International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Working of Stock Price Prediction Using LSTM

Prof. Mrs. Sampada A. Kulkarni, Asst. Prof., Modern College of Engineering
Shubham Gurav, Aditya Lahade, Deepak Gudavalekar, Nagnath Sangale

Dept. of I.T., P.E.S Modern College of Engineering, Savitribai Phule Pune University, Maharashtra, India

Abstract *:*Our proposal centers on the development of a robust tool designed to forecast stock prices by leveraging a sophisticated combination of advanced techniques. We're merging spatial features, which essentially encapsulate the distinctive attributes of individual stocks, with the powerful capabilities of long-and-short-term memory networks (LSTM). These networks excel at recognizing and interpreting patterns and trends over extended periods. By integrating these two elements, our model gains the ability to capture both short-lived fluctuations and enduring trends within stock prices.This fusion of spatial features and LSTM networks represents a significant advancement in stock price prediction methodologies. We're confident that this innovative approach will surpass existing methods in terms of accuracy and reliability. To validate our confidence, we're conducting rigorous testing and refinement of the system using diverse sets of data. This ensures that our model performs effectively across various market conditions and scenarios, thereby instilling trust among investors seeking dependable insights into stock market trends.

Key Word*: Long and Short-Term Memory, Time-Series Data, Yahoo Finance.*

I. Introduction

The stock market is a regulated marketplace where investors can buy or sell stocks publicly or privately. Companies often turn to the stock market to raise capital for business expansion, making it a popular investment option for investors. To make informed investment decisions, many investors rely on predictions based on past market trends. In recent years, the stock market has changed significantly,[4] making it crucial to anticipate its future value or price given how dynamic the market is. Predictive analysis is frequently carried out using LSTM model, a mathematical technique and deep learning method. The LSTM model, which learns patterns and relationships in historical stock price data., is largely consistent with the continuous/real values of mathematical variables. The algorithm makes the predictions in accordance with the guidance provided in the training data. Set after first training with it. The other tools and technology used in the given website are html and CSS which are used for frontend. The library named scikit learn is used as it has all machine learning algorithms and functions inbuilt in it. The framework Django which comes under Python is used at the backend to access web services and resources which is inbuilt in python programming language. To fetch the data for the website which should be reliable and accurate,[4] Yahoo finance site is used and connecting

it with API. So the prediction of stocks of all the companies which are there in yahoo finance can be done through our website.

II. Inputs

1. User Inputs:

Stock label name: This is the identifier for the stock you want to predict the price of. For example, you might input "AAPL" for Apple or "GOOG" for Google.
Number of days: This specifies the number of days into the future you want to predict the stock price for. For example, you might input "5" to predict the price for five days from now.

2. Dataset:

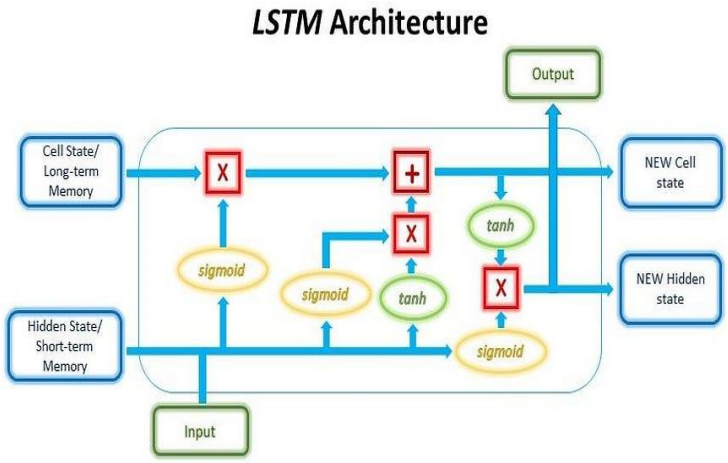
The yfinance library is a popular tool for downloading stock price data from Yahoo Finance. The dataset typically includes several key features, each representing different aspects of stock market activity. Here are the common columns available in the dataset along with their meanings:
Date: The specific trading day for the data entry. Each row corresponds to one trading day.
Open: The price at which the stock started trading at the beginning of the trading day.
High: The highest price at which the stock traded during the trading day.
Low: The lowest price at which the stock traded during the trading day.
Close: The price at which the stock finished trading at the end of the trading day. This is often used for calculating returns and technical indicators.
Adj Close: The adjusted closing price, which accounts for corporate actions like stock splits, dividends, and rights offerings. This provides a more accurate reflection of a stock's value over time.
Volume: The number of shares traded during the trading day. This indicates the trading activity and liquidity of the stock.

