International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Utilizing Long Short Term Memory (LSTM) for Stock Market Price Prediction with Machine Learning

Ramya Sree K N[1], Vidhya A[2], Jotheeswaran G[3]
[1]Department of Information Technology, Jeppiaar Engineering College, Chennai, Tamil Nadu, India- 600073
[2] Department of Information Technology, Jeppiaar Engineering College, Chennai, Tamil Nadu, India- 600073
[3] Department of Information Technology, Jeppiaar Engineering College, Chennai, Tamil Nadu, India- 600073

Abstract— *Stock market prediction plays a crucial role in financial decision-making by attempting to forecast the future value of company stocks or other financial instruments traded on exchanges. This research paper presents a Python-based approach to stock price prediction utilizing machine learning techniques. The code utilizes various libraries for data handling, preprocessing, visualization, and building a Long Short-Term Memory (LSTM) model and comparing it to SVR model. The historical stock data for Apple Inc. (AAPL) is downloaded using the yf.download function from the yfinance library, and the data is visualized through time series plots. The data is split into training and testing sets, and input sequences and output values are generated using a sliding window approach. The LSTM model is constructed using the Sequential API from Keras. The trained model is then used to predict stock prices on the test data, and the root mean square error (RMSE) is calculated to evaluate the model's performance. The predictions, along with the actual and training data, are visualized to gain insights into the model's predictive capabilities*.

Keywords— *LSTM, RMSE, Stock.*

I. INTRODUCTION

Based on plenty of studies on the subject, algorithms based on machine learning can be used to predict values for regulating and restricting an extensive collection of resources, in addition to for investing and many other operations. The term machine learning, in general, pertains to any algorithmic approach that uses computer to find patterns based simply on data and without the need for programming instructions. Many studies are now being undertaken on the subject of machine learning methods applied in finance. A number of the studies mentioned utilised tree-based models to anticipate returns on portfolios [1], whereas others used deep learning to forecast future financial asset values [2], [3]. The predicting of returns using the ADaBoost algorithm was also reviewed by certain writers [4]. Others continue to predict stock returns employing a special decision-making model for day trading investments on the stock market that was established by the authors [5], in this model, portfolio selection is done using the MV

approach and the SVM method, respectively. Deep learning techniques for smart indexing were discussed in another study [6]. According to [7] RNNs of the LSTM kind are able to solve linear problems. A deep learning method is LSTM. The learning of very lengthy sequences is imposed on LSTM Units. The gated recurrent system is shown here in a more generalised form. Because LSTMs address the evanescent gradient problem faced by [8], they are less harmful than other deep learning techniques like RNN or conventional feed forward. The objective of this study is to provide a machine learning-based technique for stock price prediction that focuses on the usage of a LSTM model. The purpose of this study is to illustrate the accuracy of LSTM in forecasting stock prices and to assess the model's performance using real-world data. Furthermore, the research seeks to provide insight into the predictive capabilities and limits of the suggested technique.

II. RELATED WORKS

In their research, Sepp Hochreiter and Jürgen Schmidhuber [9] The LSTM is a form of neural network used in artificial intelligence. LSTM features connections with feedback, as opposed to standard feed forward neural networks. This version of RNN can assess not just individual data points (such as images), but also entire data sequences (such as audio or video). The LSTM design intends to offer RNN with a short-term memory that may endure hundreds of thousands of timesteps, thereby providing "long short-term memory." According to Adil Moghar and Mhamed Hamiche [10], machine learning, which involves instructing algorithms to execute activities that would ordinarily need human intellect, is now the leading trend in academic research.To anticipate future stock market values, a model may be built using RNN and, in particular, the LSTM. Jiayu Qiu, Bin Wang, and Changjun Zhou [11] propose a forecasting framework to anticipate stocks initial prices in their study. They employed a wavelet transform on stock data and an attention-based LSTM neural network to anticipate the stock opening price, with good results. The results of the experiment reveal that our suggested model outperforms the commonly used LSTM, GRU, and LSTM neural network models with wavelet transform in terms of fitting degree and prediction accuracy. Thu Hang Nguyen, Nguyen Trung Tuan, and Thanh Thi Hien Duong [12] In addition to standard forecasting methods such as LR and ARIMA, analysts are now trying to incorporate new deep learning algorithms to anticipate stock market trends and direction in order to make more precise forecasts. RautSushrut et al. [13] proposed using automated classifiers to estimate stock price movement based on financial index data and determining their capabilities. Portfolio modelling has been used in the financial market as a computational analytical method. Manoj S Hegde *et al.* [7] researched the LSTM networks, which are a sort of RNN that has a potential for solving involute problems, there is also an argument on using RNN to anticipate share prices. M. Roondiwala *et al.* [14] proposed that the most prevalent RNN design is Long ShortTerm Memory. LSTM introduced a

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

memory cell in the encrypted network layer; a device for processing that supplants traditional artificial neurons.

