2018 IEEE International Conference on Data Mining Workshops (ICDMW)

Stock Market Prediction Analysis by Incorporating Social and News Opinion and Sentiment

Zhaoxia WANG, Seng-Beng HO
Social and Cognitive Computing (SCC)
Institute of High Performance Computing (IHPC);
Agency for Science, Technology and Research (ASTAR)
Singapore
wangz@ihpc.a-star.edu.sg, hosb@ihpc.a-star.edu.sg

Zhiping LIN
School of Electrical and Electronic Engineering (EEE)
Nanyang Technological University
Singapore
ezplin@ntu.edu.sg

***Abstract*— The price of the stocks is an important indicator for a company and many factors can affect their values. Different events may affect public sentiments and emotions differently, which may have an effect on the trend of stock market prices. Because of dependency on various factors, the stock prices are not static, but are instead dynamic, highly noisy and nonlinear time series data. Due to its great learning capability for solving the nonlinear time series prediction problems, machine learning has been applied to this research area. Learning-based methods for stock price prediction are very popular and a lot of enhanced strategies have been used to improve the performance of the learning based predictors. However, performing successful stock market prediction is still a challenge. News articles and social media data are also very useful and important in financial prediction, but currently no good method exists that can take these social media into consideration to provide better analysis of the financial market. This paper aims to successfully predict stock price through analyzing the relationship between the stock price and the news sentiments. A novel enhanced learning-based method for stock price prediction is proposed that considers the effect of news sentiments. Compared with existing learning-based methods, the effectiveness of this new enhanced learning-based method is demonstrated by using the real stock price data set with an improvement of performance in terms of reducing the Mean Square Error (MSE). The research work and findings of this paper not only demonstrate the merits of the proposed method, but also points out the correct direction for future work in this area.**

Keywords—Machine learning; stock market prediction; sentiment analysis; enhanced learning-based method; time series data prediction

I. INTRODUCTION

Financial time series are the nosiest and most complicated signals in the world [1] and predicting stock price is one of the most challenging tasks [2]. This makes many economists believe in the Efficient Market Hypothesis (EMH) and Random Walk Theory (RWT) [3] [4] [5], which states that the financial market is unpredictable as it follows a random walk. However, in the last few decades, the efforts of various researchers all over the world have shown that stock market is predictable. The only issue is that the existing prediction methods are not good enough. Thus, stock market prediction is still one of the most difficult topics in this research area and has always been one of the most popular research topics in the world [2].

Financial data is complex in term of its parameters. There are a lot of parameters to be considered, such as opening and closing prices, product sales, political issues and so on. Based on different parameters, the financial analysis can be divided into two types of methods - fundamental analysis and technical analysis [6]. Fundamental analysis analyzes the stock price based on the physical nature of the company by considering product sales, infrastructure of the company, etc. Technical analysis is based on the stock price movements. It is assumed that the market moves in trends and the price movements follow certain pattern.

Some researchers have leveraged the hybrid methods which combine fundamental analysis and technical analysis as they believe that only one type of information is not enough for prediction [7]. Other researchers find that news articles and social media data are also very useful in financial prediction [8][9][10]. However, the consideration of these text data makes the analysis of financial market even more complex [11] [12].

In order to deal with this unstructured, non-stationary, noisy and nonlinear time series data, learning-based algorithms are widely used in this field. Among various learning-based algorithms, Neural Networks (NN) is one of the most recognizable algorithms because of its great learning ability for solving prediction, classification and regression problems.

In this research work, we propose an enhanced learning-based method for stock price prediction by leveraging NN as the initial learning-based method and applying it to analyze the stock price time series data. The result shows that the Mean Square Error (MSE) could be reduced when effects of the events or sentiments relevant to the stock market are considered.

The rest of the paper is organized as follows. The related works are discussed in Section II. Section III describes the

2375-9259/18/$31.00 ©2018 IEEE
DOI 10.1109/ICDMW.2018.00195

stock market data as well as the news article data used in this paper. Section IV focuses on the proposed methodology followed by Section V that analyzes and discusses the results. Conclusion and future work are presented in Section VI.

II. Related Work

Stock market price prediction has become a good topic for researchers. NN is a mathematical model inspired by the biological neurons and it can be used as function approximation which means that it can automatically approximate whatever nonlinear functional form to best characterize the data involved and can be applied to solve prediction problems [13].

