

OPINION MINING OF CUSTOMER REVIEWS: FEATURE AND SMILEY BASED APPROACH

I R Jayasekara and W M J I Wijayanayake

Department of Industrial Management,
University of Kelaniya, Sri Lanka

ABSTRACT

With the rapid growth in ecommerce, reviews for popular products on the web have grown rapidly. Although these reviews are important for making decisions, it is difficult to read all the reviews. Automating the opinion mining process was identified as a solution for the problem. Although there are algorithms for opinion mining, an algorithm with better accuracy is needed. A feature and smiley based algorithm was developed which extracts product features from reviews based on feature frequency and generates an opinion summary based on product features.

The algorithm was tested on downloaded customer reviews. The sentences were tagged, opinion words were extracted and opinion orientations were identified using semantic orientation of opinion words and smileys. Since the precision values for feature extraction and both precision and recall values for opinion orientation identification were improved by the new algorithm, it is more successful in opinion mining of customer reviews.

KEYWORDS

Opinion mining, feature and smiley based approach, data mining

1. INTRODUCTION

Today the usage of the Internet increases at a rapid rate across a wide variety of fields. With that the usage of World Wide Web in different industries such as businesses, sports, social media, education, fashion & clothing, retailing etc. has gone up in a higher rate. Data are being collected accumulated and stored at a dramatic pace [1].

Most web data which are in semi structured format contain lot of useful information. With the rapid expansion of ecommerce more and more products are sold on the web. More and more people do shopping on the web. Not only this, people also tend to share their experience about products on the web. They use weblogs, twitter and other similar web sites to express their feelings, share the experiences with others. Web sites have become more popular for social interactions. In order to enhance customer satisfaction and shopping experience, it has become a common practice for online merchants to enable their customers to review or to express opinions of the products that they have purchased. As the number of online shoppers increases the number of reviews expressed on the web also increases in a rapid rate. Some products have hundreds and thousands of reviews. Understanding the consumers' idea about products is very useful for both merchants and the customers who are willing to buy those products in the future. Reading all the reviews one by one is not so efficient when the number is large. The review content also sometimes makes confusions. Most product reviews contain lot of long sentences. Very few of them actually give the opinion. So it is harder to read and understand the meaning of comments.

If someone reads only few numbers of reviews and come to a decision then the decision could be biased. Because of these reasons having a better data mining technique to mine these product reviews which are in semi structured format is very important. Not only product reviews but reviews about some places, sports and movies are also important if they are mined effectively to extract their true opinion. [2]

This problem has been studied by many researchers in the past few years. The research area is called opinion mining and sentiment analysis. There are two main tasks of this research area. They are (1) finding product features that have been commented on by reviewers and (2) deciding whether the comments are positive or negative. Both tasks are very challenging and different researches have been conducted focusing on them. [3]

Although both tasks are covered through various research approaches there are some areas to be improved. Some of them are identifying verbs & verb phrases and some special conditional sentences as product features and orientations. Exploiting useful signals such as 'pros' & 'cons' sentiments, enhancing the usage of smileys in opinion mining, improving the existing techniques through integration are some of the future work that have been identified.

Emotions are our subjective feelings and thoughts. Emotions have been studied in multiple fields as they are closely related to sentiments. The strength of a sentiment or opinion is typically linked to the intensity of certain emotions as [4].

In social media people used to express their emotion using different 'smileys'. It has become a trend today. Thus using both sentiment lexicon and emotion expressing smileys together in an algorithm to mine the opinion of customer reviews on the web will be more successful.

Considering all the potentials data mining synergy, developing data mining algorithms for opinion mining of unstructured web content can be identified as a really important research area. Our research was conducted using feature based opinion mining in customer reviews while utilizing the smileys used in reviews.

2. RELATED WORK

According to [4], opinion mining has been investigated mainly at three levels: Document Level, Sentence Level and Entity Aspect Level. The task at document level is to classify whether a whole opinion document expresses a positive or negative sentiment. This level of analysis assumes that each document expresses opinions on a single entity (e.g., a single product). Thus, it is not applicable to documents which evaluate or compare multiple entities. The task at sentence level goes to the sentences and determines whether each sentence expressed a positive, negative, or neutral opinion. Neutral usually means no opinion.

