Science and Information Conference 2015
July 28-30, 2015 | London, UK

Metaheuristic Algorithms for Feature Selection in Sentiment Analysis

Siti Rohaidah Ahmad
Department of Science Computer
Faculty of Defence Science & Technology,
Universiti Pertahanan Nasional Malaysia,
Sungai Besi, Kuala Lumpur, Malaysia
sitirohaidah@upnm.edu.my

Azuraliza Abu Bakar[1], Mohd Ridzwan Yaakub[2]
Data Mining and Optimization Research Group
Centre for Artificial Intelligence Technology,
Faculty of Information Science & Technology,
University Kebangsaan Malaysia, Selangor, Malaysia
azuraliza@ukm.edu.my[1], ridzwanyaakub@ukm.edu.my[2]

***Abstract*—Sentiment analysis functions by analyzing and extracting opinions from documents, websites, blogs, discussion forums and others to identify sentiment patterns on opinions expressed by consumers. It analyzes people's sentiment and identifies types of sentiment in comments expressed by consumers on certain matters. This paper highlights comparative studies on the types of feature selection in sentiment analysis based on natural language processing and modern methods such as Genetic Algorithm and Rough Set Theory. This study compares feature selection in text classification based on traditional and sentiment analysis methods. Feature selection is an important step in sentiment analysis because a suitable feature selection can identify the actual product features criticized or discussed by consumers. It can be concluded that metaheuristic based algorithms have the potential to be implemented in sentiment analysis research and can produce an optimal subset of features by eliminating features that are irrelevant and redundant.**

Keywords—feature selection; sentiment analysis; opinion mining; metaheuristic algorithms

I. INTRODUCTION

Sentiment analysis (SA) functions by extracting consumers' sentiments or opinions by analyzing documents in text form [1]. At present, humans use media channels such as websites, blogs, social network websites, discussion forums and others to express their views and opinions on certain matters. Therefore, SA technology is very much needed to analyze the information or content in media. This technology is very important and useful to consumers, for example in trading of products, whereby new consumers are able to evaluate the quality of a product based on comments and opinions from previous consumers who had bought or tried the product. These comments help them to make decisions in buying the product. Meanwhile, organizations gain benefits from SA technology in product quality and service improvement based on the analysis of comments by consumers [2-3]. SA can be implemented not only to products but in many fields such as politics, services, organizations, events and issues. SA is a combination of text mining technology, natural language programming (NLP) and text classification [4]. SA is closely related to data in the form of text which involves text processing. Its function is to classify information found in the text documents such as comments on products, films, services and others, whether the documents contain positive or negative sentiment [5-7]. The method used to classify sentiments is the machine learning method combined with linguistic method to identify the type of sentiment on features found in the content of document in text form [7]. There are four important steps in SA which are: data processing, selection of features, identifying the relationship with features and sentiment words, and sentiment classifications. Each step in SA plays an important role to produce accurate classification of good sentiments.

Problems in SA involve text classification due to its relevance to the text [4]. Text classification in sentiment analysis is different from traditional text classification found in text mining. Text classification in text mining focuses on identifying topics in documents. For SA, text classification is based on sentiment on certain topics expressed by the writer, whether it is a positive or negative sentiment element [4]. According to [8], the classification step is one of the important steps in data mining, which functions to classify features based on classified training data set. They proposed dimension reduction strategy to reduce the size of large capacity training data set. The function of this strategy is to detect and eliminate irrelevant, poor and overlapping features in increasing the accuracy of data classification. This strategy can be applied in sentiment classification because the only difference between SA and text mining is on the type of data used.

The major challenges in sentiment classification are large size dimension, irrelevant and overlapping features [6, 7, 9]. Feature selection (FS) is one of the important steps in SA and functions by selecting the optimum subset feature from the real feature list without changing the initial data content. It is also a process of selecting the optimum subset from initial feature list, and the list is evaluated based on certain criteria [10]. The technique of identifying high quality features with suitable quantity of features is an issue that has always been discussed in sentiment classification based on machine learning method. Suitable FS plays an important role in SA [11]. Meanwhile, [2, 12] viewed that the accurate feature selection method can eliminate irrelevant features from the feature vector. This will reduce the size of the feature vector and increase the accuracy of sentiment classification. To deal with this issue, there are many FS techniques that have been proposed from literature study. The objective of this paper is to determine the role of FS and identify the appropriate metaheuristic algorithm for FS in SA.

