Survey On Stock Market Forecast Using Machine Learning And Text Mining With Cloud's CaaS

M Therasa* S Sarveshwar* M Madhusudhaanan* V Naveen Kumar*

*Department of Computer Science and Engineering,
Panimalar Institute of Technology,
Chennai-600123

ABSTRACT

The big data techniques in conjunction with data mining are rapidly increasing its domination in the data analysis paradigm. These techniques can be effective when combined with the robust applications of cloud computing which can be used in the forecasting of volatile share market. The current existing process of share market predictions uses various existing algorithms but does not take into account the hidden variables that could affect the share market. Our proposed system, uses multi-agent module which can handle multitude of scenarios in which the investment can be favourable. This system is deployed in the cloud platform which can provide high performance computing as well as secured data service. Initially, the system collects the data of the company from social media like twitter and facebook through REST API and performs text mining to retrieve the data that are in correlation with the financial market. After the correlation analysis, the system uses diverged quantitative trading algorithms and performs forecasting analysis. The probability of buying or selling of the stocks are calculated based on the values acquired though forecasting analysis. The reliability factor associated with the quantitative algorithms and the social media posts is modified based on the probability of success in the process of back testing and live share market. Thus the prediction of stock price with company related posts from social media and the quantitative trading algorithms reduces the risk of investment in the volatile share market.

Index-Terms: Stock Market, Data Mining, Cloud Computing, Big Data, Trading Algorithms, Investment Recommendation System, survey of stock market prediction, Cloud Container Service.

I. INTRODUCTION

The stock market is a platform where the stocks of the public companies are enlisted. The stocks defines the ownership of the client over the public company. The market trends depends on the supply and demand of stocks for target companies. The Stock market price prediction has always been an area for research and development. Its mainly due to the fact that the market is non-linear, volatile, dynamic and unpredictable. In the past few decades, the researchers had proposed multiple techniques that could be viable in the astute prediction of the moment to buy/sell the stocks and attain a profit or minimize the loss. Some of the most common techniques suggested by the researchers are technical analysis on the market trend namely mean average, momentum, oscillators based on the historical data. Most of the techniques does not provide an astute result of the market trend. In the recent years, the trending technology named machine learning and deep neural networks has been proposed for analysis of the market trends. These machine learning algorithms in combination with the traditional techniques has proved to be efficient in the prediction of stock prices. Having said that, the financial news and social media namely twitter, facebook are also playing a vital role in the prediction of market trend. The social media being the fastest platform in transmission of information provides an insight in developing the business. Based on these information we have developed a system in which we will combine the best features of machine learning, traditional techniques, social media and the financial statements of the target company and assist the achieving profit or minimize the loss. This system is to be deployed in the light weight containers of cloud computing to achieve maximum efficiency and performance.

II. LITERATURE SURVEY

During the literature Survey, we have collected the following information for the stock market price prediction using machine learning, social media, traditional techniques and cloud computing.

1. Using Structural Events to predict stock price movement:

The need for stock market prediction has be deliberately increasing in the past few years. The technique of analysing the stock price using multitude traditional algorithms and graphs has not always generated the astute results. So it is necessary to algorithms for accurate prediction. It has been proved that the stock price movements are influenced by the news. The news has served to be a prime factor in the investment decisions. The traditional finance news analysis was carried out by identifying the noun phrases, named entities which does not properly capture the entity-relation of the events. The technique that was employed in this paper is Open IE Technology that enables the extraction of the structured entity-relation between the words in the news. The Linear and Non-Linear models were used to investigate the complex relationships between the events and the news of the corresponding event. This method can outperform the bag-of-words based methodologies and has a potential to increase the stock price prediction[1].

2. Deep Learning for Event-Driven Stock Prediction:

The use of deep learning in the prediction of the stock prices has been trending in the past few years. Several researchers are involved in improving the accuracy of the deep learning algorithms. The paper extracts the news title from a particular financial news source (Bloomberg) represented as dense vectors and trained using novel neural tensor networks. This methodology is followed for the prediction of the stock prices for both short and long term. The extraction of the news data from a single source does not always provide an astute result. Moreover, In this cited paper the news is analysed based on the title which could sometimes be misleading. There are always other factors besides the news that provides a significant contribution to the stock price of the company which has not been taken into account in this cited paper [2].