Both the document level and the sentence level analyses do not discover what exactly people liked and did not like. Aspect level performs finer-grained analysis. Aspect level is also called feature level (feature-based opinion mining and summarization) [2]. Instead of looking at language constructs (documents, paragraphs, sentences, clauses or phrases), aspect level directly looks at the opinion itself. It is based on the idea that an opinion consists of a sentiment (positive or negative) and a target (of opinion). An opinion without its target being identified is of limited use. Realizing the importance of opinion targets also helps us understand the sentiment analysis problem better. Opinion targets are the product features.

A technique to mine and to summarize all the customer reviews of a product is proposed in [2] based on data mining and natural language processing methods. The objective is to provide a

feature-based summary of a large number of customer reviews of a product sold online. Experimental results indicate that the proposed techniques are very promising in performing their tasks.

In [5] feature-based opinion mining model is used. In some domains nouns and noun phrases that indicate product features also imply opinions. In many such cases, these nouns are not subjective but objective. Their involved sentences are also objective sentences and imply positive or negative opinions. This research tries to identify such nouns and noun phrases for effective opinion mining in these domains. Research study [3] focuses on customer reviews of products to determine the semantic orientations (positive, negative or neutral) of opinions expressed on product features in reviews. Most existing techniques for opinion mining utilize a list of opinion words (opinion lexicon) for the purpose. Opinion words are words that express desirable (e.g., great, amazing, etc.) or undesirable (e.g., bad, poor, etc.) states. In this paper, a holistic lexicon-based approach is used which allows the system to handle opinion words that are context dependent, which cause major difficulties for existing algorithms. It also deals with many special words, phrases and language constructs which have impacts on opinions based on their linguistic patterns. It also has an effective function for aggregating multiple conflicting opinion words in a sentence. A system, called Opinion Observer, based on the proposed technique has been implemented.

In [6], it is produced an opinion summary of song reviews similar to that in [2], but for each aspect and each sentiment (positive or negative) they first selected a representative sentence for the group. In [7], blog opinion summarization is produced as brief and detailed summaries, based on extracted topics (aspects) and sentiments on the topics. For the brief summary, their method picks up the document/article with the largest number of positive or negative sentences and uses its headline to represent the overall summary of positive-topical or negative-topical sentences. Research [8] is conducted focusing the task (1) and proposed a method to deal with the problems of the state of the art double propagation method for feature extraction. Research [9] presents a practical system that deals with two related problems in applications of data mining such as mining entities discussed in a set of posts and assigning entities to each sentence. These are very important because without solving them, any opinion discovered from the user-generated content is of limited use. Research study [5] identifies noun product features that imply opinions. Likewise several researches have been conducted performing the task (1).

To perform the task (2) focusing on the orientation of semi structured web content several researches have been conducted. The study [10] studies sentiment analysis of conditional sentences with the aim of determining whether opinions expressed on different topics in a conditional sentence are positive, negative or neutral. A machine learning approach which is different from rule based or statistical techniques is proposed in [11]. This approach naturally integrates multiple important linguistic features into automatic learning. Study [12] has tested and evaluated two approaches for opinion extraction. The first one consists in building a lexicon containing opinion words while the second method consists in using a machine learning technique to predict the polarity of each review. In [3] a holistic lexicon-based approach to determine semantic orientation by exploiting external evidences and linguistic conventions of natural language expressions is proposed. This also capable of handling context dependent opinion words which cause major difficulties for existing algorithms and dealing with many special words, phrases and language constructs which have impacts on opinions based on their linguistic patterns.

In [13] it is proposed a supervised sentiment classification framework which is based on data from Twitter, a popular micro blogging service. By utilizing 50 Twitter tags and 15 smileys as sentiment labels, this framework avoids the need for labour intensive manual annotation, allowing identification and classification of diverse sentiment types of short texts. ASCII smileys and other

punctuation sequences containing two or more consecutive punctuation symbols were used as single-word features. In [14] a smiley based automatic method to collect a corpus for sentiment analysis and [15] introduces a system for opinion mining from the textual content of tweets and discusses the differences between tweet-level and target-oriented opinion mining.

