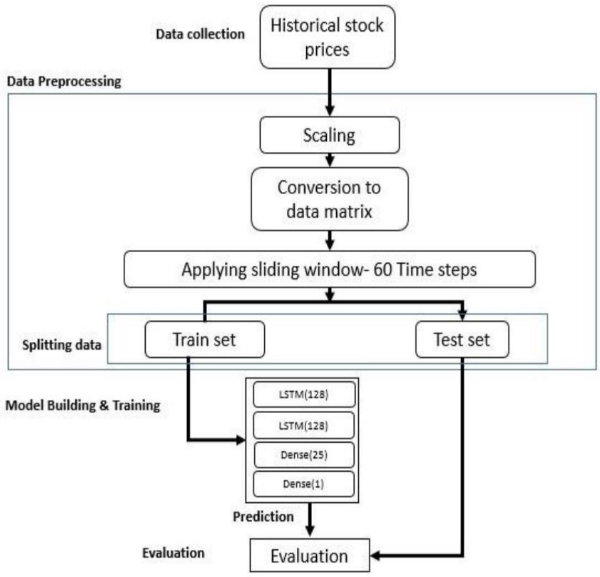

Fig. 3. Block diagram of stock prediction using LSTM

V. Results

In the current research, we used Long Short-Term Memory (LSTM) models to anticipate five Indian and five American companies' stock prices across various time periods. To assess the performance of the model, the root mean square error (RMSE) metric was employed. Batch sizes in the tests ranged from one year to twenty years.

The metrics used for performance evaluation were RMSE and R^2 values. A lower RMSE value means better model performance. R Squared (R^2) coefficient quantifies the proportion of variation in the dependent variable (y) explained by the regression line, relative to the variation explained by the mean of y. It gauges how much more accurately the regression line predicts each point's value compared to simply using the average value of y. Subsequently, detailed and comparative results were derived after testing. The formulae for calculating RMSE and R^2 are provided below:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_{pred,i}-Y_i)^2}{n}} \quad \text{...Eq.1}$$

$$R^2 = 1 - \frac{unexplained\ variation}{total\ variation} \quad \text{...Eq.2}$$

In assessing the performance of Long Short-Term Memory (LSTM) models in predicting share prices across different markets, distinct trends emerged from the analysis of Root Mean Square Error (RMSE) values. The tables presented showcase the RMSE values for Indian and American shares, delineated across varying batch sizes. Notably, the results reveal a consistent pattern where Indian shares exhibit higher RMSE errors compared to their American counterparts across all considered batch sizes. This discrepancy in RMSE values suggests a greater volatility or unpredictability in the Indian stock market compared to the American market. The higher RMSE for Indian shares underscores the complexity and inherent challenges in accurately forecasting share prices within the Indian stock market context. Conversely, the lower RMSE errors in American shares may reflect the relative stability and efficiency of the American stock market, characterized by robust regulatory frameworks, transparency, and investor confidence. Consequently, while LSTM models demonstrate efficacy in capturing trends and patterns in both markets, the higher RMSE errors in Indian shares highlight the need for further refinement and customization of predictive models to effectively navigate the unique challenges inherent in forecasting share prices within the Indian market landscape. The result tables displaying the studied trends of RMSE and R2 values have been described below:

TABLE II

RMSE values for the top five Indian companies as per NSE

Shares	1 year	3 years	5 years	10 years	20 years
Reliance Industries Ltd	105.54044	100.17441	135.62823	115.33968	86.86474
TCS Ltd	139.76412	93.26381	137.40990	137.69589	124.06299
HDFC Bank Ltd	45.57018	37.26086	48.98969	58.02315	62.00986
Bharti Airtel Ltd	106.49257	83.55592	76.02187	43.91697	35.96272
ICICI Bank Ltd	38.06724	49.48240	53.35982	41.26262	33.00196

TABLE III

RMSE values for the top five American companies as per NMS

Shares	1 year	3 years	5 years	10 years	20 years
Microsoft Corporation	18.36040	15.90182	19.83722	18.00669	11.56377
Apple Inc	4.89365	6.41479	11.33944	7.64766	5.98214
NVIDIA Corporation	99.71166	65.19945	59.07498	37.15963	24.88455
Alphabet Inc.	8.98963	5.79968	7.93105	7.04587	5.33729
Amazon.com Inc.	12.98339	6.22628	6.57836	7.90343	5.50817

