A Review of Contextual Information for Context-Based Approach in Sentiment Analysis

Nor Nadiah Yusof, Azlinah Mohamed, and Shuzlina Abdul-Rahman

***Abstract*—Online opinionated data has increased tremendously since the arrival of web 2.0. Users have the authority to generate online content expressing their sentiments or opinions regarding subjects of interest. Although these phenomena caused the problem of information overload, the opinionated data is valuable and beneficial to others. Looking at the prominent significance of this issue, many researchers grab this opportunity to further investigate sentiment analysis. The main task of sentiment analysis is to classify opinions into several orientations. However, the orientation of sentiment is highly dependent on the context of sentiment text. Thus, it is important to consider contextual information in order to correctly classify sentiments. This paper would like to facilitate researchers in capturing the contextual information of context-based approach in sentiment analysis; which includes sequence and collocation of words, negation handling, and ambiguity. Several recent studies that utilize the contextual information are also discussed. Besides, this paper presents the state of art of context based on the level, type, and representation of context. With the help of this review, researchers will be able to correctly classify sentiments based on context in improving the performance of sentiment analyzer.**

***Index Terms*—Contextual information, context-based approach, natural language processing, sentiment analysis.**

I. Introduction

Sentiment analysis analyses users' opinions in the written text either from documents or sentences. It refers to the process of evaluating the opinions, emotions, and attitudes of people towards certain topics of interest or subject matters [1]. Due to the advancement of internet technology, sentiment analysis has become a popular research area since 2001 [2]. Sentiment analysis is multidisciplinary artificial intelligence problem with the aim to minimize the gap between human and computer in natural language processing [3].

The opinions of other users are essential and have a significant effect on other users mainly in decision making such as to assist in online shopping by selecting the best products, services or events [4]. Besides, sentiment analysis can be beneficial for companies to better know their customers' preferences. The sentiments from users can be utilized for marketing campaigns and business strategies [5].

Basically, sentiment analysis is a text classification problem. Traditional text classification relies on keywords to classify documents into several themes or topics. The key features for the classification task are the list of related words about particular themes or topics. Compared with the traditional classification task, sentiment classification is a more challenging task as it requires a deeper semantic level of the reviews [6]. Sentiment analysis relies on natural language lexicons that are used to express the subjectivity of sentiment texts instead of objectivity. Subjectivity refers to the opinions expressed by users whereas objectivity is natural facts that are not to be considered in sentiment analysis task.

In most available lexicons, they are compiled as a list of keywords, rather than the meaning of the words, or also known as senses [7]. The positivity or negativity sentiment orientation of words could change depending on context and targeted entities [8]. This creates an ambiguous orientation of the sentiment texts. Moreover, out of all the opinion words, some words behave in the same manner. It means they have the same polarity in all contexts. However, some words are context dependent. It means they have different polarity in a different context [9]. Therefore, to consider the meaning of words, appropriate contextual information is required in order to identify the most relevant context for the selected words.

Currently, sentiment analysis is still one of the most challenging problems in natural language processing due to the lack of contextual information in sentiment analysis tasks [10], [11]. It is imperative to incorporate context in classifying sentiment text as content alone can be misleading. The study of context-based approach in the area of sentiment analysis is exclusively fresh and developing. The remainder of this paper is organized as follows. Section II presents the overview of context. Section III deliberates on contextual information in sentiment analysis. Related researches on context-based approach in sentiment analysis are discussed in Section IV. Finally, several points are brought into conclusion in Section V.

II. Overview of Context

Sentiment analysis move towards content, concept and context-based text analysis of natural language text [12], [13]. Context may be defined as any information that can be used to characterize the situation of an object or entity [14]. In textual form, context considers the environment of the selected term in a sentence. Words that surround the selected term are important to explain the meaning of the sentence.

