ELSEVIER
Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia Engineering 53 (2013) 453 – 462
Procedia Engineering
www.elsevier.com/locate/procedia

Malaysian Technical Universities Conference on Engineering & Technology 2012, MUCET 2012
Part 4 - Information And Communication Technology

Opinion Mining of Movie Review using Hybrid Method of Support Vector Machine and Particle Swarm Optimization

Abd. Samad Hasan Basari[a,]*, Burairah Hussin[a], I. Gede Pramudya Ananta[a], Junta Zeniarja[b]

[a]*Centre for Advanced Computing Technology (C-ACT), Faculty of Information and Communication Technology Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya, 76100 Durian Tunggal, Malacca, Malaysia*
[b]*Faculty of Information and Communication Technology, University of Dian Nuswantoro, Semarang, Indonesia*

Abstract

Nowadays, online social media is online discourse where people contribute to create content, share it, bookmark it, and network at an impressive rate. The faster message and ease of use in social media today is Twitter. The messages on Twitter include reviews and opinions on certain topics such as movie, book, product, politic, and so on. Based on this condition, this research attempts to use the messages of twitter to review a movie by using opinion mining or sentiment analysis. Opinion mining refers to the application of natural language processing, computational linguistics, and text mining to identify or classify whether the movie is good or not based on message opinion. Support Vector Machine (SVM) is supervised learning methods that analyze data and recognize the patterns that are used for classification. This research concerns on binary classification which is classified into two classes. Those classes are positive and negative. The positive class shows good message opinion; otherwise the negative class shows the bad message opinion of certain movies. This justification is based on the accuracy level of SVM with the validation process uses 10-Fold cross validation and confusion matrix. The hybrid Partical Swarm Optimization (PSO) is used to improve the election of best parameter in order to solve the dual optimization problem. The result shows the improvement of accuracy level from 71.87% to 77%.

Selection and peer-review under responsibility of the Research Management & Innovation Centre, Universiti Malaysia Perlis

Keywords: opinion; opinion mining; sentiment; sentiment analysis; SVM; SVM-PSO.

1. Introduction

Today, the growth of the Internet, especially Online Social Networks (OSN) is getting a lot of people's interest in the world. Many persons enthusiastic about OSN because much detail about public sentiment or opinion are General guidelines for the preparation of your text being used to research trends and estimate the products [1]. Micro blogging systems (like Twitter) are used by different individuals to show their sentiment about many subjects, thus it is a useful resource of individuals' sentiment [2].

Opinion mining or sentiment analysis comes into play when such lots of data makes it difficult to evaluate them personally [3]. Opinion mining learns individuals' views, tests, behavior, and feelings toward people, individuals, issues, activities, subjects and their features. Opinion are considerable because they are primary influences of our behaviors [4].
The subject of movies is of significant interest among the social networking individual community, recognized both by the vast number of individuals talking about movies, as well as a significant difference in their sentiments [5]. Opinion mining

* Corresponding author. *E-mail address:* abdsamad@utem.edu.my

Selection and peer-review under responsibility of the Research Management & Innovation Centre, Universiti Malaysia Perlis
doi:10.1016/j.proeng.2013.02.059

of movie reviews is measured more challenging than the opinion mining of other categories of reviews, that is product reviews [6], [7].

About some entity of interest is an important piece of info for most users during the decision making process ascertaining what others think [8]. Opinion mining is not only useful for clients, but also helps organizations to evaluate opinions and behavior of clients towards their corporation and its product, the organizations can get reviews about its product straight from clients in social networking such as Tweets.

Several opinion mining methods were released either to determine whether feedback sentences in a natural language are subjective or objective, or whether they are positive or negative. The first paper on such an end-to-end opinion mining research was already released [9]. These methods were released based on real-life needs such as movie reviews, which lead to a commercial success of this research, and most of the main companies in product fabrication based on this opinion mining methods.

