Technical Approach in Text Mining for Stock Market Prediction: A Systematic Review

Mohammad Rabiul Islam[1], Imad Fakhri Al-Shaikhli[2], Rizal Bin Mohd Nor[3], Vijayakumar Varadarajan[4]

[1,2,3] Department of Computer Science, International Islamic Univesity Malaysia, Kuala Lumpur, Malaysia.
[4] School of Computing Science and Engineering, VIT University Chennai, India.

Article Info	ABSTRACT
	Text mining methods and techniques have disclosed the mining task throughout information retrieval discipline in the field of soft computing techniques. To find the meaningful information from the vast amount of electronic textual data become a humongous task for trading decision. This empirical research of text mining role on financial text analysing in where stock predictive model need to improve based on rank search method. The review of this paper basically focused on text mining techniques, methods and principle component analysis that help reduce the dimensionality within the characteristics and optimal features. Moreover, most sophisticated soft-computing methods and techniques are reviewed in terms of analysis, comparison and evaluation for its performance based on electronic textual data. Due to research significance, this empirical research also highlights the limitation of different strategies and methods on exact aspects of theoretical framework for enhancing of performance.

Corresponding Author:

Mohammad Rabiul Islam,
Department of Computer Science,
International Islamic University Malaysia.
50728, Gombak, Kuala Lumpur, Malaysia.
Email: rabiul.islam@live.iium.edu.my

1. INTRODUCTION

Since last decade, economic textual news categorization immensely influence on stock prediction in terms changing our socio economy. In general, stock price trends and decision taken through data mining techniques from historical price decorations, human behaviour and indicators. Within the text mining techniques, stock predictive model (SPM) need to improve by the rank search method with the inclusion of gains and ratios along with forward selection methods by integrating dimensionality reduction techniques e.g. principle component analysis (PCA) for huge quantity of characteristics text [4]. Throughout this model, enormous characteristics textual data need the degradation of accurateness with less efficiency for its appropriate classification. Various executive methods in soft-computing, e.g. text classification, text document summarization, taxonomy creation, clustering and other computational approach boost the interdisciplinary field of data mining activities [1]. Basically, the study of the research focused on different methods and techniques based on the proposed PCA techniques for usability of textual and numerical data on the stock price trend.

2. ELECTRONIC TEXT MINING

Electronic forms of text document become an important field of research due to availability of web text and text mining categorization [3]. Over the last few decades, historical large amount of data has been

stored electronically and extend the level of volume in future [12]. Theoretical, conceptual and categorized based electronic text mining forms able to bring the significance research values for its advantages that briefly described blow.

2.1. Theoretical Methods in Text Mining

Historic data is well founded that holds the essential information for predicting the future direction of stock market that bring out the advantages of Knowledge Discovery in Text (KDT) which deal with extraction of pattern from textual data [9]. Beside of this theory of Random Walks is based on the efficient-market hypothesis (EMH) that hold the best trading strategy "buy and hold" [10]. In this case, text mining of news articles gives the intrinsic value for trading decision through analytical technique. On the other hand, Support Vector Machine (SVMs) is well-founded for theoretical understanding and analysis [11]. Theoretically data mining is well founded due to historical data that holds key facts for prediction [12].

2.2. Conceptual Text Mining

Concept based text model analyzed and followed by three components [2]. First, statistical analyser able to analysis through each term of the sentence and levels of documents which enable to set up the weighting scheme based on important or non-important terms. Second is Conceptual Ontological Graph (COG) that represents the concept according to the sentence meaning. Third is conceptual extractor that could be distinguished based on COG and denote important concepts. Other popular term on text mining is bag-of-words (BOW) that hold semantic conceptual information to classify the text based on categories from the external knowledge repositories [3].
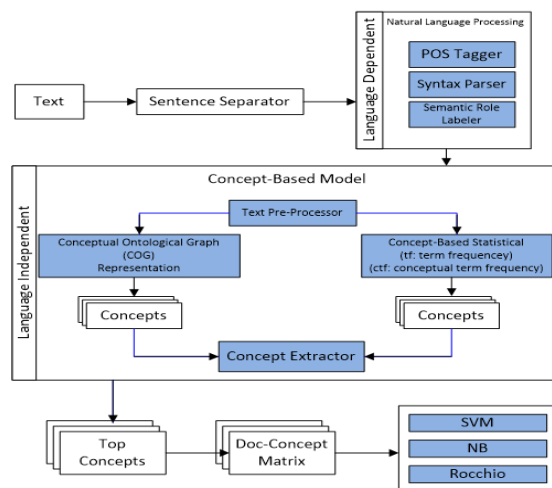