Context is useful in natural language processing to solve several tasks such as spelling error correction and sense disambiguation. Context-sensitive spelling error tasks address the problem of correcting spelling error mistakes in a sentence. The correct word spelling cannot be identified by looking at the term individually. Thus, it requires investigation of

Manuscript received May 2, 2018; revised July 9, 2018.

Nor Nadiah Yusof, Azlinah Mohamed and Shuzlina Abdul Rahman are with Faculty of Computer & Mathematical Science, Universiti Teknologi MARA, Malaysia (e-mail: nornadiah.yusof@gmail.com, {azlinah, shuzlina}@tmsk.uitm.edu.my).

doi: 10.18178/ijmlc.2018.8.4.719

neighbouring terms to detect the correct word spelling. Sense disambiguation deliberates the problem of identifying the correct senses for multiple meaning words. Similarly, the neighbouring term is required to identify the most suitable meaning of a term in the correct context.

In sentiment analysis, context plays an important role to determine the orientation of sentiment polarity. Word in a certain domain might carry different orientation when the context changes. For example, the word high may be positive when it refers to high salary. However, the orientation may be negative when it refers to high price. Fig. 1 shows several aspects of context concerning the level, type and representation of context.

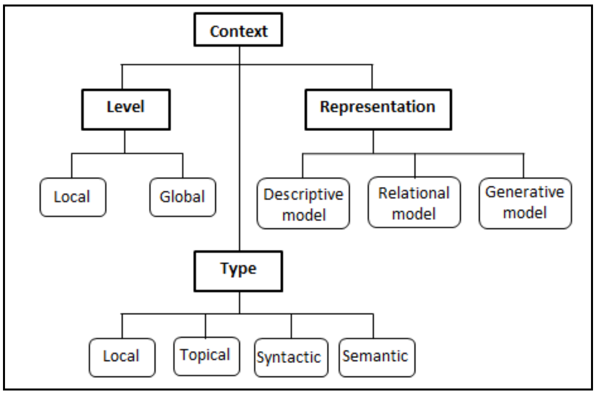

Fig. 1. Overview of context.

Level of context can be categorized as global or local context [15], [16]. Local context captures different local features that surround the targeted word or sense. Features in the context of sentiment analysis are words or terms that strongly express the opinion polarity as either positive or negative sentiment. It can also be referred to the same theme or topic in a review [17].

Besides, local context can capture features relations in a sentence, for example by using n-grams or chi-squared approach. In sentiment analysis, local context is widely used in word sense disambiguation and valance shifters [15]. Global context, on the other hand, exploits the context features as an overall understanding of the sentiment. It analyses the sentences in sentiment as a whole opinion. For example, global context is used to identify implicit opinions in sentiment analysis.

According to Bhala and Abirami [18], the type of context can be categorized into local, topical, syntactic or semantic. Local context in this state is generally quite similar as mentioned above. Local context represents several features surrounding the targeted word with representation approaches such as part-of-speech tagging and positions of words in a sentence. In contrast to local context, topical context represents a more general context by defining a common theme or topic of discussion in a text. Normally, the bag-of-words approach is used as a mode of representation. Syntactic context represents the word usage based on structure and grammar rules. It refers to syntactic clues on how words are being formed to produce a meaningful sentence. Sometimes, these words might be outside the local context [19]. Semantic context represents semantic information that depends on additional knowledge. For example, semantic context depends on formerly known senses of words in a context which is based on domain ontologies or glossaries.

Kwong [20] identifies three types of context representation; which are descriptive model, relational model and generative model. Descriptive model is the most common model for defining senses. It is a model for the general purpose of sense representation, which based on the dictionary definition. Relational model recognizes senses based on semantic relation such as synonym, antonym and hyponym. Generative model is a dynamic model for representing senses, which the encoding is left open-ended.