III. METHDOLOGY

A. Data Retrieval and Preprocessing

The first topic is concerned with data retrieval and pre-processing. During this phase, historical stock data for Apple Inc. (AAPL) is retrieved within a defined date range using the yfinance library. The downloaded data is then saved for subsequent examination. The matplotlib.pyplot module is used to visualise historical stock prices in order to obtain insights into the data. The date is plotted on the x-axis, while the prices ($) are plotted on the y-axis. Furthermore, the stock's closing prices are retrieved from the downloaded data. To guarantee that the models are properly trained, the closing prices are scaled using MinMaxScaler, which normalises values between 0 and 1.

B. Model development and Training

This section focuses on model creation and training. The training data for both the LSTM and SVR models has been produced. The sliding window method is used to generate input sequences and associated output values. The last 60 closing prices are represented by each input sequence, while the output value indicates the next closing price. Furthermore, the input sequences (x_train) are rearranged to fit the LSTM and SVR models' needed dimensions. The LSTM model is built with Keras' Sequential API, with extra fully connected layers (Dense) for final prediction. Simultaneously, an SVR model with a radial basis function (RBF) kernel is built. To discover the patterns and correlations in the stock price data, both models are trained using prepared training data.

C. Prediction, Evaluation, and Evaluation

This topic focuses on predicting, evaluating, and visualising results. Using the testing data (x_test), the trained models are utilised to forecast stock prices. After that, the forecasted prices are inverse translated back to their original scale. To assess the models' performance, the root mean squared error (RMSE) is produced by comparing projected and actual prices (y_test). This assessment score gives a quantifiable measure of the models' ability to forecast stock prices. Furthermore, the findings are given graphically by charting the original stock prices during the training period and the anticipated values during the testing period. This visualisation gives a full knowledge of the model's performance and important insights into the stock price prediction system's accuracy and dependability. The original stock prices (train) and forecasted prices (validation ['Predictions']) throughout the training period are displayed. The figure displays the training data, as well as the actual stock prices during the testing period and the anticipated stock prices. To distinguish between the training, validation, and projected pricing, a legend is provided.

IV. IMPLEMENTATION

A. Support Vector Regression (SVR)

SVR is a ML approach for performing regression analysis. SVR finds a function that minimises errors in prediction by approximating the relationship between input variables and a continuous target variable. SVR looks for the hyperplane that suits the data points in uninterrupted space most effectively as opposed to SVMs, which are used for classification problems. Performance is achieved by translating the supplied parameters to a feature space with high dimensions and selecting the hyperplane with the lowest prediction error while minimising the margin (distance) between each hyperplane and the nearest data points.

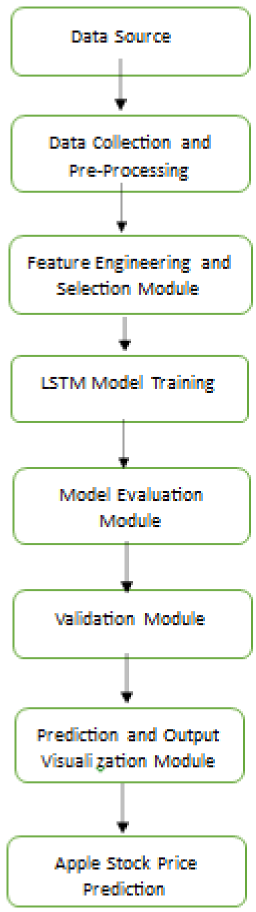

Fig. 1. System Architecture

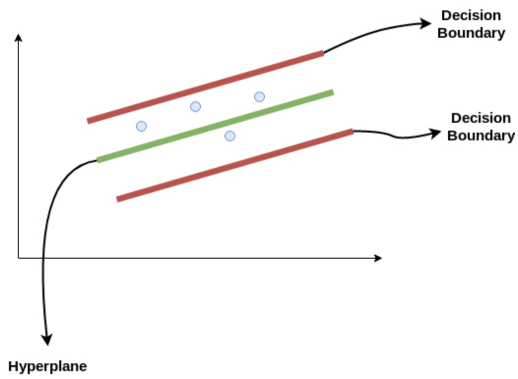

Fig. 2. Hyperplane concept for SVR model

By converting the data to a higher-dimensional space with the use of a kernel function, SVR is able to deal with non-linear