Opinion mining is a latest study in the part of Text Mining (TM) that has been specific by different conditions like sentiment analysis, subjectivity analysis, or sentiment orientation [10]. Sentiment classification can be considered as a binary-classification process in traditional period [6], [11], [12]. Turney [6] suggested to figure out the direction of conditions by bootstrapping from a couple of two minimal set of seeds conditions by keeping track of the number of hits refunded from search engine. Pang et al. [11] discovered out that machine learning studying overtakes human-proposed baselines. They used Naive Bayes, Maximum Entropy Classification, and Support Vector Machine (SVM) to execute sentiment classification process on movie review records. In relation to their research, SVM maintained to do the best, and unigram with existence information changes out to be the most beneficial function.

SVM have been used efficiently in many text classification study due to their major benefits such as they are robust in high perspective areas, any function is appropriate, robust when there is a sporadically set of sample, and most text classification issues are linearly independent. Moreover, SVM has obtained great results in opinion mining and this methods has overwhelmed other machine learning methods [10].

SVM is a novel machine learning technique depending on the statistical learning concept, which resolves the issue of over-fitting, local optimal solution and low convergence ratio endured in ANN and has outstanding generalization capability in the scenario of minor sample. On the other hand, the practicality of SVM is impacted due to the problems of choosing suitable SVM parameters. Particle Swarm Optimization (PSO) is an optimization technique that very simple to apply and there are few parameter to modify [13], [26]. The aim of this study is to classify sentiment of movie reviews from Twitter data using Support Vector Machine with Particle Swarm Optimization (SVM-PSO). The expected result is the accuracy of SVM-PSO is better than using SVM.

Opinion mining (sentiment analysis, sentiment mining, sentiment classification, subjectivity analysis, review mining or appraisal extraction, and in particular situations polarity classification) offers with the computational treatment of opinion, sentiment, and subjectivity in text [8]. The goal is to determine the attitude or opinion of a presenter or author with regard to a certain topic or target. The attitude could indicate his/her reasoning, opinion or assessment, his/her affective condition (how the author wants to impact the reader). Moreover, it should be mentioned that in this perspective "subjective" does not mean that something is not real [14].

Much work has recently been undertaken in opinion mining over the last few years. Prabowo and Thelwall [15] concentrated the study in opinion mining for a joint technique. They declares that includes rule-based category, supervised learning and machine learning into a new joint technique. This technique is examined on movie reviews, product reviews and MySpace feedback. Moreover, they recommend a semi-automatic, contrasting strategy in which every classifier can give rise to other classifiers to accomplish a good level of efficiency. The results illustrate that a hybrid classification can enhance the category efficiency with regards to micro- and macro-averaged F_1. F_1 is evaluated that takes both the precision and recall of classifier's efficiency into account.

Asur and Huberman [5] illustrative that how to estimate real-word outcomes uses online social networking. They use the chatter in tweets in the subject of movie to prediction box-office earnings. The outcomes have shown that fashionable from social networking can be accurate display of upcoming outcomes. Additionally, they used the Hollywood Stock Exchange (HSX), is a popular play cash market, where the values for movie bonds can perfectly estimate actual box office outcomes, to compare with their tweet-based prototype. The fact that an easy straight line regression prototype considering only the amount of twitter posts on movie is capable of doing better than artificial money market, features the power of social networking.

Zhu et al. [16] states the viewpoint that uses personal prototype (i-model) based on artifial neural network (ANNs) to decide text sentiment classification. The person style contains expressive features, feature weight and prior expertise. Trial results show that the accuracy of personal style is higher than that of SVM and hidden Markov model (HMM) classifiers on movie review corpora.

Bollen et al. [17] used tweets to forecasts the stock markets with evaluate the writing content of regular tweets for by two feelings monitoring resources, that is OpinionFinder which measures positive vs. adverse feelings and Google-Profile of Mood States (GPOMS) which measures feelings with regards to 6 measurements (calm, alert, sure, vital, kind and happy). The accuracy is 86.7% in forecasting the daily up and down changes in the ending principles of the Dow Jones

Industrial Average (DJIA) and a decrease of the Mean Average Percentage Error (MAPE) by more than 6%.