Fundamental Research Grant Schema (FRGS), Kementerian Pendidikan Malaysia under Grant no FRGS/1/2013/ICT02/UKM/01/1.

II. AN OVERVIEW OF FEATURE SELECTION IN SENTIMENT ANALYSIS

SA is closely related to data in the form of text. Normally, there is an abundance of data in the form of text in websites. A text classification process is needed in order to obtain features and sentiments that are closely related. Bag-of-Words (BOW) concept represents text document used in sentiment classification using machine learning method [2, 6]. The words present in the text documents will produce a feature vector. Usually, this feature vector will cause the size of vector to become larger and this can affect the accuracy of value performance in sentiment classification [2, 6]. Machine learning method is able to administer this problem using suitable FS to eliminate overlapping and irrelevant features, and having high dimension areas [2, 6-7]. The main objective in FS is to produce feature subset from the original feature set by eliminating irrelevant and overlapping features. According to [13-14], feature searching process is divided into two which are: feature selection and feature extraction. The aim of feature selection is to select an individual feature from a group of large size features in increasing the accuracy of the classification process. Meanwhile, feature extraction means a process using different weighting schemes combined with linear features to reduce available features and produce new non-correlated features. Among the examples are Principal Component Analysis (PCA) and Linear Discrimination Analysis (LDA).

Based on the literature study, FS in SA can be divided into two categories which are the FS based on NLP and FS based on modern method. NLP is a combination of computer science, artificial intelligence and linguistic. It is an interaction between human and computer using NLP and is not a programming language. In this study, the NLP method is divided into basic and topic modeling methods. Among the techniques in NLP that use the basic method to extract features from text is Part-of-Speech (POS) tagging, document frequency, dictionary, feature weightage and others. Meanwhile, the topic modeling method is a generative probability model using distribution of vocabulary in searching for topics with text elements [15]. This method is intended to identify the topic from a collection of documents. Most topics represent features of certain products. Based on topic probability, the researcher identifies topics related to a specific document. For example, a collection of documents on comments about laptops; probable relevant topics would be battery, purchase assurance, price, and others.

FS based on modern techniques is divided into three and they are filtration, wrapping and hybrid techniques [10, 16-17]. Filtration technique is a pre-processing step that does not depend on the machine learning algorithm. During the filtration process, relevant scores for feature will be calculated, and the feature with low score value will be eliminated. Feature subset obtained is the input to algorithm classification [17]. However, the filtration technique has its weakness where it does not consider the interaction with other features and classifiers [17]. Wrapping technique has an interaction with learning algorithm and is used in evaluating features [10]. This technique's weakness is the intensive computations and high risk for over fitting [7, 17]. Therefore, the study by [10, 17] proposed the combination of filtration and wrapping techniques to overcome the prevailing weaknesses, which is known as a hybrid technique. The advantages of the hybrid technique are its interaction with classification model, intensive reduction in cost (computationally) and ability to handle a large sized data set. Wrapping technique needs intricate computation when involving complex data. This study proposed to initially reduce searching dimension areas using the filtration technique before the wrapping technique is used. A broad overview of the FS is presented in Fig. 1.

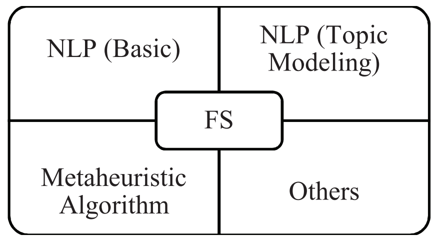

Fig. 1. FS in Sentiment Analysis

III. THE IMPORTANCE OF FEATURE SELECTION IN SENTIMENT ANALYSIS

Adnan and Song [18] mentioned that previous studies using NLP approach in identifying features had a few difficulties such as:

1) Feature selection process requires twenty manhours because the process involves manpower [19].