3. Leverage Financial news to predict stock price movements using Word Embedding and Deep Neural Network.

The financial news provides us some useful information about the public companies and the stock market trend. The most popular word embedding techniques and deep neural networks are used for leveraging the financial news to predict stock price movements in the market. The neural network model is trained using the historical price data of the target company over a period of 5 days. The financial news feature is extracted using the bag of keywords technique and the polarity score of the relation between the keyword and the stock price of the company is calculated. In order to uncover the impact of the specific events on the stock price, the events are classified into various categories. Some of the category tags include events related to new-product release, acquisition, price-rise, law-suit, fiscal-report, government, analyst highlights. The neural network model is tested on training set and it has proved to be providing valid results under supervised environments [3].

4. Placement architecture for a container as a service in cloud environment:
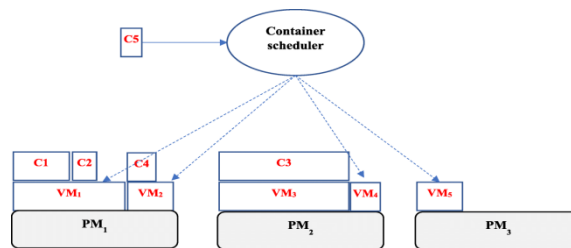

Figure 1: **Container Architecture**

The Cloud Computing techniques has vividly contributed to the evolution of server side deployment of management of files. However inorder to improve the services provided by the cloud, the technique of using containers has been proposed. Unlike the traditional VM, a container is an emerging VM technique that encapsulates the tasks and the library dependencies for execution. The traditional cloud computing approaches leads to the over utilization of the hypervisor or under utilization of the PM. This paper uses the placement techniques like best fit, Max Fit which evaluates the wasted resources in the VM and in the PM. The optimal allocation of application in the containers has let to the increased and efficient utilization of the resources [4].

5. A Survey on Stock Market Prediction Using SVM:

In the recent studies it has been proved that the regression algorithms are inefficient due to parameter instability and model uncertainty. The studies has provided considerable evidences that the traditional strategies used in the prediction of the market trend does not assure to solve the problem of accurate prediction. The Support Vector Machine consists of a kernel, decision function and sparsity of the solution. It has been evident that the SVM has provided better and accurate results than the traditional methodologies and regression algorithms. The Interconnection between the stock price and the market index is calculated using correlation analysis [5].

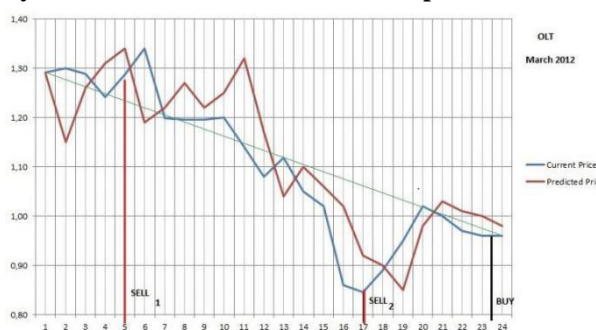## 6. Stock Market Multi-Agent Recommendation System Based on Elliott Wave Principle:

Figure 2: Market prediction using Elliot Wave Principle.

This paper provides a hybrid stock recommendation system by combining the traditional trading techniques with Elliott wave principle. The multi-agent architecture improves the profitability in short and long term investments. Considering that ,The stock market patterns reflects the behaviour of the human, The Elliot wave principle defines two waves namely impulse and corrective wave. The impulse wave is a five wave sequence whereas the corrective wave is a three wave sequence in the direction opposite to the impulse wave. These waves are used in the prediction of the prices for both short and long term. This multi-agent system combines the best features of Fibonacci Series and Elliott wave principle along with some special technical analysis methods (such as Gap Analysis, Breakout System) and Deep Neural Networks. This incorporation of multiple algorithms has proved to be efficient in the forecasting of the stock market [6].

7. Indian stock market prediction using artificial neural networks in tick data:

The Stock market is a platform that is driven by supply and demand. The Stock market is one of the arising field. The market trend is evaluated using three techniques namely technical analysis, traditional time series forecasting and machine learning method.

In the premature stages of stock market the classical regression algorithms like Linear Regression, polynomial regression were used to predict the trend of the market. In addition to that, the traditional trading models such as moving average, smoothing and ARIMA were used to predict the prices. However, the current prediction techniques uses support vector machine and artificial neural networks. These models are trained with historical data and then tested on the testing data. The models allocate adaptive weights based on its training on the historical data. The models perform cross validation of the data and calculates the correlation between the stock price and the market. These models yield astute results up to a time period of 15 minutes. The SVM and the artificial neural network are proved to be efficient in prediction of the prices for a certain time period. However, these models does not take into account the external factors that affect the stock price of the target company [7].

