International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 12 | December - 2022 Impact Factor: 7.185 ISSN: 2582-3930

JPX Japan Stock Exchange Prediction

Zarinabegam Mundargi
Artificial Intelligence and Data Science
Vishwakarma Institute of Technology
Pune, India
zarinabegam.mundargi@vit.edu

Chetanya Rathi
Artificial Intelligence and Data Science
Vishwakarma Institute of Technology
Pune, India
chetanya.rathi20@vit.edu

Harsh Dhabekar
Artificial Intelligence and Data Science
Vishwakarma Institute of Technology
Pune, India
harsh.dhabekar20@vit.edu

Harshit Mundhra
Artificial Intelligence and Data Science
Vishwakarma Institute of Technology
Pune, India
harshit.mundhra20@vit.edu

Pushkar Helge
Artificial Intelligence and Data Science
Vishwakarma Institute of Technology
Pune, India
pushkar.helge20@vit.edu

Abstract*— It is crucial to make an extremely accurate forecast of a future trend in a market that is financially volatile, such as the stock market. In the context of the economic crisis and the need to make money, a trustworthy estimate of stock values is crucial. Advanced machine learning methods are needed to predict another model. Thousands of people make stock market investments every business day. Some of these investors experience financial gains or losses. On any given trading day, the gain or loss is unexpected. Stock market research is required because of the high demand for stock price forecasts. Effective forecasting systems help traders in an indirect way by providing useful information like price position in the future. In this study, a stacked LSTM-based system for stock forecasting is built. Finding sound investments is a requirement for success in any financial sector. These financial decisions were done manually by experts, technology has opened up new possibilities for retail investors. Numerous quantitative trading initiatives are currently being employed to study financial markets and develop investment plans. Such a technique needs real-time and historical data, which is difficult to come by, especially for individual investors. We have considered both of these factors in this model based on the dataset provided by Japan Exchange Group, Inc. (JPX).

Key words : LSTM , JPX , ML , Real time prediction.

I. INTRODUCTION

A successful stock forecast can result in enormous gains for the seller and the broker. It is frequently stated that predictions are chaotic rather than random, meaning they may be made by carefully studying the history of the relevant stock market. Such processes can be effectively represented using machine learning. By predicting a market value that is somewhat near to the tangible worth, it improves accuracy. Due to its effective and precise measures, machine learning's introduction to the field of stock prediction is very attractive and also has gained more attention amongst many researchers.

All the economic activities of contemporary society are significantly influenced by stock markets. There is no evidence showing that a large number of market participants value the appraisal of the stock market index. The Tokyo Stock Exchange, run by the Japan Exchange Group, has the biggest and most well-known Japanese companies with a global presence, including Toyota, Honda, and Mitsubishi. On January 1st, 2013, Tokyo Stock Exchange Group and Osaka Securities Exchange merged to form Japan Exchange Group, Inc. (JPX). JPX acquired Tokyo Commodity Exchange, Inc. on October 1 to broaden its operations into trading commodity derivatives.

Additionally, the Tokyo Stock Exchange offers market statistics, real-time and historical index quotations, and particular trade information.

To give market participants dependable venues for trading listed stocks and derivatives instruments, JPX runs financial instruments exchange marketplaces. JPX provides market infrastructure and market data in addition to brokerage and settlement services via a central counterparty and trading oversight to maintain the integrity of the markets. We continue to make every effort to ensure dependable markets and increase convenience for all market users as we collaborate as an exchange group to provide a wide variety of services.

For this hypothesis, LSTM models are utilized individually. In order to recall the data and outcomes over the long term, LSTM entails reducing error. Finally, graphs are produced showing how prices have changed over time and how they have differed from predictions (using an LSTM-based model). The following is the remainder of the paper: The related work is discussed in Section II. The two models utilized and the techniques employed in them are presented in detail in Section III. The result generated using various graphs for both models is covered in length in Section IV. The last part deals with the references, whereas Section V deals with future scope.

II. LITERATURE REVIEW

On the basis of a literature survey, information from all the prevailing stock price prediction models are considered and it is being utilized in making a better model with greater accuracy and with low bias. The existing models have several approaches which includes statistical methods like Linear Discriminant Analysis(LDA) and regression algorithms. Pattern Recognition in which traders catch signals based on the candlesticks chart patterns that helps in predicting the future of the stock price. Machine learning which is one of the most used in stock price prediction models with different algorithms. In past few years, as we have seen in covid crash which nobody can predict there were sentiments which gave stock markets a different direction and sentiment analysis came into existence which utilizes news, media, social trends to predict the stock prices.

According to the Efficient Market Hypothesis (EMH), stock market prices correctly represent all currently available information. Making forecasts therefore need information from several sources, which may be loosely split into two categories: (1) quantitative data (such as previous prices and turnover rates) and (2) qualitative descriptions (such as annual reports, announcements, news articles, and social media posts). Since qualitative data is typically unstructured and can be challenging to extract key signals from, dealing with it can be challenging.