2) Using manpower in feature selection process and very low sentiment classification accuracy outcome, i.e. between 56% and not more than 76% [20-22]; the low accuracy outcome is may be due to unsuitable words during the classification process.

3) *According to [23], a large size feature will cause slow classification of text document due to large number of words to be processed and can cause less accuracy in sentiment classification.*

According to [7], a large size n-gram features in text document can cause intricate computation process and will affect the implementation performance of stemming process during feature set cleansing process. Besides that, n-gram feature area at high level can increase in size with additional and larger number of features. Having excess features is a major problem as it involves many features, as well as the inclination of n-gram traits to have an excess number of features. Abbasi *et al*. [7] also proposed suitable FS in handling large feature areas generated by the usage of n-gram with heterogeneous traits. Feature selection process is a difficult optimization task because it involves feature areas for large size sentiment words, i.e. 2^n for candidate feature [24]. They chose genetic algorithm to obtain optimum feature list that can produce high value during sentiment classification accuracy. Therefore, FS plays an important role in sentiment analysis to ensure the analysis conducted will produce the best result and assist consumers to evaluate and make decisions on the products on sale.

IV. Type of Feature Selection in Sentiment Analysis

This section presents the result of past researches conducted on text feature selection (FS) in SA.

A. Feature Selection based on Natural Language Processing (NLP)

The study by [25] and [26] are early researches in SA. They used POS tagging concept in identifying features in SA. Their study used association rule, i.e. Apriori algorithm to generate item set in sequence based on nouns extract by Part-of-Speech (POS) tagging. Noun is an indicator for features commented by consumers, whereas adjective is an indicator for sentiment words. Popescu and Etzioni [27] developed an unsupervised information system, i.e. OPINE, which functions to extract product features and sentiment words from the comments by consumers. OPINE extracts noun from comments and retains the words that contain high frequency value based on the given threshold value. Noun is valued by calculating the score of Point-wise Mutual Information between words and meronymy discriminators. Meronymy is a semantic relation used in linguistic. For example, "mouth" is a meronymy of "face" because mouth is a part of the face. The word "part of" represents meronymy in knowledge representation language. Their study used a manual method to extract sentiment words. Although the study by [28] used POS tagging to extract nouns as feature, they used term frequency (TF) and inverse document frequency (IDF) to obtain the list of features collection of a product. On the other hand, the study by [29] used a tokenization concept, i.e. dividing the words to word pieces known as token and eliminating punctuation marks in the text. The tokens produced would go through a stemming process, a process that would change every token to a root word, e.g. fishing, fished, and fisher changed to the root word of fish.

Khan *et al*. [30] used the method of comments by consumers to identify product features based on auxiliary verb (is, was, are, were, has, have, had). The results of the study showed that 82% were features and 85% were sentiment words using auxiliary verbs in sentences. Their study did not consider the structure of every feature that existed in every consumers' comment [31]. Meanwhile, [32] utilized the semantic relation between product feature word and sentiment word. Their approach identified product feature words and sentiment words based on synthesis and semantic information by applying dependency relations and ontology knowledge with probability model base. The study by [23] and [33] used lexicon approach based on SentiWordNet to obtain features related to sentiment. Their study applied a combination of noun, verb and adverb to identify features in the text document.

TABLE I. Text Feature Selection in Sa Using Nlp Approach

Author	Methods
[28], [34], [35]	Feature Weighting Methods
[23], [33]	Lexicon Approach
[25], [27], [30], [32], [36]–[38]	Part-of Speech (POS) Tagging

Based on Table I, previous studies of based NLP approach were mostly using POS tagging as a FS in SA.

B. Feature Selection based on Modern Methods (Metaheuristic and Other Algorithms)

Based on Table II, only Genetic Algorithm (GA) has been used as FS. Therefore, metaheuristic algorithms have the potential to be implemented in the SA research to produce an optimal subset of features by eliminating features that are irrelevant and redundant.