3. METHODOLOGY

The proposed technique of this research is a feature and smiley based data mining technique for opinion mining in product reviews. The new technique is proposed based on orientation identification of opinion sentences. To identify the orientation of the opinion sentences a feature and smiley based approach was used. A summary of the reviews is given after identifying the orientation of opinion sentences. Figure 1 gives the architectural overview of our opinion summarization system.

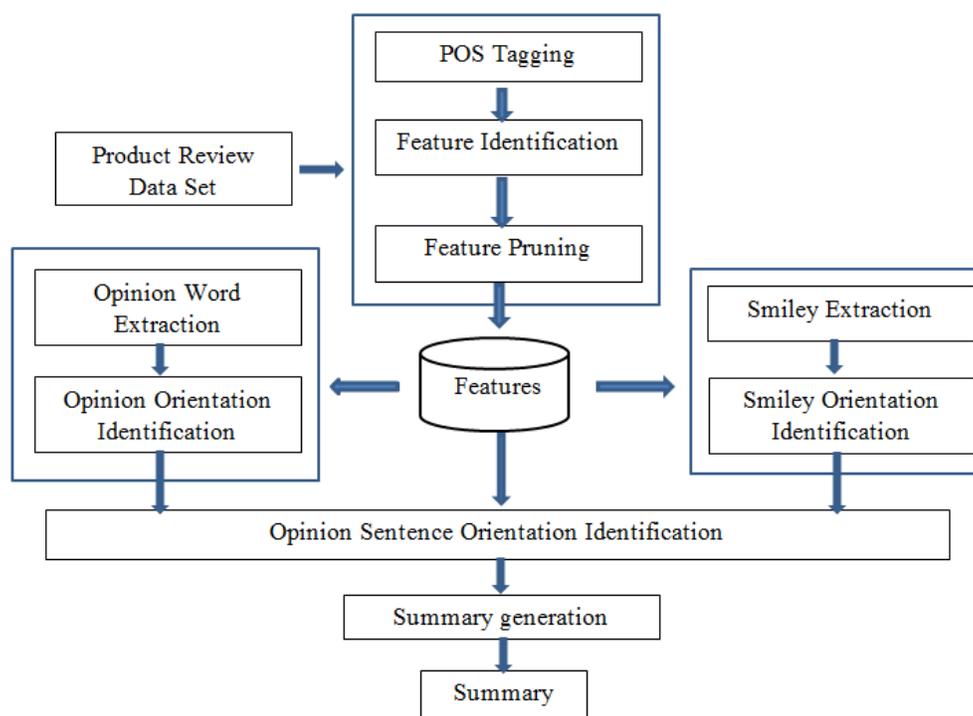


Figure 1. Architectural Overview of the proposed Technique

The inputs to the system are a product name and a set of downloaded data set. The same data set that has been used in [2] was used in our technique in order to improve the effectiveness of validation process. The data set contains reviews of five products including two digital cameras, a cellular phone, an Mp3 player and a DVD player. In research study [2] it uses crawled and downloaded data set of first 100 reviews for each product. The reviews had been collected from Amazon.com and Cnet.com. The dataset was downloaded from [16]. Since the novel technique was developed on a smiley based approach, the data set was updated by adding some smileys to some product reviews.

The system performed the summarization in three main steps

1. Mining product features that have been commented on by customers.

2. Identifying opinion sentences in each review and deciding whether each opinion sentence is positive or negative.
3. Summarizing the results.

These steps were performed in multiple sub-steps. Then features that people had expressed their opinions were found. Features are product attributes. Ex: Battery size, Picture quality. After finding the features, the opinion words were extracted using the resulting features. Opinion words are the words that are primarily used to express subjective opinions.

Semantic orientations of the opinion words were identified with the help of SentiWordNet. SentiWordNet is a large lexical database of English. It is a lexical resource for opinion mining. SentiWordNet assigned each word one of three sentiment scores such as positive, negative, or neutral.