Figure 1: Text Mining in Concept-Based Model [4]

The main factor of conceptual document and sentence are depending on weighting captured that extracted from it. Based on weighting scheme, sentence semantics is assigned through conceptual statistical analyser and conceptual ontological graph [4]. Figure 1 represents the concept-based model. This textual based conceptual model brings analysis of raw text document from concept-based statistical analyser, concept extractor and conceptual ontological graph representation. In this manner, concept-based statistical term run over the sentence or document levels to find the statistical terms instead of single term [4]. Basically, this model brings the advantages of statistical concept-based weighting scheme that enhance the concept to COG document level that able to bring out more accurate result in text mining [4].

2.3. Categorized Base Text Mining

Text categorization weighting schemes on term documents followed by statistical informative through into two categories: unsupervised and supervised that approached by multilingual text categorization [3]-[6]. The advantages of this text categorization are brought by genetic algorithm that used to optimize and build user template [5]. In such case, best text categorization could be obtained by using local dictionaries and local features [7]. Moreover, Instant based learning algorithm help to categorize the closest feature space from training set that basically mapped into multi-dimensional feature space [8]. Text categorization performance

has improved based on corpus-based thesaurus and WordNet, where k-Nearest Neighbour (k-NN) algorithm with back propagation neural network algorithms able to be achieved text categorization result [16].

3. TECHNICAL APPROACH OF TEXT MINING

Technical approach on financial text data used to analyse through many classification tasks. The study of this article briefly presents performance of technical approach on stock market.

3.1. Analytical Approach

Textual data in financial market are applied through two analytical approaches [13]. First: Applied maximum entropy text classification for the prediction of whole body in text article. Second: Applied the genetic algorithm for learning simple rules based on numerical data of trading volume. Clustering and classification algorithms are applied over the features extracted through text mining approaches on the stock market news and time series techniques [14]. In this case, frequent term-based text clustering (FTC) and (HFTC) Hierarchical clustering is to be applied for text clustering [15]. The advantages found by FTC that covering whole database and HFTC covering the graph-structuring clustering.

3.2. Contextual Approach

The advantages of contextual approach help to define the sentiment clue that is syntactically engaged to sentiment topic in a sentence. In this case, the contextual feature part is assisting by extracted sentiment clues of polarities with the large amount of training data which is contextually included that identified by Bootstrapping algorithm [17].

3.3. Text Document Approach

In the traditional term weighting schemes, textual document analyzed through tf-idf method that only exploited by statistical information terms within documents [3]. In the Text document approach TF-IDF equation used in the information retrieval field.

$$TF\text{-}IDF(t_i,d_j) = count(t_i,d_j) \times log$$

Aforementioned equation, d_j indicate document, t_i indicate term in the documents with term frequency (*tf*) and corpus refers documents in the corpus with count_doc (t_i, corpus). In terms of features selection, the moderate number of distinct terms as collection of text documents operated by this method. Moreover, text features selection also improves the classification accuracy via this method [18].

3.4. Text Document Approach

Features based text mining in stock market prediction attempt to approach through various analytical techniques as briefly described here.

3.4.1. Pattern Based

Pattern based discovering data focus on high-level language that directly used by human like semantic query optimizer and expert system [24]. The only disadvantages of pattern based features is discovered the knowledge from database that could be tautological or uninteresting.

3.4.2. Dictionary Based

The advantages of data dictionary in text mining defined as the syntax of database. By systematic approach data can be stored in the data dictionary or manually by expert [24]. E.g. Retrieval system from clinical database to know patient major diagnostic categories and result may come as error or correct categories from data storage.

3.4.3. Weighting Schemes Based

In the field of information retrieval, weighting schemes features able to enhance classification accuracy in where *document representation* act on implicit syntactic indicators [25], favourability measures [26], stylistic and syntactic feature [27] that may select from diverse sources.