NN is also known to be a self-adjusting learning method that uses training data to solve prediction problems without any model constraints [14] [15]. In view of the non-linear relationship between stock market parameters and target stock price, many researches have applied NN for this prediction [16] [17] [18][19].

An NN learns the situations affecting stock market in a given environment and store the knowledge inside the networks which are then used for the prediction. Abhishek et al. have used various technical parameters related to stock market as inputs to NN with trainlm and traingdx as training functions [16]. The results showed that NN is able to predict the trend of the testing data with a relatively high accuracy of 99% using trainlm. Naeini et al. compared Multi-Layer Perceptron (MLP) NN with Elman recurrent network and simple regression on stock market prediction [17]. The inputs are also the same set of technical parameters. The result showed that MLP NN has the lowest error on predicting the amount of value changed. This indicates that NN adapts well to the dynamic nature of the stock market by providing the lowest error rate [17].

NN was also compared with the adaptive exponential smoothing (AES) method by E. L. de Faria et al. on predicting the Brazilian stock market [18]. The training for NN was performed through a windowing technique. The result showed that NN outperforms AES with a proper number of windows selected. Besides technical parameters, there are many other types of information that can be used as inputs to NN. Sutheebanjard & Premchaiswadi have considered the SET index (Thailand), the Straits Times index (Singapore), the Dow Jones index (New York), the Nikkei index (Japan), the domestic Minimum Loan Rate (MLR), the Hang Seng index (Hong Kong), and the domestic Gold prices to predict the stock exchange of Thailand index as these inputs are related to the SET index [15]. The result shows the model can obtain less than 2% Mean Absolute Percentage Error. There are also some techniques to improve the performance of NN on stock market prediction. Genetic algorithm (GA) is one of the methods. Kim & Han proposed the GA approach for feature discretization [19]. As feature discretization needs relevant and rational discretizing thresholds, the GA used can optimize both connection weights for NN and the thresholds for feature discretization [19].

Fuzzy inference system and fuzzy logic, which were initially proposed by Zadeh [20], have been leveraged to enhance the performance of NN for prediction problems [21] [22]. Adaptive Network based Fuzzy Interference System (ANFIS) is an algorithm that combines NN and fuzzy interference system. Boyacioglu & Avci examined whether ANFIS is capable of predicting the stock market return accurately [22]. Six macroeconomic variables are used as input variables for predicting the return on the stock price index of the Istanbul Stock Exchange (ISE). ANFIS was found to be rapid, easy to operate and not expensive in the research effort needed [22].

Similarly, GA can also be leveraged by ANFIS to optimize the weights of rules to improve the performance of the methods [23] [24]. The combined method is being verified against traditional fuzzy time-series model and conventional ANFIS model. The experimental results in ANFIS models showed a lower RMSE as compared to that of the traditional fuzzy time-series model, and the GA-enhanced ANFIS model performed better as compared to conventional ones [24].

There are researchers who leveraged an improved Levenberg Marquardt (LM) training algorithm to compare NN and ANFIS, and they found that the performance of NN trained using the improved training algorithm was better than that of ANFIS with lower Root Mean Squared Error (RMSE) [25]. Besides NN and ANFIS, researchers also reviewed and investigated other machine learning methods for stock market prediction problems, such as Support Vector Machine (SVM) [26], K-Nearest Neighbor (kNN) [27], and Naïve Bayes (NB) [28] based trend prediction systems. Each has its own strengths and weaknesses depending on the characteristics of the stock market datasets and situations. This implies that among learning-based prediction methods, no one algorithm is always better than the other, and the method will perform well when the problem is identified and optimized successfully.

News contents are one of the important factors that have influence on the stock market. By considering the influence of the news contents, the better performance of the predictors is expected by researchers. Nikfarjam et al. investigated the impact of financial news on stock market prediction [29]. Schumaker & Chen used financial news articles to predict the trend as well as the stock price value 20 minutes after the news is released [30]. Besides news articles, it is found that social media data is also a significant factor for stock market movements. Gomide et al. discussed whether the emotions in tweets will affect the stock market [3]. They have compared the trend of stock market with different mood of sentiment and found that only certain mood of the public sentiment is relevant to the improvement of the prediction accuracy.

