OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

STOCK PRICE PREDICTION USING MACHINE LEARNING

Soham K Patil[1], Saiprasad A Gavhane[2], Himanshi Daharwal[3], Prof. Prachi R Salve[4]

Computer Engineering, Dr. Dy Patil School of Engineering Ambi, Pune[1]

Computer Engineering, Dr. Dy Patil School of Engineering Ambi, Pune[2]

Computer Engineering, Dr. Dy Patil School of Engineering Ambi, Pune[3]

Asst. Professor Computer Engineering, Dr. Dy Patil School of Engineering Ambi, Pune[4]

***Abstract*: This report would look at both conventional and modern stock market forecasting methods. We use three distinct methodologies to solve the problem: fundamental analysis, technical analysis, and machine learning. We find support for a weak version of the Efficient Market Hypothesis, which states that while historical prices are meaningless, out-of-sample data can predict future prices. We show how Fundamental Analysis and Machine Learning can assist investors in making better investment decisions. We show how Technical Analysis has a flaw and provides insufficiently usable data. Based on our findings, Quantopian is used to develop and model algorithmic trading programmes. This article makes use of the concepts of *stock* estimation, data analysis, natural language processing, and machine learning.**

Keywords:-*Stock Prediction, Data Analysis, Natural Language Processing, Machine Learning.*

----------∴ ∴ ∴----------

I INTRODUCTION

One of the most common avenues for middle-class investors to prosper is by stock investment. Then there's the real market sector, which involves high-end customers and traders. The most critical point for investors is the company's share price, which is constantly fluctuating up and down. To avoid losing money and, of course, to make money, one must always keep an eye on the stock market's live price. To do so, you must first study the company's financial past and strategic objectives. You can invest only after doing a rigorous corporate and organisational study. As a result, you must define the criteria of the sample, because no one can promise that the test and analysis are right. The key points of Stock Rate are the company's industry understanding, proclivity to retain consumers during any time of slack, regulations, and announcements. To be a good investor, you must have a great deal of knowledge in this area.

II LITERATURE SURVEY

Nonita Sharma, Akanksha Juneja, "Combining of Random Forest Estimates using LSboost for Stock Market Index Prediction"[1]. — The aim of this research is to predict future stock market index values using historical data. The empirical estimate is based on ten years of data from two Indian stock market indexes, the CNX Nifty and the S&P BSE Sensex. For the next one to five, fifteen, thirty, and forty days, forecasts are made. The article proposes combining the predictions/estimates of a Random Forest ensemble of trees using LSboost (i.e. LS-RF). The proposed model's prediction accuracy is comparable to that of the well-known Support Vector Regression. Technological parameters inform each prediction model. The average value is calculated using the stock's most recent sale price. The proposed scheme outperforms Support Vector Regression in terms of precision and can be used to construct mathematical models for market price prediction.

Du Peng, "Analysis of Investor Sentiment and Stock Market Volatility Trend Based on Big Data Strategy" [2]. — The aim of this paper is to investigate the fundamental mechanism by which investor sentiment affects stock market volatility. It conducts a comparative analysis using Pollet and Wilson's volatility decomposition theorem and the volatility origins. This article constructs an insight database using data from the online news emotion index, site search volume, social network emotion index, and social network heat index. It eliminates variables with a direct connection to the stock market and incorporates them into forecasting studies after doing correlation analysis and Granger causality tests. The formula produces a market volatility index and analyse the connection between consumer trust and equity price volatility. In laboratory analysis, the variance between sale price and value is used as an explanatory element, and stock price uncertainty is quantified using the logarithmic return on the stock. In comparison to the stock market confidence index, the sharemarket volatility index has a higher predictive potential for the share market volatility turning point in a more volatile setting, especially for the one- to two-day decline turning point ahead of time, and is therefore critical for forecasting stock price volatility and risk management.

Muhammad Firdaus, Swelandiah Endah Pratiwi, Dionysia Kowanda, Anacostia Kowanda." Literature review on Artificial Neural Networks Techniques Application for Stock Market Prediction and as Decision Support Tools"[3— The aim of this study is to examine the application of Artificial Neural Network (ANN) techniques to stock market forecasting. Content analysis was used as an experimentation tool in this design. This paper was compiled using data from ProQuest electronic databases. Evaluation techniques: We used core terms and phrases associated with Artificial Neural Network Stock Market Prediction from 2013 to 2018. Out of 129 scholarly articles reviewed, only four stock market papers follow the inclusion criteria. The forecasting research and estimation were conducted using six ANN derivatives techniques. As a result of the studies' research, it was discovered that ANN stock market predictions are highly accurate in both of them. Two tests indicate an accuracy of greater than 90%, and two indicate an accuracy of greater than 50%. According to this report, ANNs' ability to predict stock prices correctly is consistent. Four techniques for predicting stock markets are over 95% accurate. The Signal Processing/Gaussian Zero-Phase Filter (GZ-Filter) achieved the highest prediction precision of 98.7 percent.

