



Analysis of the stock market prediction techniques and its impact in the equity domain

Cerene Mariam Abraham¹, M. Sudheep Elayidom², T. Santhanakrishnan³, Manoj P. Michel⁴

¹Cochin University of Science And Technology, India, cma@cusat.ac.in

²Cochin University of Science And Technology, India, sudheep@cusat.ac.in

³Defence Research & Development Organization, India, santhannpol@gmail.com

⁴Cochin Stock Exchange, India, payyappillys@yahoo.co.in

ABSTRACT

A stock market is the place that processes large scale of complex information which is affected by a wide range of socioeconomic factors. Parameters like inflation, return on equity, price to earnings ratios and valuation multiple are responsible for the changing trends in the stock price. Several researches are going on in this field to find out the factors that are responsible for the up down trends in the equity market. The motivation behind this paper is to analyze different stock market techniques which play a key role in predicting the price of the stock and its impact in the equity domain.

Keywords: GPOMS, OpinionFinder, AZFinText, Particle swarm optimization, Least square support vector machine.

1. INTRODUCTION

Now a days it is a common notion that huge amounts of capital are exchanged through the Stock Markets all around the world. The economy of a country are strongly linked and steadily influenced by the execution of their Equity Markets. The feature that all Stock Markets have in common is the uncertainty that is related with its short and long term destiny. This character is not acceptable for the dealer but it is also inevitable whenever the Stock Market is chosen as the investment instrument. The best method that one can do is to try out to decrease this uncertainty. Stock Market Forecasting or Stock Market prediction is one gadget in this operation. Prediction of stock price has been at centre for years since it can return significant financial gain. Forecasting the stock market is not a plain thing, mainly as a result of the close to stochastic behaviour of equity time series.

Equity market or stock market is the place where stocks in listed companies are bought and sold, whether physical or electronic. It may be a public traded company or private or a non-profit company. This market provides a managed place where investors and companies may gather to make transactions on neutral ground. The stocks are recorded and traded on exchanges which are instances of a mutual organization or corporation focused in the business of assembling sellers and buyers of the organizations to a listing

of securities and stocks together. Securities are the financial instruments by which companies can raise their finance to run its operations.

2. OBJECTIVE

The prediction of the market is without any confusion a captivating job. It is the act of testing to evaluate the future value of a financial instrument traded or stock of the company on a stock exchange. The successful forecasting of a future price of the stock could yield notable profit.

The EMH (efficient-market hypothesis) tells that stock prices mirror all currently available information and any change in the price are not based on newly discovered information and thus are unpredictable by default. The EMH asserts that price of a stock is a function of processed data and rational expectations, and that newly unveiled data about the prospects of a company is almost immediately showed in the current price of the stock. This would implicit that all generally known information about a company, which obviously includes its price history of stock, would already be evinced in the current price of the stock. Accordingly, alteration in the stock price manifest delivery of new information reshapes in the market generally, or arbitrary movements around the price that exhibits the existing data set. There are several researches are happening in this field to accomplish this task. Those methods use different approaches, ranging from highly informal ways to more formal ways.

They are categorized as follows:-

- Artificial intelligence methods
- Statistical approaches
- Mathematical approaches

Artificial intelligence has mostly significant algorithms that deal with machine learning and optimization. Genetic algorithms, SVMs (support vector machines) and neural networks all take financial securities as input parameters and give predictions as output which is based on the hidden patterns within the historical data. Most of these techniques have been restricted to either recognizing the parameters, which are relevant, or evaluating stock data in terms of trend of the market.

Statistical approaches use probability methods such as Monte Carlo simulations and game theory. In Monte Carlo simulations, it is too difficult to implement directly the problem of price prediction, so input parameters are given as series of suitable random numbers and are evaluated for how close they reach at the predicted value [4]. In game theory, the players in the game will observe the strategies of other players and elect a stand that will gain them the best payoff.

Mathematical approaches borrow heavily from the areas of econophysics and applied mathematics. To model future prices, this uses more complex mathematical formalisms, such as log-periodic oscillations, wavelet transforms and percolation methods.