However, even though there are so many enhancement strategies to enhance the performance of the learning-based method, NN, there is still no existing works that covers the deep analysis on why the public sentiment can enhance the prediction of the methods. In addition, when considering the effect of social event or public opinions, these papers only used news articles or twitter moods as inputs. They did not consider the inclusion of the technical parameters.

In this paper, we use NN as an initial learning-based case method and propose a new enhanced method which not only takes the technical parameters as inputs, but also considers news articles in a unique way.

III. DATASETS USED

Nowadays, there are various ways to obtain datasets: for example, using application programing interface (API) provided by related companies, buying data from data companies, and also obtaining them from some open source communities. In this research, two types of datasets are necessary. The first one is the stock market price data. The second one is the news articles from mainstream media.

For financial market value dataset, daily DJIA data was downloaded from Yahoo! Finance [31]. It contains 'Date', 'Open', 'High', 'Low', 'Close', 'Volume', and 'Adj Closed' (Adjusted Closed) - six columns in total. The time interval taken was from 1/1/2007 to 31/12/2016 - ten years in total.

For news articles dataset, New York Times news articles were used in this paper [32]. The dataset was obtained by using NY Times Archive API [33]. The API python libraries can be downloaded from github [34]. The articles, downloaded by using the NY Times Archive API, are based on timing and are packaged by month. The time interval taken was from 1/1/2007 to 31/12/2016 - also ten years in total. The dataset was split into two parts: 80% of the data is used as training data, and 20% of the data is used as testing data.

IV. THE PROPOSED METHODOLODGY

For implementation of the enhanced learning-based method – enhanced NN - for stock price prediction, the stock market data downloaded from the web will be processed first before they are used as inputs to the NN method.

After preprocessing of the stock market data, sentiment analysis of the news articles is performed by three different sentiment analysis methods, and a voting method is used to select the final sentiment scores. The news articles downloaded from New York Times need to be filtered before sentiment analysis is performed on them to filter and select news articles which are relevant to the stock market. Different training models were also investigated in this research. We investigated four different training models to select the best one. The descriptions of each step of the method are detailed in this section.

A. Data Preprocessing

In order to obtain an accurate sentiment score, the news articles dataset needs to be preprocessed and the analyzing & filtering function can be implemented by leveraging python codes, which can be downloaded from the website [35]. In this paper, articles were filtered by the section name. Only sections named as 'Business', 'National', 'World', 'U.S.', 'Politics', 'Opinion', 'Tech', 'Science', 'Health' and 'Foreign' will be kept. The rest of the sections are ignored. After the filter preprocessing, approximately 400,000 articles were selected from 1 million articles.

B. Performing Sentiment Analysis

After filtering and preprocessing the news articles dataset, sentiment analysis is applied. The Natural Language Toolkit Package (NLTK) [36] is used to perform sentiment analysis. It is a suite of open source Python modules, datasets and tutorials supporting research and development in Natural Language Processing. SentiMo [37][38][39] as well as Vader Sentiment Analyzer [40] are also used to obtain sentiment analysis results.

Vader Sentiment Analyzer is another open source tool used in this research. It is a simple rule-based model for general sentiment analysis. With the help of these two open source libraries as well as SentiMo software through voting methods, sentiment scores are obtained. The scores obtained on a daily basis contain four parts: mixed score, negative score, neutral score and positive score. The sentiment scores will be used in the proposed enhanced method to analyze the changes and the trend of stock prices for improving the performance of the learning-based prediction method.

C. Implementation of The Method

For implementation of the enhanced learning-based method – enhanced NN system for stock price prediction - different training models or functions are studied. In this paper, we investigated four different training models: Bayesian regularization backpropagation (trainbr) [41], Levenberg-Marquardt backpropagation (trainlm) [42] [43], Conjugate gradient backpropagation with Powell-Beale restarts (traincgb) [44], and Gradient descent with adaptive learning rate backpropagation (traingda) [45] to find the best one for this prediction problem.

Two scenarios are implemented for the comparisons. In the first scenario, only stock market price time series data is used and the influence of the news articles sentiments are not considered. In the second scenario, we use both stock market price time series data as well as the news article sentiment results, which are obtained by using the sentiment analysis methods, to study the performance of the prediction and the influence of the news sentiments. For the stock prices data, 'Close' values are used in this research work. The news sentiments data used are: negative score and positive score.

Stock market values as well as news article sentiment analysis are used in this research work. Our research results demonstrated that the incorporation of sentiment analysis enhances the performance of the learning-based methods, and the results and findings are provided in the next section.