4.   INFORMATION RETRIEVAL (IE) METHOD IN THE FIELD OF TEXT MINING

Last few years, many innovative methods of text mining have been used for text analysis in different purpose as briefly describe here with its advantages [1]. Large amount of Meaningful corpus data is also performed by IE [55].

4.1. News Classification

Automated textual news classification method applied on the financial field of stock price prediction with the semantics text and positive or negative feedback of stock price [19].

4.2. Time Series Prediction

Time series analysis techniques applied in the field of information retrieval for stock market prediction which is another approach that retrieve the financial informative text through classification or clustering algorithm with tf-idf and signal processing methods over the features extracted [20]. The time series prediction brings the advantages with its own internal model which is original approach by artificial neural networks (anns). This model also offers on ANN for qualitative methods that usually applied on weather, stock marker, medical, economic and business in where traditional methods failed to provide [21].

4.3. Fuzzy Methods

Multi variable fuzzy forecasting is also applicable to fuzzy rule interpolation and fuzzy clustering techniques [23]. This proposed method applied on Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) data. The experimental results showed that the proposed method produces forecasting results better than several existing methods [22].

5.   VARIOUS ALGORITHM AND APPLICATION TECHNIQUES IN TEXT MINING.

Different algorithms and techniques are applicable on big database. Machine learning methods are mostly used for enhancing better deal with the issues to discoverable database [24]. Other mining techniques which are also applicable based on requirement for retrieval are briefly given below.

5.1. Text Mining Via Sparse Matrix Factorization

Most elements in sparse matrix is zero, if most of the elements are nonzero then its consider as dense, in this case zero valued elements divided by the total number of elements that known as sparsity of the matrix. In terms of text mining, sparse matrix factorization techniques incorporate within three unified frameworks [28]. (1) Correlation among different stocks, (2) historical stock prices and (3) newspapers content to predict stock price movement. Advantages of this techniques is latent factor model which can be characterized the stock price based on certain day function of the latent features.

5.2. Polarization Technique

Structural sentence based sentiment analysis, opinions, emotions and sentiments expressed in text that may focused on sentiment polarity classification which determining the opinion of text and hold positive or negative sentiment [29],[30]. The advantages of sentiment classification in polarity form that depends on subjectively classification instead of objectively form of sentences. But disadvantages of such classification are that the sentiment polarity produce ambiguous in sentiment literature as it considers objective text or label for sentiment that lies between positive or negative form.

5.3. Textual Data Mining Through Machine Learning

Different machine learning (M. Learning) techniques are available to use for stock market prediction [36]. E.g. GA/TDNN, ANFIS, ICA-BPN, GA/ATNN. Text classification via machine learning considerable good method but performance of very large training corpus is concerned for its inefficiency [33]. Beside this performance, M. Learning also combine the concept of classifiers that create new direction for the improvement of individuals' classifiers. Number of researchers proved that the combination of different classifiers can improve the accuracy of classification [34], [35].

5.4. Deep Learning for Large Scale Data Classification

Deep learning is one of the most popular methods in machine learning. The new area of deep learning is able to approach variety of applications. E.g. speech recognition [39], natural language processing [41], object recognition [40] and others. The advantages of deep learning method bring the meaningful representation from an unsupervised fashion [37]. Beside this technique, deep learning is also useful for event-

driven stock price movement prediction that basically modelled with the combination of long-term and short-term events [38].

6. TEXT MINING IN WEB-BASED APPLICATION

Web-based application much attention through web intelligence that analysed the news filtering, summary of web and news recommendation system [31]. Among these, system news recommendation is much classifier than others. E.g.

- Content-based recommendation
- Utility-based recommendation
- Collaborative recommendation
- Knowledge-based recommendation
- Demographic recommendation

As seen from the research news recommendation is potential field that helped through user's news interest, filtering and summarization of personalized web news [31]. Same as the forecasting data obtain through decision support system that help by providing investment decision for traders in forex market [32] as shown below in Figure 2 that produced by BPNNFRFS.

Figure 2. General Integrated Framework for Web

Web-based trading decision support system (WFTDSS) constructed based on three-tier structure. E.g. Model Base (MB) that can handle models, Knowledge Base (KB) used to judge stock decision or determination and Database (DB) help to repository of historical data which also provide data as required for the KB and MB [32].