III. Contextual Information in Sentiment Analysis

Existing sentiment analysis systems rely on the bag-of-concept model to classify sentiments into several orientations. It is basically based on the frequency of occurrences of words and ignores the order of words in sentiment text. Thus, the meaning of the word is far to achieve. Context-based sentiment analysis highly concerns about the order and collocation of words in a sentiment text. Word in isolation hinders the meaning of the whole sentiment. To understand the context of a word, the neighbouring target word is exclusively imperative.

Therefore, one of a good tactic to consider the neighbours of the targeted word is using n-grams approach. N-grams is a sequence of N words and it can be constructed into higher level such as bigrams and trigrams. It is important in language processing to estimate the probabilities of next words and also for whole sequences [21]. Another approach to consider the neighbouring target words is using graph-based approach. A context graph links the target word with its neighbour words in by considering their co-occurrences [22]. The co-occurrences are measured based on the location of neighbouring words from the target word, normally for the two words preceding and two worlds following the target word. Later, they are represented by a network with words as nodes and the co-occurring frequencies as edges. The aim of this approach is to model the contextual use of words in a sentence or document. The approach was proposed by [23] to extract strong relation between words, subsequently to identify the most appropriate context.

Moreover, it is incapable to understand the semantic meaning if the existence of negation as the orientations can be oppositely assigned [12]. A simple negation is not only reversing a word meaning but can also change the whole sentence [24]. Negation can exist in two ways, either local (such as **not** beautiful) or involve long-distance dependencies (such as **does not** look very beautiful) [25]. The first computational model that accounts for negation in a model is proposed by Polanyi and Zaenen [26]. They discussed Applications interesting factors of several circumstances that can cause the values of lexical resources to shift or flip the values with the existence of negation. Handling negation is crucial as it is the most obvious shifters for sentiment polarity.

Natural language exists in some criteria that make its processing complicated. Every word can have one or multiple meanings. This created vague and uncertain conditions of the text. Words that change their polarity depending on the context are known as ambiguous words [27]. The ambiguity could be resolved when the context is identified. The ambiguity resolution process is referring to the capability of identifying the meaning of a word in context. To undergo this process, all these components must be incorporated; the definition of words or senses, to include external knowledge resources, context and concept representation and algorithm for disambiguation [18]. A word may be represented in one or more senses, thus the definition of words or senses is needed to reduce ambiguity. The possible ambiguity types are determined once the definition of word or senses is established. In sentiment analysis, several word lexicons can be utilized such as WordNet, WordNet-Affect, ConceptNet and SentiWordNet. Although each type of these lexical resources has their own specialized scopes, these resources are capable of representing richer contextual information.

Moreover, the most complex situation exists if to deal with sarcasms and puns [28]; [29]. Sarcasms basically take the idea of opposite meaning and create confusion about the actual message that needs to deliver [30]. Paronomasia or more generally known as puns is a form of wordplay, where a word is used to evoke several independent meanings [28]. It is normally carried out to have more than one meaning and uncertainty which one is the intended meaning. To overcome sarcasms and puns, it is compulsory to reduce the ambiguity and have a good understanding the context of situations.

IV. Related Researches of Context-Based Approach in Sentiment Analysis

Various literature has reported different approaches in sentiment analysis, ranging from lexicon-based to machine learning approaches [31]. However, not many kinds of literature deliberate on context-based approach for sentiment analysis. Tang *et al*. [32] utilized contextual information by combining context and sentiment level evidences before embedded in neural network algorithm. They mapped neighbouring words that encode sentiment texts in continuous word representation to word level and sentence level sentiment classification and also building sentiment lexicon. The proposed lexicon is based on sentiment embedding that is useful for improving lexical level tasks in finding similarities between words. Similarly, Ghiassi *et al*. [33] also proposed an approach to supervised feature reduction using n-grams and statistical analysis to develop a lexicon for Twitter. They demonstrated the effectiveness of the proposed lexicon by developing comparable sentiment classification model using Support Vector Machine (SVM) algorithm.

