Review Article

International Journal of Advanced Computer Research, Vol 9(44)
ISSN (Print): 2249-7277 ISSN (Online): 2277-7970
http://dx.doi.org/10.19101/IJACR.PID90

A review of feature selection in sentiment analysis using information gain and domain specific ontology

Ibrahim Said Ahmad[1*], Azuraliza Abu Bakar[2] and Mohd Ridzwan Yaakub[3]
Lecturer, Department of Information Technology, Bayero University Kano[1]
Professor, Department of Information Science and Technology, Universiti Kebangsaan Malaysia[2]
Senior Lecturer, Universiti Kebangsaan Malaysia[3]

Received: 18-September-2018; Revised: 2-January-2019; Accepted: 30-January-2019

Abstract
There is a continued interest in understanding people's interest through the contents they share online. However, the data generated is massive, characterized by textual jargons and tokens that contain no sentiment or opinion value. One way of reducing the data dimension and pruning of irrelevant features is feature selection. However, the existing approaches of feature selection are still inefficient. Two prominent feature selection methods in sentiment analysis are information gain and ontology-based methods. Information gain has the disadvantage of not considering redundancy between features while ontology-based approach requires a lot of human intervention. The aim of this paper is to review these two methods. The review of these two methods shows that using the two methods in a two-step approach can overcome their limitations and provide an optimal feature set for sentiment analysis.

Keywords
Sentiment analysis, Feature selection, Information gain, Ontology.

1.Introduction
The contemporary web technology allows people to generate an unlimited amount of data online. People share their opinion, mostly in the form of writing, images or videos about facts, events or things through the web technologies. This is possible with the development of several social media platforms including Facebook and Twitter. The user- generated content on these platforms provides an important data that can be used to understand people's opinion about virtually every subject of discussion. Understanding peoples' sentiments and opinion using this data became an important field of research called sentiment analysis. Research on sentiment analysis began over a decade ago. Early works include [1–5]. An important concept in sentiment analysis is feature selection. Feature selection is a process of identifying and selecting most relevant features from a noisy data, thereby pruning irrelevant features and reducing the data dimension [6,7]. It is important in sentiment analysis due to the nature of the sentiment analysis data which is characterized by textual jargon.

The purpose of this paper is to review feature selection in sentiment analysis based on two widely used methods, ontology-based and information gain-based approaches. The research question that this study intends to address is: can Information gain and ontology-based feature selection methods be used together to provide a more accurate feature selection method? Based on this research question, the aim of this paper is to review information gain and ontology-based feature selection methods in sentiment analysis. The following are some of the contributions of this paper:

1. The problem of feature selection in sentiment analysis was reviewed, and how it relates to feature selection in topical text classification.
2. Information gain and ontology-based feature selection methods in sentiment analysis were extensively reviewed with the aim of identifying their strengths and weaknesses.
3. A proposal on how information gain and ontology-based feature selection methods can be used together for a more accurate feature selection was presented.

*Author for correspondence

This paper is organized into five related sections. The first section explores feature selection, levels of feature selection, approaches of feature selection, and the applications of feature selection. The second section discusses feature selection in sentiment analysis and the two categories of approaches to feature selection in sentiment analysis. The third section provides a review of information gain and ontology-based feature selection methods. The fourth section contains the proposed approach. The last section contains the conclusion and future work.

2.Research approach

This study is based on literature review. Two prominent feature selection methods: Information gain and ontology-based approaches were reviewed. Research studies from ACM Digital Library, IEEE Explore, Scopus, and Google Scholar were extracted. The search strings that were used for retrieving the research papers was formed by a combination of the following key phrases: "*Information gain feature selection", "*Ontology-based feature selection", "Information gain feature selection*", *Ontology-based feature selection*". Studies between 2002 and 2017 were included in the research while those of unknown conferences were excluded.

3.Overview sentiment analysis

Sentiment Analysis is a novel field of research in natural language processing (NLP) that deals with the identification and classification of people's opinion and sentiments about products and services contained in a piece of text, usually in web data [8]. The existence of these data can be attributed to the recent tremendous growth in the web technologies. There are a lot of platforms through which people can now share their opinion and sentiments about issues, products and services. These platforms include the social media, blogs, and reviews. The process of identifying and grouping the text is called sentiment classification while the positive, negative or neutral orientation of the text is called the polarity orientation of the text [9].