Technical analysis was the focus of earlier efforts in this field. These research made predictions using simply technical data by utilizing various statistical methods and artificial intelligence algorithms.

This strategy has a drawback since the market responds to information from outside sources that was not present in the historical data utilised to derive the technical information.

Many writers suggest using text mining and machine learning approaches to examine textual data and extract information that may be pertinent to the forecasting process, drawing inspiration from basic analysis. Thanks to improved processing power and the capacity to handle massive databases, it is now feasible to use more complicated machine learning models, such as deep learning models, which outperform more traditional Natural Language Processing (NLP) tasks. The best deep learning models are CNN, RNN, and Recurrent Convolutional Neural Network, notably the LSTM architecture (RCNN).

Weakly supervised learning takes the form of the multiple instance learning (MIL) paradigm. Bags or groups refer to training instances that are organized into sets. For entire groups, not just specific instances, a label is given. Positive groups contain at least one

positive case, whereas negative groups do not have any positive occurrences.

III. METHODOLOGY

After Considering different models we found out LSTM as the most promising model for predicting the stock price data. Long Short-Term Memory is recurrent neural network and it is used in the field of machine learning and artificial intelligence because of the ability to store the past data which is very useful for predicting the stock price. Every prediction moel uses historical data of. In our model we used stock price data to find the patterns of the stock technically as well as fundamentally and by that way we can get the information of the financial health of a particular company very easily.

Fig 1.1

1. Collection of data and preparation of dataset -:

For prediction, we have collected time series-related 60 columns in the dataset which include the date and many more variables including closing price, day high,day low, and volume traded . We will be using the opening and closing statistics to investigate time series using LSTM. This is the historical data of stocks which will be very useful to predict the prices in the future by considering all of these factors.

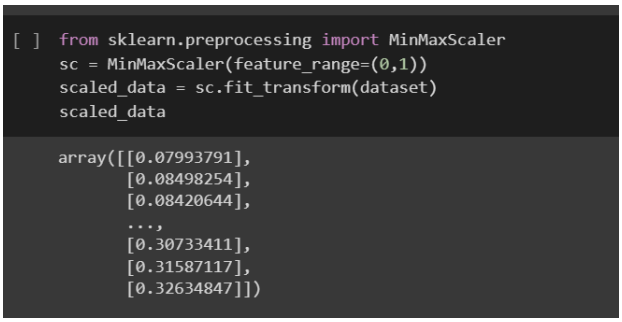
Fig 1.2

2. Visualizing Stock Prices History -:

Data visualization is a very important part of machine learning as it allows us to visualize the dataset and to select the features which should be taken into account for training the model and predicting the right result. Without data visualization, one will not be able to get the idea of the whole dataset.

3. Data Preprocessing -:

We must pre-process this data before using the LSTM to determine the stock price. To transform the values in our data, we use the fit transform function. To ensure that all of the pricing values are on the same scale. Scaling of data is done using min max scaler. Scaling is important for the stock price data to suppress the outliers because of the volatility of stock there are times when a stock gives sudden up move or down move which can hamper our model accuracy. Then, using 80% of the total data, we split the data into training and testing halves.



```
[ ] from sklearn.preprocessing import MinMaxScaler
    sc = MinMaxScaler(feature_range=(0,1))
    scaled_data = sc.fit_transform(dataset)
    scaled_data

    array([[0.07993791],
           [0.08498254],
           [0.08420644],
           ...,
           [0.30733411],
           [0.31587117],
           [0.32634847]])
```

Fig 1.3

4. Preparation of training set -:

Training set : 80%
Testing set : 20%

5. Training LSTM Model -:

We use dropout regularization between the two LSTM layers in our model. Fifty units are assigned in the LSTM parameter with a 10% dropout rate. Mean squared error is the loss function that Adam Optimizer uses to optimize the problem. Since it relates to time-series data, our LSTM network employs mean absolute error as the metric.

6. Visualizing the Predicted Prices -:

We can utilize our model to provide predictions after fitting the data with it. We must use a number of circumstances in order to use the modified function to restore the original value. Using these statistics, we can now see the forecast. Data science is a very useful mix of machine learning, statistics, and mathematical formulas which help us to visualize the trends very easily.

Fig 1.4

IV. CONCLUSION

▾ Date 2021-12-06 (first date of supp files)

Prediction :2438.73

Actual : 2418.0

In this paper the algorithm which is used is LongShortTermMemory(LSTM) algorithm because it is very useful in predicting stock prices and also used to learn order dependance. Firstly, we collected the data from the Japanese stock markets of the last ten years, Visualized the prices from history. After that we have done the data preprocessing. We prepared a training dataset and an LSTM model. Then we visualized the predicted prices. This was all the process which is used in the prediction of the JPX Stock exchange.