7. MOST EFFICIENT TECHNIQUES AND COMPARISON

Stock market prediction followed by many application techniques for the numerical or textual data. But due to large text, major techniques usually follow for better performance. Below are some major techniques with discussion of advantages and disadvantages.

Table 1. Evaluation of Different Techniques and Methods

No	Text Mining Methods	Advantages and Disadvantages of Different Methods	
		Advantages	Disadvantages
1	Genetic Algorithms.	Genetic Algorithm runs into a single financial predictor for better performance [46].	But it can't perform on financial data, because multiple data sources and techniques is key to progress [46].
2	Deep Learning	Deep learning method performs better than traditional machine learning methods [43].	This method still unable to explorer streaming data, distributed computing, scalability computing [42].
3	Machine Learning	Naïve Bayes classification approach relatively lower than other discriminative algorithms [44].	ANN capable of learning complex nonlinear relationships but difficult to finely model for human reasoning of logical process [45].
4	Apriori-like algorithm	It can only perform well in database consisting of short frequent sequences [51].	This algorithm is time-consuming for generating of nTerms sequences in the framework [52].
5	Fuzzy Algorithm	Neuro-Fuzzy framework has emerged by combining learning ability in neural network and functionality of the fuzzy expert system [53].	Fuzzy system have the insufficient learning ability and difficultly for its time consuming to determine for the correct set of parameters [54].

8.   TEXT MINING FOCUS ON FINANCIAL MARKET

Since financial market is continuous news feeding techniques so most researchers try to present the impact of news items by analysis or market movement. Text categorization component provide through the result of positive, neutral and negative categories text based on financial daily news [47]. Financial textual information and numerical data help to analysis for trading decision. In some case, text and data bring not only the effect but also find the reason of happed [50].

8.1. Financial Textual Information

Mining textual form of financial news is able to assist in trading sector for guessing probability [48]. Indeed, it has been proved that regression and technical analysis with mining technology that taken as input textual information in the framework for economical results analysis, political news those consequently influence on bankers and politicians [49].

8.2. Financial Numerical Data

It could be unable to manage large amount of numerical data by human, so knowledge discovery textual data apparently effect on all of us [49]. Instead of this, numerical time series data also utilized by model that able to forecast the financial market [50].

9.   CONCLUSION

Many soft-computing methods have basically been developed for analytical purpose on the large-scale data set. From the above discussion within pros and cons, it could be identified that featured based weighting schemes still need to work for enhancing the framework performance in terms of classification accuracy. Every now and then, soft-computing methods and techniques are usable to be applied on big data analysis based on the requirements of research area. E.g. Financial text mining early related enough research to prove through automatic news article analysis techniques that able to predict stock market price [47].

REFERENCES

[1]   M. Shrabanti and P.Anita. "New approach of Text Mining in **R**", *GESJ: Computing Science and Telecomunication*. Volume: 1, Issue: 1, Pages: 30-36, 2015
[2]   S. Shehata, August 12-15, *"A Concept-based Model for Enhancing Text Categorization",* International conference on Knowledge discovery and data mining-KDD. Sun Jose, California, USA. 2007. Page: 629.
[3]   Q. Luo, *et al*. "A semantic term weighting scheme for text categorization". *Expert Systems with Application*, Vol: 38, Issue: 10, 2011, Pages: 12708-12716.
[4]   Case Study: Thailand Stock Exchange. "Stock Price Trend Prediction using Artificial Neural Network Techniques" 2016 IEEE, ISBN: 9781509044207.
[5]   B. Baharudin, *et al.* "A Review of Machine Learning Algorithms for Text-Documents Classification". *Journal of Advances in Information Technology*, Vol: 1, Issue: 1, Feb 2010. Pages: 4-20.
[6]   Chung-Hong Lee and Hsin-Chang Yang, "Construction of supervised and unsupervised learning systems for multilingual text categorization". *Expert System with Application*, Vol: 36, Issue: 2, March 2009, Page:2400-2410.
[7]   B.C How and W. T. Kiong. *"An examination of the feature selection frameworks in the text categorization".* Asia Information Retrieval Symposium. Vol: 3689, In AIRS. 2005 Page: 558-564.
[8]   E.H. Sam*, et al* ; *"Text Categorization Using Weighting Adjusted k-Nearest Neighbor Classification"*, 5th Pacific-Asia Conference on Knowledge Discovery and Data Mining. Page: 5-65, 2001.
[9]   R. Desai, "Stock Mrket Predicdtion Using Data Mining". *International Journal of Engineering Development and Research.* 2014, Vol: 2, Issue: 2, Pages: 2780-2784.
[10] K. Aase and P. Ozturk. "Text Mining of News Articles for Stock Price Predictions". *Departmetn of Computer and Information Scienc*. Vol:3 Msc, Issue: 6, Pages: 82. June 2011.
[11] Thorsten Joachims, *"Text Categorization with Support Vector Macines: Learning with Many Relavent Features"*: ECML-98 10th European Conference on Machine Learning, Pages: 137-142, 1998.
[12] P. Sekar, *et al*. "Financial Stock Market Forecasting using data mining techniques" *International MultiConference of Engineering and Computer Scientists*. Volume: 1, Pages: 5. March 17-19, 2010.
[13] J. Thomas and K. Sycara. "Integrating Genetic Algorithms and Text Learning for Financial Prediction". *Data mining with Evolutionary Algorithms.* Pages: 72-75. 2000.
[14] S. Seker, *et al*. "Time Series Analaysis on the Stock Market for Text Mining Correlation of Economy News". *International Journal of Social Science and Humanity*. Vol: 6. Issue: 1. Pages: 23, 2014.

