International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 09 | Sep 2019 www.irjet.net p-ISSN: 2395-0072

Data Visualization and Stock Market and Prediction

Ashutosh Sharma[1], Sanket Modak[2], Eashwaran Sridhar[3]

[1,2,3]*Student, Department of Computer Science, Terna Engineering College, Nerul, India*

Abstract - *Stock price forecasting is a popular and important topic in financial and academic studies. Share Market is an untidy place for predicting since there are no significant rules to estimate or predict the price of a share in the share market. Many methods like technical analysis, fundamental analysis, time series analysis, and statistical analysis, etc. are all used to attempt to predict the price in the share market but none of these methods are proved as a consistently acceptable prediction tool. In this paper, we will attempt to implement, predict and analyse stock market prices. Artificial Neural networks and Machine Learning are very effective tools for the implementation of forecasting stock prices, returns, and stock modelling. With the help of statistical analysis, the relation between the selected factors and share price is formulated which can help in forecasting accurate results. Although, share market can never be predicted due to its vague domain, this paper aims at applying the concept of prediction and analysis of data for forecasting the stock prices.*

Key Words: **Forecasting; Predicting; Modelling; Analysis; Machine Learning; Artificial Neural**

1. INTRODUCTION

Investment firms, hedge funds and even individuals have been using financial models to have a better understanding of the market behaviour and make a profitable investment into the trades. A lot of information about stock data fluctuations in present for analysis and processing.

Is predicting stock prices using machine learning really an efficient choice? Investors take calculated guesses by analyzing data. They read the news, study the company history, industry trends and other lots of variables that go into making a prediction. The prevailing theories is that stock prices are totally random and unpredictable. This raises the question why top firms like Morgan Stanley and Citigroup hire quantitative analysts to build predictive models.

This paper seeks to utilize Deep Learning models, Long-Short Term Memory (LSTM) Neural Networks, to predict stock prices. For data with time-frames recurrent neural networks (RNNs) come in handy but recent researches have shown that LSTM, networks are the most popular and useful variants of RNNs. A business may become vulnerable to market fluctuations beyond your control - including market sentiment, economic conditions or developments in your sector.

2. RELATED WORK

Traditional approaches to stock market analysis and stock price prediction include fundamental analysis, which looks at a stock's past performance and the general credibility of the company itself, and statistical analysis, which is solely concerned with number crunching and identifying patterns in stock price variation.

Then predictions were achieved with the help of Genetic Algorithms (GA) or Artificial Neural Networks (ANN's), but these fail to capture correlation between stock prices in the form of long-term temporal dependencies. Another major issue with using simple ANNs for stock prediction is the phenomenon of exploding / vanishing gradient, where the weights of a large network either become too large or too small (respectively), drastically slowing their convergence to the optimal value. This is typically caused by two factors: weights are initialized randomly, and the weights closer to the end of the network also tend to change a lot more than those at the beginning.

An alternative approach to stock market analysis is to reduce the dimensions of the input data and apply feature selection algorithms to shortlist a core set of features (such as GDP, oil price, inflation rate, etc.) that have the greatest impact on stock prices or currency exchange rates across markets. However, this method does not consider long term trading strategies as it fails to take the entire history of trends into account; furthermore, there is no provision for outlier detection.

2.1 Proposed System

We proposed an online web-based application using learning model for predicting the price of a given stock. The challenge of this project is to accurately predict the future closing value of a given stock across a given period of time in the future. For this project we will be using a Long Short-Term Memory network – usually just called "LSTMs" to predict the closing price of the S&P 500 using a data set of past prices.

3. PROPOSED ANALYTIC MODEL

We have used Keras to feed a LSTM model to predict the stock prices using historical closing price and trading volume and visualize both the predicted price, values over time and the optimal parameters for the model. The model predicts 30 data points based on the test data set and the last data point is pushed as the output. This model was set as a backend for a website with input data integration functionality.

Initially we have tried linear regression and then we have used Long Short-Term Memory networks – usually just called "LSTMs" to predict the closing price of the S&P 500 using a data set of past prices. We have used Mean Squared Error (MSE) as a performance measure and Root Mean Squared Error (RMSE) calculated as the difference between predicted and actual values of the target stock at adjusted close price and the delta between the performance of the benchmark model (Linear Regression) and our primary model (Deep Learning).

3.1. Exploring the stock prices:

The data used in this paper is of the Google from January 1, 2005 to March 20, 2019. This is a series of data points indexed in time order or a time series. Our goal was to predict the closing price for any given date after training. For ease of reproduction and re usability, all data was pulled from the Yahoo finance Python API. The prediction has to be made for Closing (Adjusted closing) price of the data. Since Yahoo finance already adjusts the closing prices for us, we just need to make prediction for "CLOSE" price.

The data-set is of following form: The whole data can be found out in 'Google.csv' in the project root folder.