Meanwhile, Tripathy *et al*. [34] applied various supervised machine learning algorithms to classify movie review documents of IMDb dataset and further implemented n-grams approach. The investigated the level of n-grams in representing the contextual meaning of words and observed that the increase 'n' in n-grams decreases the classification accuracy. Sidorov *et al.* [35] proposed a concept of syntactic n-grams which are constructed using paths in syntactic trees to extract relevant related terms. The idea is to introduce contextual linguistic information into the statistical method of machine learning. Liu and Jin [36] utilized focus sentence and context in classifying Chinese comment text. They assumed that the first sentence of a text with sentimental tendency is a focus sentence. Different location of a sentence has a different impact on the sentiment orientation of the whole text. The context of dynamic sentiment words is calculated based on the respective score in SentiGram. The proposed method is proven with a better result from the baseline approach through series of experiments.

Farooq *et al*. [37] investigate the problems of identifying the scope of negation and proposed a negation handling method based on linguistic features. They successfully improve the accuracy of both negation scope identification and overall sentiment analysis. Sharif *et al.* [38] proposed an approach of negation identification for consumer reviews which does not rely on the direct presence of words not, no, never etc. They used universal dependency of Stanford Parser to identify a grammatical relationship in a sentence. The correct polarity of reviews is recognized by calculating dependencies of negation words. Similarly, Diamantini *et al.* [39] suggested a negation handling algorithm based on dependency based parse trees. Apart from that, the sequence of words is also taken into consideration as n-grams approach is coupled with negation handling to better identify the sentiment polarity.

Katz *et al*. [40] introduced a model of key terms in contexts to generate features of context terms for supervised classification by using rotation forests in decision tree. The context terms are used to model the relations between the identified key terms to identify the most suitable context for each key term. Bhuvan *et al.* [41] proposed a context-specific grammar model of semantics for movie reviews domain where features are obtained from matching semantics patterns. The proposed model is tested using various machine learning algorithms such as Naïve Bayes, SVM, Logistic Regression and Sequential Minimal Optimization. Weichselbraun *et al.* [11] attempt to resolve the ambiguity of words and consider the context of sentiment terms by using machine learning approach. They developed their own contextualized lexicons by collecting context terms for each ambiguous term identified. They evaluate their proposed method using user reviews in products and hotels domains and also with benchmark data on movies review by Pang and Lee [42]. Weichselbraun *et al.* [43] expand their previous work by increasing the lexicons coverage and derive concept information. They utilized knowledge resources from SenticNet to recognize ambiguous terms and extract contextual information. Later, they dug up structured resources of WordNet and ConceptNet to obtain conceptual knowledge.

The contextual information is imperatively important for better sentiment classification results. Table I shows the studies as discussed above, which are mapped according to the respective contextual information. By utilizing at least one or two types of contextual information, most researches have shown good performance and successfully improve the accuracy of their sentiment analyzers from baseline approach. Thus from this review, we found that sentiment analyzers with more enrich contextual information such as sequence and collocation of words, negation handling, and ambiguity management plays a big role in sentiment analysis as they are capable to produce better sentiment classification results.

TABLE I: STUDIES OF CONTEXT-BASED APPROACH IN SENTIMENT ANALYSIS WITH RESPECTIVE CONTEXTUAL INFORMATION

Author	Order/sequence	Negation	Ambiguity
[32]	✓		✓
[33]	✓		✓
[34]	✓		
[35]	✓		
[36]	✓		
[37]		✓	
[38]		✓	
[39]	✓	✓	
[40]			✓
[41]			✓
[11]			✓
[43]			✓

