IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

SYSTEMATIC ANALYSIS AND REVIEW OF STOCK PRICE PREDICTION USING MACHINE LEARNING

[1]Saiteja Kobbaji, [2]Varun Kammadanam, [3]Kishore Bhukya, [4]Akshaya Koripelly, [5]CH China Subbareddy

[1]Student, [2]Student, [3]Student, [4]Student, [5]Assistant Professor
[1]Computer Science Engineering,
[1]Joginpally BR Engineering College, Hyderabad, India

Abstract: Predicting stock market trends is one of the most complex and widely studied challenges, attracting researchers from fields like economics, finance, mathematics, and computer science. The highly volatile nature of the stock market makes traditional methods like time-series analysis or simple regression less effective. This paper takes a closer look at various stock market prediction models, including techniques like Recurrent Neural Networks (RNN) and Long Short-Term Memory Network (LSTM). Mainly focusing on equity market like shares of Tata, Adani from the National Stock Exchange, this study provides valuable insights into building more effective and accurate prediction systems.

I. INTRODUCTION

In this project, we'll use a **Long Short-Term Memory (LSTM)** model to predict stock prices. LSTM is a type of **Recurrent Neural Network (RNN)** that's really good at learning from patterns over time, especially when it comes to data that changes with time, like stock prices. Since stock movements are influenced by past trends, LSTM can look at the history of prices, highs, lows, and trading volume to understand long-term patterns and make better predictions about where the stock might go next. This approach helps us get more accurate forecasts for future stock price prediction.

In this project, we're boosting the accuracy of our stock price prediction by using the **50-day**, **100-day**, and **200-day exponential moving averages**. These averages help smooth out daily price ups and downs, making it easier to spot the overall trend. The 50-day average focuses on short-term trends, while the 100-day and 200-day averages capture longer-term movements. By adding these to our LSTM model, we're giving it more useful context to better understand the market's direction and improve the accuracy of the price predictions. These moving averages helps us to spot the trend and to know the bias of the market this will help the traders to trade smoothly.

This project focuses on predicting stock prices using an LSTM (Long Short-Term Memory) model, which is great for analyzing time-based data like stock prices. We look at daily stock data, including the open, close, high, low, adjusted close, and trading volume to predict future price movements. To make the predictions more accurate, we also use key technical indicators like the 50-day, 100-day, and 200-day exponential moving averages, which help us understand both short-term and long-term trends in the market. The goal is to create a reliable model that can predict stock prices more accurately.

II. NEED OF THE STUDY.

The study of stock market prediction is essential for enabling better investment decisions by helping investors predict trends, reduce risks, and improve returns. It addresses the challenge of market volatility by providing tools to navigate unpredictable price fluctuations. Advanced technologies like AI and sentiment analysis uncover patterns that traditional methods often miss, offering deeper insights into market behavior. By integrating real-time data from news, social media, and historical prices, these models allow faster and smarter decision-making. This not only benefits individual investors but also supports financial stability and economic growth. Additionally, such predictions empower policymakers and institutions to maintain balance in the financial ecosystem. The study bridges the gap between raw data and actionable insights, making trading more reliable and efficient. Ultimately, it's about transforming data into smarter investments for everyone.

2.1 Data and Sources of Data

The project uses **historical stock data** like prices and volumes from Yahoo Finance and **real-time data** from TimesofIndia.com using Python scripts. **News articles and social media posts** are analyzed for sentiment (positive, negative, or neutral). This combination of structured data (stock prices) and unstructured data (sentiments) creates a robust dataset for accurate predictions.

2.2 Theoretical framework

The framework combines **historical stock data** (like prices and volumes) with insights from **news and social media** to understand market sentiment. **Machine learning models** like neural networks and decision trees use this data to spot trends and make predictions. We also fetch **real-time data** to keep the analysis up-to-date. By mixing hard data with public sentiment, we get more accurate predictions. This approach aims to help investors make smarter, more informed decisions.