Ramya Sree K N, Vidhya A, and Jotheeswaran G, "Utilizing Long Short Term Memory (LSTM) for Stock Market Price Prediction with Machine Learning," *International Research Journal of Advanced Engineering and Science*, Volume 8, Issue 4, pp. 113-117, 2023.

correlations between the input and objective variables. As a result, it is an excellent tool for regression issues containing complex relationships between input and target variables. SVR employs identical ideas as SVM to address regression-related problems. Let's take a few moments to comprehend the concept of SVR.

These 2 lines in red represent the decision boundary, while the line in green represents the hyperplane. The purpose of SVR is to merely assess the elements that are within the scope of the boundary line. The most suited line is the hyperplane with the most points. The initial phase in understanding this is identifying the decision limit (the danger red line shown above!). Assume these lines are at any distance from the hyperplane, let's say 'a'. So they're the lines formed from the hyperplane at '+a' and '-a' distances. This 'a' is referred to as epsilon throughout the text.

The hyperplane's equation is as described below:

$$Y = wx+b$$

Decision boundary is given by:

$$Wx+b= +a$$
$$Wx+b= -a$$

Therefore, any hyperplane that meets our SVR model should also meet:

$$-a < Y- wx+b < +a$$

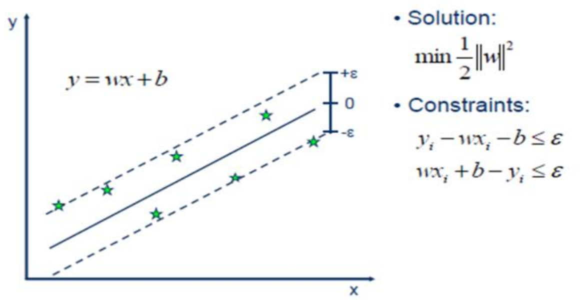

Fig. 3. Equation of hyperplane

B. Long Short-Term Memory

LSTMs are deep learning, sequential neural networks that allow knowledge to be retained. It is a form of Recurrent Neural Network that can manage the vanishing gradient problem that RNN faces. Hochreiter and Schmidhuber created LSTM to tackle the difficulty produced by classical rnns and machine learning methods. The Keras package is used to implement LSTM in Python.

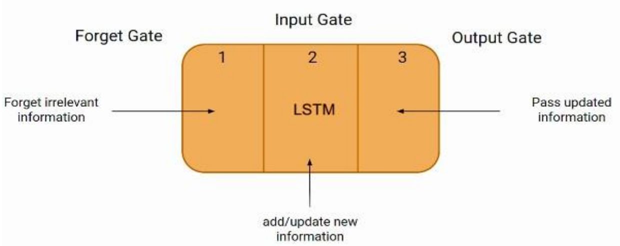

Fig. 4. LSTM architecture

At a high level, LSTM functions similarly to an RNN cell. The LSTM network's internal operation is shown below. The LSTM network design is divided into three pieces, as indicated in the picture below, and each portion serves a specific purpose. These three components of an LSTM unit are referred to as gates. They regulate the flow of information into and out of the memory cell, also known as the LSTM cell. The first gate is known as the Forget gate, the second as the Input gate, and the final as the Output gate. An LSTM unit made up of these three gates and a memory cell, or LSTM cell, can be thought of as a layer of neurons in a classic feed forward neural network, with each neuron having a hidden layer and a current state.

Forget Gate - First step in an LSTM neural network cell is to select whether to keep or discard the information from the previous time step. Here is the forget gate equation.

$$f_t = \sigma (x_t * U_f + H_{t-1} * W_f)$$

X_t: input to the current timestamp.

U_f: weight associated with the input.

H_{t-1}: The hidden state of the previous timestamp

W_f: It is the weight matrix associated with the hidden state.

