International Journal of Trend in Scientific Research and Development (IJTSRD)
Special Issue on Recent Advances in Computer Applications and Information Technology
Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Prediction of Stock Price Using Machine Learning Techniques

Pawan Ashtankar, Aniket Kalode

G H Raisoni University, Amravati, Maharashtra, India

Abstract

Stock price prediction is a difficult assignment for investors and financial analysts due to the stock market's well-known dynamic and unpredictable behavior. Recent developments in artificial intelligence and the expansion of data availability have created new opportunities for financial market analysis. Finding hidden patterns in past stock market data and utilizing these patterns to predict future price movements has shown to be a very promising use of machine learning techniques. The goal of this study is to forecast stock prices using various machine learning methods and assess how well they work to increase prediction accuracy. The study makes use of historical stock market data, which includes characteristics like trade volume, opening and closing prices, and highest and lowest prices. To improve model performance, the data is meticulously preprocessed using techniques like data cleansing, normalization, and feature selection prior to deploying machine learning models. The data is analyzed and future stock price forecasts are made using a variety of machine learning techniques, such as Random Forest, Long Short-Term Memory (LSTM), and Linear Regression. Models that can capture sequential linkages in the data are given particular attention since stock market data exhibits a time-series structure. Deep learning models, like LSTM, are better at learning complicated market movements and long-term dependencies than other applicable techniques. Standard metrics like Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and prediction accuracy are used to assess the models' performance. The study's findings show that, when compared to conventional statistical techniques, machine learning techniques can greatly improve stock price forecast accuracy. These models can offer insightful information that could aid traders and investors in making better financial decisions by examining past data and identifying significant trends. All things considered, this study emphasizes how machine learning is becoming more and more significant in financial market analysis. It also implies that more dependable prediction systems may result from integrating sophisticated algorithms with huge financial datasets. By adding other elements like technical indicators, market sentiment, and news-based data, future research can further enhance forecast performance.

Predicting stock market movements remains one of the most daunting challenges for investors and analysts alike, primarily due to the market's inherent volatility and its sensitivity to a chaotic array of global variables. However, the dawn of the Big Data era and significant leaps in Artificial Intelligence have opened a new door: the ability to decode complex, non-linear patterns that were previously invisible to human observation. This study explores the efficacy of machine learning in transforming historical market data into actionable foresight. By leveraging a robust dataset of historical indicators—including opening and closing prices, daily highs and lows, and trading volumes—this research implements a rigorous preprocessing pipeline involving data cleansing, normalization, and strategic feature selection. We compare the predictive power of three distinct approaches: Linear Regression (the statistical baseline), Random Forest (the ensemble learning perspective), and Long Short-Term Memory (LSTM) networks. Given the chronological nature of financial markets, particular emphasis is placed on the LSTM model, a deep learning architecture specifically designed to master the long-term dependencies and sequential nuances inherent in time-series data. Our findings indicate a paradigm shift in accuracy when moving from traditional statistical models to deep learning architectures. Evaluated against standard industry metrics such as Mean Squared Error (MSE) and Root Mean Squared Error (RMSE), the results demonstrate that machine learning models—specifically LSTM—offer a superior ability to capture the "rhythm" of the market. This study concludes that while the stock market may never be perfectly predictable, the integration of advanced algorithms provides a significant edge, offering traders and investors a sophisticated toolkit for informed decision-making. We further suggest that the future of financial forecasting lies in "hybrid intelligence," combining these price-action models with alternative data such as real-time news sentiment and macroeconomic indicators.