V. RESULT AND DISCUSSION

As mentioned in the methodology section, different training algorithms can be used to optimize the performance of the NN methods and we investigated 4 training algorithms: trainbr [41], trainlm [42] [43], traincgb [44], and traingda [45]. the NN method trained by using trainlm algorithms obtained the best performance for the prediction problem in this research work, compared with the other 3 training models.

It is common sense that the effect of the event or sentiments on the stock market sometimes will last for a few days, even several weeks or months. Different events will have different effects lasting for different number of days. Analyzing the sentiment analysis results, it is discovered that the value difference between positive sentiment scores and negative sentiment scores are in the range from -0.203 to 0.088. In order to simplify the problem, we have assumed that the news at particular date will have an effect only on the next *n* days. After performing many tests, we find that for predicting the stock price time series data downloaded for this research, the suitable value for *n* is 7.

We also define a sentiment effect parameter, α, to represent the strength or percentage of the effect of the news on stock market prices. We assume that the effect is strongest on the next day and weakens over time.

During the testing process, a windowing method is used in this work. For example, window size = 3 means that we are using three consecutive day's values to predict the next day's value. We have changed the window size from 2 days to 10 days.

The prediction results obtained by using NN without considering the effects of the news sentiments are shown in Table I. The best performance of the prediction by using NN is obtained when the window size is set to 6 days as shown in Table I. This means when the stock market values of the past 6 days are used to predict the value of the seventh day, the best result is obtained for this stock price prediction problem.

Different values for sentiment effect parameter, α, are tested to analyze the effects of the sentiments. In this paper, α is set from 0% to 10% with an increment of 0.001%. Through testing the prediction model described above, different MSE values are obtained for different scenarios as shown in Table II.

The lowest MSE value has been highlighted in Table II. The result shows that when a proper value of sentiment effect parameter, α was chosen, the MSE is reduced compared with that of the method without considering the effect of sentiments. This observation is consistent with the previous work that sentiments have an effect on stock market values [46].

Fig. 1 shows, in a graphical form, the prediction results of using the last two years' data as testing data and the previous 8 years data as training data. It is observed that the prediction results using the proposed method are almost a perfect overlap with the original data. This also demonstrates the effectiveness of the proposed method.

It is also observed that if the selection of sentiment effect parameter, α, is not suitable, the performance of the method can become worse. This observation is consistent with the common sense that the selection of the sentiment effect parameter is critical.

We also encounter some phenomena which are not easy to explain: when the window size is set larger than 8 days, such as 10 days, the MSE is larger than that of 6 days. It appears that there is an optimal point at 6 days and the MSE increases after that.

We believe the key to enhancing the results that we obtain above is a sufficient amount of news sentiment data as well as stock market price data. Also, different stock price time series data may have its own natural characteristics – as we see above, there is an optimal point for the selection of different window size and different sentiment effect parameter. How many days the effect of the news sentiments will last is also an interesting question which is worthy of further research.

.

TABLE I. MSE OBTAINED USIING DIFFERENT WINDOW SIZE WITHOUT CONSIDERATION OF SENTIMENTS

MSE Obtained by Using Different Window Size without Consideration of Sentiments									
Window Size	***2 days***	***3 days***	***4 days***	***5 days***	***6 days***	***7 days***	***8 days***	***9 days***	***10 days***
MSE	4.91E-05	4.88E-05	8.40E-05	4.93E-05	**3.72E-05**	5.95E-05	5.55E-05	6.13E-05	6.51E-05

TABLE II. MSE OBTAINED USIING BEST WINDOW SIZE WITH CONSIDERATION OF SENTIMENTS

MSE Obtained by Using Best Window Size with Consideration of Sentiments									
Sentiment Effect Coefficient α	***0.001%***	***0.005%***	***0.01%***	***0.05%***	***0.1%***	***0.5%***	***1%***	***5%***	***10%***
MSE	4.53E-05	5.96E-05	**3.57E-05**	3.85E-05	4.17E-05	3.98E-05	4.19E-05	2.02E-04	5.32E-04