V. CONCLUSION

Context plays an important role in sentiment classification. Without context, the opinions of people can be misunderstood and lead to different meanings. Consequently, the sentiment texts are incorrectly classified. The main problem in sentiment analysis is the subjectivity of words or phrases could depend on their context. Incorporating appropriate contextual information could provide the right context to classify sentiments in a correct means. The overview of context is well discussed to give a better understanding of the definition of context and subsequently to deliberate relevant contextual information in sentiment analysis. Common approaches in sentiment classification tend to ignore the semantic meaning of sentiments by using bag-of-words approach and lose the word order. However, the context-based approach in sentiment analysis greatly concerns about the order and collocation of words in sentiment text. The existence of negation in sentiment texts may invert the polarity of all words in a sentence, thus it needs to be well handled. Natural language conditions may occur in complex situations such as sarcasm and puns creates confusion as many meanings can be interpreted. Thus, identification of contextual information in resolving ambiguities can help to identify the sentiment orientations. Several related studies that utilized contextual information in classifying sentiments have shown good performance as compared to baseline approach which is the classification without context information. Sentiment analysis greatly relies on contextual information in order to identify the correct word meanings or senses according to respective context to acquire better accuracy in the classification task of sentiment text.

ACKNOWLEDGMENT

The authors would like to thank Research Management Centre, Universiti Teknologi MARA, Selangor, Malaysia for LESTARI research grant (reference number: 600-RMI/MYRA 5/3/LESTARI (19/2016)) and Ministry of Higher Education, Malaysia for the research support.

REFERENCES

[1] B. Liu, *Sentiment Analysis and Opinion Mining*, Morgan & Claypool Publishers, 2012.

[2] 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.

[3] A. Kaur and V. Gupta. (2013). A survey on sentiment analysis and opinion mining techniques. *Journal of Emerging Technologies in Web Intelligence*. [Online]. 5(4). pp. 367–371. Available: https://pdfs.semanticscholar.org/81ac/64634d9db9e5691273da87ad2578d78249bc.pdf

[4] V. S. Jagtap and K. Pawar. (2013). Analysis of different approaches to Sentence-Level Sentiment Classification. *International Journal of Scientific Engineering and Technology.*[Online]. 2(3). pp. 164–170. Available: https://www.researchgate.net/profile/Vandana_Jagtap/publication/286272892_Analysis_of_different_approaches_to_Sentence-Level_Sentiment_Classification/links/5768dedd08aef6cdf9b40d11.pdf

[5] A. S. Ab Rahman, S. Abdul-Rahman, and S. Mutalib, “Mining textual terms for stock market prediction analysis using financial news,” in *Proc. International Conference on Soft Computing in Data Science*, 2017, pp. 293–305.

[6] C. Zhao, S. Wang, and D. Li, “Fuzzy sentiment membership determining for sentiment classification,” in *Proc. 2014 IEEE International Conference on Data Mining Workshop*, 2014, pp. 1191-1198.

[7] C. Akkaya, J. Wiebe, A. Conrad, and R. Mihalcea, “Improving the impact of subjectivity word sense disambiguation on contextual opinion analysis,” in *Proc. the Fifteenth Conference on Computational Natural Language Learning*, 2011, pp. 87–96.

[8] H. Saif, M. Fernandez, Y. He, and H. Alani, “SentiCircles for contextual and conceptual semantic sentiment analysis of twiter,” in *Proc. European Semantic Web Conference*, 2014, vol. 8465, pp. 580–594.

[9] H. Kansal and D. Toshniwal, “Aspect based summarization of context dependent opinion words,” in *Proc. the 18th International Conference in Knowledge Based and Intelligent Information and Engineering Systems*, 2014, pp. 166–175, vol. 35.

[10] E. Lefever and V. Hoste, “SemEval-2013 Task 10: Cross-lingual word sense disambiguation,” in *Proc. Second Joint Conference on Lexical and Computational Semantics (*SEM)*, 2013, pp. 158–166, vol. 2.

[11] A. Weichselbraun, S. Gindl, and A. Scharl, “Extracting and grounding contextualized sentiment lexicons,” *IEEE Intelligent Systems*, vol. 28, pp. 39–46, 2013.

[12] S. Poria, E. Cambria, G. Winterstein, and G. Bin Huang, “Sentic patterns: Dependency-based rules for concept-level sentiment analysis,” *Knowledge-Based Systems*, vol. 69, no. 1, pp. 45–63, 2014.