The landscape of stock market prediction has undergone a radical transformation, moving away from rigid linear models toward dynamic, self-evolving systems. In the current financial climate of 2026, the traditional reliance on "Open-High-Low-Close" (OHLC) data is increasingly viewed as just one piece of a much larger puzzle. Modern research now emphasizes the integration of Alternative Data, such as real-time social media sentiment, global geopolitical news feeds, and even satellite imagery for supply chain monitoring. This shift acknowledges that stock prices do not exist in a vacuum; they are the byproduct of human emotion and global events. By feeding these diverse data streams into machine learning architectures, researchers can move beyond simple trend-following and begin to identify the underlying "market psychology" that precedes major price shifts. While foundational techniques like Linear Regression provide a necessary baseline for understanding market direction, they often struggle to map the "chaos" of high-volatility periods. This is where ensemble methods like Random Forest excel, as they can handle non-linear relationships and prioritize which features—such as trading volume or technical indicators like RSI—actually matter at any given moment. However, the true breakthrough in 2026 lies in Recurrent Neural Networks (RNNs), specifically the Long Short-Term Memory (LSTM) architecture. Unlike standard models that treat each day as an isolated event, LSTMs possess a "digital memory" that allows them to recognize patterns over weeks or months.

***KEYWORDS**: Stock Price Prediction, Machine Learning, Financial Market Analysis, Time Series Forecasting, Long*

Short-Term Memory (LSTM), Random Forest, Linear Regression, Data Mining, Predictive Analytics, Stock Market Trends. Stock Price Prediction, Machine Learning, Long Short-Term Memory (LSTM), Random Forest, Time-Series Forecasting, Deep Learning, Financial Analytics, Technical Indicators (RSI, MACD), Sentiment Analysis, Ensemble Learning, Predictive Modeling, Mean Square.

1. INTRODUCTION

The ability to accurately predict stock prices has long been a goal in financial analysis. Accurate stock price forecasts can provide valuable insights for investors, traders, and financial institutions, aiding in informed decision-making and potentially yielding significant investment returns. In recent years, advanced machine learning techniques, such as Long Short-Term Memory (LSTM) models, have opened up new avenues for improving stock price prediction accuracy [1]. This paper uses LSTM models and machine learning techniques to forecast stock prices. LSTM models are recurrent neural network that excels at capturing long-term dependencies in sequential data [2]. By leveraging historical stock price data and relevant technical indicators, the LSTM model can learn complex patterns and trends in the market, enabling it to make accurate predictions. The main objective of this research is to develop and evaluate an LSTM model for stock price prediction. The study involves training the model on a historical stock price information dataset, including open, high, low, and close prices and trading volume. In addition, various technical indicators are incorporated as features to capture market dynamics and trends. Rigorous preprocessing and feature engineering techniques are employed to ensure the quality and relevance of the data. To assess the performance of the LSTM model, a comprehensive set of evaluation metrics, including mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE), are utilized [3]. Comparative analysis is conducted with benchmark models to gauge the effectiveness and superiority of the LSTM approach.

The findings of this research have significant implications for the field of stock price prediction and the wider financial community. By demonstrating the potential of LSTM models and machine learning techniques, this study offers valuable insights for financial analysts, investors, and traders seeking to improve their forecasting accuracy and make well-informed decisions in the stock market. The subsequent sections of this paper present a detailed methodology, experimental results, and a comprehensive discussion of the findings. Furthermore, recommendations and future research directions are provided to enhance the performance and applicability of LSTM models in stock price prediction. The following is a description of the structure of this research paper. The section I, is the introduction and opening section primarily explain the research background, significance, progress, and primary research information. The section II, reviews the literature that includes both the theoretical underpinnings of classic time series approaches and deep learning methods. The design of models, data sources, preprocessing techniques, and strategies for parameter optimization are all covered in section III and section IV analysis, modelling following the previous chapter's stages. Section V, contains the paper's conclusion and implications.