These features collectively provide a comprehensive view of a stock's daily trading activity. The opening and closing prices help understand market trends and investor sentiment. The high and low prices indicate the volatility and trading range for the day. The adjusted close ensures that historical price comparisons are accurate, accounting for corporate actions. Finally, the volume shows the level of trading activity, which can signify the strength or weakness of a price move. Using this detailed data, models like LSTM can effectively learn patterns and predict future stock prices.

3. Data Model: This step involves creating a mathematical representation of the stock price data. This typically involves converting the data into a series of numerical features that can be fed into the LSTM algorithm. The data model for stock price prediction using Long Short-Term Memory (LSTM) networks involves structuring historical stock price data into sequences that the model can learn from. The dataset typically includes features such as the opening, high,

low, and closing prices, as well as trading volume. These features are normalized to ensure that they are on a similar scale, enhancing model performance. The data is then segmented into windows or time steps, where each sequence represents a fixed number of previous trading days (e.g., 60 days) used to predict the stock price for the next day. This time-series format allows the LSTM to capture temporal dependencies and patterns in the stock prices. The model is trained on these sequences, using the LSTM layers to retain and learn from long-term dependencies, ultimately providing accurate predictions for future stock prices based on historical trends.

III. Working and Architecture of Model



Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) known for its ability to learn from sequential data and capture long-term dependencies. In the context of stock price prediction, in the following manner LSTM work.

1. Temporal Data Sequencing: The historical stock price data obtained from Yahoo Finance is inherently sequential, where each data point represents the price of a stock at a specific time. LSTM is adept at processing such sequential data, as it can learn from past price movements to predict future trends. Temporal data sequencing is a crucial step in preparing stock price data for prediction using Long Short-Term Memory (LSTM) networks. This process involves organizing historical stock price data into a series of overlapping windows or sequences, enabling the LSTM model to capture and learn from temporal dependencies in the data. Each sequence, or time step, consists of a fixed number of previous trading days' data points used as input features to predict the stock price for the subsequent day. The data for each sequence includes relevant features such as opening price, closing price, high price, low price, and trading volume. These features are normalized to ensure they are on a similar scale, enhancing the model's performance and convergence. The data is then reshaped to fit the LSTM input requirements, typically in the format of [samples, time steps, features], where 'samples' represent the number of sequences, 'time steps' represent the length of each sequence, and 'features' represent the number of attributes used in each time step. Creating sequences involves slicing the historical stock price data into overlapping windows.

2.Feature Selection and Preprocessing: Before feeding the data into the LSTM model, it undergoes preprocessing steps. This typically involves selecting relevant features (e.g., closing prices) and normalizing the data to a common scale. Normalization ensures that all features contribute equally to the model's learning process, regardless of their original scales.

The closing price is selected as a key feature for stock price prediction because it represents the final price at which a stock trades during regular market hours, reflecting the cumulative effect of all market activities, news, and investor sentiments throughout the day. It is widely regarded as the most reliable and significant price point of the trading day, often used as a benchmark for analyzing historical performance and making future predictions. The closing price is crucial for technical analysis, forming the basis for various indicators like moving averages and trends, which are essential for identifying market patterns and forecasting future price movements. Consequently, using the closing price as a feature helps in capturing the essential market dynamics needed for accurate stock price prediction.

3.Long-Term Dependency Learning: LSTM is well-suited for capturing long-term dependencies in sequential data, which is crucial for stock price prediction. It achieves this by incorporating memory cells within its architecture, allowing the model to retain important information over extended periods. In the context of your project, these memory cells enable the LSTM model to remember significant past price movements, market trends, and other relevant factors that influence stock prices.