There were some opinion sentences which had opinion words that cannot be identified by SentiWordNet due to various reasons such as spelling mistakes. The smiley based approach was used to identify the orientation of such opinion words. The sentences which were not classified as positive or negative using opinion words were taken and smiley extraction was done for them. Using the smileys and their orientations the orientation of the opinion sentences towards each feature was found. Smileys were used for orientation mining purpose.

Finally the orientations of opinion sentences were identified and a final summary was produced. To find opinion features the part-of-speech tagging (POS tagging) which is from natural language processing, was used.

3.1. Part-of-Speech Tagging (POS)

Usually product features are nouns or noun phrases in review sentences. Therefore, identifying the relevant nouns or noun phrases correctly was crucial for the effectiveness of the proposed technique. For this purpose we used part-of-speech tagging of Stanford Log-linear Part-Of-Speech Tagger to split text into sentences and to produce the part-of-speech tag for each word [17]. The word can be a noun, verb, adjective etc. Simple noun and verb groups were identified through this mechanism.

Each line in the data set was tagged using Stanford POS tagger version 3.5.0. Nouns and noun phrases were identified as product features. Other components of the sentence were unlikely to be product features. Some pre-processing of words was also performed to remove unnecessary words or symbols.

3.2. Feature Identification

In this step product features were identified which customer had expressed their opinions on. Since the system aimed to find what customer like and dislike about a given product, finding the product features that customer talk about was more important. Due to the difficulty of natural language understanding this was a hard task to accomplish. Some sentences explicitly mention the product features while some sentences used implicit ways. For an example in the sentence “The pictures are very clear.” the feature is mentioned explicitly, saying that user is satisfied with the picture quality. But in the sentence “While light, it will not easily fit in pockets.” product feature is not explicitly mentioned. The size of the product is implicitly mentioned in the sentence. In this research, it is focused only on finding explicitly mentioned features that appear as nouns or noun phrases in the reviews.

It is common that a customer review contains many things that are not directly related to product features. However, when customers comment on product features, they normally use same set of words. Thus finding frequent word sets is appropriate because those frequent word sets are likely to be product features. However, not all results generated are relevant features. Pruning was done to remove the unlikely features. This procedure identified features that were meaningless and redundant have been removed.

3.3. Opinion Words Extraction

Opinion words are the words that are primarily used to express subjective opinions. Presence of adjectives is useful for predicting whether a sentence is subjective or expressing an opinion. Adjectives were used as opinion words in our study. The opinion words extraction was limited to the sentences that contain one or more product features, as we were only interested in customers' opinions on these product features. If a sentence contains one or more product features and one or more opinion words, then the sentence is called an opinion sentence. We extracted opinion words from opinion sentences within this sub step.

3.4. Orientation Identification for Opinion Words

Within this sub step the semantic orientation of each opinion word was identified. It was used to predict the semantic orientation of each opinion sentence. Words that encode a desirable state (e.g., beautiful, awesome) have a positive orientation, while words that represent undesirable states have a negative orientation (e.g., disappointing). There are some adjectives which have no orientation as well (e.g., external, digital). In this research, only the adjectives which have positive or negative orientations were considered. The orientations of the adjectives were identified using SentiWordNet.

3.5. Smiley Extraction

Smiley based technique was added to the system for completeness. Thirteen smiley types were used in this research to identify the opinions of opinion sentences. They were clearly identified as positive or negative.

In this step the opinion mining algorithm was completed using smileys. Modern customers used to type smileys when they give product reviews on the web. It has become a trend. These symbols were used in order to identify the opinions of opinion sentences which were filtered and missed from the above mentioned process. For an example if a sentence contained a product feature correctly but the opinion word could not be identified correctly using SentiWordNet, then that sentence was rejected. But in this algorithm the orientation of that sentence was identified using smileys if the sentences contain any to make the results more complete and accurate.