[15] F. Beil, M. Ester, X. Xu *et al*, *"Frequent term-Based Text Clustering"*, KDD'02: Proceedings of the eighth ACM SIGKDD international conference on knowledge discovery and data minin. Pages: 436-442, 2002

[16] C. Li, *et al.* "Text Categoricaton algorithms using semantic approaches courpus-based thesaurus and wordnet". *Expert System with Applicaions*. Vol: 39, Issue: 1, Pages: 765-772, 2012

[17] Y. Choi, *et al.* "Domain-Specific Sentiment Analaysis using Contextual Feature Generation", *Scicne and Technology*. Pages: 37-44, Issue: July, 2009.

[18] A. Uysal and S. Gunal. "Knowledge-Based System A novel probabilistic feature selection method for text classification", *Knowledge Based Systems*. Volume: 36, Pages: 226-235. 2012

[19] M. Hagenau, *et al*. "Automated news reading: Stock prie prediction based on financial news using context-capturing features", *Decision Support Systems*. Vol: 55, Issue: 3. Pages: 685-697. 2013

[20] S. Seker, et al. "Time Series Analysis on Stock Market for Text Mining Corelation of Economy News". *International Journal of Social Sciences and Humanity*. Vol: 6, Issue: 1, Pages: 23. 2014.

[21] M. Saini, "Forecasting Stock Exchange Market and Weather Using Soft Computing", Vol: 4, Issue: 5, Pages: 877-882, 2014.

[22] D.K. Kirange, *et al*. "Sentiment Analysis of News Headlines for Stock Sentiment Analysis of News Headlines for Stock Price Prediction" COMPUSOFT, *"An international journal of advance computer technology"*, Vol:V, ISSUE-III, March-2016.

[23] Shyi-Ming Chen and Y u-Chuan Chang, "Multi- Variable Fuzzy Forecasting Based on Fuzzy Clustering and Fuzzy Rule Interpolation Techniques", *Information Sciences*, Vol.180, No.24, Pages: 4772- 4783, 2010.

[24] W. Frawley & G.Piatetsky-Shapiro. "Knowledge Discovery in Database: An Overview"*AI Magazine*. Pages: 127,1992

[25] T. Wilson, *et al.* "Recognizing contextual polarity in phrase- level sentiment analysis". *Computational Linguistics*, Vol: 35, Issue: 3, September 08, Pages: 399-433, 2009.

[26] Tony Mullen and Nigel Collier. *"Sentiment analysis using support vector machines with diverse information sources"*. Proceedings of EMNLP, pages 412– 418, Association for Com- putational Linguistics. July 2004.

[27] Ahmed Abbasi, et al. "Sentiment analysis in multiple languages: Feature selection for opinion classification in web forums". *ACMTrans.* Vol: 26, Issue:3, Jun 1, 2008.