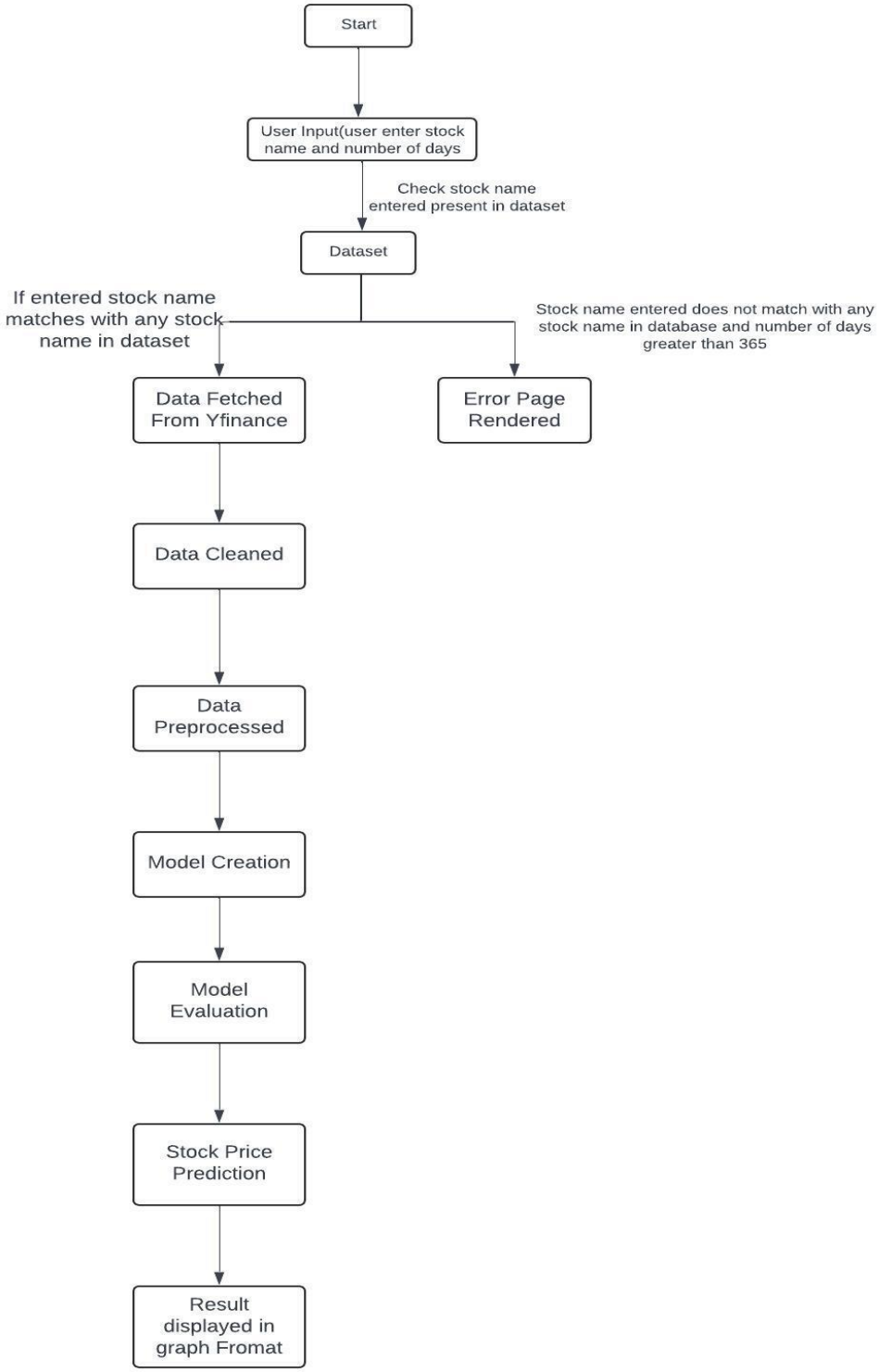
4.Gating Mechanisms: LSTM employs gating mechanisms, including input, forget, and output gates, to regulate the flow of information within the network. These gates control how much information from past price data is retained or discarded at each time step, enabling the model to focus on relevant patterns while filtering out noise and irrelevant signals. This mechanism is particularly valuable in stock price prediction, where the model needs to discern meaningful trends amidst the inherent volatility and randomness of financial markets.

5.Sequence Learning and Prediction: The LSTM model processes the sequential data, learning patterns and relationships between past and future stock prices. By analyzing historical price movements, the model learns to recognize recurring patterns, market trends, and other factors that influence stock prices. Leveraging this learned knowledge, the LSTM model generates predictions about future stock prices, aiming to provide insights into potential market trends and investment opportunities.