Contrary to the divergence observed in RMSE values, the comparison of R-squared (R^2) scores between Indian and American shares reveals a nuanced similarity. Despite the notable differences in RMSE, the R^2 scores exhibit a relative parity between the two markets across various batch sizes. This observation suggests that while the LSTM models may struggle with higher error rates in predicting Indian share prices, they nonetheless demonstrate a comparable level of explanatory power for both Indian and American markets. The consistency in R^2 scores underscores the robustness of the LSTM models in capturing underlying trends and patterns within the respective stock markets, despite the distinct challenges posed by the Indian market's higher RMSE values. This convergence in R^2 scores suggests that the LSTM models possess a similar ability to elucidate the relationships between input variables and share price movements across both Indian and American markets, despite the disparities in predictive accuracy reflected in the RMSE values.

TABLE IV

R2 values for top five Indian companies as NSE

Shares	1 year	3 years	5 years	10 years	20 years
Reliance Industries Ltd	0.81456	0.83722	0.92456	0.97891	0.98685
TCS Ltd	0.76183	0.88448	0.95361	0.97933	0.98834
HDFC Bank Ltd	0.73298	0.85667	0.94346	0.97961	0.98649
Bharti Airtel Ltd	0.48580	0.77532	0.88262	0.96423	0.97117
ICICI Bank Ltd	0.59287	0.85007	0.94443	0.97543	0.98331

TABLE V

R2 values for top five American companies as NMS

Shares	1 year	3 years	5 years	10 years	20 years
Microsoft Corporation	0.75280	0.90674	0.93255	0.97321	0.98742
Apple Inc	0.66625	0.87861	0.93092	0.98266	0.98852
NVIDIA Corporation	0.70237	0.89479	0.90868	0.95351	0.96695
Alphabet Inc.	0.46792	0.90489	0.93293	0.96803	0.98327
Amazon.com Inc.	0.62240	0.95371	0.95516	0.97715	0.99027

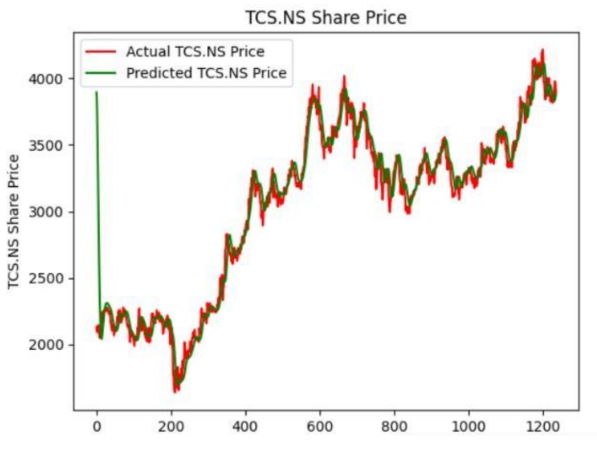

Fig. 4. Price prediction over a period of 5 years for TCS

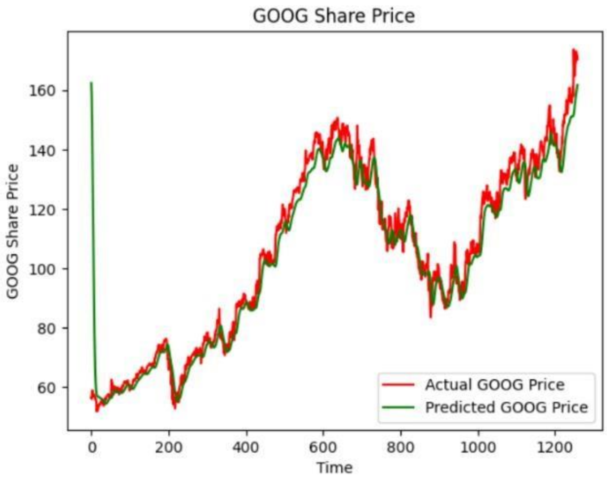

Fig. 5. Price prediction over a period of 5 years for Google

The results indicated notable variations in the prediction accuracy across different stocks and timeframes. Specifically, we observed that the LSTM model performed differently for Indian and American stocks, suggesting potential differences in the underlying patterns and factors influencing their price movements. Moreover, as the prediction horizon increased, the RMS error tended to de-escalate, reflecting the inherent challenges of forecasting stock prices over longer time intervals.

