IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 6, June 2021

DOI: 10.17148/IARJSET.2021.8630

Stock Market Analysis Using Deep Learning

Arun Aswin Gopinath[1], Ganeshkumar G[2], Mahalakshmi G[3]

[1] UG – Information Technology, College of Engineering Guindy, Anna University, Chennai, Tamil Nadu

[2] UG - Information Technology, College of Engineering Guindy, Anna University, Chennai, Tamil Nadu

[3] Teaching Fellow, Information Technology, College of Engineering Guindy, Anna University, Chennai, Tamil Nadu

Abstract: Predicting and analysing are the two most powerful uses of deep learning. Credits go to the tremendous development in Artificial Intelligence. Analysis and prediction have been providing us with highly accurate results along with saving huge amounts of processing time and power. Tasks have become simpler and more accurate. Predicting and analysing the stock market's behaviour is one such application. By predicting the Stock market, we can help people reap benefits. This is done using historical stock data and then predicting the future value of a stock. Here the stock data refers to historical stock prices news headlines, events which affect the stock's behaviour. Stock markets are highly volatile and predicting its behaviour can only be done very experienced investors and traders. Thus, it can be difficult for an average investor to make the right decisions. Using the historical stock price data, we use the LSTM model for analysis and using the news events we implement sentiment analysis for predicting and analysis the stock market's behaviour. Our aim is to understand the deep learning and sentiment analysis methods, implement it to predict the stock market. The final goal of this project is to provide an average/inexperienced trader/investor citizen an insight into the stock exchange and to attain maximum profit with the right amount of investment.

Keywords: Stock Market, Deep Learning, Sentiment Analysis.

I. INTRODUCTION

The stock market is known for its volatility, randomness, and unpredictability. It is a chaotic place with an unbelievably huge continuously changing stream of data which makes predicting and acting on those predictions to make a profit very hard. It is actually one of the most challenging tasks in times series forecasting.

This project's main goal is to study and apply deep learning techniques and sentiment analysis to the stock market in order to predict stock behaviour and thus act on those predictions to avoid investment risk and generate profit. This project will be a helpful tool that aims to help beginner traders make better decisions.

The goal is to be able to understand the deep learning models and sentiment analysis methods and adapt it to the trading market. The importance of this project lies in the possibility of giving your average investor, an informed insight onto the stock market in order to at least make better choices than random decisions.

II. EXPERIMENTAL METHODS OR METHODOLOGY

2.1 LSTM Model

Long Short Term Memory - is a special kind of RNN (Recurrent Neural Network) capable of learning and analysing long term dependencies. Remembering information for long periods of time is their default behaviour. LSTM has feedback connections.

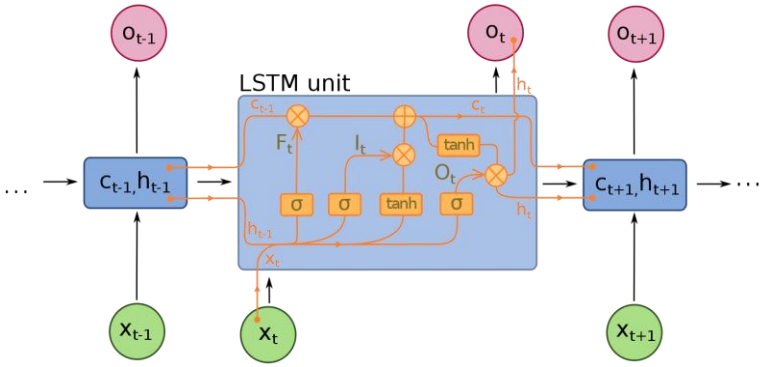


Fig. Impact of DOS Attack on Intruder and Authorized User

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

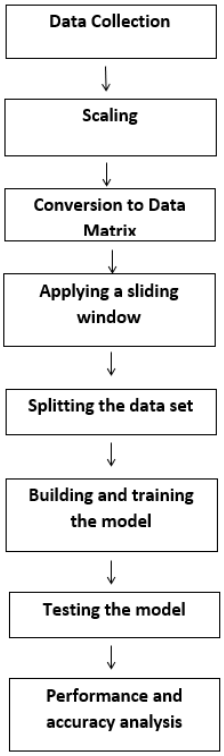
International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 6, June 2021