Dev Shah, Haruna Isah, Farhana Zulkernine, "Predicting the Effects of News Sentiments on the Stock Market"[4]. The usefulness of stock market forecasting in the preparation of economic operations could be overstated. Many scholars from various fields, including computer science, mathematics, economics, banking, and operations analysis, have been drawn to stock price prediction. Recent research has found that the overwhelming amount of publicly available online content, such as Wikipedia user patterns, mass media news coverage, and social media discussions, may have a measurable impact on investors' attitudes toward financial markets. Since the stock market is so vulnerable to the economy and can directly result in financial failure, the accuracy of statistical models for stock market prediction is critical. The results of news emotions on the stock market were retrieved, extracted, and reviewed in this article. The creation of a financial sentiment analysis dictionary, the development of a dictionary-based sentiment analysis model, and the assessment of the model for gauging the effect of news sentiments on stocks for the pharmaceutical industry are among our key contributions. We were able to forecast the patterns in short-term stock price activity with a positional precision of 70.59 percent using only press sentiments.

Muhammad Waqar, Hassan Dawood,Muhammad Bilal Shahnawaz,, "Prediction of Stock Market by Principal Component Analysis"[5]. — Machine learning models face an intriguing problem when categorising high-dimensional data, as the presence of a large number of closely associated dimensions or attributes will minimize the classification model's accuracy. The paper investigates the problem of high dimensionality in the stock market and predicts market dynamics using principal component analysis (PCA) and linear regression. PCA can help improve machine learning approaches' predictive efficiency while reducing data redundancy. Experiments were carried out on a three-dimensional spectral model of the stock exchanges in New York, London, and Karachi. Before and after PCA, the precision of a regression analysis classification protocol is compared. The experiments show that PCA can greatly increase machine learning efficiency in general if and only if the relative similarity between input features is investigated and the selected features are correctly chosen. The classification model is tested using the base mean standard errors (RMSE).