[13] C. Cambria, E. Schuller, B. Xia, and Y. Havasi, “New avenues in opinion mining and sentiment analysis,” *IEEE Intelligent Systems*, vol. 28, no. 2, pp. 15–21, 2013.

[14] J. Janek, J. Pieter, and D. Villiers, “A unified model for context-based behavioural modelling and classification,” *Expert Systems with Applications*, vol. 42, no. 19, pp. 6738–6757, 2015.

[15] F. Benamara, B. Chardon, Y. Mathieu, and V. Popescu, “Towards context-based subjectivity analysis,” in *Proc. the 5th International Joint Conference on Natural Language Processing*, 2011, pp. 1180–1188.

[16] C. Galleguillos and S. Belongie, “Context based object categorization : A critical survey,” *Computer Vision and Image Understanding*, vol. 114, no. 6, pp. 712–722, 2010.

[17] T. Kincl, M. Novák, J. Přibil, and P. Štrach, “Language-independent sentiment analysis with surrounding context extension,” *International Conference on Social Computing and Social Media*, pp. 158–168, 2015.

[18] R. V. V. Bhala and S. Abirami, “Trends in word sense disambiguation,” *Artificial Intelligence Review*, vol. 42, no. 2, pp. 159–171, 2014.

[19] R. Navigli, “A quick tour of word sense disambiguation, induction and related approaches,” *SOFSEM 2012: Theory and Practice of Computer Science*, pp. 115–129, 2012.

[20] O. Y. Kwong, *New Perspectives on Computational and Cognitive Strategies for Word Sense Disambiguation*, Springer New York Heidelberg Dordrecht London, 2012.

[21] D. Jurafsky and J. H. Martin, “Lexicons for Sentiment and Affect Extraction,” *Speech and Language Proceessing*, ch. 21, pp. 1-18, 2015.

[22] O. Serban, A. Pauchet, A. Rogozan, and J. P. Pecuchet, “Modelling context to solve conflicts in SentiWordNet,” in *Proc. 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 2013, pp. 393–398.