Wong et al. [1] analysed in the perspective of movie reviews on tweets and compare the sentiment of Twitter posts on the internet population of IMDb and Rotten Tomatoes, they illustrate that scores calculated from Twitter opinions and other websites do not necessarily convert into foreseeable box-office. Yu et al. [18] conducted a research domain in the movie sector and fix the problem of mining of forecasting revenue performance, they proposed Sentiment PLSA (S-PLSA), Autoregressive Sentiment-Aware (ARSA), Autoregressive Sentiment and Quality Aware (ARSQA) prototype to utilize opinion and excellent for forecasting revenue efficiency. Liang Liu [19] suggested design and develop a movie rate and review summarization scheme in a mobile environment, they recommend a novel strategy depending on LSA to recognize related item features.

This paper is organized as follows. Section 2 describes the research design. Section 3 presents the experimental result and discussion from a simulated dataset. Section 4 give remarks and provide a conclusion.

2. Research Design

In this paper described a flow of the research design. The detail of the research design can be shown in Figure 1.

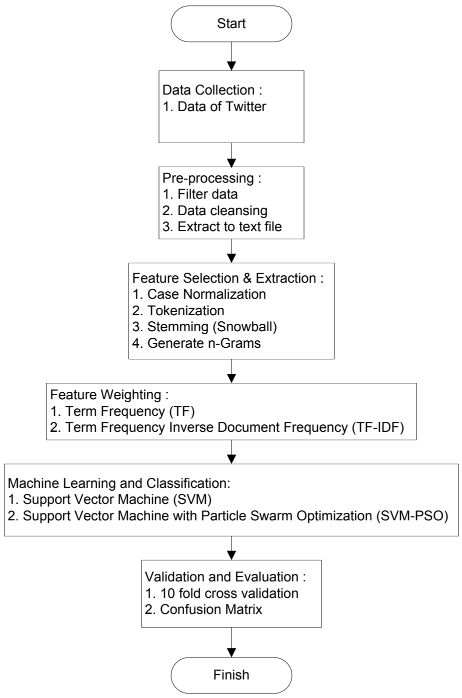

Fig. 1. Proposed Research Design

2.1 Data Collection

The data collection (data RAW) from twitter can be found in the website with the address here: http://www.stanford.edu/~alecmgo/cs224n/trainingandtestdata.zip. The data has been prepared by ensuring that the emoticons are removed off.

2.2 Preprocessing

2.2.1 Filter Data

Filter the raw data within keyword using the name of movies such as “Transformers”, “Star Trek”, “X-Men”, “The Hangover”, and “Angel and Demons”.

2.2.2 Data Cleansing

Since tweets contain several syntactic features that may not be useful for machine learning, the data needs to be cleaned such as @ (at) for link to username, url or link website (http, url, www), #(hashtag), RT(for retweet). A module that allows option of different cleaning operations was designed.

2.2.3 Extract to Text File

After cleaning the data, the message in the row data set must be extracted into a text file.

2.3 Feature Selection and Extraction

2.3.1 Case Normalization

Most English texts (and other Romance languages) are published in combined case that is, published text contains both higher and lowercase characters. The process is to turn the entire document or sentences into lowercase one.

2.3.2 Tokenization

Tokenization is splitting up the systems of text into personal terms or tokens. This procedure can take many types, with regards to the terminology being examined. For English, an uncomplicated and effective tokenization technique is to use white space and punctuation as token delimiters.

2.3.3 Stemming (Snowball)

Stemming is the procedure of decreasing relevant tokens into a single type. Typically the stemming procedure contains the recognition and elimination of prefixes, suffixes, and unsuitable pluralizations.