III. RESEARCH METHODOLOGY

The aim of this project is to study stock market trends and use machine learning models to predict future stock prices. By analyzing historical data and identifying patterns, we hope to provide valuable insights that can help investors make better decisions. The focus is on cleaning and preparing the data, building accurate models, and presenting the findings in a clear and understandable way.

3.1 Data Preprocessing

Before analyzing data, it's essential to clean and prepare it. Raw data often contains inconsistencies, missing values, or errors. During preprocessing:

- We clean and organize the data to make it usable.
- Handle missing or incorrect values.
- Use feature scaling to ensure all variables are on the same scale, which helps models work effectively.

3.2 Model Training and Validation

To build an accurate and reliable model:

- Cross-validation is used to split the data into training and validation sets, ensuring the model is tested on unseen data during training. This makes the model more reliable.
- Fine-tune the algorithm's parameters to improve its ability to learn and make predictions.

3.3 Testing and Final Adjustments

- The test data is kept separate and untouched during training to provide a fair evaluation of how the model performs on completely unseen data.
- Once predictions are made, the results are scaled back to their original range (like actual stock prices) to make them easier to interpret.

3.4 Visualization

- This process ensures the data is clean, the model is accurate, and the results are both reliable and easy to interpret.
- Finally, the results are visualized through charts or graphs, making it easier to understand the trends, patterns, and prediction derived from the data.

IV. ALGORITHMS

Algorithm: Stock prediction using LSTM
Input: Historic stock data
Output: Prediction of stock price using price variation

Step 1: Start.
Step 2: Data pre-processing after getting the historic data from the market for a particular share.
Step 3: Import the dataset to the data structure and read the open price.
Step 4: Do a feature scaling on the data so that the data values will vary from 0 and 1.
Step 5: Creating a data structure with 60 timestamps and 1 output.
Step 6: Building the RNN (Recurrent neural network) for Step 5 data set and Initialize the RNN by using sequential repressor.
Step 7: Adding the first LSTM layer and some Dropout regularization for removing unwanted values.
Step 8: Adding the output layer.
Step 9: Compiling the RNN by adding Adam optimization and the loss as mean_squared_error.
Step 10: Making the predictions and visualizing the results using plotting techniques.

IV. IMPLEMENTATION RESULT

4.1 Data Collection
In this project, we collected historical stock data, such as prices and trading volumes, from Yahoo Finance. For real-time updates, Python scripts were used to fetch live Sensex and Nifty data from websites like TimesofIndia.com. Additionally, we analyzed financial news and social media posts to gauge market sentiment, categorizing them as positive, negative, or neutral. By blending real-time stock data with sentiment analysis, we created a well-rounded dataset for more accurate predictions. This approach combines both data and emotions to better understand and forecast market trends.

4.2 Data Preprocessing
Normalize the data for LSTM.
Split the data into training and testing datasets.
Create time-series data (e.g., sequence of 60 days to predict the next day's

4.3 LSTM Model
Build an LSTM model using frameworks like TensorFlow or PyTorch.
Compile the model with a suitable loss function and optimizer.