A sigmoid function is then applied to it. This will result in foot being an integer between 0 and 1. This f_t is then multiplied by the preceding timestamp's cell state, as seen below.

$$C_{t-1} * f_t = 0 \quad \text{...if } f_t = 0 \text{ (forget everything)}$$

$$C_{t-1} * f_t = C_{t-1} \quad \text{...if } f_t = 1 \text{ (forget nothing)}$$

Input Gate - The input gate measures the significance of the new information carried by the input. Here is the input gate equation.

$$i_t = \sigma (x_t * U_i + H_{t-1} * W_i)$$

We applied the sigmoid function on it once more. As a result, at timestamp t, the value of I will be between 0 and 1.

$$N_t = \tanh(x_t * U_c + H_{t-1} * W_c) \text{ (new information)}$$

$$C_t = f_t * C_{t-1} + i_t * N_t \text{ (updating cell state)}$$

Output gate- Here is the equation for the Output gate, which is similar to the prior two gates.

$$o_t = \sigma (x_t * U_o + H_{t-1} * W_o)$$

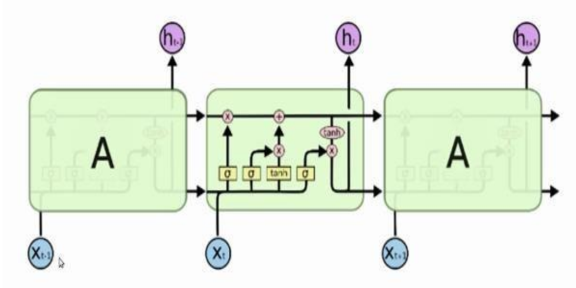

Fig. 5. Intuitive diagram of the LSTM network.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

V. EVALUATION

RMSE metrics are used to assess the effectiveness of the LSTM model for stock price prediction. The test data is used in the assessment to determine the model's ability to produce correct predictions. In the context of stock price prediction, investors and financial experts must evaluate the performance of predictive algorithms. The RMSE measure is critical in determining the accuracy of LSTM models used to forecast stock prices. Researchers and practitioners can use the RMSE to assess how well anticipated stock prices match actual values. This statistic gives a quantifiable measure of prediction performance, allowing stakeholders to assess the LSTM model's ability in identifying underlying patterns and trends in stock price movements. A lower RMSE value implies more accuracy, meaning that the LSTM model generates forecasts that closely match real stock values. A greater RMSE number, on the other hand, indicates that there are considerable disparities between the anticipated and actual values, indicating a less accurate model. Investors and analysts may make educated judgements based on the dependability and accuracy of the LSTM model's stock price forecasts by using RMSE as an assessment parameter, thereby enhancing their investing strategies and financial outcomes.