IV. Working of Stock Market Prediction Website

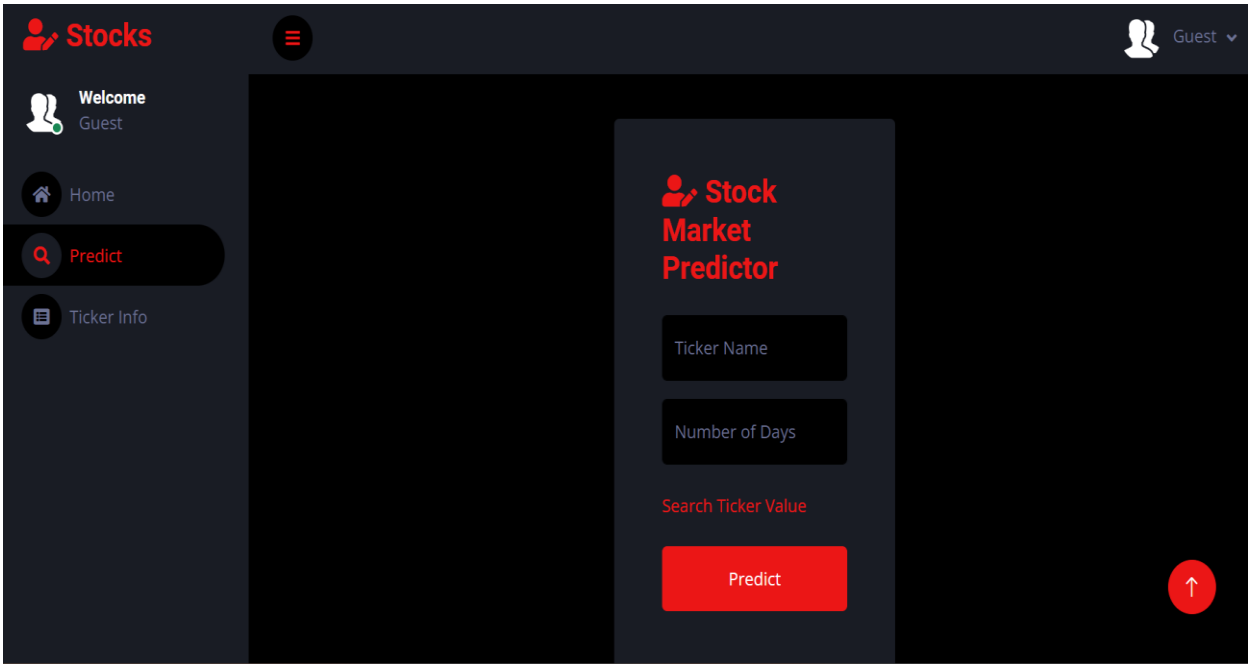
In order to develop GUI (graphical user interface) to make this predictor tool accessible to users, predictor is converted to a website which will show real time stock market prediction. The first step in this is to collect and fetch the data from reliable source site Yahoo- finance and creating a DBMS (data base management system) using SQLite. Data preprocessing is yet another important step in which inaccuracy or inconsistencies in data had been removed and data is transformed into user understandable format using pandas library in python. In the backend part numPy, plotly,

matplotlib are used for multiple operation execution. The prediction part is performed using LSTM model of deep learning. To access the deep learning tools and resources for the website, scikit learn library and Django framework are used respectively. For UI/UX part Html5 and CSS3 versions are used. The website is maintained at regular intervals and fetch new and current status of data every time.