DOI: 10.17148/IARJSET.2021.8630

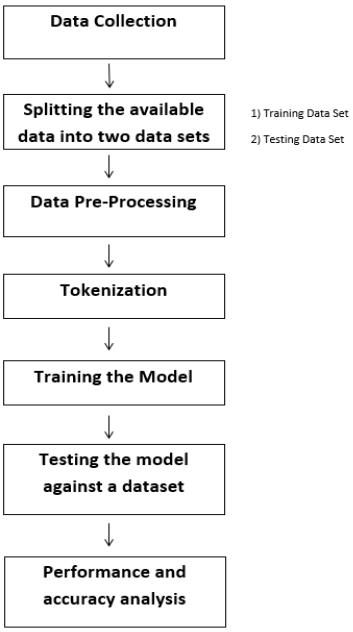
It can not only process single data points (such as images), but also entire sequences of data (such as speech or video). LSTM networks are well-suited to classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series.

The diagram shows the overall LSTM architecture and model that will be used for analysis and prediction.



2.2 Sentiment Analysis Model

Sentiment analysis is the process of detecting positive or negative sentiment in text. It's often used by businesses to detect sentiment in social data, gauge brand reputation, and understand customers. Since customers express their thoughts and feelings more openly than ever before, sentiment analysis is becoming an essential tool to monitor and understand that sentiment.



IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 6, June 2021

DOI: 10.17148/IARJSET.2021.8630

III. RESULTS AND DISCUSSION

The project implementation for both the models involves pre-processing methods respective to its own type of dataset. Next step is to train the model, both the LSTM model and various variation under the sentiment analysis model. After training, the models are tested with the data set to check their performance. The LSTM model trains and performs with a low RMSE value and when compared with the original data set values.

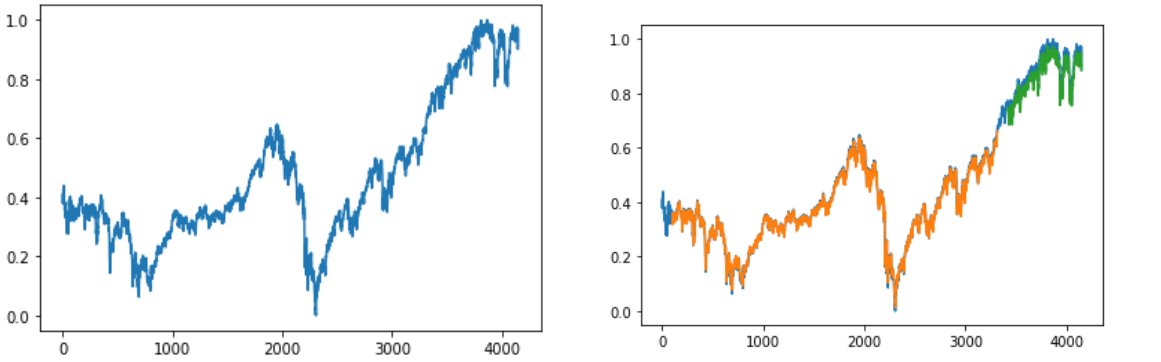


Figure . (a) Original dataset plot (b) LSTM trained and tested model plotted against the original values.

```
In [81]: import math
         from sklearn.metrics import mean_squared_error
         math.sqrt(mean_squared_error(y_train,train_predict))

Out[81]: 0.011355054725502705

In [82]: math.sqrt(mean_squared_error(ytest,test_predict))

Out[82]: 0.02704859523129325
```

Figure . shows the RMSE performnance metric of the model.