[28] F. Ming Fai Wong, *et al*. *"Stock Market Prediction from WSJ: Text Mining via Sparse Matrix Factorization"*. International Conference on Data Mining. IEEE 2014, Pages: 430-439. December 14-17, 2014.

[29] O. Hernández, *et al.* "Approaching Sentiment Analysis by using semi-supervised learning of multi-dimensional classifiers" *Neuro-computing*, Vol: 92. Pages: 98-155. 2012.

[30] B. Pang, et al. *"Sentiment classification using machine learning techniques",* Conference on Empirical Methods in Natural Language Processing (EMNLP'02), 2002, pp. 79–86.

[31] Kumar and K Muthu, "News Recommendation System Using Web Mining: A Study", *International Journal of Engineering Trends and Technology (IJETT)*. Vol: 12, ISSUE: 6. Pages: 293-299, January 2014.

[32] K. Lai, *et al.* "A neural network and web based decision support system for forex forecasting and trading." Lecture Notes in Artificial Intelligence, Pages: 243-253, 2004

[33] M. Ikonomakis, *et al,* "Text Classification using machine learning techniques" *WSEAS Transactions on computers.* Vol: 4, Issue:8; Pages: 966-974. 2005.

[34] Y. Bao and N. Ishii, "Combining Multiple kNN Classifiers for Text Categorization by Reducts", LNCS 2534, 2002, Vol: 2534, pp. 340-347, 08 November 2002.

[35] Sung-Bae Cho and Jee-Haeng Lee, "Learning Neural Network Ensemble for Practical Text Classification", Lecture Notes in Computer Science, Volume: 2690, pp. 1032 –1036. March 21-23, 2003.

[36] Z. Iqbal *et al.* "Efficient Machine Learning for Stock Market Prediction", *Journal of Engineering Research and Applications.* Vol: 3, Issue-6, Nov-Dec 2013, pp. 855-867.

[37] X. Glorot, *et al.* *"Domain Adaptation for Large Scale Sentiment Classification: A Deep Learning Approach"* Proceeding of the 28th International Conference on Machine Learning. pp. 513-520. 2011.

[38] X. Ding, *et al. "Deep Learning for Event Driven Stock Predictoin"* Proceedings of the Twenty-Forth International Joint Conference on Artificial Intelligent. pp. 2327-2333, Issue: Ijcai, 2015,

[39] P. Swietojanski *et al,* "Convolutional Neural Networks for Distant Speech Recognition," *IEEE 2014, Signal Processing Letters,* vol.21, no.9, pp.1120-1124, September 2014.

[40] X. Chen, *et al,* "Vehicle Detection in Satellite Images by Hybrid Deep Convolutional Neural Networks," *IEEE 2014, Geoscience and Remote Sensing Letters,* vol.11, no.10, pp. 1797- 1801, Octorber 2014.

[41] Q. Mao, *et al,* "Learning Salient Features for Speech Emotion Recognition Using Convolutional Neural Networks," *IEEE Transactions on Multimedia*, vol.16, no.8, pp. 2203- 2213, December 2014.

[42] M.Najafabadi, *et al.* "Deep learning applications and challenges in big data analytics" *Journal of big data*. Vol:2, Issue:1, Pages:1, 2015.

[43] Y. Yan, *et al.* "Deep Learning for Imbalanced Multimedia Data Classification" Vol. 00, pp. 483-488, 2015.

[44] B. Baharudin, *et al.* "A Review of Machine Learning Algorithm for Text-Documents Classification" *Journal of Advances in Information Technology.* Vol: 1, Issue : 1, Pages: 4-20, 2010.

[45] H. Nhu, *et al.* "Prediction of Stock Price Using an Adaptive Neuro-Fuzzy Inference System Trained by Firefly Algorithm". *International Computer Science and Engineering Conference (ICSEC),* 2013.Pages: 302-307.

[46] J. Thomas and K. Sycara. *"Integrating Genetic Algorithms and Text Learning for Financial Prediction" Data Mining with Evolutionary Algorithms*. In Proceeding of the Genetic and Evolutionary Computing Conference (GECCO) , Las Vegas, Nevada, Pages. 72-75, 2000.

[47] K. Aase and P. Ozturk, "Text Mining of News Articles for Stock Price Predictions" Department of Computer and Information Science. June 2011, Page: 82.