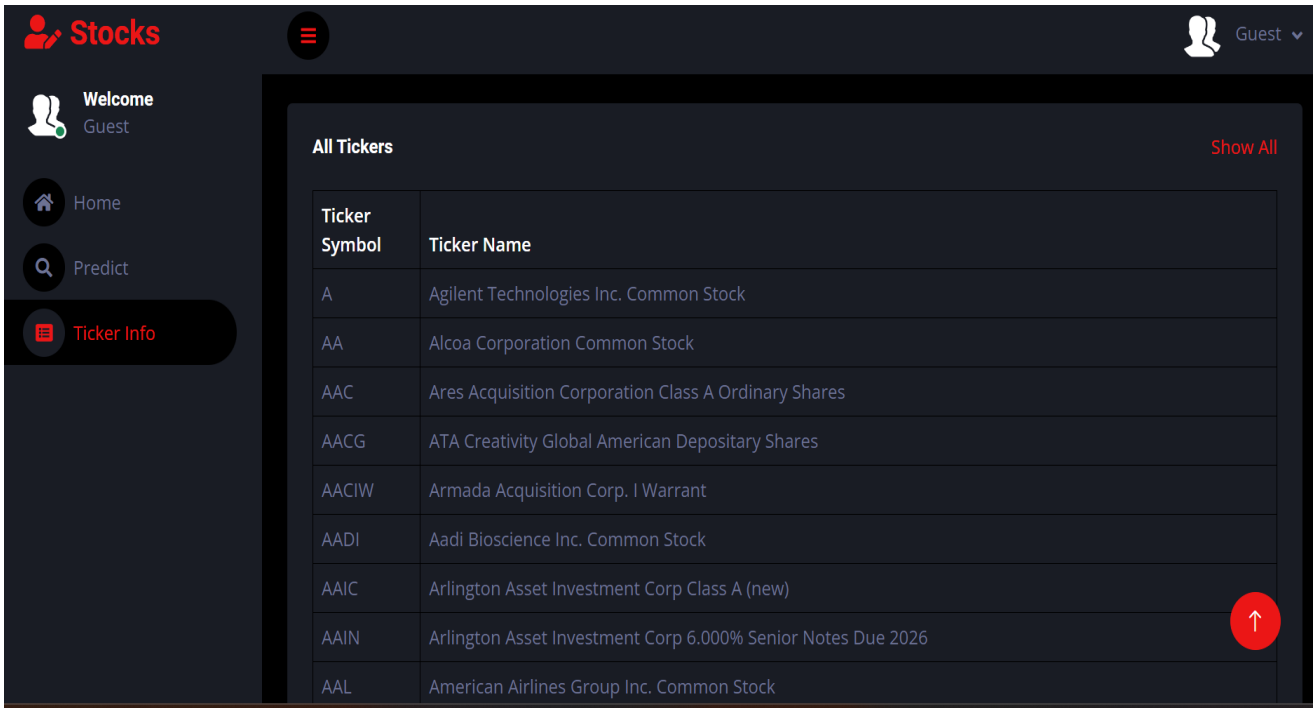


V. Screenshots

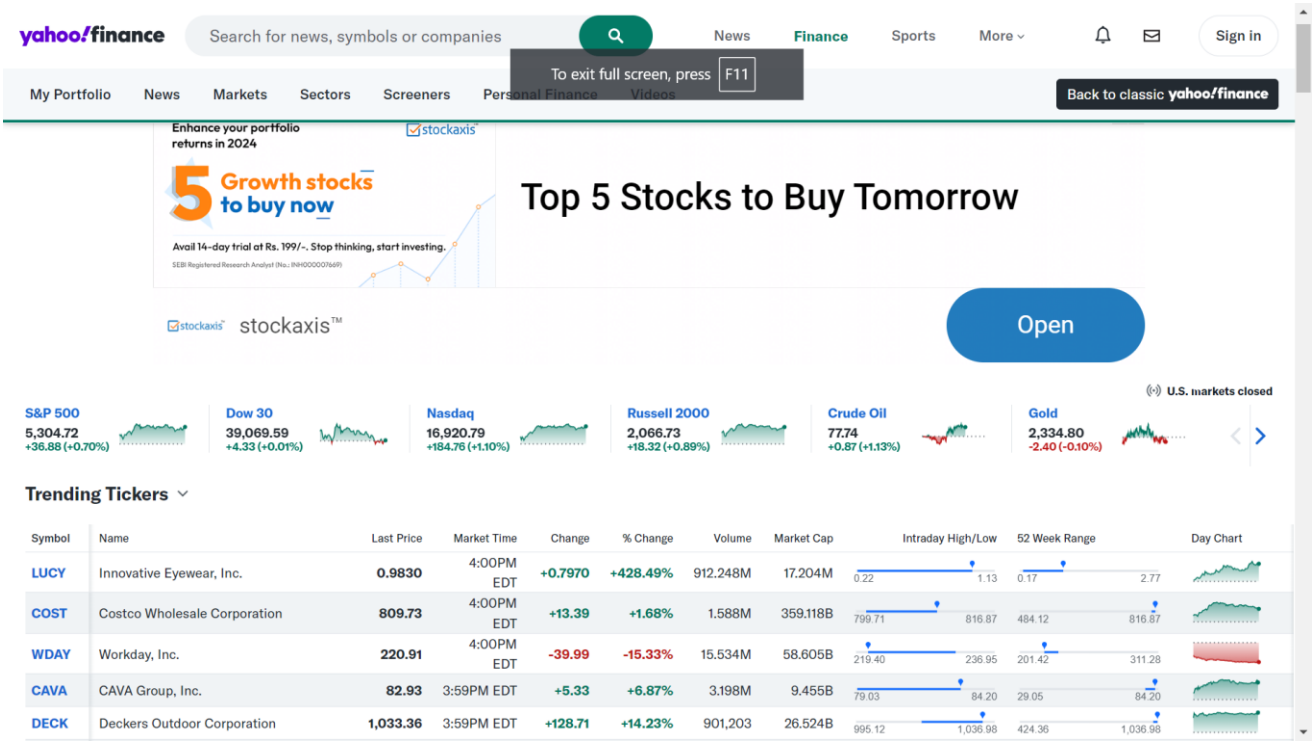
1. Home Page



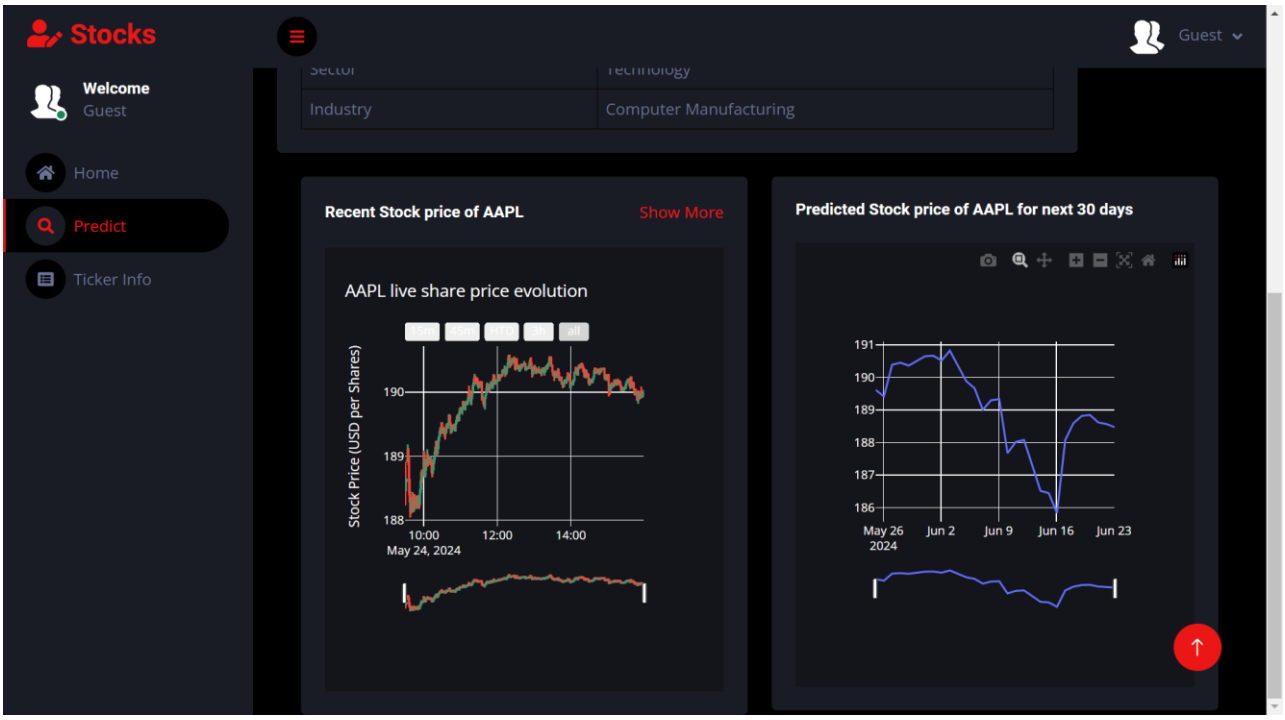
2. Predict Page



3. Stock Info Page



3. Yahoo Finance Page



4. Predicted Output Page

VI. Future Work

Hybrid Architectures: Investigate hybrid models combining the strengths of both LSTM and CNN architectures. Explore how temporal dependencies captured by LSTM and spatial features extracted by CNN can be synergistically utilized for more accurate predictions.