The Sentiment analysis also involves pre-processing steps on the dataset and then training and testing the model. Different variation of classifiers and vectorizers are used to get to the best result possible. The vectorizers used were Count Vectorizer and TFI-DF Vectorizer. The classifiers used were Random Forest Classifier and Multinomial Naïve Bayes Classifier. A total of 4 variations were tried out and the results are shown below with the best result of 85.44%.

```
[[141  45]
 [ 10 182]]
0.8544973544973545
              precision    recall  f1-score   support

           0       0.93      0.76      0.84       186
           1       0.80      0.95      0.87       192

    accuracy                           0.85       378
   macro avg       0.87      0.85      0.85       378
weighted avg       0.87      0.85      0.85       378

[[130  56]
 [  0 192]]
0.8518518518518519
              precision    recall  f1-score   support

           0       1.00      0.70      0.82       186
           1       0.77      1.00      0.87       192

    accuracy                           0.85       378
   macro avg       0.89      0.85      0.85       378
weighted avg       0.89      0.85      0.85       378

[[141  45]
 [ 18 174]]
0.8333333333333334
              precision    recall  f1-score   support

           0       0.89      0.76      0.82       186
           1       0.79      0.91      0.85       192

    accuracy                           0.83       378
   macro avg       0.84      0.83      0.83       378
weighted avg       0.84      0.83      0.83       378

[[138  48]
 [ 10 182]]
0.8465608465608465
              precision    recall  f1-score   support

           0       0.93      0.74      0.83       186
           1       0.79      0.95      0.86       192

    accuracy                           0.85       378
   macro avg       0.86      0.84      0.84       378
weighted avg       0.86      0.85      0.84       378
```

Figure . shows the Sentiment analysis model result using different varioations

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 6, June 2021

DOI: 10.17148/IARJSET.2021.8630

IV. CONCLUSION

The LSTM model uses closing prices of stocks and the sentiment analysis model uses top news headlines. Both the models use data of the Dow Jones Industrial Average from the year 2000 to 2016. The highest accuracy achieved by the Sentiment analysis model is 85.44\% where, a combination of Random Forest Classifier and Count Vectorizer was used. The LSTM model's RMSE score is 0.027. On comparing both the models, we concluded that the LSTM model performs better since it gives a more accurate prediction. The RMSE score of the LSTM model is close to 0, which indicates the model has performed well and can be used to predict a stock's price for a given period of 30 days.

This LSTM model can be implemented in the SaaS form for investors, where they can use it for efficiently buying and trading stocks. Also to further explore and extend the working of LSTM model, we can test it on newer entities in the market such as crypto currency. In addition to this, using parameters other than closing prices and news headlines, we can implement a different model and compare it against the currently tested models.

REFERENCES

1. Achyut Ghosh , Soumik Bose , Giridhar Maji , Narayan C. Debnath , Soumya Sen," Stock Price Prediction Using LSTM on Indian Stock Market" in 32nd International Conference on Computer Applications in Industry and Engineering ,2019

2. Samuel Olusegun Ojo ; Pius Adewale Owolawi ; Maredi Mphahlele ," Stock Market Behaviour Prediction using Stacked LSTM Networks" in 2019 International Multidisciplinary Information Technology and Engineering Conference (IMITEC),2020

3. Ge Li ; Ming Xiao ; Ying Guo," Application of Deep Learning in Stock Market Valuation Index Forecasting" in 2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS),2019

4. Caner Türkmen ; Ali Taylan Cemgil." An Application of deep learning for trade signal prediction in financial markets" in 23nd Signal Processing and Communications Applications Conference (SIU),2015

5. Roger Achkar ; Fady Elias-Sleiman ; Hasan Ezzidine ; Nourhane Haidar, "Comparison of BPA-MLP and LSTM-RNN for Stocks Prediction" in 6th International Symposium on Computational and Business Intelligence (ISCBI), 2018

6. D. Mondal, G. Maji, T. Goto, N. C. Debnath and S. Sen, "A Data Warehouse Based Modelling Technique for Stock Market Analysis," International Journal of Engineering & Technology, vol. 3, no. 13, pp. 165-170, 2018.

7. 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, 2017.

8. S. Selvin, R. Vinayakumar, E. A. Gopalkrishnan, V. K. Menon and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in International Conference on Advances in Computing, Communications and Informatics, 2017.

9. G. Maji, S. Sen and A. Sarkar, "Share Market Sectoral Indices Movement Forecast with Lagged Correlation and Association Rule Mining," in International Conference on Computer Information Systems and Industrial Management, Bialystok, Poland, Sprigner, 2017

10. Z. A. Farhath, B. Arputhamary and L. Arockiam, "A Survey on ARIMA Forecasting Using Time Series Model," Int. J. Comput. Sci. Mobile Comput, vol. 5, pp. 104-109, 2016.