VI. CONCLUSION

This adds to the expanding research on stock price prediction employing deep learning techniques. Our results highlight the significance of integrating temporal and geographical factors when crafting predictive models for financial markets. The diverse performance of the LSTM model across distinct stocks and timeframes underscores the necessity for customized strategies in stock price forecasting.

Furthermore, the results highlight the limitations of using a single model for predicting diverse sets of stocks. Future research could explore ensemble techniques or hybrid models to enhance prediction accuracy and robustness. Additionally, Including outside variables like market sentiment, news sentiment analysis, and macroeconomic data may enhance the models' ability to forecast future events.

In conclusion, while LSTM models show promise in stock price prediction, further refinement and exploration are necessary to fully leverage their potential in real world financial applications. The present research serves as a basis for forthcoming research initiatives that seek to augment the efficacy of prediction models within the financial markets.

REFERENCES

[1] D. G. Bhalke (2022): Stock Price Prediction Using Long Short-Term Memory. Vol. 14, Iss: Spl-2 issu, pp 271-273

[2] Malti Bansal, Apoorva Goyal, Apoorva Choudhary- Stock Market Prediction with High Accuracy using Machine Learning Techniques. © 2023 The Authors. Published by Elsevier B.V.

[3] Achyut Ghosh, Soumik Bose, Giridhar Maji, Narayan C. Debnath, Soumya Sen- Volume 63, 2019, Pages 101–110 Proceedings of 32nd International Conference on Computer Applications in Industry and Engineering.

[4] Adil MOGHAR, Mhamed HAMICHE- Stock Market Prediction Using LSTM Recurrent Neural Network. 1877-0509 © 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

[5] G. Maji, S. Sen and A. Sarkar, "Share Market Sectoral Indices Movement Forecast with Lagged Correlation and Association Rule Mining," in International Conference on Computer Information Systems and Industrial Management, Bialystok, Poland, Sprigner, 2017, pp. 327-340.

[6] Qiang Jiang, Chenglin Tang, Chen Chen, Xin Wang, Huang Sichuan (Springer, Cham), 2018 - pp 393-408.

[7] Patel Vraj, Hemang Kumar, Patel Aksh, Hemant Kumar, Patel Ansh Satishbhai, Padariya, Jagrut Kumar Kaushikbhai, Parmar Sahaj Ketan- 14 Mar 2023 - Indian Scientific Journal Of Research In Engineering And Management - Vol. 07, Iss: 03

[8] Yi Zhao- Stock Prediction based on LSTM Neural Network; VOL. 3 (2023): 4TH International Conference on Management Science and Innovative Education (MSIE 2022)

[9] Shun Lin, Yuan Feng- Research on Stock Price Prediction Based on Orthogonal Gaussian Basis Function Expansion and Pearson Correlation Coefficient Weighted LSTM Neural Network (2022); DOI: 10.23977/acss.2022.060504 ISSN 2371-8838 Vol. 6 Num. 5

[10] Ruobing Zhang- LSTM-based Stock Prediction Modeling and Analysis; Proceedings of the 2022 7th International Conference on Financial Innovation and Economic Development (ICFIED 2022.

[11] Bhanu Sridhar Mantravadi, Anuhya Karedla, G Blessitta, Ch. Ramya Krishna, Ch. Abitha- Prediction of Stock Prices based on Historical Data using Linear Regression Model through LSTM Algorithm; 2023 International Conference on Advancement in Computation & Computer Technologies, IEEE. DOI: 10.1109/InCACCT57535.2023.10141804

[12] Bathla, G., Rani, R., Aggarwal, H.: Stocks of the year 2020: prediction of high variations in stock prices using LSTM. Multimed. Tools Appl. 82(7), 9727–9743 (2023)