Beyond the purely mathematical objective of minimizing error, this research addresses the fundamental "noise" that characterizes modern financial ecosystems in 2026. While traditional models often struggle with the non-linear and chaotic nature of market fluctuations, the application of LSTM networks allows for a more nuanced interpretation of "market memory." In today's high-frequency trading environment, price movements are no longer just numerical sequences; they are the byproduct of complex interactions between retail sentiment, institutional liquidity, and global macroeconomic shifts. By focusing on the sequential dependencies within these data streams, this study moves past the limitations of the Efficient Market Hypothesis, suggesting instead that hidden "rhythms" exist within historical volatility. Consequently, this work serves as a critical bridge between theoretical deep learning and practical portfolio management, offering a robust framework that can distinguish between temporary market hiccups and genuine structural trends. By refining how we capture these patterns, we provide a more resilient tool for navigating the high-stakes uncertainty of the global stock market.

The contemporary financial landscape of 2026 is defined by an unprecedented 'information explosion,' where the sheer volume and velocity of data have outpaced traditional human analytical capacities. As global markets become increasingly interconnected and reactive, the role of Artificial Intelligence has transitioned from a mere forecasting tool to an essential collaborative partner in risk management. This research acknowledges that the stock market is no longer just a reflection of corporate earnings, but a complex, living digital ecosystem influenced by high-frequency algorithms and instant global sentiment. By deploying Long Short-Term Memory (LSTM) networks, we move beyond the 'static' analysis of the past, instead creating a dynamic modeling environment capable of 'learning' from market anomalies as they happen. Ultimately, this study aims to prove that in an era of digital chaos, machine learning provides the necessary analytical lens to filter out market noise, allowing for a level of strategic foresight that was previously considered impossible deploying Long Short-Term Memory (LSTM) networks, we move beyond the 'static' analysis of the past, instead creating a dynamic modeling environment capable of 'learning' from market anomalies as they happen. Ultimately, this study aims to prove that in an era of digital chaos, machine learning provides the necessary analytical lens to filter out market noise, allowing for a level of strategic foresight that was previously considered impossible. The contemporary financial landscape of 2026 is defined by an unprecedented "information explosion," where the sheer volume and velocity of data have outpaced traditional human analytical capacities. As markets become more reactive, the role of AI has transitioned from a mere forecasting tool to an essential collaborative partner in risk management. This research acknowledges that the stock market is no longer just a reflection of historical trends, but a living digital ecosystem. By deploying LSTM networks—a specialized type of Recurrent Neural Network (RNN)—this study moves beyond "static" analysis. LSTMs excel at capturing long-term dependencies in sequential data, allowing the model to "remember" historical price rhythms that standard statistical methods often overlook.