V. FUTURE SCOPE

We will keep some of the other stock classification creation techniques for later research. One of them entails grouping companies by industry to create a domain-specific model. Others include considering how news concerning one business may have a negative impact on the stock price of another, as well as more broad business and international news that may indicate a general market stability. The word that is more important may be determined using a word weighting technique in future work, and the sentiment score can then be applied properly in the pipeline of sentiment analysis. The word that is more important may be determined using a word weighting technique in future work, and the sentiment score can then be applied properly in the sentiment analysis pipeline. Since manually expanding domain-specific sentiment lexicons is not possible, our method would become more versatile by expanding the dictionary utilizing

well-established lexicons. Last but not least, developing a hybrid that incorporates stock prices, sentiment analysis based on social media and newspaper data, in addition to algorithms for machine learning such as the LongShortTermMemory(LSTM) neural network model and traditional statistical classification models like the Autoregressive Integrated Moving Estimate (ARIMA), may result in more precise forecasts of short and long-term movements in stocks.

VI. REFERENCES

1. **Zhang, Xi, Siyu Qu, Jieyun Huang, Binxing Fang, and Philip Yu. "Stock market prediction via multi-source multiple instance learning." *IEEE Access* 6 (2018): 50720-50728.**
2. **Idrees, Sheikh Mohammad, M. Afshar Alam, and Parul Agarwal. "A prediction approach for stock market volatility based on time series data." *IEEE Access* 7 (2019): 17287-17298.**
3. **Chen, Kai, Yi Zhou, and Fangyan Dai. "A LSTM-based method for stock returns prediction: A case study of China stock market." In *2015 IEEE international conference on big data (big data)*, pp. 2823-2824. IEEE, 2015.**
4. **Naeini, Mahdi Pakdaman, Hamidreza Taremian, and Homa Baradaran Hashemi. "Stock market value prediction using neural networks." In *2010 international conference on computer information systems and industrial management applications (CISIM)*, pp. 132-136. IEEE, 2010.**
5. **Koo, Eunho, and Geonwoo Kim. "A Hybrid Prediction Model Integrating GARCH Models with a Distribution Manipulation Strategy Based on LSTM Networks for Stock Market Volatility." *IEEE Access* 10 (2022): 34743-34754.**
6. **Vargas, Manuel R., Carlos EM Dos Anjos, Gustavo LG Bichara, and Alexandre G. Evsukoff. "Deep leaming for stock market prediction using technical indicators and financial news articles." In *2018 international joint conference on neural networks (IJCNN)*, pp. 1-8. IEEE, 2018.**
7. **Nabipour, Mojtaba, Pooyan Nayyeri, Hamed Jabani, S. Shahab, and Amir Mosavi. "Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis." *IEEE Access* 8 (2020): 150199-150212.**
8. **Mohan, Saloni, Sahitya Mullapudi, Sudheer Sammeta, Parag Vijayvergia, and David C. Anastasiu. "Stock price prediction using news sentiment analysis." In *2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService)*, pp. 205-208. IEEE, 2019.**
9. **Sethia, Divyasikha. "Stock Price Prediction Using News Sentiment Analysis." In *2022 2nd International Conference on Intelligent Technologies (CONIT)*, pp. 1-6. IEEE, 2022.**
10. **Wen, Min, Ping Li, Lingfei Zhang, and Yan Chen. "Stock market trend prediction using high-order information of time series." *Ieee Access* 7 (2019): 28299-28308.**
11. **Leung, Carson Kai-Sang, Richard Kyle MacKinnon, and Yang Wang. "A machine learning approach for stock price prediction." In *Proceedings of the 18th International Database Engineering & Applications Symposium*, pp. 274-277. 2014.**
12. **Obthong, Mehtabhorn, Nongnuch Tantisantiwong, Watthanasak Jeamwatthanachai, and Gary Wills. "A survey on machine learning for stock price prediction: algorithms and techniques." (2020): 63-71.**
13. **Nikou, Mahla, Gholamreza Mansourfar, and Jamshid Bagherzadeh. "Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms." *Intelligent Systems in Accounting, Finance and Management* 26, no. 4 (2019): 164-174.**

14. **Lee, Jae Won. "Stock price prediction using reinforcement learning." In *ISIE 2001. 2001 IEEE International Symposium on Industrial Electronics Proceedings (Cat. No. 01TH8570)*, vol. 1, pp. 690-695. IEEE, 2001.**
15. **Mehtab, Sidra, Jaydip Sen, and Abhishek Dutta. "Stock price prediction using machine learning and LSTM-based deep learning models." In *Symposium on Machine Learning and Metaheuristics Algorithms, and Applications*, pp. 88-106. Springer, Singapore, 2021.**