[23] J. Hyungsuk, S. Ploux, E. Wehrli, J. Hyungsuk, S. Ploux, E. Wehrli, L. Knowledge, H. Ji, S. Ploux, and E. Wehrli, "Lexical Knowledge Representation with Contexonyms," *Proceedings of 9th MT Summit Machine Translation*, pp. 194–201, 2003.
[24] N. Pröllochs, S. Feuerriegel, and D. Neumann, "Enhancing sentiment analysis of financial news by detecting negation scopes," in *Proc. the Annual Hawaii International Conference on System Sciences*, 2015, pp. 959–968.
[25] T. A. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing Contextual Polarity: An exploration of features for phrase-level sentiment analysis," *Computational linguistics*, vol. 35, no. 3, pp. 399–433, 2009.
[26] L. Polanyi and A. Zaenen, "Contextual valence shifters," *Computing* attitude *and affect in text: Theory and Applications*, vol. 20, no. 1, pp. 1–10, 2006.
[27] B. Agarwal, N. Mittal, P. Bansal, and S. Garg, "Sentiment analysis using common-sense and context information," *Computational Intelligence and Neuroscience*, pp. 30-39, 2015.
[28] T. Miller and M. Turković, "Towards the automatic detection and identification of English puns," *The European Journal of Humour Research*, vol. 4, no. 1, pp. 59–75, 2016.
[29] A. Moreo, J. L. Castro, and J. M. Zurita, "Handling context in lexicon-based sentiment analysis," *Communications in Computer and Information Science*, pp. 245–254, 2012.
[30] D. Maynard and M. A. Greenwood, "Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis," *Lrec*, pp. 4238–4243, 2014.
[31] N. N. Yusof, A. Mohamed, and S. Abdul-Rahman, "Reviewing classification approaches in sentiment analysis," *Communications in Computer and Information Science*, pp. 43–53, 2015.
[32] D. Tang, F. Wei, B. Qin, N. Yang, T. Liu, and M. Zhou, "Sentiment Embeddings with Applications to Sentiment Analysis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 2, pp. 496–509, 2016.
[33] M. Ghiassi, J. Skinner, and D. Zimbra, "Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network," *Expert Systems with applications*, vol. 40, no. 16, pp. 6266–6282, 2013.
[34] A. Tripathy, A. Agrawal, and S. K. Rath, "Classification of sentiment reviews using n-gram machine learning approach," *Expert Systems with Applications*, vol. 57, pp. 117–126, 2016.
[35] G. Sidorov, F. Velasquez, E. Stamatatos, and A. Gelbukh, "Syntactic n-grams as machine learning features for natural language processing," *Expert Systems with Applications*, vol. 41, no. 3, pp. 853–860, 2014.
[36] R. Liu and L. Jin, "Sentiment analysis method based on the focus sentence and context for Chinese comment text," in *Proc. 2nd International Conference of Information Technology and Electronic Commerce*, pp. 267–270, 2014.
[37] U. Farooq, H. Mansoor, A. Nongaillard, Y. Ouzrout, and M. Abdul Qadir. (2017). Negation handling in sentiment analysis at sentence level. *Journal of Computers.* [Online] 12(5). pp. 470–478. Available: http://www.jcomputers.us/vol12/jcp1205-11.pdf
[38] W. Sharif, N. A. Samsudin, M. M. Deris, and R. Naseem, "Effect of negation in sentiment analysis," in *Proc. 6th Inernatinal Conference on Innovative Computing Technology*, 2017, pp. 718–723.
[39] C. Diamantini, A. Mircoli, and D. Potena, "A negation handling technique for sentiment analysis," in *Proc. International Conference on Collaboration Technologies and Systems*, 2016, pp. 188–195.
[40] G. Katz, N. Ofek, and B. Shapira, "ConSent: Context-based sentiment analysis," *Knowledge-Based Systems*, vol. 84, pp. 162–178, 2015.
[41] M. S. Bhuvan, V. D. Rao, S. Jain, T. S. Ashwin, and R. M. R. Guddeti, "Semantic sentiment analysis using context specific grammar," in *Proc. International Conference on Computing, Communication & Automation*, 2015, pp. 28–35.
[42] B. Pang and L. Lee, "A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts," in *Proc. the 42nd Annual Meeting on Association for Computational Linguistics*, 2004, p. 271.
[43] A. Weichselbraun, S. Gindl, and A. Scharl, "Enriching semantic knowledge bases for opinion mining in big data applications," *Knowledge-Based Systems*, vol. 69, no. 1, pp. 78–85, 2014.

Nor Nadiah Yusof received her bachelor's degree in information technology with honours from Universiti Kebangsaan Malaysia in 2008 and master of science in information technology from Universiti Teknologi MARA, Malaysia in 2010. She is currently persuing her PhD in computer science at Universiti Teknologi MARA, Malaysia. Her research interest includes natural language processing, text mining, and machine learning.

Azlinah Mohamed received her bachelor's degree in computer science from Universiti Kebangsaan Malaysia in 1990, master of science in artificial intelligence from Bristol, UK in 1992 and PhD in science and system management from Universiti Kebangsaan Malaysia in 2001. She is currently working as a professor and dean at the Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA, Malaysia. Her research interest includes artificial intelligence, decision support systems, soft computing, and big data analytics.

Shuzlina Abdul Rahman received her bachelor's degree in computer science from Universiti Sains Malaysia in 1996, master of science in information technology from Universiti Utara Malaysia in 2000 and PhD in Science and system management from Universiti Kebangsaan Malaysia in 2012. She is currently working as an associate professor at Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA, Malaysia. Her research interest includes computational intelligence, machine learning and data analytics & optimization.