TABLE II. Text Feature Selection in SA Based on Modern Method (Metaheuristic Algorithm and Other Algorithms)

Authors	Methods
[24], [39]	Genetic Algorithm (GA)
[2], [39]	Information Gain (IG)
[2], [6]	Rough Set Theory (RST)
[6]	Minimum Redundancy Maximum Relevancy (mRMR)

TABLE III. Comparison Based on Advantages and Disadvantages

Authors	Techniques	Advantage	Disadvantage
[39]	IG + GA	Successfully increases sentiment classification accuracy and obtains optimum feature subset.	Data is in the form of document. Document data only considers sentiment based on the overall document without refining the document content in detail. At the document level, the SA conducted was not able to identify the likes and dislikes of consumers. Therefore, it is difficult to know in detail the feature of the product commented by consumers.
[2], [6]	IG	Can determine the importance of features in a document [2], [6].	Threshold value must be known first, and does not consider the excess between features, and there is no communication between features [2], [17].
[2], [6]	RST	Reduces the number of irrelevant, excess and noisy features. RST has an advantage of considering combination dependency trait between features [40].	a. Problem in obtaining optimum reduction of subset features (NP-hard). Therefore, metaheuristic algorithm was proposed to overcome the problem [6]. b. Process takes a long time [2], [41], [42].

Based on Table III, there are some advantages and disadvantages from each technique whereby the algorithms can be combined, refined and modified to obtain an optimal subset of features.

TABLE IV. TEXT FEATURE SELECTION IN TRADITIONAL TEXT CLASSIFICATION USING METAHEURISTIC ALGORITHM

Authors	**Methods**
[43], [44]	Ant Colony Optimization (ACO)
[14], [44], [45]	Genetic Algorithm (GA)
[47]–[49]	Particle Swarm Optimization (PSO)

Table IV shows several text feature selections in text classification selections using metaheuristic algorithm from previous studies such as ACO, GA and PSO. The result of the research [44] in traditional text classification found that ACO was able to obtain optimum feature subset compared to GA. The followings are the advantages of ACO based on the study [44]:

1) ACO has fast ability in the convergence process.

2) ACO has good finding process ability in the problem space.

3) ACO is efficient in finding minimum subset feature.

However, there are weaknesses in ACO, i.e. processing time may be affected due to dimension problem (feature total) and data size [44]. This problem can be overcome by combining ACO and RST [50, 51]. RST is a technique that can reduce feature size by eliminating overlapping features [40]. Meanwhile, in a study of SA, RST was combined with IG [2]. From the result of the research, it was determined that RST is lacking in obtaining optimum feature subset and it can also be proposed that RST should be combined with metaheuristic algorithms to obtain optimum feature reduction [41, 42]. Therefore, the advantages of ACO and RST can be potentially used as FS in SA.

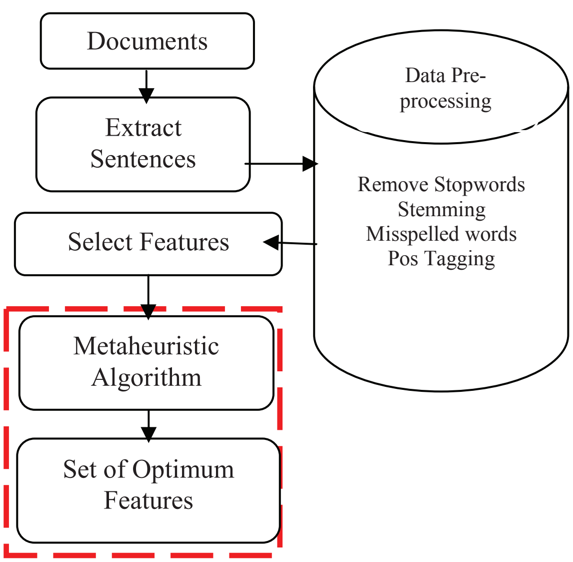

Fig. 2. Process of Product Features Selection

A general overview of the process for features selection using metaheuristic algorithms to get a set of optimum features is given in Fig. 2. The system input consists of document datasets. Firstly, we need to extract sentences from the documents and perform data pre-processing such as remove stopwords, stemming, misspelled words and POS tagging. In POS tagging, we parse the sentence, identify and extract set of features. Finally, metaheuristic algorithm is used for selecting the set of optimum features.