2.3.4 Generate n-Grams

Character n-grams are n nearby figures from a given feedback sequence. For example, a 3-gram of the phrase 'TERM' would be '_ _T', '_TE', 'TER', 'ERM', 'RM_', 'M_ _'. N-grams of dimension 1 are known as 'unigrams', of dimension 2 'bigrams', of dimension 3 'trigrams', and with a dimension more than 3 figures, they are basically termed as 'n-grams'. N-grams are mostly valuable for language identification [20] speech recognition [21]. Cavnar and Trenkle obtained a 99.8% appropriate language classification amount on Newsgroup content [20]. Also, n-grams are used to make string kernels in SVM.

2.4 Feature Weighting

2.4.1 Term Frequency (TF)

Term frequency is frequent a particular term showed up in a document while document regularity is depends on the records containing that term. Term frequency may be regarded as relatively more essential, since document regularity is depending on binary value of a term presence or lack in a document and it disregards the actual participation of a term within a document. For example, two terms having term frequencies of 10 and 100, respectively, in a document will have same document regularity of 1. This means that we cannot assess their comparative significance for a document. Term frequency instead study such information which may be beneficial in selection of essential features [22].

2.4.2 Term Frequency - Inverse Document Frequency (TF-IDF)

TF-IDF is a typical measurement used in text classification projects, but its use in opinion mining has been less extensive, and amazingly it does not look to have been used as a unigram feature weight. TF-IDF is consisting of two ratings, phrase regularity and inverse papers regularity. Term frequency is discovered by basically keeping track of frequent that a given phrase has took place in a given document, and inverse document frequency is discovered by splitting the amount of records that given term seems to be in. When these principles are increased together we get a ranking that is maximum for terms that appear regularly in a few records, and low for conditions that appear regularly in every document, enabling us to discover conditions that are essential in a document [23].

2.5 Machine Learning and Classification

2.5.1 Support Vector Machine (SVM)

SVM are often considered as the classifier that makes the greatest accuracy outcomes in text classification issues. They function by building a hyperplane with maximum Euclidean range to the nearest exercising cases [24]. Basically put, SVM signify cases as factors in area which are planned to a high-dimensional area where the planned cases of individual sessions are separated by an as large as possible tangential range to the hyperplane. New cases are planned into that same area, and based on which part of the hyperplane they are placed, they are expected to fit in with a certain category. SVM hyperplanes are completely established by a relatively small part of the training circumstances, which are known as the support vectors. The relaxations of the exercising data have no impact on the qualified classifier. SVM have been applied efficiently in text classification and in a large range of series handling programs.

Figure 2 described that how SVM does work. First, index the term of opinion ascending. Then, all the terms are weighted according to its features. If the score of weighting is greater than zero (weight>0), the term is classified as positive review. Otherwise, the term is classified as negative review.

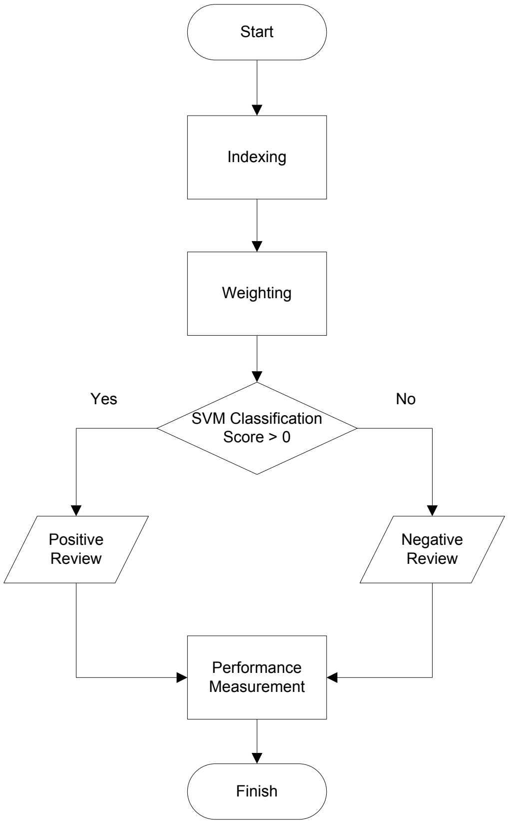

Fig. 2. Flow of SVM Model

2.5.2 Support Vector Machine with Particle Swarm Optimization (SVM-PSO)

In this paper will be explain the suggested SVM-PSO model for classification. This research originally is designed at improving the accuracy of SVM classifier by discovering the part of best useful features and calculating the best principles for regularization of kernel factors for SVM style. To experience this PSO based enhanced structure is used. A SVM-PSO technique includes two machine learning techniques by improving the parameters of SVM using PSO.