There can be some misspelled opinion words, non-English opinion words with some frequent product features. For those adjectives that SentiWordNet cannot recognize, they were discarded as they may not be valid words. There can be opinion sentences with features which have no identified orientation. But as a sentence it gives an overall opinion about a product feature. That kind of sentences also can be categorized as positive or negative by identifying the orientation of the smileys used in them.

3.6. Orientation Identification for Smileys

Orientation mining of the extracted smileys was done by giving orientations for each smiley. In [15] several important smileys have been identified with their orientations. In this research also same set of smileys was used. They are as follows,

Positive types of smileys are, :) , :) , ;) , :-) : D , =) , ;) , (:

Negative types of smileys are, :(, : (, :-(,) : ,) :

This information was given in the algorithm to identify the opinions of the opinion sentences with frequent product features. This approach completed the algorithms by increasing the number of sentences considered in the process.

3.7. Predicting the Orientations of Opinion Sentences

In this step the orientation of an opinion sentence was predicted as it was positive or negative. In general, the dominant orientation of the opinion words in the sentence was used to determine the orientation of the sentence. That is, if positive or negative opinion prevails, the opinion sentence was regarded as a positive or negative one. It was also considered whether there was a negation word such as “no”, “not”, “yet”, appearing closely around the opinion word. If so, the opinion orientation of the sentence was taken as the opposite of its original orientation. Using this kind of technique the orientation of the opinion sentences were predicted for the purpose of summary generation.

3.8. Summary Generation

For each discovered feature, related opinion sentences were put into positive and negative categories according to the opinion sentences' orientations. A count was computed to show how many reviews has given positive/negative opinions to the feature.

All features were ranked according to the frequency of their appearances in the reviews. There were several ways of ranking features. Number of reviews that express positive or negative opinions with the feature was used to rank the features.

4. EXPERIMENTAL EVALUATION

4.1. Testing Criteria

Five data sets were tested using the developed algorithm. Precision and Recall values were calculated for feature extraction and the opinion extraction processes.

Tests were done for the original data set first. Then the data sets were updated by adding smileys relevantly. After all twenty tests were conducted altogether as four per each product. Two were for the data set before adding smileys and the other two were for the dataset after adding smileys. Smileys were added randomly but relevantly for each product review data set when updating them for the testing process. Forty new smileys were added for each data set. To calculate the precision and recall values manual counts of the datasets were taken.

4.2. Evaluation and Validation

The test results of twenty tests for five data sets are shown in table1.

Table 1. Test Results.

Data set		Feature Based Algorithm		Feature + Smiley Based Algorithm	
	Smiley	Feature extraction	Opinion identification	Feature extraction	Opinion identification
(1)	No	Precision=0.6667 Recall= 0.5823	Precision=0.6829 Recall= 0.7778	Precision=0.6667 Recall= 0.5823	Precision=0.6842 Recall= 0.7945
	Yes	Precision = 0.6667 Recall = 0.5823	Precision=0.6829 Recall= 0.7778	Precision=0.6667 Recall = 0.5823	Precision=0.720 Recall= 0.8611
(2)	No	Precision=0.6507 Recall= 0.6212	Precision=0.7086 Recall= 0.694	Precision=0.6507 Recall= 0.6212	Precision=0.7320 Recall= 0.7272
	Yes	Precision=0.6507 Recall= 0.6212	Precision=0.7086 Recall= 0.694	Precision=0.6507 Recall= 0.6212	Precision=0.7687 Recall= 0.7987
(3)	No	Precision=0.6461 Recall= 0.6268	Precision=0.8562 Recall= 0.6777	Precision=0.6461 Recall= 0.6268	Precision=0.8588 Recall= 0.6919
	Yes	Precision=0.6461 Recall= 0.6268	Precision=0.8562 Recall= 0.6777	Precision=0.6461 Recall= 0.6268	Precision=0.8632 Recall= 0.7773
(4)	No	Precision=0.6747 Recall= 0.8235	Precision=0.8619 Recall= 0.7862	Precision=0.6747 Recall= 0.8235	Precision=0.8625 Recall= 0.7900
	Yes	Precision=0.6747 Recall= 0.8235	Precision=0.8619 Recall= 0.7862	Precision=0.6747 Recall= 0.8235	Precision=0.8653 Recall= 0.8092
(5)	No	Precision=0.5345 Recall= 0.6327	Precision=0.7590 Recall= 0.6517	Precision=0.5345 Recall= 0.6327	Precision=0.7619 Recall= 0.6586
	Yes	Precision=0.5345 Recall= 0.6327	Precision=0.7590 Recall= 0.6517	Precision=0.5345 Recall= 0.6327	Precision=0.7626 Recall= 0.6789

Test results were compared with manual results to validate the algorithm.