Several attempts have been made to present a means through which useful meaning can be extracted from the data, they mostly involve identifying if the text is contained positive, negative or neutral sentiment proposed a method of classifying a review as "recommended" or "not recommended" using a pointwise mutual information – information retrieval (PMI-IR) to estimate the semantic orientation of a phrase [4]. Investigated the use of sentiment analysis for customer satisfaction, identification from online clients' comment. Using comparative analysis of correlation between the sentiments and manually assigned score, they concluded that sentiment analysis can be used in customer satisfaction identification [10]. Proposed a feature- based vector model, a weighting algorithm-based TF–ITF algorithm, and an algorithm to extract sentiment six-tuple based on dependency parsing for Chinese sentiment classification [11]. proposed the use of sentiment oriented terminological ontologies to classify documents [12]. They proved that sentiment can effectively be identified using a probabilistic approach based on the Latent Dirichlet Allocation (LDA).

The term level of sentiment analysis is used to describe the manner in which the classification is achieved. Three levels of sentiment analysis have been identified, they are documented level sentiment analysis, sentence level sentiment analysis and aspect/entity level sentiment analysis. In document level sentiment analysis approach, the sentiment analysis is achieved by taking the entirety of the opinionated document into consideration and so, this method can only be able to tell the orientation of an opinionated document but cannot specify the details of the orientation. On the other hand, sections and sub-sections of an opinionated document are considered for sentiment analysis using sentence-level sentiment analysis. This approach provides a finer sentiment analysis than the document level sentiment analysis and can specify the polarity of different sections and sub-sections of an opinionated document. In aspect-level sentiment analysis, an opinionated document is broken down into pieces, to contain features and such features are used for sentiment analysis. In this way, it provides a fine-tuned sentiment analysis and it gives the details as to what exactly makes the orientation and the reason for the polarity orientation of an opinionated document.

Sentiment analysis is closely related to text classification, and as such sentiment analysis is often treated as a text classification problem [8]. This classification can be achieved in many ways, but the approaches can generally be grouped into supervised learning approach and unsupervised learning approach.

3.1Supervised learning approach

Supervised learning approach is sometimes referred to as a machine learning approach. In this approach, popular machine learning algorithms are used for the task of sentiment analysis in the same way they are

used in traditional text classification tasks [8]. Several machine learning algorithms have been used for sentiment analysis. The most widely used include naive Bayes, maximum entropy and support vector machines. Proposed a method of sentiment analysis based on support vector machine [13]. Their approach takes into account whether a post is subjective and whether the poster is credible or not. They argued that their method can be very effective in decision making when used the business domain. Proposed improvement to the naïve Bayesian classifier in sentiment classification [14]. Their improvement is to overcome a problem of reduced accuracy that can be encountered when using supervised learning approaches. Their experimental results showed an increased accuracy compared to support vector machines and traditional naïve Bayesian classifier. Assessed the performance of naive Bayes (NB), maximum entropy (ME), stochastic gradient descent (SGD), and Support Vector Machine (SVM) in the classification of sentiment contained in customer reviews [15]. They experimented with different parameters and concluded that a higher coefficient of *n* in n-gram yields lower classification accuracy result.

3.2Unsupervised learning approach

Sentiment Analysis using unsupervised learning approach, also referred to as lexicon-based approach is achieved by identifying features in the documents, then comparing them against sentiment lexicons developed, whose sentiment values have already been determined [16]. For example, the basic idea of a lexicon-based approach is to first develop a lexicon of both positive and negative lexicons used in expressions then analyze a test document to find them. If the document consists of more of the positive lexicons then it is considered as positive sentiment else as a negative sentiment. Three methods of building an opinion lexicon have been identified. They are the manual approach, dictionary-based approach and corpus-based approach. The manual approach is done manually and therefore takes longer time to complete and is mostly used to validate the automatic approaches. The dictionary based approach begins by manually collecting words with well-known orientation, then building a bigger lexicon. The lexicon is propagated by adding the synonyms and antonyms of the words included in lexicon, which are acquired from existing corpus [8]. The most commonly used corpora are WordNet (Miller, 1995 and Thesaurus [17]. Previous studies that used dictionary-based approach includes studies by [18] and [19] who both used variant types of dictionary-based approaches to build a sentiment lexicon for sentiment analysis. Corpus-based approach is similar to dictionary-based approach; however, it considers the semantic information and is used to develop a domain and context specific lexicon from a large corpus which is better suited for sentiment analysis. Presented a classical example of this approach [20]. They proposed an efficient method of expanding the corpus with more adjectives using a concept called sentiment consistency. Adopted Lexicon-based approach (Dictionary) to conduct sentiment analysis on tweets for consumer reviews on popular brands [21]. The researcher used both quantitative and qualitative approaches for the analysis, using QDA Miner for the former and R the later.