Table -1: Head of data set

Date	Open	High	Low	Close	Volume
30- Jun-17	943.9	945.0	929.6	929.6	2287662
29- Jun-17	951.3	951.6	929.6	937.8	3206674
28- Jun-17	950.6	963.2	936.1	961.0	2745568

The mean, standard deviation, maximum and minimum of the data was found to be following:

Table -2: The following table contains the statistical information about the data

Feature	Open	High	Low	Close	Volume
Mean	382.5	385.87	378.73	382.35	42057
Std	213.4	214.60	212.08	213.43	38483
Max	1005.	1008.6	1008.6	1004.2	41182
Min	87.74	89.29	86.37	87.58	52114

We can infer from this data set that date, high and low values are not important features of the data. The features such as High, Low, Volume are important but it can be clearly observed that there is a direct relation between Open and Close prices of the data set. What matters is the opening price of the stock and closing prices of the stock. If at the end of the day we have higher closing prices than the opening prices that we have some profit otherwise we see losses. Volume of share is also important. Rising market should see rising volume, i.e., increasing price and decreasing volume show lack of interest, and this is a warning of a potential reversal. A price drop (or rise) on large volume is a stronger signal that something in the stock has fundamentally changed.

Hence, we have removed Date, High and low features from data set during processing step. The mean, standard deviation, maximum and minimum of the processed data was found to be following

Table -3: The mean, std, max and min of normalized data set

Mean	Std	Max	Min
Open	0.3212	0.23261	1.0
Close	0.3215	0.2328	1.0
Volume	0.09061	0.0953	1.0

3.2 Exploratory Visualization to visualize the data

We have used the Matplotlib python package for initial graphing of data set. This is the hysterical data plotted in scale. The features are number of days and the opening price at each day.

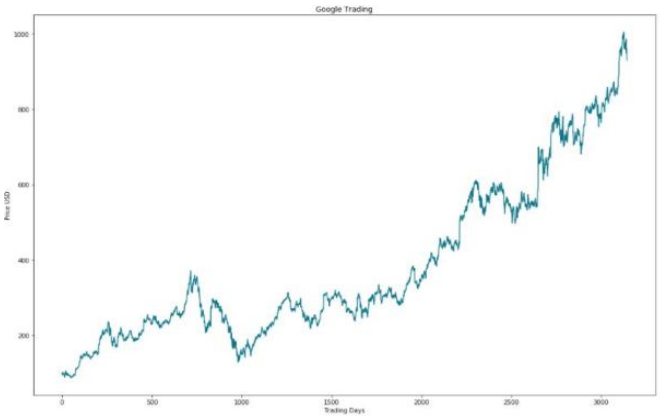

Fig -1: Visualization of processed hysterical data fetched from the API.

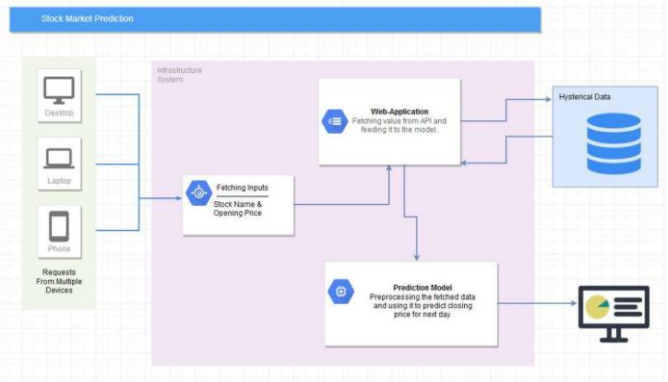

Fig 2: Flow of the system.

4. ALGORITAMS AND TECHNIQUES USED

The goal of this paper was to study time-series data and explore as many options as possible to accurately predict the Stock Price. In our study we have found that Recurrent Neural Nets (RNN) which are used specifically for sequence and pattern learning. As they are networks with loops in them, allowing information to persist and thus ability to memorize the data accurately. Recurrent Neural Nets have vanishing Gradient descent problem which does not allow it to learn from past data as was expected. The remedy of this problem was solved in Long-Short Term Memory Networks, usually referred as LSTMs. These are a special kind of RNN, capable of learning long-term dependencies.

In addition to adjusting the architecture of the Neural Network, the following full set of parameters can be tuned to optimize the prediction model:

•Input Parameters

•Preprocessing and Normalization.

•Neural Network Architecture

•Number of Layers (how many layers of nodes in the Model; used 3)

•Number of Nodes (how many nodes per layer)

•Training Parameters

•Training / Test Split (how much of data set to train Versus test model on; kept constant at 71% and 29% for benchmarks and lstm model)

•Batch Size (how many time steps to include during a Single training step)

•Optimizer Function:

Mean operation done at the end of prediction to scale The output set back to the range of expected output Set.