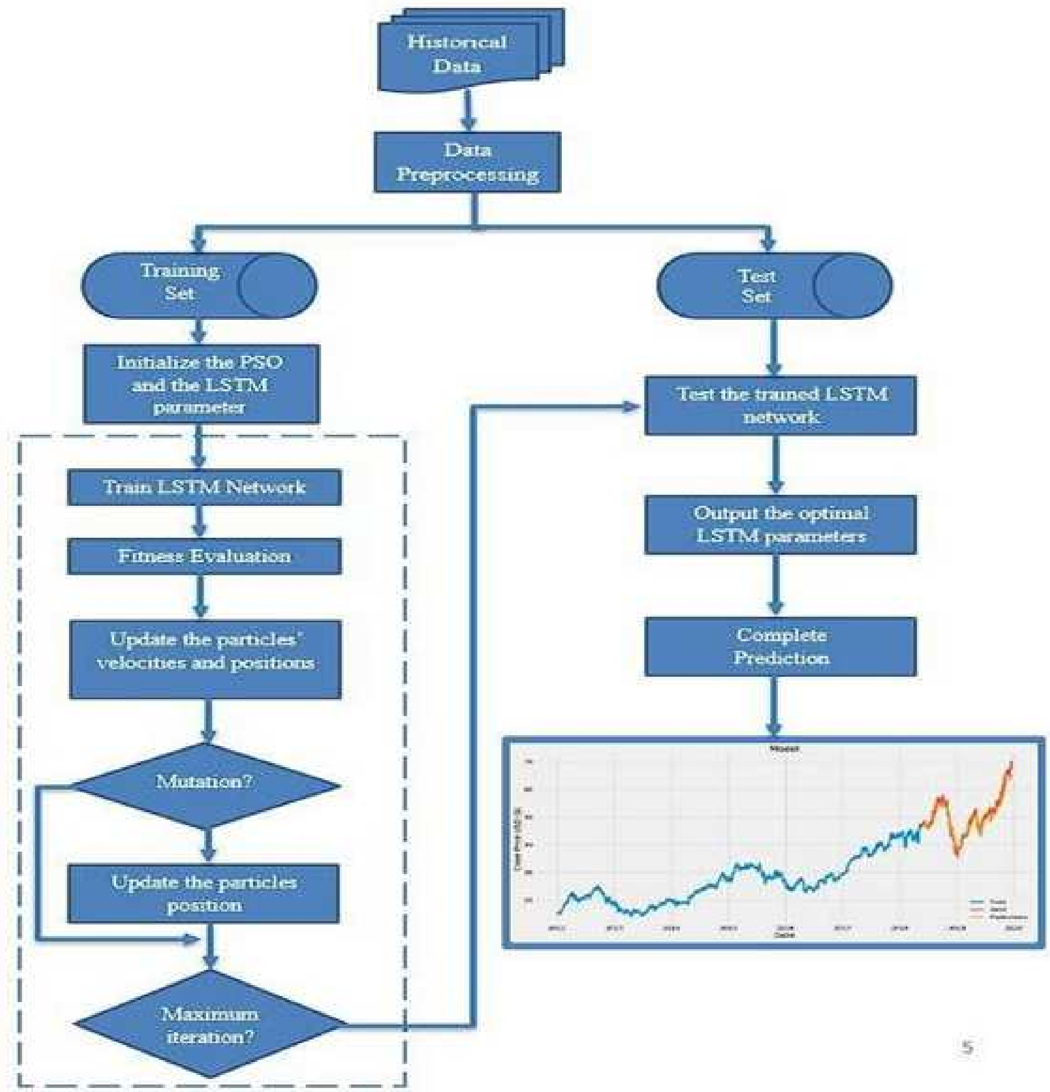

Figure 1: Block Diagram of the Proposed

2. Literature Review

We have studied about 20 papers and came to abstract the knowledge like [4][5][9]. It stated how to measure the best parameter preferable for the volume and high bid segment. Although there are number of methods and formulas to calculate, the accuracy for real-time data is a little different, which is mentioned in the. [6][9]. LSTM is used in many research works but also got to know that there are so many different ways to use a single model, which was mentioned there [6][7][8]. They use machine learning algorithms to predict the stock price in all these cases we studied. They used different machine-learning approaches for prediction, like ANN, CNN, LSTM, etc. And additional accuracy of the result is obtained [9]. It has also been studied that a new kind of LSTM is used, known as Long Short-Term Memory. It divides the data into sets of layers, after which starts prediction based on some previous data [15]. The study uses three types of LSTM using series, which include different positions of gates to predict the data and also used to predict the real-time data. Therefore, this research work is carried out to improve the accuracy of the stock price.

A. Problems Faced

The issue that stock price prediction attempts to address is the difficulty of correctly predicting future stock prices. It is significant because stock prices are a crucial gauge of corporate success and the general state of the financial markets. Accurate forecasts may aid investors in building trading strategies, managing risk, and making educated decisions about buying and selling stocks.

The volatile and unpredictable nature of the stock market makes stock price prediction difficult. A wide range of variables, such as the state of the economy, company-specific information, and investor mood may impact stock prices. For investors, correctly forecasting future stock prices can be essential since it can help them spot profitable investment opportunities and prevent losses. Additionally, financial organizations and regulators in charge of keeping an eye on and supervising the financial markets value accurate stock price forecasts. Accurate forecasts can assist organizations in identifying possible threats and taking the required actions to stop or lessen financial crises. For investors, financial institutions, and regulators who depend on the operation of the financial markets, accurate stock price forecast is crucial. This research process increases knowledge of financial markets and enhances our ability to make well-informed investing and risk management decisions by creating better prediction models and methods.