Recent studies leverage the relationship between different concepts and sentiment words contained in a piece of text for sentiment analysis. This is done by the use of ontology, which is a representation of concepts and their relationships. Proposed an ontology–based sentiment analysis model for calculating people's perception of a product. They adapted mathematical formulas for the calculation based on the importance of individual features [22]. Their model depends on product features, opinion orientation and strength of the features to be determined by existing opinion mining technique.

3.3Applications of sentiment analysis

Sentiment analysis has several applications. It provides a way of understanding and mining people's opinion through the contents they share on social media, thereby making it possible to understand their sentiments and opinion without direct or physical interaction with them. It also provides a way of accessing one of the most diverse form of data that covers different classes of people and across the wide age bracket. The huge amount of data can be about the weather, economy, products, politics, medical facilities, disease outbreaks or virtually anything. Applied sentiment analysis in the field of healthcare. They developed a model for advance warning and early detection of contagious outbreaks [23]. Backed by their experimental results, they argued that their model can effectively be used for monitoring contagious outbreaks on a global scale. A completely different application of sentiment analysis is in the politics. Investigated on how people use the social media for political discussion and if such discussions have any correlation with the election results [24]. They investigated using dataset pertaining German

federal election. Their results show that Twitter was widely used for political discussions.

4.Feature selection in sentiment analysis

People's opinion, comments, reviews, tweets or status updates shared online are mostly in the form of poorly structured text. This sometimes makes it difficult for even people to understand the overall sentiment contained in the text. One way of organizing such unstructured data is through text classification. Text classification involves arranging documents into relevant categorizes based on the occurrence of particular features. Sentiment analysis is thus often treated as a text classification problem. However, traditional text classification is concerned with features that distinguishes topics whereas sentiment analysis is concerned with features that distinguishes subjectivity [25]. Feature selection is an important field of research in sentiment analysis because the efficiency of the feature selection method used will determine the accuracy of the sentiment analysis [26] Several attempts have been made for feature selection using different approaches, however the approaches can broadly be grouped into statistical based approaches and lexicon-based approaches. Statistical approaches are fully automatic techniques while Lexicon based approaches need human intervention.

4.1Statistical based feature selection

Many statistical feature selection methods for topical text classification can also be used for sentiment analysis [24]. One of the earliest researches on sentiment analysis was by [3] and they used a simple statistical approach for the feature selection. They considered words with the highest frequency in the dataset to likely be associated with sentiment. This approach is also popular in traditional text classification problems and produces desirable results. Other approaches include that of [11] who proposed a feature selection method based on Fisher's discriminant ratio and [27] proposed a Gini Index- based feature selection method.

4.2Lexicon based feature selection

In linguistics, a lexicon is a vocabulary of person, language or a branch of knowledge. The lexicon-based approaches take advantage of the lexicon of a language for sentiment analysis. The basic idea of lexicon-based feature selection approach works by manually selecting some sets of words with well-known orientation, then bootstrapping this set through synonym detection or various online resources to create a larger lexicon [25]. Used lexicon – based approach to determine the semantic orientation, opinions expressed on product features in reviews [28]. In their approach, they also used verbs, nouns and idiom as an addition to the existing lexicon, they argued these words are also used to convey sentiments. Their experiment results showed that the method is effective . However, lexicon – based approaches can be time consuming, especially when there is need for manual annotation. Whitelaw et al. (2005) report that their feature selection process took 20 man-hours due to its dependence on human annotation [29].

5.Ontology-based feature selection

Ontology-based approach is sometimes classified as lexicon-based approach because it also uses the lexicon of a language. However, ontology is used to represent features contained in the text. According to Gómez-Pérez et al., ontology is defined as "formal, explicit specification of a shared conceptualization". Ontologies are used in defining features, the relationship between the features in a specific domain. Ontology is widely used as a tool for knowledge representation [30]. Since sentiment analysis is concerned with classifying the polarity of the texts contained in big data, and such data contains information about some domain knowledge, ontologies have the potential of being deployed for sentiment analysis. One of the most prominent and recent work on ontologies and sentiment analysis is by by [31]. They proposed a method of identifying features related to finance in a corpus. They used this approach to identify the semantic relationship between features in their dataset. [26] also adopted ontology learning technique by the use of OntoGen to analyse tweets. The limitation of their study is the use OpenDover for the sentiment classification, because they have no control over the classification. *Table 1* presents a summary of the related work on ontology-based feature selection.