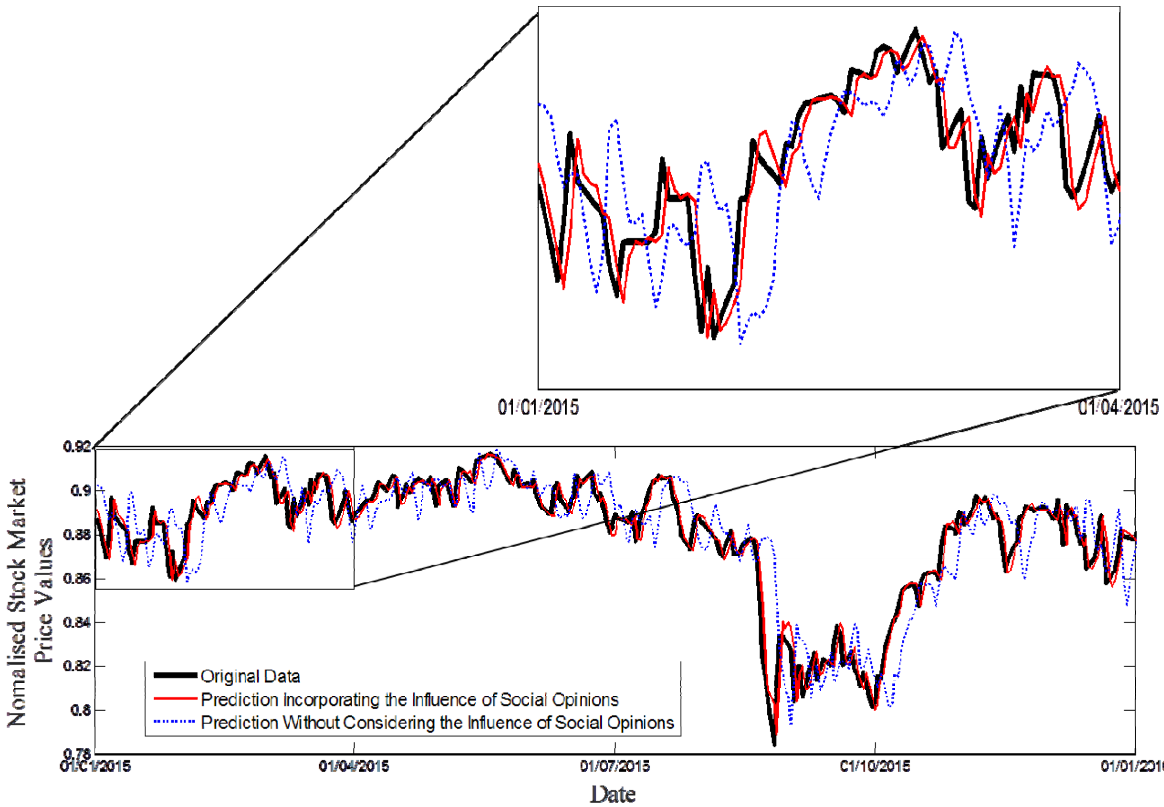

Fig. 1. The comparision between the acutual data, that predicted by the existing NN method and that predicted by our proposed method. There is a near perfect overlap between the original data and that predicted by our proposed method.

VI. Conclusion and Future Work

In this paper, the performance of learning-based method for stock market price prediction with the inclusion of news sentiment data was investigated. An enhanced learning-based method was proposed by considering the effects of news sentiments which are relevant to stock market. Using NN as a case learning-based method, the results of this research work demonstrated that the news sentiments relevant to stock market can be used to improve the performance of the learning-based predictor. The effectiveness of the enhanced method has been demonstrated using the real stock price time series data as well the corresponding news articles data.

This research also indicates that a better performance of NN method can be obtained when the proper window size is selected. The performance of learning-based method is further enhanced when the effects of the news sentiments are considered. How do researchers select the suitable sentiment effect parameters for solving this prediction problem is an unsolved issue. Therefore, investigation on the effect of the news sentiments on the stock market and the selection of sentiment effect parameters will be our future work.

Acknowledgment

The authors would like to thank Dr. Boon Kiat Quek, Dr. Kenneth Kwok and the team of project “AI Opinion-based Stock Market Trending Analysis” for their valuable discussion. The authors would also like to thank the intern from NTU, Adrian Tan, FYP students from NTU, Zhihuan Chen and Yizhou Liu, for their contributions to this research work.

References

[1] Y. S. Bu-Mostafa and A. F. Atiya, “Introduction to financial forecasting,” *Appl. Intell.*, vol. 6, no. 3, pp. 205–213, 1996.