[48] M. Hagenau, *et al*. "Automated news reading: Stock price prediction based on financial news using context-capturing features." *Desicon Support Systems*. 2013, Vol:55, Issue:3, Pages: 685-697.

[49] B. Wuthrich, *et al*. "Daily Stock Market forecast from textual web data" International Conference on System, Man and Cybernetics. Vol:3, Pages: 1-6. 14 Oct 1998.

[50] S. Seker, *et al*. "Time Series Analysis on Stock Market for Text Mining Correlation of Economy News" *International Journal of Social Science and Humanity*. 2014, Vol: 6, Issue:1 Pages:23.

[51] J. Pei, *et al*. *"PrefixSpan: Mining sequential patterns efficiently by prefix- projected pattern growth,"* Proceedings of Int. Conf. on Data Engineering (ICDE'02), Heidelberg, Germany, 2001, pp. 215-224.

[52] S. Wu, *et al*. *"Automatic Pattern-Taxonomy Extraction for Web Mining."* International conference on web Intelligence. IEEE/WIC/ACM (WI 2004) 20-24 Sept. 2004, pp.242-248.

[53] Chin-Shien Lin, "Can the neuro fuzzy model predict stock indexes better than its rivals?" 2002 Depart of Finance, Graduate School of Business Administration. Providence University. 2002.

[54] H. Nhu, *et al*. *"Prediction of stock Price Using an Adaptive Neuro-Fuzzy Inference System Trained by Firefly Algorithm"* International Computer Science and Engineering Conference, ICSEC 2013. Page: 302-307.

[55] Ramzan Talib and Muhammad Kashif. "Text Mininig: Techniques, Applications and Issues" *International Journal of Advanced Computer Science and Applications. IJACSA* 2016, Vol: 7, Issue No:11, 2016, Page: 414-418,

BIOGRAPHIES OF AUTHORS

Mohammad Rabiul Islam, he received his BSc-degree in Business Information Technology and later he received his Master's Degree in computer science sequentially from LUCT and Universiti Tun Hussein Onn Malaysia. Currently he is a Ph.D. candidate at International Islamic University Malaysia. The area of his research interest lies in data mining, sentiment analysis and improving text categorization in the field of soft-computing techniques. He also rolling his research over the forex and stock market within computer science. He has several conference and journal papers based on emerging economy.

Imad Al-shaikhli is a professor and the head of research at IIUM (International Islamic University Malaysia). He is also a lecturer at the Faculty of Information and communication Technology.
He is a IEEE senior member, obtained his BSc (Hon) in Mathematics, MSc in Computer Science from Iraq, and Ph.D degree from Pune University, India, 2000. He has been the editor in chief of International journal on Advanced Computer Science and Technology Research since 2011 now, and the general chair of the international conference on Advanced Computer Science Applications and Technologies since 2012 till now. He obtained a US patent for his work with his Ph.D student on smart traffic light with accident detection system on 2nd December 2014. Prof. Imad has published more than 100 papers, journals and book chapters in addition to three books.

Dr. Rizal Mohd Nor completed his PhD in Computer Science from Kent State University, Kent, OH, USA 2012 and Master in Business Administration from International Islamic University Malaysia, 2004. He also hold the double degree in Electrical Engineering and Computer Engineering from Johns Hopkins Unviersity, MD, USA 2000. He is a researcher in cryptocurrency with in-depth knowledge of the crypto-currency ecosystem and profound understanding of its industry. Currently, he has been involved in blockchain projects, forums and seminars in fintech and blockchain applications. He advises several companies and NOGs on their fintech and blockchain implementation as well as regulatory issues pertaining to crypto-currency relating to Shariah.

Prof. Dr. Vijayakumar Varadarajan. Currently, he is a Professor of school of computer science and Engineering at VIT University, Chennai, India. He has more than 16 years of experience including industrial and academic. His research interests span in computational areas covering grid computing, cloud computing, computer networks and big data. He has completed BE, CSE and MBA HRD with First Class. He has also completed ME, CSE and MBA HRD with First Class. He completed his PhD from Anna University in 2012. He ia a reviwer in IEEE Transactions, Inderscience and Springer Journals. Has has initiated a number of international research collaboration with university in Eurpoe, Australia, Africa and North America.