Table 1 Related work on ontology-based feature selection

Author(s)	Year	Aim	Findings
[31]	2017	Use ontology to detect features concerning financial news.	Results show high and balanced precision and recall in differentiating correct

Author(s)	Year	Aim	Findings
			statements and errors.
[32]	2016	Proposed a feature selection method based on SVM and Fuzzy domain ontology.	Accuracy: 82.7%
[33]	2015	Used ConceptNet ontology to determine domain specific concepts which serves as features. The polarity of the feature is determined by contextual polarity lexicon and context information of a word.	Accuracy: 80.1%
[34]	2015	Ontology-based Sentiment Analysis Process for Social Media Content	Poor results
[35]	2014	Presented ontology-based model for identifying cyber-security threats, estimating their goals, and assessing their risks based on sentiment analysis.	Accuracy: 86%
[36]	2017	Proposed an ontology approach to aspect extraction in product sentiment summarization.	Accuracy: 88.73%
[37]	2017	Proposed an improve sentiment classification approach based on SVM using ontology-based feature selection.	Their method improves the performance of the SVM for classification of aspect sentiments and reduces the reliance on training data.
[38]	2016	Proposed a model for feature-based sentiment analysis using lexical resources and ontology of a car.	Their model performs better than dictionary method.
[39]	2016	Proposed a hybrid feature selection approach based on ontology and machine learning techniques.	High & Balanced precision and recall
[40]	2015	Proposed a novel ontology-based feature selection method that can be used across different domains.	Accuracy: 80%

5.1Feature selection using common-sense knowledge ontology

Common-sense knowledge can be seen as the most basic knowledge about facts, events and the world as a whole shared by most people. Common-sense knowledge is a knowledge than an average person is expected to be aware of. For example, when a person says I'm going to the library, it can simply be concluded that he intends to study or borrow a book and he will eventually be back or return the book. In Artificial Intelligence, common-sense knowledge is considered as one of the most key elements necessary for machines to understand natural language semantics [41]. For this reason, it is necessary to collect, represent and store common-sense knowledge in a formal machine-readable representation. Several attempts have been made to store and represent this knowledge, one of the earliest being Cyc and WordNet in which the knowledge is collected manually. However, recent developments have seen to automated data-driven methods of common-sense knowledge extraction and representation. One of such automated methods is ConceptNet, which is a semantic network of common-sense knowledge built from English sentences of the open mind common sense (OMCS) corpus, through an automatic process [42].

A number of investigations on feature selection using common sense ontology have been conducted. One of the early and prominent research is the use of Sentic computing [43] for the sentiment analysis. In Sentic computing, AI and semantic web techniques are utilized for efficient and effective sentiment analysis. The task sentiment analysis is accomplished using common sense reasoning and domain specific ontology. A sentiment analysis resource tool called SenticNet [44] based on Sentic computing was published. Two other versions of this tool have also been published which are SenticNet 2 [45] and SenticNet 3 [46]. Another study by [33] investigated the effect domain specific ontology, the importance of the features, and contextual information in determining the overall sentiment of the text. They used ConceptNet to extract and develop a domain specific ontology which they use to produce domain specific important features. They evaluated their method across different parameters, and got a better accuracy of 80.1% on the software review dataset.

6.Information gain-based feature selection

Information gain feature selection method is one of the most important and most popular feature selection methods. Its history can be dated back to 1948, attributed to the contributions of Claude Shannon, who invented the basic concepts of information theory, entropy and information gain. IG is used to select important features with respect to class attribute [47]. Using IG, the importance of a feature is calculated in relation to a general class. If the importance calculated surpasses a certain threshold, it is selected. Therefore, it is said to be used in dimension reduction for efficient classification. Information gain of a term can be calculated by using equation [48].

$$G(D,t) = -\sum_{i=1}^{m} P(C_i) log P(C_i) + P(t)\sum_{i=1}^{m} P(C_i|t) log P(C_i|t) + P(\bar{t})\sum_{i=1}^{m} P(C_i|\bar{t}) log P(C_i|\bar{t})$$

From Equation, C represents the document collection. $P(C_i)$ represents the probability of the i^{th} category. $P(t)$ and $P(\bar{t})$ represents the probabilities that the term t appears or does not in the document respectively. $P(C_i|t)$ and $P(C_i|\bar{t})$ represents conditional probability of the i^{th} class value given the term t appears or does not appears in the document respectively.