3. LITERATURE REVIEW

Here we analyze and evaluate the approaches used in three different papers. The input parameters for the prediction of up-down trends in stock price are dissimilar in nature. The first paper adopts twitter data as its input parameter and evaluates the effect of its mood on stock market forecasting. Financial news is taken as the input value in the second paper. Third paper chooses financial technical indicators as its input. The methods used for the forecasting are explained in each subsection.

A. Twitter mood data to predict stock Market values

Fear and greed is the factor that drives the stock market, in other words, it is by both psychological and financial factors. Forecasting market behavior needs understanding the parameters that shape investors' individual as well as collective behavior. Investor sentiment theory and behavioral finance have firmly fixed that investors' behavior can be fashioned by whether they feel bullish (optimistic) or bearish (pessimistic) about future stock values. The ascend of social networking environments has paved the way to the development of scalable, reliable and rapid assessments of the public behaviour [1].

This work uses large-scale Twitter data to compute global public mood states at short time intervals and the resulting Twitter mood data is used to speculate stock market values, in particular the daily price fluctuations of DJIA (Dow Jones Industrial Average). GPOMS and OpinionFinder are the two emotion analysis tools to collect daily tweets. GPOMS is a proprietary tool that computes six different dimensions of mood, which are often ignored by traditional methods of sentiment tracking system. OpinionFinder is a widely used tool for sentiment analysis that categorizes texts in terms of its positive versus negative sentiment and is useful for cross checking.

Granger causality analysis is applied to dictate whether public sentiment affects the values of stock market. The result shows that public mood and DJIA daily time series values overlapped in the same direction regularly. This pattern is stable with a basic assumption of behavioral finance that emotions contribute a major role in financial markets. Twitter proposes a potentially rich source to chase public sentiment in countries including the Portugal, Japan, Spain and UK (Portuguese, Japanese, Spanish and English are the most prevalent languages on Twitter), to probe the footprint of mood on their stock markets.

This prognostic correlation between Twitter mood and DJIA values stays a pivotal area for future research. Researchers might need to take into consideration social and cognitive impacts in which individual agents are furnished with the potential to learn from their past experiences and can tune their trading behaviour accordingly.

B. Prediction Engine based on Financial news

One of the difficulties in prediction comes from the complexities associated with market dynamics, where parameters are not fully defined and are constantly shifting. Details from breaking news stories or quarterly reports can dramatically affect share price of a security. Implementing computational techniques to this textual data shapes the basis of financial text mining[5]. The disadvantage of this method is that it cannot respond to unexpected events other than historical norms.

The Arizona Financial Text system (AZFinText)[2] is a unique type of quant trader that emphasis on making discrete numeric forecasting which depends on the combination of stock price quotes and financial news articles. While previous textual financial research has counted on tracking price direction alone, AZFinText influences statistical learning techniques to induce numeric price predictions and then build trading decisions from them. Because of its use of financial text as one of its key means, this system differs from other quants in terms of foreseeing stock price movement.

Working of AZFinText system is as follows:-

- Extract all article terms from every article
- Terms are recognised by their parts of speech.
- Entire set of proper nouns is denoted in binary as either present or not.
- Value of stock at the time the article was released is adjoined to each news article at the model building stage.
- Then machine learning takes place.

The results shows that success of AZFinText came mostly from making predictions from stock quotes and financial news articles, whereas other approaches used complex mathematical models on a large data set of financial variables. This helps to identify a promising research direction in financial text mining.

C. Machine Learning Model for Prediction

Fundamental and technical analyses were the first two methods used to forecast stock values. Artificial Neural network (ANN) is the most commonly used technique in machine learning for prediction [3]. ANNs suffer from over-fitting problem in most of the cases because of the huge number of parameters to analyse and the little prior knowledge of the user about the significance of the input parameters in the analysed problem.

This work integrates two methods such as Particle swarm optimization (PSO) and least square support vector machine (LS-SVM). The PSO algorithm is employed to reform LS-SVM to foretell the price of daily stock. PSO will be used to get the best combination of parameters for LS-SVM. The financial technical indicators used are relative strength index, money flow index, exponential moving average, stochastic oscillator and moving average convergence/divergence. The PSO is employed continuously as global optimization algorithm to optimize LS-SVM for forecasting the stock market price.