[2] S. D. Patel, D. Quadros, V. Patil, M. Pawale, and Harsha Saxena, “Stock prediction using neural networks,” *Int. J. Eng. Manag. Res.*, vol. 7, no. 2, pp. 490–493, 2017.

[3] J. Bollen, H. Mao, and X. Zeng, “Twitter mood predicts the stock market,” *J. Comput. Sci.*, vol. 2, no. 1, pp. 1–8, 2011.

[4] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, “Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques,” *Expert Syst. Appl.*, vol. 42, no. 1, pp. 259–268, 2015.

[5] S. Dutta and R. Rohit, “Stock market prediction using data mining

techniques with R," *Int. J. Eng. Sci. Comput.*, vol. 7, no. 3, pp. 5436–5441, 2017.

[6] S. Agrawal, M. Jindal, and G. N. Pillai, "Momentum analysis based stock market prediction using adaptive Neuro-Fuzzy Inference System (ANFIS)," in *The international multiconference of engineers and computer scientists*, 2010, vol. I, pp. 526–531.

[7] S. Barik, S. Das, and S. K. Sahoo, "A hybrid forecasting model for stock value prediction using soft computing skill," *Int. J. Comput. Sci. Eng.*, vol. 5, no. 4, pp. 40–45, 2017.

[8] Z. Wang, A. Tan, F. Li, and S.-B. Ho, "Comparisons of learning-based methods for stock market prediction," in *The 4th International Conference on Cloud Computing and Security (ICCCS 2018)*, 2018.

[9] F. Z. Xing, E. Cambria, and R. E. Welsch, "Natural language based financial forecasting: a survey," *Artif. Intell. Rev.*, vol. 50, no. 1, pp. 49–73, 2018.

[10] F. Z. Xing, E. Cambria, and R. E. Welsch, "Intelligent asset allocation via market sentiment views," *IEEE Comput. Intell.*, vol. 13, no. 4, pp. 1–20, 2018.

[11] T. A. Bohn, "Improving long term stock market prediction with text analysis," The University of Western Ontario, Electronic Thesis and Dissertation Repository, 2017.

[12] X. Li, C. Wang, J. Dong, F. Wang, X. Deng, and S. Zhu, "Improving stock market prediction by integrating both market news and stock prices," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6861 LNCS, no. PART 2, pp. 279–293, 2011.

[13] T. Hill, L. Marquez, M. O'Connor, and W. Remus, "Artificial neural network models for forecasting and decision making," *Int. J. Forecast.*, vol. 10, no. 1, pp. 5–15, 1994.

[14] K. Abhishek, A. Khairwa, T. Pratap, and S. Prakash, "A stock market prediction model using Artificial Neural Network," in *2012 3rd International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2012.

[15] P. Sutheebanjard and W. Premchaiswadi, "Stock exchange of Thailand index prediction using back propagation neural networks," *2nd Int. Conf. Comput. Netw. Technol. ICCNT 2010*, pp. 377–380, 2010.

[16] K. Abhishek, A. Khairwa, T. Pratap, and S. Prakash, "A stock market prediction model using Artificial Neural Network," in *2012 3rd International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2012, no. July.

[17] M. P. Naeini, H. Taremian, and H. B. Hashemi, "Stock market value prediction using neural networks," *Comput. Inf. Syst. Ind. Manag. Appl. (CISIM), 2010 Int. Conf.*, pp. 132–136, 2010.

[18] E. L. de Faria, M. P. Albuquerque, J. L. Gonzalez, J. T. P. Cavalcante, and M. P. Albuquerque, "Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods," *Expert Syst. Appl.*, vol. 36, no. 10, pp. 12506–12509, 2009.

[19] K. Kim and I. Han, "Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index," *Expert Syst. Appl.*, vol. 19, no. 2, pp. 125–132, 2000.

[20] L. A. Zadeh, "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic," *Fuzzy sets Syst.*, vol. 90, no. 2, pp. 111–127, 1997.

[21] Z. Wang, T. Hao, Z. Chen, and Z. Yuan, "Predicting nonlinear network traffic using fuzzy neural network," in *Information, Communications and Signal Processing, 2003 and Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference on*, 2003, pp. 1697–1701.

[22] M. A. Boyacioglu and D. Avci, "An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: The case of the Istanbul stock exchange," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 7908–7912, 2010.