V. CONCLUSION

In this paper, we have reviewed FS in SA to identify their advantages and disadvantages. Metaheuristic algorithms can potentially be used as FS in SA. We proposed metaheuristic algorithms for selected optimum features from the customer reviews. In future work, we plan to identify the characteristics and evaluate appropriate metaheuristic algorithms as FS in SA. Future work would require more experiments and investigations to evaluate appropriate metaheuristic algorithms as FS in SA. Finally, we will try to build a suitable metaheuristic algorithm which could work as FS in SA.

ACKNOWLEDGEMENT

The authors gratefully acknowledge Kementerian Pendidikan Malaysia as well as the Fundamental Research Grant Scheme for supporting this research project through grant no FRGS/1/2013/ICT02/UKM/01/1.

REFERENCES

[1] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis," Found. Trends® Inf. Retr., vol. 2, no. 2, pp. 1–135, 2008.

[2] B. Agarwal and N. Mittal, "Sentiment Classification using Rough Set based Hybrid Feature Selection," in Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis (WASSA 2013), 2013, no. June, pp. 115–119.

[3] B. Agarwal and N. Mittal, "Optimal Feature Selection for Sentiment Analysis," in Computational Linguistics and Intelligent Text Processing, 2013, vol. 7817, pp. 13–24.

[4] B. Seerat and F. Azam, "Opinion Mining: Issues and Challenges (A Survey)," Int. J. Comput. Appl., vol. 49, no. 9, 2012.

[5] Y. Y. Chen and K. V. Lee, "User-Centered Sentiment Analysis on Customer Product Review," World Appl. Sci. J. 12 (Special Issue Comput. Appl. Knowledgement Manag., pp. 32–38, 2011.

[6] H. Arafat, R. M.Elawady, S. Barakat, and N. M.Elrashidy, "Different Feature Selection for Sentiment Classification," Int. J. Inf. Sci. Intell. Syst., vol. 1, no. 3, pp. 137–150, 2014.

[7] A. Abbasi, S. France, Z. Zhang, and H. Chen, "Selecting Attributes for Sentiment Classification Using Feature Relation Networks," IEEE Trans. Knowl. Data Eng., vol. 23, no. 3, pp. 447–462, 2011.

[8] J. Han, M. Kamber, and Jian Pei, Data Mining: Concepts and Techniques, 3rd ed. The Morgan Kaufmann Series in Data Management Systems, Third Edit. Morgan Kaufmann Publishers, 2011.

[9] Vinodhini G. and R. Chandrasekaran, "Effect of Feature Reduction in Sentiment Analysis of Online Reviews," Int. J. Adv. Comput. Eng. Technol., vol. 2, no. 6, pp. 2165–2172, 2013.

[10] H. Liu and L. Yu, "Toward Integrating Feature Selection Algorithms for Classification and Clustering," IEEE Trans. Knowl. Data Eng., vol. 17, no. 4, pp. 491–502, 2005.

[11] P. Koncz and J. Paralic, "An Approach to Feature Selection for Sentiment Analysis," in International Conference on Intelligent Engineering System (INES 2011), 2011, pp. 357–362.

[12] A. Sharma and S. Dey, "A Comparative Study of Feature Selection and Machine Learning Techniques for Sentiment Analysis," in Proceeding RACS '12 Proceedings of the 2012 ACM Research in Applied Computation Symposium, 2012, pp. 1–7.

[13] S. C. Yusta, "Different Metaheuristic Strategies to Solve the Feature Selection Problem," Pattern Recognition Letters, vol. 30, no. 5. pp. 525–534, 2009.

[14] H. Uğuz, "A Two-Stage Feature Selection Method for Text categorization by Using Information Gain, Principal Component Analysis and Genetic Algorithm," Knowledge-Based Systems, vol. 24, no. 7. pp. 1024–1032, 2011.