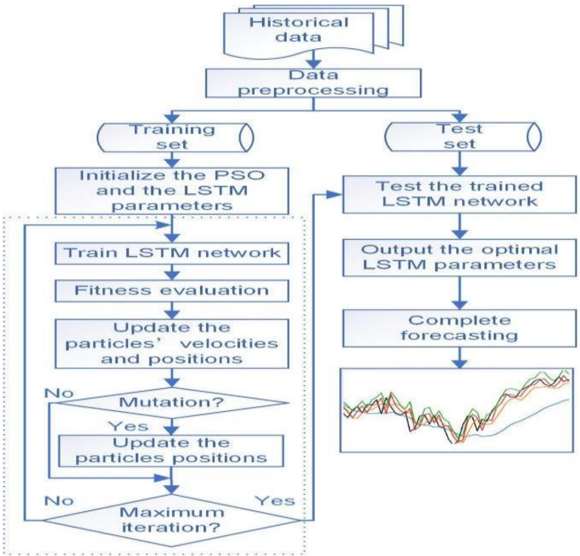

Fig. 6. Data flow diagram

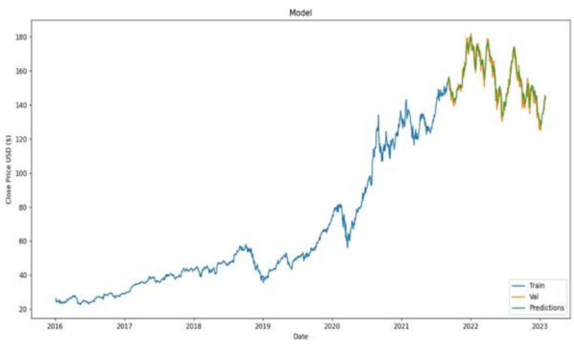

Fig. 7. Actual stock price vs predicted price for LSTM model

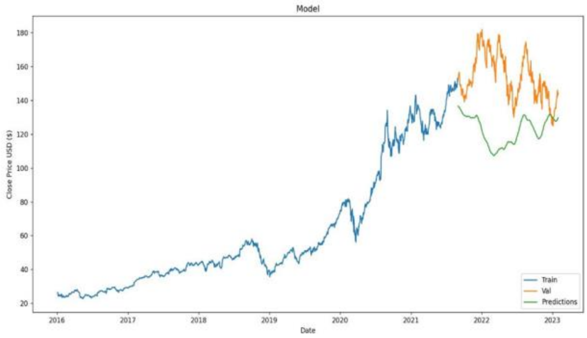

Fig 8. Actual stock price vs predicted price for SVR model

From fig 7 and 8 it is clear that the LSTM model predicts the price of stock more accurately than SVR model.

VI. CONCLUSION

Finally, the model predictions show that the LSTM model beats the SVR model in terms of accuracy for stock price prediction. The visualisations of anticipated values vs actual values clearly show that the LSTM model's predictions are considerably closer to the real stock prices than the SVR model's. This shows that the LSTM model captures the underlying patterns and trends in stock price movements better. The capacity of LSTM to simulate nonlinear interactions is one of its primary benefits over SVR. Stock prices are impacted by a variety of complicated elements and display nonlinear dynamics, making linear-based models like SVR difficult to effectively capture these patterns. In contrast, the design of LSTM allows it to manage nonlinear dependencies and record extensive data interactions. This adaptability allows the LSTM model to possibly deliver more accurate forecasts by incorporating the stock market's complicated dynamics. Furthermore, the capacity of LSTM can handle sequences of varying lengths is useful in stock price prediction. Financial data is frequently distinguished by varied time spans and inconsistent data availability. Because of its intrinsic flexibility to process variable length sequences and adapt to changing data availability without considerable preparation, LSTM is well-suited for modelling stock price data.

Currently, LSTM models for stock price prediction use historical price data as their primary input. Incorporating other data sources such as news sentiment, economic indicators, social media data, or company-specific information, on the other hand, might give useful context and perhaps enhance forecasts. These extra characteristics can be used as supplemental inputs in the LSTM model to capture a larger variety of variables impacting stock prices.

REFERENCES

[1] Moritz, B., & Zimmermann, T. (2016). Tree-based conditional portfolio sorts: The relation between past and future stock returns. Available at SSRN 2740751.

[2] Takeuchi, L., & Lee, Y. Y. A. (2013). Applying deep learning to enhance momentum trading strategies in stocks. In Technical Report. Stanford University.

[3] Batres-Estrada, B. (2015). Deep learning for multivariate financial time series.

[4] Wang, S., and Y. Luo. 2012. "Signal Processing: The Rise of the Machines." Deutsche Bank Quantitative Strategy (5 June).

[5] Paiva, F. D., Cardoso, R. T. N., Hanaoka, G. P., & Duarte, W. M. (2018). Decision-Making for Financial Trading: A Fusion Approach of Machine Learning and Portfolio Selection. Expert Systems with Applications.
[6] Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning for finance: deep portfolios. Applied Stochastic Models in Business and Industry, 33(1), 3-12.
[7] M. S. Hegde, G. Krishna and R. Srinath, "An Ensemble Stock Predictor and Recommender System," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Bangalore, 2018, pp. 1981-1985.
[8] Jordan Prosky, Andrew Tan, Xingyou Song, Micael Zhao. "Sentiment Predictability for Stocks" - arXiv:1712.05785v2 [cs.CL], 18 Jan 2018.
[9] Sepp Hochreiter; Jürgen Schmidhuber (1997). "Long short-term memory". Neural Computation.
[10] Adil Moghar, Mhamed Hamiche Stock Market Prediction Using LSTM Recurrent Neural Network (April-6 2020)
[11] Jiayu Qiu, Bin Wang, Changjun Zhou Forecasting stock prices with long-short term memory neural network based on attention mechanism (January 3, 2020).
[12] Nguyen Trung Tuan, Thu Hang Nguyen & Thanh Thi Hien Duong Stock Price Prediction in Vietnam Using Stacked LSTM (23 August 2022).
[13] Raut Sushrut Deepak, Shinde Isha Uday, Dr. D. Malathi, “Machine Learning Approach In Stock Market Prediction”, IJPAM 2017.
[14] M. Roondiwala, H. Patel and S. Varma, "Predicting stock prices using LSTM," International Journal of Science and Research (IJSR), vol. 6, no. 4, pp. 1754-1756, 20