4.4 Training
Train the model on the training dataset.

4.5 Evaluation
Predict stock prices on the test set and calculate metrics like Mean Absolute Error (MAE).

4.6 Visualization
Plot actual vs. predicted prices for evaluation.

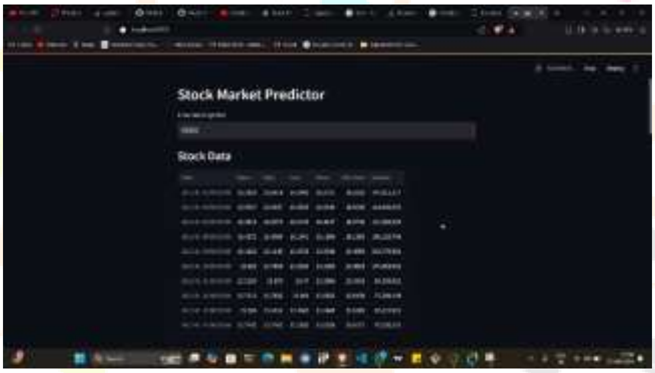

Figure 1: Stock Data Table

Figure 2 : Graph showing Price vs MA50

Figure 3 : Graph showing Price vs MA50 vs MA100

V. CONCLUSION

This project successfully demonstrated the potential of hybrid neural network modxels, specifically combining CNN and LSTM layers, for stock price prediction. The CNN effectively extracted key features, while the LSTM captured temporal patterns in the data. By incorporating dropout layers between the CNN and LSTM components, the model achieved improved robustness and minimized overfitting. After testing various configurations, the single-layer bi-directional CNN-LSTM and double-layer unidirectional CNN-LSTM models emerged as the most effective, showcasing superior prediction accuracy and generalization capabilities. Among these, the single-layer bi-directional CNN-LSTM stood out due to its balance of high accuracy and computational efficiency, making it a practical solution for real-time stock market analysis. This project highlights the value of hybrid CNN-LSTM architectures in financial forecasting, offering reliable and efficient tools for predictive analytics in dynamic market environments.

This project successfully demonstrated the potential of hybrid neural network modxels, specifically combining CNN and LSTM layers, for stock price prediction. The CNN effectively extracted key features, while the LSTM captured temporal patterns in the data. By incorporating dropout layers between the CNN and LSTM components, the model achieved improved robustness and minimized overfitting. After testing various configurations, the single-layer bi-directional CNN-LSTM and double-layer unidirectional CNN-LSTM models emerged as the most effective, showcasing superior prediction accuracy and generalization capabilities. Among these, the single-layer bi-directional CNN-LSTM stood out due to its balance of high accuracy and computational efficiency, making it a practical solution for real-time stock market analysis. This project highlights the value of hybrid CNN-LSTM architectures in financial forecasting, offering reliable and efficient tools for predictive analytics in dynamic market environments.

VI. ACKNOWLEDGMENT

REFERENCES

[1].Mehtab, S., Sen, J., Dutta, A. (2021). Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models. In: Thampi, S.M., Piramuthu, S., Li, KC., Berretti, S., Wozniak, M., Singh, D. (eds) Machine Learning and Metaheuristics Algorithms, and Applications. SoMMA 2020. Communications in Computer and Information Science, vol 1366. Springer, Singapore.
[2]. Sen J, Chaudhuri TD. Stock price prediction using machine learning and deep learning frameworks. In Proceedings of the 6th International Conference on Business Analytics and Intelligence, Bangalore, India 2018 Dec 20 (pp. 20-22).
[3]. Soni, P., Tewari, Y., & Krishnan, D. (2022). Machine learning approaches in stock price prediction: a systematic review. In Journal of Physics: Conference Series (Vol. 2161, No. 1, p. 012065). IOP Publishing.
[4]. Jeevan, B., Naresh, E. and Kambli, P., 2018, October. Share price prediction using machine learning technique. In 2018 3rd International Conference on Circuits, control, communication and computing (i4c) (pp. 1-4). IEEE.
[5]. Mokalled, W. E. H. M., & Jaber, M. (2019, September). Automated stock price prediction using machine learning. In Proceedings of the Second Financial Narrative Processing Workshop (FNP 2019) (pp. 16-24).
[6]. Shahi TB, Shrestha A, Neupane A, Guo W. Stock price forecasting with deep learning: A comparative study. Mathematics. 2020 Aug 27;8(9):1441. Shahi TB, Shrestha A, Neupane A, Guo W. Stock price forecasting with deep learning: A comparative study. Mathematics. 2020 Aug 27;8(9):1441.
[7]. Milosevic N. Equity forecast: Predicting long term stock price movement using machine learning. arXiv preprint arXiv:1603.00751. 2016 Mar 2.
[8]. Tsai CF, Wang SP. Stock price forecasting by hybrid machine learning techniques. In Proceedings of the international multiconference of engineers and computer scientists 2009 Mar 18 (Vol. 1, No. 755, p. 60).