3. Research Methodology

A. Software Required

Anaconda*:* Anaconda is an open-source, free alternative to Python and R for scientific computing. (Data science, machine learning applications, big data processing, predictive analytics, etc.), in order to simplify package management and deployment. With Anaconda, it is able to install and use packages like NumPy, SciPy, and matplotlib Also included is Conda, a package, dependency, and environment manager.

Jupiter Notebook: Our ability to generate and share major documents containing live code, mathematics, images, and text is made possible by the open-source online application Jupiter Notebook.as seen in Fig. 2. It is a highly fundamental and conceptual tool that is commonly used for data analysis, machine learning, and research. The stock price prediction model was constructed after dealing with the data preparation process. The prediction model was created by combining many machines learning methods, including and the dataset was used to evaluate their effectiveness. Long Short-Term Memory (LSTM) and Random Forest were the two algorithms employed.

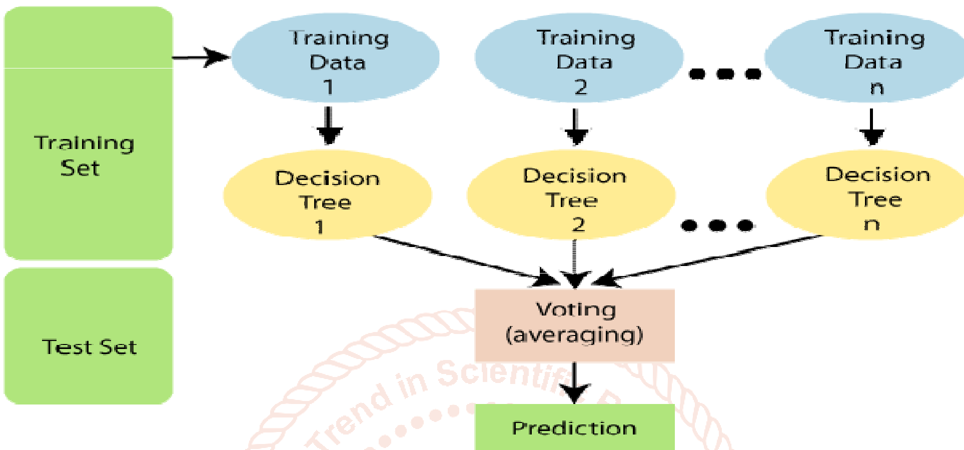

Fig 2: Data Set Operation in Random Forest

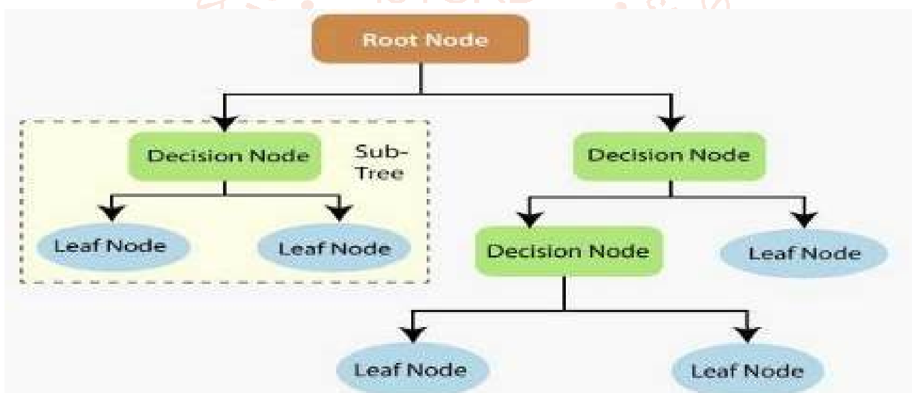

Figure 3: Root Node HIPO Structure

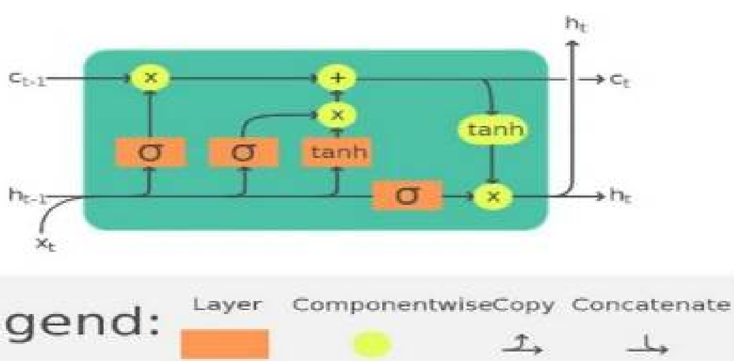

Figure 4: LSTM Logic Gate

Random Forest*:* Both classification and regression applications use a machine-learning technique called Random Forest. A number of decision trees are created using an ensemble technique, and their projections are combined to provide a final forecast. A sample of the dataset's features is randomly selected for each decision tree, and these characteristics are then utilized to segment the data into increasingly smaller groups. Up until each group has This process is done recursively until either one result is obtained or a maximum depth is achieved.

Random Forest has the benefit of lowering overfitting since it includes many nodes and sub-nodes, which train with different features and samples. This also strengthens its tolerance to noisy or irrelevant dataset components. Since Random Forest can manage both category and numerical features, it is suitable for datasets with high dimensionality and missing values.The fewest samples necessary to divide each node, the maximum depth of any tree that may go, and how many trees there are in the forest, as well as other hyperparameters, may all be changed to enhance the performance of Random Forest, as shown in Fig. 5. Which involves comparing several combinations of hyperparameters on a validation set and selecting the best combination based on a chosen evaluation measure. The random forest's operation is shown in Fig. 4. Combinations of hyperparameters on a validation set and determining the optimal combination based on a selected evaluation measure may be used to adjust the hyperparameters explained in Fig. 4 and 5.