Information gain has widely been used as a feature selection method in many data mining tasks. In sentiment analysis, one of the earliest works is by [49]. They investigated the accuracy of standard information gain feature selection of topical applications like sentiment analysis. They showed that the standard information gain is not able to identify discriminatory features. They therefore proposed a probability redistribution procedure (PRP) to counter this problem. Their experimental results on three datasets showed an increased improvement in classifier accuracy using the PRP approach. Later works include that of [50]. They proposed an improved information gain feature selection method to address two deficiencies they identified in the existing approaches, these deficiencies are limiting document frequency's word frequency (LDFWF) and distribution information (DI). Their experimental results showed an increase in classification accuracy. *Table 2* presents a summary of related work on feature selection based on information gain.

Table 2 Related work on information gain-based feature selection

Author(s)	Year	Aim	Findings
[51]	2017	Propose an improved feature selection method based on information gain and document frequency.	The proposed approach constructs sub-features that reach better performance in classification by selecting features that have high frequency the dataset and is relevant to the output class.
[52]	2017	Proposed an approach based on information gain and word embedding for feature selection.	Experiment on Chinese text classification showed an improved result.
[53]	2016	Proposed a feature ranking method for feature selection based on information gain for aspect-level sentiment analysis.	Showed that selecting only few features during feature selection does not significantly affect the accuracy.
[54]	2016	Select most important features using information gain and K-means clustering.	Using their feature selection method, the clustering algorithm error ration was reduced from 44.48% to 21.42%

Author(s)	Year	Aim	Findings
[55]	2015	Proposed a feature selection method based on information gain that take into account the sparsity of the feature vector. The method is able to use less features to obtain a targeted performance level.	Showed that the approach is able to improve the performance of sentiment classification.
[56]	2016	Proposed an information gain-based feature selection method using a unique entropy formula in breast cancer treatment.	Experimental results of proposed method show that the proposed approach is able to select the most informative features.
[57]	2014	Proposed an improved information gain-based feature selection approach based on term frequency information and balance factor.	Experimental results of proposed approach showed a better classification accuracy than the compared approaches.

7.Proposed approach

The proposed approach is in two steps. First the information gain of all the features will be calculated and features with information gain above 0 will be selected. The domain specific ontology will then be used to further fine-tune the features and have an optimal feature subset. This will help in removing redundantly features likely to be selected by information gain. The aim is to have an approach that will counter the disadvantages of the two approaches. A feature is said to be important depending on how relevant and redundant it is. A feature is relevant if it can be used to predict the class while it represents a feature is redundant if it there are other features similar to it in the feature set. The purpose of feature selection is to select features that are highly relevant but not redundant [8]. Information Gain is used to calculate how important a feature is in a document; however, it does not take redundancy into consideration which is a major limitation, and also a threshold value is needed prior which is generally unknown [50]. This might result to the method returning a large number of features when a massive number of documents are to be considered. Another approach of feature selection is to use domain specific ontology to identify a feature set. On the other hand, knowledge-based approaches are highly dependent on context and perform poorly with indirect expressions like sarcasm [51]. *Figure 1* below represents the steps of the proposed approach.

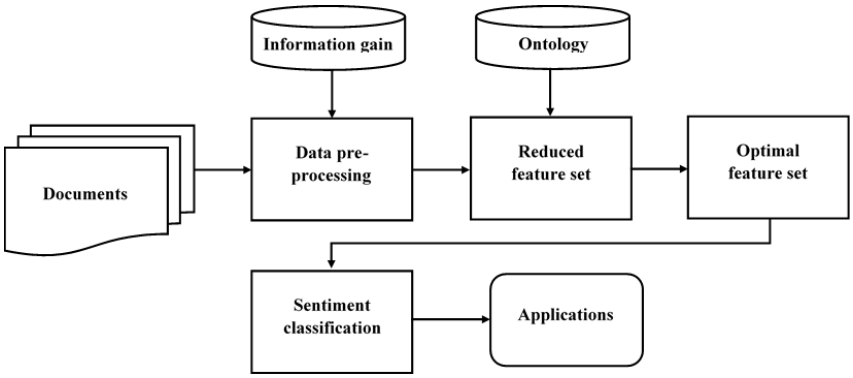

Figure 1 Proposed approach

8.Conclusion

Sentiment analysis has become an important field of research due to the increasing amount of data generated on social media. An important task in sentiment analysis is feature selection which is used in the data reduction. In this paper, we reviewed feature selection in sentiment analysis, specifically information gain and ontology-based approaches. We proposed a new 2 step feature selection approach based on information gain and domain specific ontology. The aim is to have an optimal approach that will overcome the limitations of the two approaches. In future work, we plan to perform experiments and evaluate the approach across different datasets.