Existing algorithm and the new developed algorithm were compared using the precision and recall values of them. This needed the manual counts of the data sets.

4.3. Statistical Analysis and Comparison

A comparison was done according to the average precision and recall values of two algorithms. Five datasets for five different products were tested using the developed algorithm. According to the results, the following observations were made.

When considering feature extraction the algorithm resulted better precision for all the data sets. Recall was increased only in two data sets. The results were the same for smiley added data sets. Feature + Smiley based algorithm gave the same results as the feature based algorithm for feature extraction. This is because the feature+ smiley based algorithm also use the same mechanism for feature extraction. In the opinion identification process, precision and recall were better in the feature based algorithm than the existing algorithm. Four datasets gave better results than the results they gave for the existing algorithm. When considering the feature + smiley based algorithm for customer review opinion mining all the five datasets gave better results than the existing algorithms for all precision and recall values in all the tests.

Average Precision and recall values for the existing system.

Feature extraction:	Precision = 0.56	Recall = 0.68
Opinion Identification:	Precision= 0.642	Recall = 0.693

Average Precision and recall values for the new algorithm (Data set without smileys).

Feature extraction:	Precision = 0.6345	Recall = 0.6573
Opinion Identification:	Precision = 0.7752	Recall = 0.7273

In feature extraction precision is improved while the recall value is decreased marginally. The new algorithm extracted features better than the existing algorithm.

Average Precision and recall values for the new algorithm (Data set with smileys).

Feature extraction:	Precision = 0.6345	Recall = 0.6573
Opinion Identification	Precision = 0.7960	Recall = 0.7850

The new algorithm has improved precision and recall values for both feature extraction and opinion identification. This implies that the new developed algorithm works better with data sets with smileys.

5. CONCLUSIONS

The objective of the research was to develop a more accurate data mining algorithm for opinion mining of customer reviews on the web. The research was conducted as an experimental study. A new algorithm was developed which enables feature and smiley based approach for opinion mining in customer reviews on the web. Product feature extraction, opinion word identification, opinion orientation identification of opinion words, smiley extraction, smiley orientation identification and opinion summary generation were the main components of the developed algorithm. Stanford POS tagging and SentiWordNet were used for tagging sentences and meaning identification of opinion words respectively.

The algorithm gives better precision values for all the datasets in every test. Although recall values were not improved in feature extraction, when considering the ultimate objective, opinion orientation identification recall values are improved by the new algorithm. For opinion identification new algorithm is better than the existing algorithm for all the datasets tested.

As future work developed algorithm can be improved to identify infrequent features. By introducing infrequent feature identification component to the algorithm, the recall values could be improved. In this study orientation of opinions were identified only as positive or negative. As future work weights can be allocated for these opinion words. Then the summary could be generated with weighted values.

Identifying the misspelled words was done to some extent in this research. But it is in preliminary stage. Misspelled word identification can be improved in future work. This will be helpful in extracting features and adjectives correctly and relevantly. This algorithm only works for the explicitly mentioned opinions and the product features. The algorithm can be improved as it can extract implicitly mentioned features and opinions. There can be situations that customers put smileys unnecessarily and sometimes smileys are put which have an opposite idea to the text. These situations can be addressed as future work by identifying the real meaning of smileys comparing with texts.