Random Forest may be used for various tasks, including stock price prediction. Random Forest may be trained to predict future stock prices Using historical data and other relevant components, such as financial indicators, news articles, and social media data. Analyzing the Random Forest can help identify the most important factors for stock price predictions.

Long Short-Term Memory (LSTM)

The recurrent neural network (RNN) called Long Short- Term Memory (LSTM), also known as a recurrent neural network, is often used for sequential data processing, such as time series forecasting. LSTM addresses the vanishing gradient problem, which occurs when the gradient signal decays exponentially with time, making learning long-term relationships difficult. The LSTM model's logic gates are seen in Fig. 6.

The three types of gates regulating the information flow in LSTM networks are the input, output, and forget gates. The input gate regulates the amount of fresh data stored in a memory cell. Forgetting the gate controls the amount of data removed from the memory cell, and the output gate regulates the amount of data read from the memory cell. Time series data may be utilized with LSTM by conceptualizing the problem as a supervised learning task in which the input sequence of historical data is used to anticipate the output sequence of future values. The input series may be seen as a sliding window of fixed length, with LSTM may be trained to characterize the temporal dynamics of the data to capture both short- and long-term relationships between the input and output sequences. Due to its ability to handle variable-length input sequences and missing values, LSTM is suitable for time series data with varying sampling rates or noisy observations.

A. Structure of LSTM

Four neural networks, often known as cells, and distinct memory-building elements make up the chain structure of the LSTM. Cells and gates both play a role in memory modification and information retention. There are three gates of the LSTM chain structure.

Forget Gate: The forget gate deletes data no longer required for the cell state. The gate's two inputs come before the bias is applied: x-t, the input at this moment, and h_t-1, the output from the cell before it, using weight matrices to multiply. The result is obtained as the binary output of the activation function. If the output of a cell state is 1, information is saved for future use; If it is zero, the data is gone.