[23] L. Y. Wei, "A hybrid model based on ANFIS and adaptive expectation genetic algorithm to forecast TAIEX," *Econ. Model.*, vol. 33, pp. 893–899, 2013.

[24] L. Y. Wei, "A GA-weighted ANFIS model based on multiple stock market volatility causality for TAIEX forecasting," *Appl. Soft Comput. J.*, vol. 13, no. 2, pp. 911–920, 2013.

[25] M. Billah, S. Waheed, and A. Hanifa, "Stock market prediction using an improved training algorithm of Neural Network," in *Computer & Telecommunication Engineering (ICECTE)*, 2016, no. December, pp. 8–10.

[26] W. Huang, Y. Nakamori, and S.-Y. Wang, "Forecasting stock market movement direction with support vector machine," *Comput. Oper. Res.*, vol. 32, no. 10, pp. 2513–2522, 2005.

[27] K. Alkhatib, H. Najadat, I. Hmeidi, and M. K. A. Shatnawi, "Stock price prediction using K-nearest neighbor algorithm," *Int. J. Business, Humanit. Technol.*, vol. 3, no. 3, pp. 32–44, 2013.

[28] B. B. Nair and V. P. M. N. R. Sakthivel, "A genetic algorithm optimized decision tree- SVM based stock market trend prediction system," *Int. J. Comput. Sci. Eng.*, vol. 2, no. 9, pp. 2981–2988, 2010.

[29] A. Nikfarjam, E. Emadzadeh, and S. Muthaiyah, "Text mining approaches for stock market prediction," *2010 2nd Int. Conf. Comput. Autom. Eng. ICCAE 2010*, vol. 4, pp. 256–260, 2010.

[30] R. Schumaker and H. Chen, "Textual analysis of stock market prediction using financial news articles," *AMCIS 2006 Proc.*, p. paper 185, 2006.

[31] "Yahoo! Finance," *https://sg.finance.yahoo.com/quote/%5EDJI/history?p=^DJI [cite on 4 Dec. 2017]*. .

[32] "New York Times," *https://www.nytimes.com [cite on 4 Dec. 2017]*. .

[33] "NT Times Article Search API," *https://developer.nytimes.com/article_search_v2.json [cite on 4 Dec. 2017]*. .

[34] "Python-newsapi," *https://github.com/cyriac/python-newsapi [cite on 4 Dec. 2017]*. .

[35] "StockPredictions/Preparing data (Analysing & filteration).py," *https://github.com/dineshdaultani/StockPredictions [cite on 4 Dec. 2017]*. .

[36] "Natural Language Toolkit Package," *http://www.nltk.org/book [cite on 4 Dec. 2016]*. .

[37] Z. Wang, R. Goh, and Y. Yang, "SentiMo - A method and system for fine-grained classification of social media sentiment and emotion patterns"," *Singapore Pat. Appl. No. 10201407766R, Provisional Pat. filed 24 Nov.*, 2014.

[38] Z. Wang, R. Goh, and Y. Yang, "A method and system for sentiment classification and emotion classification," U.S. Patent Application 15/523,201, filed, 2017.

[39] "SentiMo," *https://www.etpl.sg/innovation-offerings/tech-for-licensing/tech-offers/2087 [cite on Dec. 2016]*. .

[40] "Vader Sentiment Analyzer," *https://github.com/cjhutto/vaderSentiment [cite on 4 Dec. 2016]*. .

[41] "Trainbr," *https://www.mathworks.com/help/nnet/ref/trainbr.html [cite on 4 Dec. 2016]*. .

[42] S. Roweis, "Levenberg-Marquardt Optimization," in *Notes, University Of Toronto*, 1996.

[43] "Trainlm," *https://www.mathworks.com/help/nnet/ref/trainlm.html [Cited 4 Sep. 2016]*. .

[44] "Traincgb," *https://www.mathworks.com/help/nnet/ref/traincgb.html [Cited 4 Sep. 2016]*. .

[45] "Traingda," *https://www.mathworks.com/help/nnet/ref/traingda.html [cite on 4 Dec. 2016]*. .

[46] B. Shri and Angelina Geetha, "Sentiment analysis for effective stock market prediction," *Int. J. Intell. Eng. Syst.*, vol. 10, no. 3, pp. 146–154, 2017.