Acknowledgment
The authors gratefully acknowledge Universiti Kebangsaan Malaysia and Fundamental Research Grant Scheme (FGRS) for supporting this research project through grant number FRGS/1/2016/ICT02/UKM/01/2.

Conflicts of interest
The authors have no conflicts of interest to declare.

References
[1] Dave K, Lawrence S, Pennock DM. Mining the peanut gallery: opinion extraction and semantic classification of product reviews. In proceedings of the international conference on World Wide Web 2003 (pp. 519-28). ACM.
[2] Nasukawa T, Yi J. Sentiment analysis: capturing favorability using natural language processing. In proceedings of the international conference on knowledge capture 2003 (pp. 70-7). ACM.
[3] Pang B, Lee L, Vaithyanathan S. Thumbs up? sentiment classification using machine learning techniques. In proceedings of the ACL-02 conference on empirical methods in natural language processing 2002 (pp. 79-86). Association for Computational Linguistics.
[4] Turney PD. Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. In proceedings of the annual meeting on association for computational linguistics 2002 (pp. 417-24). Association for Computational Linguistics.
[5] Yi J, Nasukawa T, Bunescu R, Niblack W. Sentiment analyzer: extracting sentiments about a given topic using natural language processing techniques. In IEEE international conference on data mining 2003 (pp. 427-34). IEEE.
[6] Ahmad SR, Bakar AA, Yaakub MR. Metaheuristic algorithms for feature selection in sentiment analysis. In science and information conference (SAI) 2015 (pp. 222-6). IEEE.
[7] Zheng L, Wang H, Gao S. Sentimental feature selection for sentiment analysis of Chinese online reviews. International Journal of Machine Learning and Cybernetics. 2018; 9(1):75-84.
[8] Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: a survey. Ain Shams Engineering Journal. 2014; 5(4):1093-113.
[9] Ravi K, Ravi V. A survey on opinion mining and sentiment analysis: tasks, approaches and applications. Knowledge-Based Systems. 2015; 89:14-46.
[10] Miranda MD, Sassi RJ. Using sentiment analysis to assess customer satisfaction in an online job search company. In international conference on business information systems 2014 (pp. 17-27). Springer, Cham.
[11] Wang S, Li D, Song X, Wei Y, Li H. A feature selection method based on improved fisher's discriminant ratio for text sentiment classification. Expert Systems with Applications. 2011; 38(7):8696-702.
[12] Colace F, De Santo M, Greco L, Moscato V, Picariello A. Probabilistic approaches for sentiment analysis: latent dirichlet allocation for ontology building and sentiment extraction. In sentiment analysis and ontology engineering 2016 (pp. 75-91). Springer, Cham.
[13] Li YM, Li TY. Deriving market intelligence from microblogs. Decision Support Systems. 2013; 55(1):206-17.
[14] Kang H, Yoo SJ, Han D. Senti-lexicon and improved naïve Bayes algorithms for sentiment analysis of restaurant reviews. Expert Systems with Applications. 2012; 39(5):6000-10.
[15] Tripathy A, Agrawal A, Rath SK. Classification of sentiment reviews using n-gram machine learning approach. Expert Systems with Applications. 2016; 57:117-26.
[16] Vohra SM, Teraiya JB. A comparative study of sentiment analysis techniques. Journal JIKRCE. 2013; 2(2):313-7.
[17] Mohammad S, Dunne C, Dorr B. Generating high-coverage semantic orientation lexicons from overtly marked words and a thesaurus. In proceedings of the 2009 conference on empirical methods in natural language processing: 2009 (pp. 599-608). Association for Computational Linguistics.
[18] Liu H, Lieberman H, Selker T. A model of textual affect sensing using real-world knowledge. In proceedings of the 8th international conference on intelligent user interfaces 2003 (pp. 125-32). ACM.
[19] Tsai AC, Wu CE, Tsai RT, Hsu JY. Building a concept-level sentiment dictionary based on commonsense knowledge. IEEE Intelligent Systems. 2013; 28(2):22-30.
[20] Hatzivassiloglou V, McKeown KR. Predicting the semantic orientation of adjectives. In proceedings of the annual meeting of the association for computational linguistics and eighth conference of the European chapter of the association for computational linguistics 1997 (pp. 174-81). Association for Computational Linguistics.
[21] Mostafa MM. More than words: social networks' text mining for consumer brand sentiments. Expert Systems with Applications. 2013; 40(10):4241-51.
[22] Abdel-Hafez A, Xu Y. Ontology-based product's reputation model. In proceedings of the IEEE/WIC/ACM international joint conferences on web intelligence (WI) and intelligent agent technologies (IAT)2013 (pp. 37-40). IEEE Computer Society.
[23] Garcia-Herranz M, Moro E, Cebrian M, Christakis NA, Fowler JH. Using friends as sensors to detect global-scale contagious outbreaks. PloS one. 2014; 9(4).
[24] Tumasjan A, Sprenger TO, Sandner PG, Welpe IM. Predicting elections with twitter: what 140 characters reveal about political sentiment. In fourth international AAAI conference on weblogs and social media 2010:178-85.