III PROBLEM STATEMENT

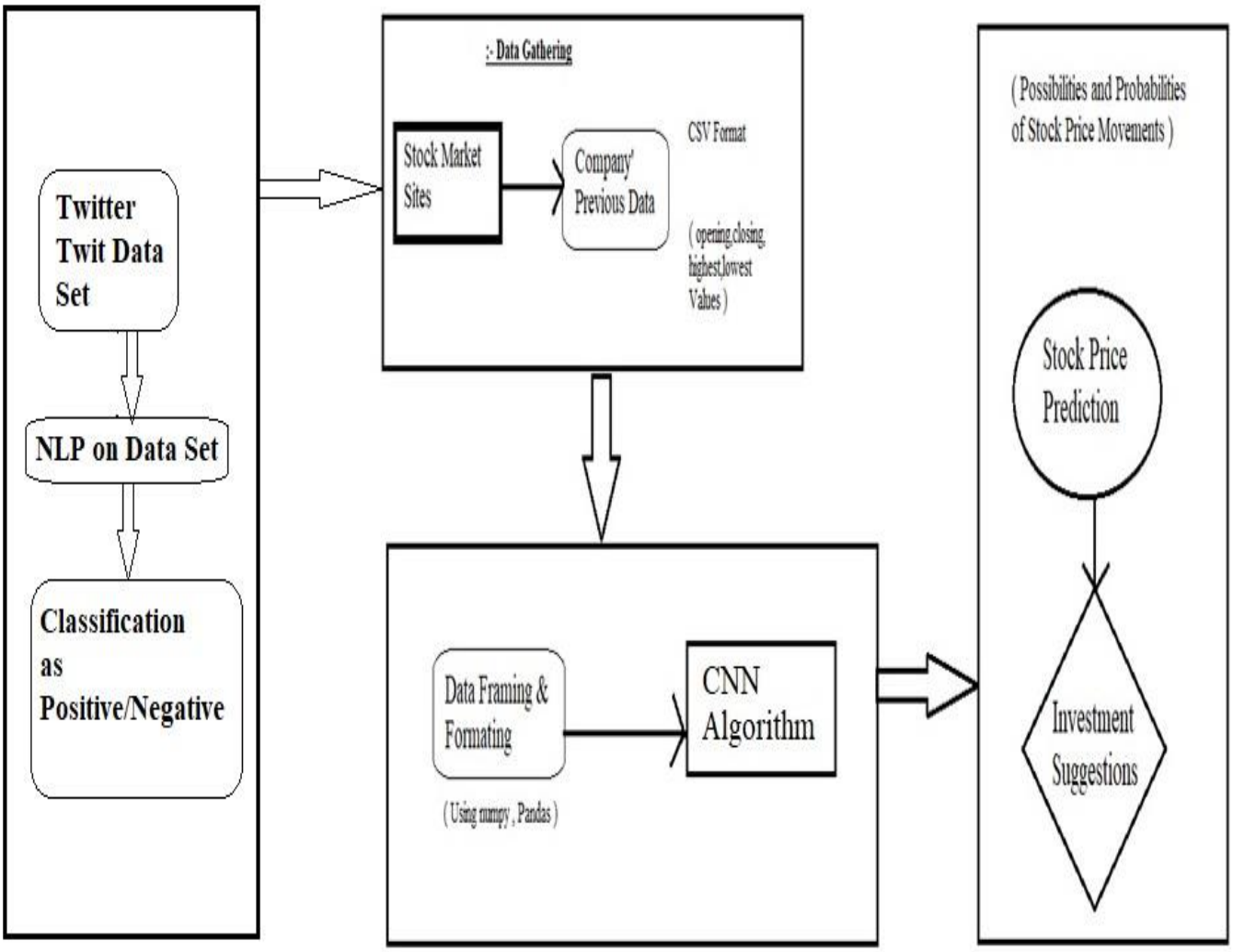
Financial analysts who exchange stocks are often blind to market patterns. They struggle with investing because they are unsure about which stocks to buy or sell in order to maximise earnings. Both stock market information is readily available in today's world. Analyzing all of this data individually or manually is very complex. As a consequence, it is essential to simplify the process.

IV MOTIVATION

We get motivated by disadvantages of existing system. The purpose of system is to build so that it will analyze by itself

and express predictions. Stock price prediction is a classic and important problem. With a fruitful model for stock forecast, we can pick up knowledge about market conduct after some time, spotting patterns that would somehow or another not have been taken note. With the inexorably computational intensity of the PC, AI will be an efficient technique to take care of this issue. Nonetheless, people in general stock dataset is unreasonably restricted for some, AI calculations to work with, while requesting more highlights may cost a huge number of dollars regular. In this paper, we will present a structure in which we coordinate client expectations into the present AI calculation utilizing open verifiable information to improve our outcomes. The spurred thought is that, in the event that we know all data about today's stock exchanging (of all specific merchants), the value is predictable. Thus, if we can obtain just a partial information, we can hope to improve the present expectation parcel. With the development of the Internet, informal organizations, and online social associations, getting day by day client forecasts is practical job1. In this way, our inspiration is to structure an open help consolidating chronicled information and clients forecasts to make a more grounded model that will benefit everybody

V SYSTEM ARCHITECTURE



VI PROPOSED SYSTEM

We suggest a scheme for stock market recognition that begins with a csv dataset and an image dataset, followed by pre-processing in which we delete superfluous images, convert NLP to Grey conversion, perform data extraction to validate the image's characteristics, segment the image into subparts, and finally classify with the lstm algorithm.