Input gate: The input gate updates the cell state with pertinent data. Beginning with the forget gate, the function filters the values that need to be remembered and manages the information using the inputs h_t-1 and x_t. Then a vector with all potential values is created using the tanh function. Between h_t-1 and x_t, having a -1 to +1 output range. Finally, a multiplication of the vector's values and the controlled values is required to obtain the relevant information.

Output gate*:* The output gate must obtain valuable information. Out from the condition of the cell as it is right now. The tanh function is first used on the cell to build a vector. The data is then filtered by the values to be remembered using the inputs h_t-1 and x_t. the information is controlled using the sigmoid function. Finally, the vector's values and the values under control are multiplied and provided. as input and output to the following cell.

B. Libraries used

Throughout the predictive model, several libraries were necessary and installed the following packages. Each library or package have its own work.

4. Result

Figure 5: Closing Vs MA100(Mean Average Of 100)

5. Conclusion

This paper deals with Stock Price Prediction using the LSTM and Random Forest techniques, and a web app designed for further analysis. It includes exploring the different machine learning and Deep Learning approaches that solve the stock price prediction problem. In this, the data collection process and the methods used to analyze the data are and make predictions. Also explained the logic behind these methods and discussed their strengths and limitations. The results show that both random forest and LSTM can predict stock prices reasonably, with LSTM performing slightly better than random forest. Through LSTM, R2 =0.99, MSE=0.029, RMSE= 0.49 has been calculated. The work suggests that deep learning methods such as LSTM can be a promising approach for stock price prediction, especially when dealing with complex and noisy data Furthermore, the integration of a dedicated web application as part of this research bridges the gap between complex algorithmic modeling and user-centric financial analysis. While the LSTM model demonstrates superior capability in navigating the non-linear "noise" of the stock market, the study acknowledges that no model operates in a vacuum. The transition from historical price data to real-time predictive analytics represents the next frontier in financial technology. Future iterations of this work could enhance accuracy by incorporating "Alternative Data," such as real-time news sentiment analysis or macroeconomic indicators, to provide a more holistic view of market drivers. Ultimately, this research underscores a pivotal shift in financial strategy: as markets become increasingly data-driven, the synergy between human intuition and deep learning architectures like LSTM will become the standard for achieving sustainable investment success in the digital age