REFERENCES

- [1] U. Fayyad, G. Shapiro and P. Smyth, 'Knowledge discovery and data mining: Towards unifying framework', in Knowledge discovery and data mining (KDD-96), Portland, Oregon, 1996.
- [2] M. Hu and L. B. 'Mining and summarizing customer reviews', in tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 168-177, ACM, 2004.
- [3] X. Ding, B. Liu and P. Yu, 'A holistic lexicon- to opinion mining', in International Conference on Web Search and Data Mining, pp. 231-240 ACM, 2008
- [4] B. Liu, Sentiment analysis and opinion mining. San Rafael, Calif.: Morgan & Claypool, 2012.
- [5] L. Zhang and B. Liu, 'Identifying noun product features that imply opinions', in 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers- Volume 2, pp. 575-580, Association for Computational Linguistics, 2011.
- [6] S. Tata and B. Di Eugenio, 'Generating fine-grained reviews of songs from album reviews', in 48th Annual Meeting of the Association for Computational Linguistics, pp. 1376-1385, Association for Computational Linguistics, 2010.
- [7] L. Ku, Y. Liyang and H. Chen, 'Opinion Extraction, Summarization and Tracking in News and Blog Corpora', in Computational Approaches to Analyzing Weblogs (Vol. 100107), 2006.
- [8] L. Zhang, B. Liu, S. Lim and E. O'Brien-Strain, 'Extracting and ranking product features in opinion documents', Association for Computational Linguistics, 2010.
- [9] D. Xiaowen, B. Liu and L. Zhang, 'Entity discovery and assignment for opinion mining applications', in 15th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 1125-1134. ACM, 2009.
- [10] R. Narayanan and B. Liu, 'Sentiment analysis of conditional sentences', in Empirical Methods in Natural Language Processing: Volume 1-Volume 1 (pp. 180-189), Association for Computational Linguistics, 2009.
- [11] W. Jin, H. Ho and R. Srihari, 'Opinion-Miner: a novel machine learning system for web opinion mining and extraction', in 15th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 1195-1204). ACM, 2009.
- [12] D. Poirier, C. Bothorel and M. Boulle, 'Two possible approaches for opinion analysis in film reviews: statistic and linguistic', in EMOT-2008: LREC 2008 Workshop on Sentiment Analysis: Emotion, Metaphor, Ontology, 2008.
- [13] D. Davidov, O. Tsur and A. Rappoport, 'Enhanced sentiment learning using twitter -hashtags and smileys', in the 23rd International Conference on Computational Linguistics: Posters (pp. 241-249), Association for Computational Linguistics, 2010.

- [14] A. Pak and P. Paroubek, 'Twitter as a Corpus for Sentiment Analysis and Opinion Mining', in LREC, 2010, 2010.
- [15] V. Hangya and F. Richard, 'Target-oriented opinion mining from tweets', in Cognitive Informatics (CogInfoCom), IEEE, 2013.
- [16] B. Liu, 'Sentiment Analysis and Opinion Mining', Synthesis Lectures on Human Language Technologies, vol. 5, no. 1, pp. 1-167, 2012.
- [17] Nlp.stanford.edu, 'The Stanford NLP (Natural Language Processing) Group', 2015.
[Online]. Available: <http://nlp.stanford.edu/software/tagger.shtml>. [Accessed: 13- Jul- 2015].

AUTHORS

Ms. Iroosha Jayasekara completed her BSc. in Management and Information Technology (Special) degree with a first class from University of Kelaniya. Currently she is working as a Technical consultant at Attune lank Pvt. Ltd. Her research interest is in the area of business intelligence and data mining and conducting her research study on data mining techniques for opinion mining of customer reviews.



Janaka I Wijayanayake received a PhD in Management Information Systems from Tokyo Institute of Technology Japan in 2001. He holds a Bachelor's degree in Industrial Management from the University of Kelaniya, Sri Lanka and Master's degree in Industrial Engineering and Management from Tokyo Institute of Technology, Japan. He is currently a Senior Lecture in Information Technology at the department of Industrial Management, University of Kelaniya Sri Lanka. His research findings are published in prestigious journals such as Journal of Information & Management, Journal of Advances in Database, Journal of Business Continuity & Emergency Planning and many other journals and international conferences