Relative Strength Index is a technical momentum measure that checks the magnitude of recent gains to recent losses in a venture to evaluate overbought and oversold conditions of a security. MFI, Money Flow Index calculates the strength of money in and out of a security. Exponential Moving Average or EMA is the pointer that gives the exponential moving average of a field over a given period of time. The stochastic oscillator is explained as a rate of the change in the current closing price of a stock and its lowest low price, with respect to its highest high price for a given period of time. MACD or the Moving Average Convergence or Divergence is a function which computes dissimilarity between a short and a long term moving average for an asset.

Particle swarm optimization is a heuristic search technique which is originated from the behavior of social groups like fish swarms or bird flocks. PSO moves from a set of points to another set in a single loop with likely improvement using a combination of probabilistic and deterministic rules. Because of its ability to effectively solve highly nonlinear, mixed integer optimization problems that are typical of complex engineering systems, intuitiveness and ease of implementation, it has been popular in academia and industry. Even though the principle, survival of the fittest, is

not used in particle swarm optimization, it is usually rated as an evolutionary method. Optimization is attained by giving each individual in the search space a memory for its earlier successes information and paving a way to blend this processed information into the movement of the individual.

The model, LS-SVM-PSO converges to global minimum. Also, it is able to get over the over-fitting problem which was in ANN during the time of fluctuations in stock market sector. In Algorithm PSO-LS-SVM, the parameters that are used can be easily tuned. The performance of the model, LS-SVM-PSO is better than LS-SVM and other compared algorithms. LS-SVM-PSO attains the lowest error rate followed by single LS-SVM, while algorithm ANN-BP has the worst one.

4. DISCUSSIONS

Here we have taken three papers on the area of stock market prediction and discussed the methods used in those papers which is mentioned in table 1. The input parameters chosen for the forecasting process is different. The first paper predicts the stock price based on twitter data and shows the Twitter mood data can fluctuates the price of the stock. Also it is proved by taking DJIA daily time series values and public mood through Twitter data.

Table 1: Methods used for equity market analysis

| Parameter used for analysis | Methods |
|------------------------------------|-----------------------|
| Twitter mood Data | GPOMS & OpinionFinder |
| Financial news | AZFinText |
| Financial technical indicators | LS-SVM-PSO |

The next paper forecasts the trends in equity sector by considering the financial news as the input parameter. It uses Arizona Financial Text system and proves the combination of stock quotes and financial news articles has an effect on trends in the market price. The third paper choses relative strength index, money flow index, exponential moving average, stochastic oscillator and moving average convergence and divergence as the parameters. This uses Particle swarm optimization and least square support vector machine algorithm and proved that it is more efficient compared to simple least square support vector machine and other compared algorithms.

5. CONCLUSION

With the rise of emerging technologies, researchers can execute computational methods to stock market data to model pattern in financial markets more efficiently and on a scale well beyond the boundary of traditional controlled experiments.

Stock market prediction is a field with lot of possibility of research. Researchers are confident that there is a window of slots where weak prediction prevails before the stock market corrects itself to equilibrium. Using this small window of slot (which may be in hours or minutes), the probability occurs to capitalize the movements of the stock price before human traders can act.

REFERENCES

1. Johan Bollen and Huina Mao. **Twitter Mood as a Stock Market Predictor**, *IEEE. Computer*, pp. 91-94, Oct 2011.
2. Robert P. Schumaker, Hsinchun Chen. **A Discrete Stock price Prediction Engine based on Financial News**, *IEEE. Computer*, pp. 51-56, Jan. 2010.
3. Osman Hegazy, Omar S.Soliman and Mustafa Abdul Salam. **A Machine Learning Model for Stock Market Prediction**, *International Journal of Computer Science and Telecommunications*, vol. 4, issue 12, pp.16-23, Dec 2013.
4. G. Cai and P. Wurman Salam. **Monte Carlo Approximation in Incomplete Information: Sequential Auction Games**, *Decision Support Systems*, vol. 39, no. 2, pp.153-168, 2005.
5. Michael Hagenau, Matthias Hauser, Michael Liebmann and Dirk Neumann. **Reading all the news at the same time: Predicting mid-term stock price developments based on news momentum**, *IEEE Computer Society*, DOI 10.1109/HICSS.2013.460, pp.1279-1290, 2012.