8. Extracting Semantic Knowledge from Twitter:

In the last 10 years, the social media has been playing a vital role and every business sector. As of now, it has been recorded that about 140 million tweets are posted on an average each day. The tweets are messages with a character limit of 140 which expresses the opinions of the individuals over a particular event. The twitter has been acting as a hidden factor in the business performance. This cited paper focuses on performing semantic analysis on the tweets and analyse the emerging topics. The NLP, machine learning and statistics plays a major role in the semantic analysis. The twitter analysis has resulted in improving the business [8].

Disadvantages of the Existing System:

1. The stock prices are unpredictable using the traditional classifiers.
2. The current system reports astute prediction of the stock prices in confined environments(Time period)

3. The current system does not take into account the reliability of the information source.

4. Multiple data sources are not exploited.

5. The current system leads to over utilization of the hypervisor or under-utilization of the physical machine.

6. The traditional semantic analysis does not provide deeper insights in entity relationship.

III. PROPOSED SYSTEM

The existing methodology has its own advantages and limitations. In this paper we propose to use the linear and non-linear models of machine learning that can be used in the prediction of stock prices. The ensemble algorithms of machine learning is to be combined with the feature of text mining of financial news in twitter and face book, analysis of financial statement and balance sheets of the target company to predict the price of the stocks. This system is deployed in Cloud's Container Service to provide enhanced and seamless performance (storage and speed). By analysing the target company in every perspective our proposed system reduces the risk involved in investing in Stock Market.

IV. CONCLUSION

To conclude, though there exists various methodologies for analysing the stock market movement, each of those has its own dominance and limitations. The currently followed technical analysis methods proves to provide the results with better accuracy. The libraries of machine learning, text mining has a great developer community and its always being updated with latest patches. Though the fundamental analysis of stock market provides an insight about the price movement, when combined with the technical and new analysis the accuracy in prediction of stock price of the target company increases deliberately, thus reducing the risks involved in stock investments.

REFERENCES

1. Ding, Xiao & Zhang, Yue & Liu, Ting & Duan, Junwen. (2014). Using Structured Events to Predict Stock Price Movement: An Empirical Investigation**.**

2. Deep Learning for Event-Driven Stock Prediction
   Xiao Ding, Yue Zhang, Ting Liu, Junwen Duan

3. Peng, Yangtuo & Jiang, Hui. (2016). Leverage Financial News to Predict Stock Price Movements Using Word Embeddings and Deep Neural Networks.

4. Hussein, M.K., Mousa, M.H. & Alqarni, M.A. A placement architecture for a container as a service (CaaS) in a cloud environment. *J Cloud Comp* **8,** 7 (2019).

5. Sachin Sampat Patil1 , Prof. Kailash Patidar2, Asst. Prof. Megha Jain. A Survey on Stock Market Prediction Using SVM

6. Tirea M., Tandau I., Negru V. (2012) Stock Market Multi-Agent Recommendation System Based on the Elliott Wave Principle. In: Quirchmayr G., Basl J., You I., Xu L., Weippl E. (eds) Multidisciplinary Research and Practice for Information Systems. CD-ARES 2012. Lecture Notes in Computer Science, vol 7465. Springer, Berlin, Heidelberg

7. Selvamuthu, D., Kumar, V. & Mishra, A. Financ Innov (2019). Indian stock market prediction using artificial neural networks in tick data. https://doi.org/10.1186/s40854-019-0131-7

8. Teufl P., Kraxberger S. (2011) Extracting Semantic Knowledge from Twitter. In: Tambouris E., Macintosh A., de Bruijn H. (eds) Electronic Participation. ePart 2011. Lecture Notes in Computer Science, vol 6847. Springer, Berlin, Heidelberg

AUTHORS:

First Author - S.Sarveshwar, UG Scholar, Panimalar Institute of Technology, sarveshwar3@gmail.com

Second Author - M.Madhusudhaanan, UG Scholar, Panimalar Institute of Technology, madhus1521@gmail.com

Third Author - V.Naveen Kumar, UG Scholar, Panimalar Institute of Technology, naveengv97@gmail.com

Correspondence Author - M.Therasa, Associate professor, Panimalar Institute of Technology, therasamic@gmail.com