PSO begins with n-randomly chosen particles and queries for the optimal particle iteratively. Every particle is a m-dimensional vector and symbolizes an applicant solution. SVM classifier is built for each selection solution to assess its efficiency through the cross validation technique. PSO algorithm controls the selection of possible subsets that lead to best forecast accuracy. The algorithm uses the maximum suitable particle to give rise to the next creation of n-candidate particle. Therefore, on the typical, each subsequent inhabitant of selection particles fits better than forerunner. This process carries on until the efficiency of SVM converges.

PSO is used to discover maximum feature subsets by finding the best feature mixtures as they fly within the issue area from the prepared datasets. The procedure explaining suggested SVM-PSO method is as follows.

Figure 3 show that the preparation of PSO with population size, inaction weight and generations without improves. Then, the fitness of each particle will be evaluated. After that, the fitness functions and the local best and global best parameters will be compared. Once finished, the velocity and position of each particle will be updated until the value of the fitness task converges. After converging, the global best particle in the swarm is fed to SVM classifier for training. Finally the SVM classifier will be trained.

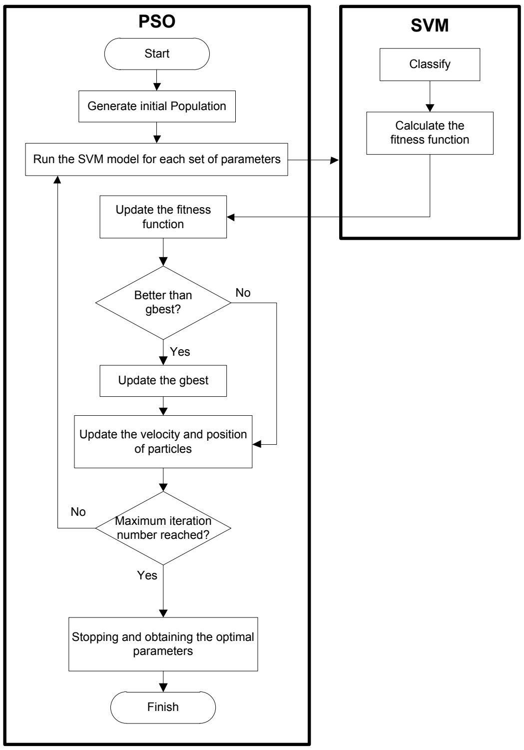

Fig. 3. Flow of SVM-PSO Model [25]

2.6 Validation and Evaluation

Process validation and evaluation in this study using k-fold cross validation and confusion matrix [27]. In this research, the value of *k* is set to 10. Thus, the dataset was divided into 10 areas, with each aspect of the information giving the same percentage of each type of data. Nine data areas were used in the training method, while the staying one was used in the testing procedure. The system was run 10 periods to allow each piece of data to take a convert as the testing data. The proportion of accuracy in category of this experiment was calculated by summing the single accuracy level for each run of testing, and then splitting the overall by 10. Since the variety of data in each category is not several of 10, the dataset cannot be portioned quite. But, the rate of the variety of data in the validation set was managed as carefully as possible to 9:1.

A confusion matrix reveals the variety of correct and mistaken forecasts made by the classification design as opposed to actual results (target value) in the data. The matrix is NxN, where N is the variety of focus on principles (classes). Efficiency of such designs is generally analyzed with the records in the matrix. The resulting table reveals a 2x2 confusion matrix for two sessions (Positive and Negative).

3. Results & Discussion

This experiment will be compared three parts, there are the comparison results of SVM using N-grams and feature weighting, the comparison result of SVM and Support Vector SVM-PSO without data cleansing, and the comparison result of SVM and SVM-PSO with data cleansing.