Data Enhancement Strategies: Explore methods for incorporating alternative data sources to enrich the analysis.

Investigate techniques to handle missing or noisy data for improved robustness.

Interdisciplinary Collaboration: Promote collaboration between finance experts, machine learning researchers and domain-specific knowledge holders. Aim to identify new features and indicators that can enhance the effectiveness of stock price prediction models.

Ethical Considerations: Address ethical considerations in the application of predictive models in financial settings. Develop frameworks for responsible and fair use, focusing on transparency and accountability.

Explainability and Interpretability: Enhance model explainability for better understanding of the reasoning behind predictions. Explore techniques to make the models more interpretable, especially in complex financial scenarios.

Optimization Strategies: Investigate optimization strategies to fine-tune and improve the performance of CNN and LSTM models. Focus on reducing overfitting and improving generalization to unseen data.

VII. Conclusion

Many individuals aim to predict future stock prices to enhance their wealth, given the ever-changing nature of the stock market. However, the market's volatility has made it challenging for existing solutions utilizing advanced technologies like Deep Learning, AI, and Neural Networks to achieve precise predictions. Therefore, this review focuses on analyzing current methods for stock market prediction using Long-Short Term Memory (LSTM) a deep learning model, learns patterns and relationships in historical stock price data. Research has highlighted the importance of selecting an appropriate dataset for effective stock market prediction with LSTM. This paper presents a comprehensive approach for stock price prediction using a LSTM architecture. This method uses opening price, highest price, lowest price, closing price, adjacent close. LSTM is used to learn the extracted feature data and predict the closing price of the stock the next day. We demonstrated the effectiveness of this approach through a case study, providing insights into the advantages of this deep learning technique for stock price prediction. The paper contributes to the existing literature by offering a comprehensive review and case study of this deep learning approach.

VIII. References

1. Sampada A. Kulkarni, Shubham Gurav, Aditya Lahade, Deepak Gudavalekar, Nagnath Sangale (2024) - "Stock Price Prediction Using LSTM" https://fdrpjournals.org/indjcst/archives?paperid=1024737848930253539

2. J. Sen and T. Datta Chaudhuri, "An alternative framework for time series decomposition and forecasting and its relevance for portfolio choice - a comparative study of the Indian consumer durable and small-cap sector," Journal of Economics Library, vol. 3, no. 2, pp. 303 - 326, 2016.

3. J. Sen and T. Datta Chaudhuri, "Understanding the sectors of Indian economy for portfolio choice," International Journal of Business Forecasting and Marketing Intelligence, vol. 4, no. 2, pp. 178-222, 2018

4. Stock Price Prediction Website Using Linear Regression - A Machine Learning Algorithm by Sonali Antad, Saloni Khandelwal, Anushka Khandelwal, Rohan Khandare, Prathamesh Khandave, Dhawal Khangar, Raj Khanke Department of Engineering, Sciences and Humanities (DESH)

5. S. Mehtab and J. Sen, “A time series analysis-based stock price prediction using machine learning and deep learning models”, Technical Report, No: NSHM_KOL_2020_SCA_DS_1, NSHM Knowledge Campus, Kolkata, INDIA. DOI: 10.13140/RG.2.2.14022.22085/2.

6. Y. Hu, "Stock market timing model based on convolutional neural network– a case study of Shanghai composite index," Finance& Economy, vol. 4, pp. 71–74, 2018.

7. Leung CKS, MacKinnon RK, Wang Y (July 2014) A machine learning approach for stock price prediction. In: Proceedingsofthe18thinternationaldatabaseengineering&applications symposium, pp 274–277.

8. T. Zheng, J. Farrish, and M. Kitterlin, "Performance trends of hotels and casino hotels through the recession: an ARIMA with intervention analysis of stock indices," Journal of Hos pitality Marketing & Management, vol. 25, no. 1, pp. 49–68, 2016.

9. S. Mehtab and J. Sen, "Stock price prediction using convolutional neural network on a multivariate time series", In Proceedings of the 3rd National Conference on Machine Learning and Artificial Intelligence (NCMLAI' 2020), New Deli, India, 2020.

10. Wall-Streeter Review 2024: A Free and Sleek Data Platform by Dillon Jacobs Updated on Jan 19, 2023.