[15] H. D. Kim, K. Ganesan, P. Sondhi, and C. Zhai, "Comprehensive Review of Opinion Summarization," Illinois Environ., pp. 1–30, 2011.

[16] I. Guyon and A. Elisseeff, "An Introduction to Variable and Feature Selection," J. Mach. Learn. Res., vol. 3, pp. 1157–1182, 2003.

[17] Y. Saeys, I. Inza, and P. Larrañaga, "A Review of Feature Selection Techniques in Bioinformatics," Bioinformatics, vol. 23, no. 19, pp. 2507–2517, 2007.

[18] D. Adnan and F. Song, "Feature Selection for Sentiment Analysis Based on Content and Syntax Models," in Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, ACL-HLT 2011, 2011, pp. 96–103.

[19] C. Whitelaw, N. Garg, and S. Argamon, "Using appraisal groups for sentiment analysis," Proc. 14th ACM Int. Conf. Inf. Knowl. Manag. - CIKM '05, p. 625, 2005.

[20] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment Classification using Machine Learning Techniques," in Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2002, pp. 79 – 86.

[21] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis," in HLT '05 Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, 2005, pp. 347–354.

[22] S.-M. Kim and E. Hovy, "Determining the Sentiment of Opinions," Proc. 20th Int. Conf. Comput. Linguist. - COLING '04, p. 1367–es, 2004.

[23] T. O'Keefe and I. Koprinska, "Feature Selection and Weighting Methods in Sentiment Analysis," in Proceedings of the fourteenth Australasian Document Computing Simposium, 2009, pp. 67–81.

[24] J. Zhu, H. Wang, and J. T. Mao, "Sentiment Classification using Genetic Algorithm and Conditional Random Field," in Information Management and Engineering (ICIME), 2010 The 2nd IEEE International Conference on, 2010, pp. 193–196.

[25] M. Hu and B. Liu, "Mining Opinion Features in Customer Reviews," in Proceeding AAAI'04 Proceedings of the 19th national conference on Artifical intelligence, 2004, vol. 21, no. 2, pp. 755–760.

[26] M. Hu and B. Liu, "Mining and Summarizing Customer Reviews," Proc. 2004 ACM SIGKDD Int. Conf. Knowl. Discov. data Min. KDD 04, vol. 04, no. 2, p. 168, 2004.

[27] A.-M. Popescu and O. Etzioni, "Extracting Product Features and Opinions from Reviews," in Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing HLT 05, 2005, vol. 5, no. October, pp. 339–346.

[28] M. Abulaish, Jahiruddin, M. N. Doja, and T. Ahmad, "Feature and Opinion Mining for Customer Review Summarization," PATTERN Recognit. Mach. Intell. Proc., vol. 5909, pp. 219–224, 2009.

[29] A. S. H. Basari, B. Hussin, I. G. P. Ananta, and J. Zeniarja, "Opinion Mining of Movie Review using Hybrid Method of Support Vector Machine and Particle Swarm Optimization," Procedia Eng., vol. 53, pp. 453–462, 2013.

[30] K. Khan, B. Baharudin, A. Khan, and F. Malik, "Automatic Extraction of Features and Opinion-Oriented Sentences from Customer Reviews," in Engineering and Technology, 2010, pp. 457–461.

[31] Vinodhini G. and R. Chandrasekaran, "Sentiment Analysis and Opinion Mining: A Survey," Int. J. Adv. Res. Comput. Sci. Softw. Eng., vol. 2, no. 6, pp. 282–292, 2012.

[32] G. Somprasertsri and P. Lalitrojwong, "Mining Features-Opinion in Online Customer Reviews for Opinion Summarization," J. Univers. Comput. Sci., vol. 16, no. 6, pp. 938–955, 2010.

[33] V. K. Singh, R. Piryani, A. Uddin, and P. Waila, "Sentiment Analysis of Movie Reviews: A new Feature-based Heuristic for Aspect-Level Sentiment Classification," in 2013 International Mutli-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013, pp. 712–717.