[25] Duric A, Song F. Feature selection for sentiment analysis based on content and syntax models. Decision Support Systems. 2012; 53(4):704-11.
[26] Kontopoulos E, Berberidis C, Dergiades T, Bassiliades N. Ontology-based sentiment analysis of twitter posts. Expert Systems with Applications. 2013; 40(10):4065-74.
[27] Manek AS, Shenoy PD, Mohan MC, Venugopal KR. Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier. World Wide Web. 2017; 20(2):135-54.
[28] Ding X, Liu B, Yu PS. A holistic lexicon-based approach to opinion mining. In proceedings of the international conference on web search and data mining 2008 (pp. 231-40). ACM.
[29] Whitelaw C, Garg N, Argamon S. Using appraisal groups for sentiment analysis. In proceedings of the ACM international conference on information and knowledge management 2005 (pp. 625-31). ACM.
[30] Gómez-Pérez A, Corcho O. Ontology languages for the semantic web. IEEE Intelligent Systems. 2002; 17(1):54-60.
[31] Salas-Zárate MD, Valencia-García R, Ruiz-Martínez A, Colomo-Palacios R. Feature-based opinion mining in financial news: an ontology-driven approach. Journal of Information Science. 2017; 43(4):458-79.
[32] Ali F, Kwak KS, Kim YG. Opinion mining based on fuzzy domain ontology and support vector machine: a proposal to automate online review classification. Applied Soft Computing. 2016; 47:235-50.
[33] Agarwal B, Mittal N, Bansal P, Garg S. Sentiment analysis using common-sense and context information. Computational Intelligence and Neuroscience.2015.
[34] Thakor P, Sasi S. Ontology-based sentiment analysis process for social media content. Procedia Computer Science. 2015; 53:199-207.
[35] Lundquist D, Zhang K, Ouksel A. Ontology-driven cyber-security threat assessment based on sentiment analysis of network activity data. In international conference on cloud and autonomic computing 2014 (pp. 5-14). IEEE.
[36] Marstawi A, Sharef NM, Aris TN, Mustapha A. Ontology-based aspect extraction for an improved sentiment analysis in summarization of product reviews. In proceedings of the international conference on computer modeling and simulation 2017 (pp. 100-4). ACM.
[37] Schouten K, Frasincar F, De Jong F. Ontology-enhanced aspect-based sentiment analysis. In international conference on web engineering 2017 (pp. 302-20). Springer, Cham.
[38] Yadav N, Chowdary CR. Feature based sentiment analysis using a domain ontology. In proceedings of the international conference on natural language processing 2016 (pp. 90-8).
[39] Gutierrez F, Dou D, Fickas S, Wimalasuriya D, Zong H. A hybrid ontology-based information extraction system. Journal of Information Science. 2016; 42(6):798-820.
[40] Alexopoulos P, Wallace M. Creating domain-specific semantic lexicons for aspect-based sentiment analysis. In international workshop on semantic and social media adaptation and personalization 2015 (pp. 1-6). IEEE.
[41] Blanco E, Cankaya H, Moldovan D. Commonsense knowledge extraction using concepts properties. In twenty-fourth international FLAIRS conference 2011(pp. 222-7).
[42] Shangfeng H, Kanagasabai R. Learning commonsense knowledge models for semantic analytics. In international conference on semantic computing 2016 (pp. 400-3). IEEE.
[43] Cambria E, Hussain A, Havasi C, Eckl C. Sentic computing: exploitation of common sense for the development of emotion-sensitive systems. In development of multimodal interfaces: active listening and synchrony 2010 (pp. 148-56). Springer, Berlin, Heidelberg.
[44] Cambria E, Speer R, Havasi C, Hussain A. Senticnet: a publicly available semantic resource for opinion mining. In AAAI fall symposium series 2010 (pp.14-8).
[45] Cambria E, Havasi C, Hussain A. SenticNet 2: a semantic and affective resource for opinion mining and sentiment analysis. In international FLAIRS conference 2012 (pp. 202-7).
[46] Cambria E, Olsher D, Rajagopal D. SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis. In AAAI conference on artificial intelligence 2014(pp.1515-21).
[47] Verdu S. Fifty years of Shannon theory. IEEE Transactions on Information Theory. 1998; 44(6):2057-78.
[48] Lee C, Lee GG. Information gain and divergence-based feature selection for machine learning-based text categorization. Information Processing & Management. 2006; 42(1):155-65.
[49] Mukras R, Wiratunga N, Lothian R, Chakraborti S, Harper D. Information gain feature selection for ordinal text classification using probability re-distribution. In proceedings of the textlink workshop at IJCAI 2007.
[50] Wu G, Xu J. Optimized approach of feature selection based on information gain. In international conference on computer science and mechanical automation 2015 (pp. 157-61). IEEE.
[51] Pratiwi AI. On the feature selection and classification based on information gain for document sentiment analysis. Applied Computational Intelligence and Soft Computing. 2018.
[52] Zhu L, Wang G, Zou X. Improved information gain feature selection method for Chinese text classification based on word embedding. In proceedings of the international conference on software and computer applications 2017 (pp. 72-6). ACM.
[53] Schouten K, Frasincar F, Dekker R. An information gain-driven feature study for aspect-based sentiment analysis. In international conference on applications of