3.1 The Comparison Result of SVM using N-grams and Feature Weighting

In this experiment, the SVM method using n-grams is compared with feature weighting. There are three types of n-grams is used in this experiment. Those are unigram, bigram and trigram. And the feature weighting are TF and TF-IDF.

Table 1 shows the accuracy of SVM using n-gram method with several feature weighting. The result is unigram is the best n-grams method for this experiment with the accuracy is 73.13% (TF) and 72.20% (TF-IDF).

Table 1. The comparison result of svm

N-grams	Feature Weighting	
	TF	TF-IDF
Unigram	**73.13**	72.20
Bigram	**72.53**	71.87
Trigram	68.40	**68.93**

Figure 4 describes the comparison among unigram, bigram and trigram with feature weighting. The blue bars show the term frequency (TF) and the red bars show Term Frequency – Inverse Document Frequency (TF-IDF). Based on the graph above, unigrams is the best n-grams method for this experiment with the accuracy is 73.13% (TF) and 72.20% (TF-IDF).

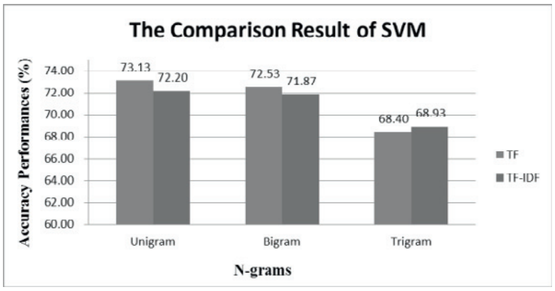

Fig. 4. The Comparison Result of SVM

3.2 The Comparison Result of SVM and SVM-PSO without Data Cleansing

This experiment shows accuracy, precision and recall for comparing both SVM and SVM-PSO without data cleansing. Based on the Table 2, SVM-PSO gives better solution for the accuracy and precision. In contrast, SVM is better value of recall than SVM-PSO.

Table 2. The comparison result of svm and svm-pso without data cleansing

	SVM	SVM-PSO
Accuracy	71.87	**77.00**
Precision	68.81	77.56
Recall	**81.87**	76.13

Figure 5 describes the comparison result of SVM and SVM-PSO without data cleansing. The blue bars show SVM without data cleansing and the red bars show SVM-PSO without data cleansing. Based on Figure 5, SVM-PSO is better solution for Accuracy and Precision. On the other hand, SVM is better value of recall than SVM-PSO.

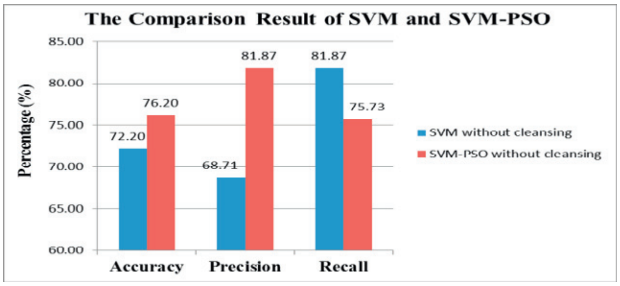

Fig. 5. The Comparison result of SVM and SVM-PSO without data cleansing

3.3 The Comparison Result of SVM and SVM-PSO after data cleansing

In this experiment shows accuracy, precision and recall for comparing both SVM and SVM-PSO with data cleansing. Based on Table 3, SVM-PSO gives better solution for the accuracy and precision. In contrast, SVM is better value of recall than SVM-PSO.

Table 3. The comparison result of svm and svm-pso with data cleansing

	SVM without cleansing	SVM-PSO without cleansing
Accuracy	72.20	**76.20**
Precision	68.71	**81.87**
Recall	**81.87**	75.73

Figure 6 describes that the comparison result of SVM and SVM-PSO with data cleansing. The blue bars show SVM with data cleansing and the red bars show SVM-PSO with data cleansing. Based on Figure 6, SVM-PSO is better solution for accuracy and precision. On the other hand, SVM is better value of recall than SVM-PSO.