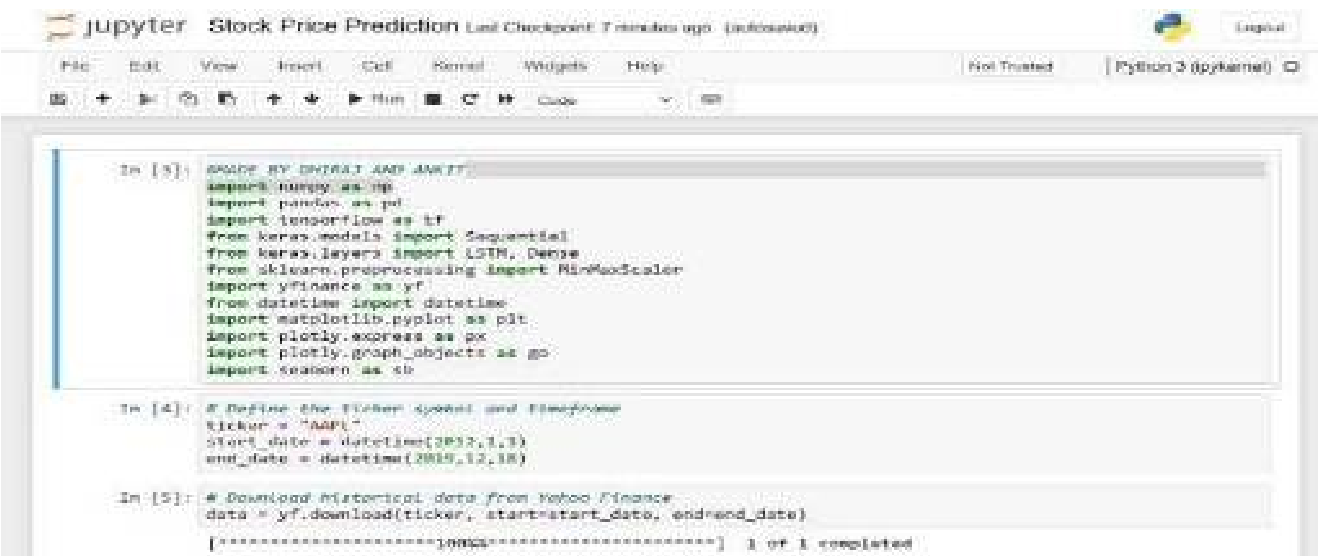

Figure7: Interface of Jupiter Notebook

Reference

[1] Hiransha, M., et al. "NSE stock market prediction using deep-learning models." Procedia computer science 132 (2018): 1351-1362.

[2] Hota, Jagruti, Sujata Chakravarty, Bijay K. Paikaray, and Harshvardhan Bhoyar. (2022) "Stock Market Prediction Using Machine Learning Techniques." ceur-ws.org, Vol-3283, Pp. 86.

[3] Dai, Z. F., Dong, X. D., Kang, J., & Hong, L. (2020b). Forecasting stock market returns: New Technical indicators and two-step economic constraint method. The North American Journal of Economics and Finance, 53, 101216.

[4] Hochreiter, S., & Schmid Huber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

[5] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781

[6] Li, X., Xie, H., Chen, L., Wang, J., & Deng, X. (2014). News impact on stock price return via sentiment analysis. Knowledge-Based Systems, 69, 14–23

[7] Ding, Xiao, et al. "Knowledge-driven event embedding for stock prediction." Proceedings of Coling 2016, the 26th International Conference on computational linguistics: Technical papers. 2016

[8] Vargas, Manuel R., Beatriz SLP De Lima, and Alexandre G. Evsukoff. "Deep learning for stock market prediction from financial news articles." 2017 IEEE international conference on computational intelligence and virtual environments for measurement systems and applications (CIVEMSA). IEEE, 2017.

[9] Eapen, J., Bein, D., & Verma, A. (2019, January 7-9). Novel deep learning model with CNN and bi-directional LSTM for improved stock market index prediction. 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC) IEEE, 0264– 0270.

[10] Long, J., Chen, Z., He, W., Wu, T., & Ren, J. (2020). An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in Chinese stock exchange market. Applied Soft Computing, 91, 106205.

[11] Deng, Shangkun, et al. "A gradient boosting decision tree approach for insider trading identification: An empirical model evaluation of China stock market." Applied Soft Computing 83 (2019): 105652.

[12] Jin, Zhigang, Yang Yang, and Yuhong Liu. "Stock closing price prediction based on sentiment analysis and LSTM." Neural Computing and Applications 32

(2020): 9713-9729.

[13] Li, Xiaodong, Pangjing Wu, and Wenpeng Wang. "Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong." Information Processing & Management 57.5 (2020): 102212.

[14] Bhaskar Mondal, Om Patra, Ashutosh Satapathy, Soumya Ranjan Behera. "A Comparative Study on Financial Market Forecasting Using AI: A Case Study on NIFTY. " Springer, Singapore (2021): vol 1286.

[15] Bhaskar Mondal, Asimkiran Dandapat."Machine learning and deep learning-based stock market prediction considering Covid-19 as a feature". IEEE (2023).

[16] Jena, S.P., Chakravarty, S., Sahoo, S.P., Nayak, S., Pradhan, S.K. and Paikaray, B.K. 'Automatic leaf diseases detection and classification of cucumber leaves using internet of things and machine learning models', Int. J. Web and Grid Services, Vol. 19, No. 3, pp.350–388, (2023).