[34] A. S. Hasan Basari, B. Hussin, I. G. P. Ananta, and J. Zeniarja, "Opinion Mining of Movie Review using Hybrid Method of Support Vector Machine and Particle Swarm Optimization," Procedia Eng., vol. 53, pp. 453–462, 2013.

[35] K. Saraswathi and A. Tamilarasi, "A Modified Metaheuristic Algorithm for Opinion Mining," Int. J. Comput. Appl., vol. 58, no. 11, pp. 43–47, 2012.

[36] S. Nadali, M. A. A. Murad, and R. A. Kadir, "Sentiment classification of customer reviews based on fuzzy logic," Inf. Technol. (ITSim), 2010 Int. Symp., vol. 2, 2010.

[37] S. Mukherjee and B. Pushpak, "Feature Specific Sentiment Analysis for Product Reviews," in 13th International Conference, CICLing 2012, 2012, pp. 475–487.

[38] A. Bagheri, M. Saraee, and F. de Jong, "Care more about customers: Unsupervised domain-independent aspect detection for sentiment analysis of customer reviews," Knowledge-Based Syst., vol. 52, pp. 201–213, 2013.

[39] A. Abbasi, H. Chen, and A. Salem, "Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web forums," ACM Trans. Inf. Syst., vol. 26, no. 3, pp. 1–34, 2008.

[40] R. Jensen and Q. Shen, "Fuzzy-Rough Sets Assisted Attribute Selection," IEEE Trans. Fuzzy Syst., vol. 15, no. 1, 2007.

[41] R. Jensen and Q. Shen, "Semantics-Preserving Dimensionality Reduction: Rough and Fuzzy-Rough-based Approaches," IEEE Trans. Knowl. Data Eng., vol. 16, no. 12, 2004.

[42] R. Jensen and Q. Shen, "New Approaches to Fuzzy-Rough Feature Selection," IEEE Trans. Fuzzy Syst., vol. 17, no. 4, 2009.

[43] M. H. Aghdam, N. Ghasem-Aghaee, and M. E. Basiri, "Application of Ant Colony Optimization for Feature Selection in Text Categorization," in Evolutionary Computation, 2008. CEC 2008. (IEEE World Congress on Computational Intelligence). IEEE Congress on, 2008, pp. 2867–2873.

[44] M. H. Aghdam, N. Ghasem-Aghaee, and M. E. Basiri, "Text Feature Selection using Ant Colony Optimization," J. Expert Syst. with Appl. An Int. J., vol. 36, no. 3, pp. 6843–6853, 2009.

[45] W. Zhao and Y. WAng, "An Improved Genetic Algorithm For Text Feature Selection," in International Conference on Intelligent Computing and Cognitive Informatics, 2010, pp. 7–10.

[46] H. Chen, W. Jiang, C. Li, and R. Li, "A Heuristic Feature Selection Approach for Text Categorization by Using Chaos Optimization and Genetic Algorithm," Math. Probl. Eng., vol. 2013, pp. 1–6, 2013.

[47] H. K. Chantar and D. W. Corne, "Feature Subset Selection for Arabic Document Categorization using BPSO-KNN," in Nature and Biologically Inspired Computing (NaBIC), 2011 Third World Congress, 2011, pp. 546–551.

[48] Y. Jin, W. Xiong, and C. Wang, "Feature Selection for Chinese Text Categorization Based on Improved Particle Swarm Optimization," in Natural Language Processing and Knowledge Engineering (NLP-KE), 2010, pp. 1–6.

[49] B. M. Zahran and G. Kanaan, "Text Feature Selection using Particle Swarm Optimization Algorithm," World Appl. Sci. J. 7(Special Issue Comput. IT), pp. 69–74, 2009.

[50] R. Jensen and Q. Shen, "Finding Rough Set Reducts with Ant Colony Optimization," Proc. 2003 UK Work. vol. 1, no. 2, pp. 15–22, 2003.

[51] R. Jensen and Q. Shen, "Webpage Classification with ACO-Enhanced Fuzzy-Rough Feature Selection," in In the Proceedings of the 5th international conference on Rough Sets and Current Trends in Computing (RSCTC 2006), 2006, pp. 147–156.