natural language to information systems 2016 (pp. 48-59). Springer, Cham.
[54] Fahrudin TM, Syarif I, Barakbah AR. Feature selection algorithm using information gain-based clustering for supporting the treatment process of breast cancer. In international conference on informatics and computing 2016 (pp. 6-11). IEEE.
[55] Ong BY, Goh SW, Xu C. Sparsity adjusted information gain for feature selection in sentiment analysis. In international conference on big data 2015 (pp. 2122-8). IEEE.
[56] Gao Z, Xu Y, Meng F, Qi F, Lin Z. Improved information gain-based feature selection for text categorization. In international conference on wireless communications, vehicular technology, information theory and aerospace & electronic systems (VITAE) 2014 (pp. 1-5). IEEE.
[57] Luo K, Luo J, Yin M, Li J. IG-C4. 5: an improved feature selection method based on information gain. In international conference on mechatronics, electronic, industrial and control engineering (MEIC-14) 2014. Atlantis Press.

Ibrahim Said Ahmad isa lecturerr in Department of Information Technology, Bayero University Kano. He received his BSc degree in Computer Science from Bayero University Kano, Nigeria, in 2011 and his MSc degree in Information Technology from The University of Nottingham, UK, in 2014. He is currently a PhD candidate in Universiti Kebangsaan Malaysia. He has published a number of journal articles and attended many conferences.
Email: isahmad.it@buk.edu.ng

Azuraliza Abu Bakar is a Professor in Data Mining at Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia since 2011. She received her PhD in Artificial Intelligence from Universiti Putra Malaysia in 2002. Her main research areas are Data Analytics and Artificial Intelligence specifically in Rough Set Theory, Feature Selection Algorithms, Nature Inspired Computing and Sentiment Analysis. She served as advisor for Data Mining and Optimization Lab and member of the Sentiment Analysis Lab. She has led 13 research projects (including 3 in progress) and member of 42 research projects. She is a member of the IEEE Computational Intelligence Society. She has also served on roughly thirty conference and workshop program committees and served as the program chair.
Email: Azuraliza@ukm.edu.my

Mohd Ridzwan Mohd Yaakub is a Senior Lecturer in Universiti Kebangsaan Malaysia. He received his PhD degree from Queensland University of Technology, Australia, in 2015. His research interest is in Databases, Data Mining And Artificial Intelligence. He is currently the head of the Sentiment Analysis Lab in Center for Artificial Intelligence Technology, Universiti Kebangsaan Malaysia. He has led a number of research projects, supervised MSc students and co-supervised PhD students. He has published several journal articles and attended several refereed conferences.
Email: ridzwanyaakub@ukm.edu.my