Predict output: Stock market detect or Not

VII ALGORITHM

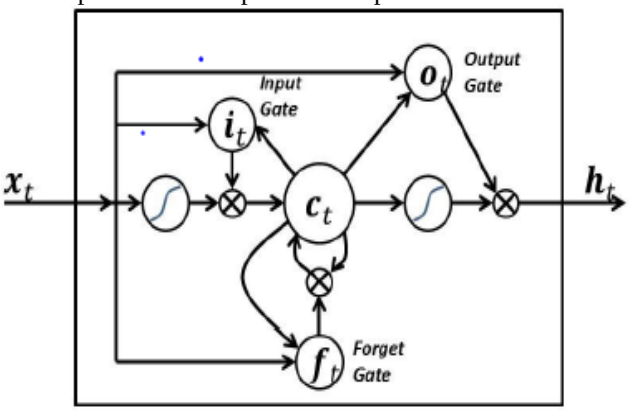
A sequential troupe of daily trading details of stocks over a set time period of N days is defined in the LSTM model. These daily details in sequence describe the trends of the stock with the attributes like the day high/low, open price, closing price and trade volume on a specific day within the N days. Comparing in sequence the closing prices of 3 consecutive trading sessions with that of the last day, the earning rates were calculated. The model comprises of two layers namely, an input layer which consists of number of cells equal to the sequence learning attributes that one sequence may hold, the LSTM layers, a compact layer, and an output layer consisting of similar number of cells. There are 4 types of learning features that could be given to the LSTM models. They are:

1. The historical trade details.
2. The technical analysis derived from these
Historical trade details.
3. The movement of the market indices.
4. The economic fundamentals.

First and the third type of data is essential for the Forecasting of the prices of different stocks.

The LSTM model has an edge over the other models as it has improved memory cells. These cells are linked to each gateway. It also contains Forget gates.

The connection to the memory stick is controlled by the forget gates. They are also responsible for remembering the error as per requirement and scale the feedback by each step [9]. It can also store information of any size. The input to the LSTM model is provided in the form of vectors. This vector comprises of elements such as the previous closing price, open price, high and low prices, the trade volumes of the current session. The vector also contains the results of the news analysis. The output layer of the model gives the final predicted price over the specific time periods.



VIII APPLICATIONS AND LIMITATIONS

Application:-

- Share market Suggestions Agencies.
- National Stock Analysis for General Purpose.
- Investment Banking
- International Trade Analyst
- Government Authorities
- National Economy Planning
- Study of Data Analysis of investment culture all over the world

Limitation:-

- Stock market cannot be accurately predicted. The future, like any complex problem, has far too many variables to be predicted.
- The stock market is a place where buyers and sellers converge. When there are more buyers than sellers, the price increases.

IX.OBJECTIVE

A stock exchange's primary purpose is to assist businesses in raising capital. It was created to provide a country's businesses with the resources they need. To do this, a private corporation's stake is transferred to the public in the form of equity certificates. The proceeds from stock sales go into the firm's capital formation.

X.CONCLUSION

We learned how to predict stock market fluctuations using machine learning technology. An person cannot read and comprehend a detailed graph of a company's stock price. In practice, multiple organisations need us to view data on a wide scale. As a consequence, we will use Machine Learning algorithms to make far more intelligent predictions. We're going to use the LSTM algorithm. As a result, we will be able to make even more precise estimates.

XI. FUTURE SCOPE

We are here covering multiple aspects in terms of getting better and updated results from our research based system. In future we can add more features. Multiple Country data can be tracked for international investments and multinational banking. Mutual Funds also can be tracked and can be given suggestion regarding mutual fund investments. Every activity , National Level effective Decisions and Decision makers can be tracked for early prediction.

REFERENCE

1) H. Isah, "Social Data Mining for Crime Intelligence: Contributions to Social Data Quality Assessment and Prediction Methods," University of Bradford, 2017.

2) P. Wei and N. Wang, "Wikipedia and stock return: Wikipedia usage pattern helps to predict the individual stock movement," in Proceedings of the 25th Interna tional Conference Companion on World Wide Web, 2016, pp. 591-594: Interna tional World Wide Web Conferences Steering Committee.

3) E. Chong, C. Han, and F. C. Park, "Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies," Expert Systems with Applications, vol. 83, pp. 187- 205, 2017.

4) J. Zhang, S. Cui, Y. Xu, Q. Li, and T. Li, "A novel data-driven stock price trend prediction system," Expert Systems with Applications, vol. 97, pp. 60-69, 2018.

5) L. S. Malagrino, N. T. Roman, and A. M. Monteiro, "Forecasting stock market index daily direction: A Bayesia Network approach," Expert Systems with Ap plications, vol. 105, pp. 11-22, 2018.

6) M. B. Patel and S. R. Yalamalle, “Stock Price Prediction Using Artificial Neural Network” International Journal of Innovative Research in Science, Engineering and Technology, vol. 3, pp.13755- 13762, June 2014.

7) Jie Wang, Jun Wang, Wen Fang. Financial Time Series Prediction Using El man Recurrent Random Neural Networks[J]. Computational Intelligence Neuro science, 20162016(12):1-14.