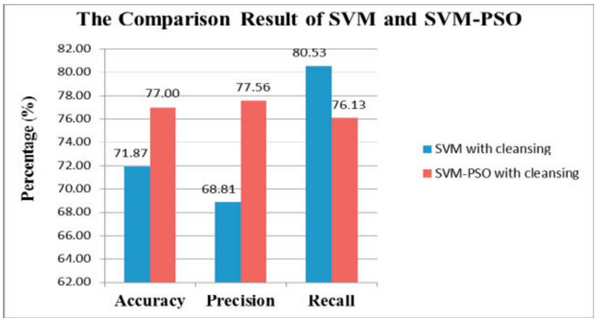

Fig. 6. The Comparison result of SVM and SVM-PSO with data cleansing

In this experiment, N-grams method effects to the accuracy level. It is showed by unigram gives the best result, which is 73.13% for term frequency and 72.20% for term frequency-inverse document frequency. Moreover, the testing result proof that using SVM-PSO gives better accuracy than SVM stands alone in the sentiment classification. On the other hand, the effect of data cleansing in the sentiment classification especially for Twitter Data is the improvement of accuracy level for SVM-PSO. However, SVM-PSO gives longer computation than computing SVM stands alone.

4. Conclusion

This study has shown that PSO affect the accuracy of SVM after the hybridization of SVM-PSO. The best accuracy level that gives in this study is 77% and has been achieved by SVM-PSO after data cleansing. On the other hand, the accuracy level of SVM-PSO still can be improved using enhancements of SVM that might be using another combination or variation of SVM with other optimization method.

This research also has done for sentiment classification using SVM with a comparison among n-grams (unigram, bigram and trigram) method and feature weighting (TF and TF-IDF). In future work the use of more combination of n-grams and feature weighting that gives a better accuracy level than this study will be considered. The work done in this research is only related to classification sentiment into two of the classes (binary classification) that is a positive class and negative class. In the future development, a multiclass of sentiment classification such as positive, negative, neutral and so on might be taken into consideration.

Acknowledgement

This paper is part of research works for Master of Computer Science (Software Engineering and Intelligence) in the Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM).

References

[1] F. M. F. Wong, S. Sen, and M. Chiang, “Why Watching Movie Tweets Won’t Tell the Whole Story?,” *Arxiv preprint arXiv:1203.4642*, no. c, p. 6, Mar. 2012.

[2] A. Pak and P. Paroubek, “Twitter as a corpus for sentiment analysis and opinion mining,” in *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC’10)*, 2010, pp. 1320-1326.

[3] L. Pan, “Sentiment Analysis in Chinese,” Brandeis University, 2012.

[4] B. Liu, *Web Data Mining*, Second Edi. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011.

[5] S. Asur and B. A. Huberman, “Predicting the Future with Social Media,” in *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, 2010, pp. 492-499.

[6] P. D. Turney, “Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews,” in *Proceedings of the 40th Annual Meeting Association for Computational Linguistics (ACL)*, 2002, no. July, pp. 417-424.

[7] P. Chaovalit and L. Zhou, “Movie Review Mining: a Comparison between Supervised and Unsupervised,” in *Proceedings of the 38th Hawaii International Conference on System Sciences*, 2005, vol. 00, no. C, pp. 1-9.

[8] B. Pang and L. Lee, “Opinion Mining and Sentiment Analysis,” *Foundations and Trends® in Information Retrieval*, vol. 2, no. 1–2, pp. 1-135, 2008.

[9] V. Hatzivassiloglou and K. R. Mckeown, “Predicting the Semantic Orientation of Adjectives,” in *35th Annual Meeting of the Association for Computational Linguistics*, 1997, pp. 174-181.

[10] M. Rushdi Saleh, M. T. Martín-Valdivia, A. Montejo-Ráez, and L. a. Ureña-López, “Experiments with SVM to classify opinions in different domains,” *Expert Systems with Applications*, vol. 38, no. 12, pp. 14799-14804, Nov. 2011.