• Epochs (Number of times for the training process)

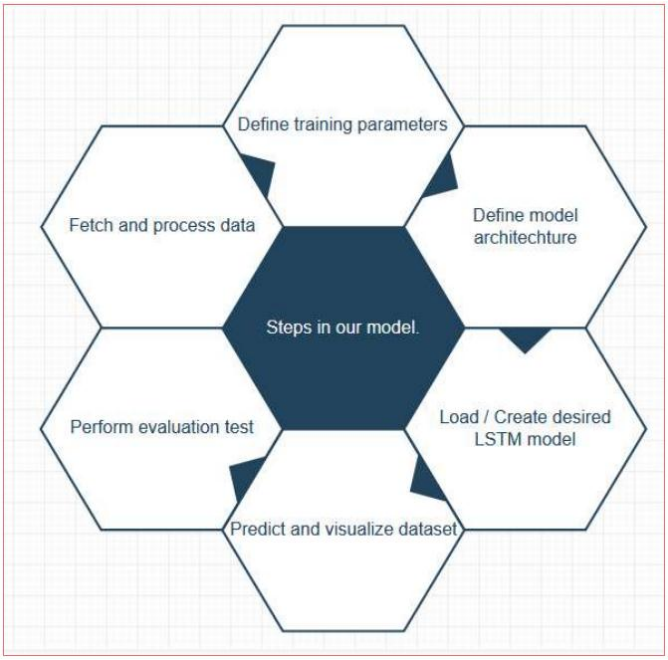

Fig -3: Steps in our model.

5. RESULTS

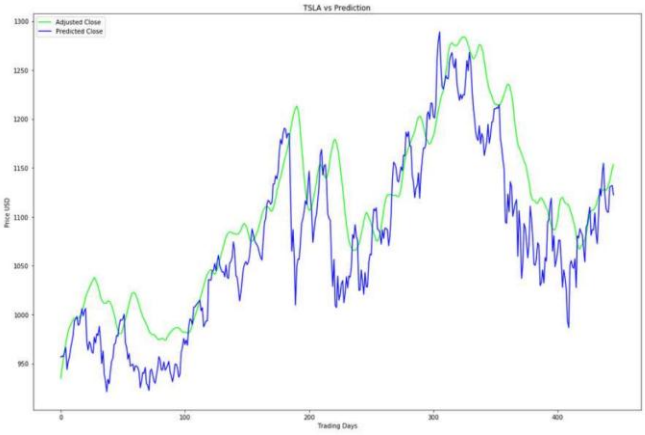

Fig -4: Output graph showing the pattern predicted by Our model and the actual pattern observed in the Dataset of closing prices.

Robustness Check: For checking the robustness of final model we used an unseen data, i.e., data of Facebook. From Jan 1, 2011 to April 16, 2019. On predicting the values of unseen data, we got a decent result for the data. The results are as follows:

Predicted closing price for Facebook on 17 April, 2019: $178.27

Actual closing price for Facebook on 17 April, 2019:

$178.87

Model Evaluation outputs:

Training set: Root Mean Square Error: 0.0417, Mean

Square Error: 0.001739

Test set: Root Mean Square Error: 0.1061, Mean

Square Error: 0.011277

6. CONCLUSIONS & FUTURE WORK

Comparing the benchmark model - Linear Regression to the final improved LSTM model, the Mean Squared Error improvement was significant.

The mean balancing done over processed LSTM helped us get better results and more accurate patterns over hysterical data sets. Predicting stock market prices is a risky trend and can often lead to inaccurate value predictions mainly because of how many factors it depends upon. This project can be extended and modified in future by training the model on more features and including some important non-numerical features as well with the help of a subject matter expert.

REFERENCES

[1] Poonam Somani, Shreyas Talele, Suraj Sawant, (2014)" Stock market prediction using Hidden Markov Model,". IEEE 7th Joint International Information Technology and Artificial Intelligence Conference.

[2] Dinesh Bhuriya, Girish Kaushal, Ashish Sharma, Upandra Singh, (2017)"Stock market predication using a linear regression". International Conference on Electronics, Communication and Aerospace Technology ICECA

[3]Dinesh Bhuriya, Girish Kaushal, Ashish Sharma, Upandra Singh, (2017)"Survey of Stock Market Prediction Using Machine Learning Approach". International Conference on Electronics, Communication and Aerospace Technology ICECA

[4]Rajat-Dhyani, Stock Price Predictor, Github.

[5] V Kranthi Sai Reddy, (2018) "Stock Market Prediction Using Machine Learning", International Research Journal of Engineering and Technology (IRJET), Volume 05 Issue: 10

[6]Alice Zheng, Jack Jin. Stanford University, (2017) "Using AI to Make Predictions on Stock Market", Proceedings on the International Conference on Artificial Intelligence (ICAI).

[7]Prof. S .P. Pimpalkar, Jenish Karia, Muskaan Khan, Satyam Anand, Tushar Mukherjee.(2017) "Stock Market Prediction using Machine Learning", International Journal of Advance Engineering and Research Development, Volume 4,

[8]Taewook Kim, Ha Young Kim.(2019) "Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data", PLOS ONE Journals,

[9]Mruga Gurjar, Parth Naik, Gururaj Mujumdar.(2018) "STOCK MARKET PREDICTION USING ANN", International Research Journal of Engineering and Technology (IRJET),

[10]Nirbhey Singh Pahwa, Neeha Khalfay, Vidhi Soni, Deepali Vora.(2017) "Stock Prediction using Machine Learning a Review Paper", International Journal of Computer Applications (0975 – 8887) Volume 163 – No 5.