[11] B. Pang, L. Lee, and S. Vaithyanathan, “Thumbs up? Sentiment Classification using Machine Learning Techniques,” in *ACL-02 Conf. Empirical Methods Natural Lang. Process*, 2002, pp. 79-86.

[12] K. Dave, S. Lawrence, and D. M. Pennock, “Mining the peanut gallery,” in *Proceedings of the twelfth international conference on World Wide Web - WWW ’03*, 2003, p. 519.

[13] S.-W. Fei, Y.-B. Miao, and C.-L. Liu, “Chinese Grain Production Forecasting Method Based on Particle Swarm Optimization-based Support Vector Machine,” *Recent Patents on Engineering*, vol. 3, no. 1, pp. 8-12, Jan. 2009.

[14] Y. Mejova, “Sentiment Analysis: An Overview Comprehensive Exam Paper,” 2009.

[15] R. Prabowo and M. Thelwall, “Sentiment analysis: A combined approach,” *Journal of Informetrics*, vol. 3, no. 2, pp. 143-157, Apr. 2009.

[16] J. ZHU, C. XU, and H.-shi WANG, “Sentiment classification using the theory of ANNs,” *The Journal of China Universities of Posts and Telecommunications*, vol. 17, no. July, pp. 58-62, Jul. 2010.

[17] J. Bollen, H. Mao, and X. Zeng, “Twitter mood predicts the stock market,” *Journal of Computational Science*, vol. 2, no. 1, pp. 1-8, Mar. 2011.

[18] X. Yu, Y. Liu, X. Huang, and A. An, “Mining Online Reviews for Predicting Sales Performance: A Case Study in the Movie Domain,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 4, pp. 720-734, Apr. 2012.

[19] C.-liang Liu, W.-hoar Hsaio, C.-hoang Lee, G.-chi Lu, and E. Jou, “Movie Rating and Review Summarization in Mobile Environment,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 3, pp. 397-407, May 2012.

[20] W. B. Cavnar and J. M. Trenkle, “N-gram-based text categorization,” in *Proceedings of SDAIR-94, 3rd Annual Symposium on Document Analysis and Information Retrieval*, 1994, pp. 161-175.

[21] D. Linares, J.-M. Benedí, and J.-A. Sánchez, “A hybrid language model based on a combination of N -grams and stochastic context-free grammars,” *ACM Transactions on Asian Language Information Processing*, vol. 3, no. 2, pp. 113-127, Jun. 2004.

[22] N. Azam and J. Yao, “Comparison of term frequency and document frequency based feature selection metrics in text categorization,” *Expert Systems with Applications*, vol. 39, no. 5, pp. 4760-4768, Apr. 2012.

[23] T. O’Keefe, “Feature selection and weighting methods in sentiment analysis,” *ADCS 2009*, 2009.

[24] E. Boiy, P. Hens, K. Deschacht, and M.-francine Moens, “Automatic Sentiment Analysis in On-line Text Concepts of Emotions in Written Text Concept of Emotions,” in *Proceedings of the Conference on Electronic Publishing (ELUB-2007)*, 2007, no. June, pp. 349-360.

[25] C.-yu H. C.-hung Yang and Y.-chih C. M.-chian Tsai, “A PSO- SVM Lips Recognition Method Based on Active Basis Model,” in *2010 Fourth International Conference on Genetic and Evolutionary Computing*, 2010, pp. 743-747.

[26] S.F. Pratama, A.K. Muda, Y.H.Choo and N.A. Muda, " PSO and Computational Inexpensive Sequential Forward Floating Selection in Acquiring Significant Features for Handwritten Authorship," in *2011 11th International Conference on Hybrid Intelligent Systems (HIS),* 2011, pp. 358-363.

[27] S. L. Salzberg. On comparing classifiers: Pitfalls to avoid and a recommended approach. *Data Mining and Knowledge Discovery*, vol. 1, pp. 317–328, 1997.