Current Applied Science and Technology Vol. 20 No. 2 (May-August 2020)

Opinion Mining for Laptop Reviews using Naïve Bayes

Pakawan Pugsee[1]* and Thanapat Chatchaithanawat[2]

[1]Innovative Network and Software Engineering Technology Laboratory, Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok, Thailand
[2]Stream I.T. Consulting Co., Ltd., Bangkok, Thailand

Received: 6 March 2020, Revised: 9 March 2020, Accepted: 23 March 2020

Abstract

This research is to develop an opinion mining application which allows users to clarify what the reviews on the laptop mentioned. The aim of the research is to analyze user's opinions from laptop reviews on popular online communities. The proposed methodology is composed of four essential processes: preparing data for analysis, detecting subjective text paragraphs, identifying the aspects and classifying the sentiments of text paragraphs. The subjective textual contents are determined by detecting subjective words occurred in the sentences of text paragraphs. Then, only the subjective paragraphs might be classified into specific aspects using comparisons with the vocabularies of aspect domains. Finally, the paragraph sentiments will be categorized into positive or negative opinions using the Naïve Bayes classifier. The experimental results with the performance evaluation showed that the accuracy and precision of the subjective detection of text paragraphs are greater than 90%. In addition, the accuracy and precision of sentiment classification are more than 70%. Therefore, this tool can help consumers in categorizing laptop review paragraphs into aspects and sentiment groups for making selections before purchasing laptops.

Keywords: opinion mining, review analysis, laptop reviews, Naïve Bayes
DOI 10.14456/cast.2020.16

1. Introduction

Currently, the market of portable computer or laptops has become more competitive marketing with tablets and mobile devices. There are several whole manufactures within the portable computer business and that they frequently produce many new laptop models to contend one another. For this reason, consumers have many choices in making decision for buying laptops. Although there are many laptop-review forums on the Internet such as online communities and blogs, customers must take time to scan and explore for too much data. It is very useful if there is a tool that facilitates customers to choose the laptops that they want, gather review information from varied review forums and analyze the helpful data for them.

*Corresponding author: Tel..: +66 22 18 5170 Fax: +66 22 55 2287
E-mail: pakawan.p@chula.ac.th

Opinion mining or sentiment analysis is the methodology that tries to recognize people's mind or opinions by analyzing information from text data, e.g. user comments, blogs, or reviews. One objective of opinion mining is to differentiate the opinion of a supply text into the positive or negative opinions. The opinions or sentiments stated intentions, emotions, decisions, evaluations, needs or desires [1]. Moreover, opinion mining is often used to analyze customer reviews to examine consumers' satisfaction. Opinion mining tools, or sentiment analysis systems can assist users who are customers or consumers to get useful information about interesting products and services. Furthermore, these tools or systems can be used for investigating market trends and surveying customer desires to improve the product qualities and the potency of consumer service.

Most reviews on online communities regarding laptops, e.g. notebookcheck.net, notebookreview.com, cnet.com and laptopmag.com consist of information about laptops in performance, design (style) and features (options). Therefore, this article studied on opinion analysis about laptops' reviews in 3 aspects that are the performance, the design and the features of a product. In addition, this opinion mining tool is the implementation of the framework [2] with a few modifications for improvement.

2. Materials and Methods

2.1 Background knowledge and related works

The goal of opinion mining or sentiment analysis is to distinguish comments or the attitude on various topics in the natural language, so that this analysis can classify the emotional aspects of communication. The research in this field is about grouping of words or messages as the positive attitude or the negative attitude. Some sentences or phrases can express opinions or attitudes, positive or negative. These sentences or phrases also help identify the groups of reviews or comments more easily. Therefore, Pang *et al*. [3] and Turney [4] developed two approaches in the sentiment analysis to identify comment or opinion messages on a social network into the positive or the negative groups.

Sentiment analysis of text statements needs some techniques of natural language processing. Sentiment analysis with natural language processing of product reviews has been utilized in widespread applications to enhance consumer retention and business processes [5]. The natural language processing is the study of computer science, artificial intelligence and linguistics in term of the interaction between humans and computers. It is composed of standard methods to make computers understand natural language or human language involving natural language comprehension and making computers understand human or natural language input. There are three main processes which are syntactic analysis, semantic analysis and pragmatic analysis. First, syntactic analysis will check the grammatical structures and the position of various groups of words that make up the sentence. Secondly, semantic analysis is the accuracy verification in term of the meaning of the sentence. The grammatical sentences normally have the exact meaning. However, some grammatical sentences considered in this field might have ambiguous meaning or no meaning at all. Lastly, pragmatic analysis is the situation needed to be considered to interpret these sentences because sometimes the sentences might not be able to interpret directly. In this case, the sender, the receiver and the content must be in the same situation in order to have the same comprehension. In addition, there are some lexicons containing only sentiment words that are used to classify the sentiments of words in semantic analysis, such as the MPQA (Multi-Perspective Question Answering) subjectivity lexicon [6] and SentiWordNet [7]. The MPQA subjectivity lexicon and SentiWordNet are a publicly available lexical for opinion mining. The MPQA Subjectivity Lexicon can be used to score words or phrases of words to determine whether they are positive or negative. For every entry, the lexicon creates a result to indicate if an entry is positive, neutral or negative in its opinion. SentiWordNet assigns to each synset of WordNet with three sentiment scores: positivity, negativity, objectivity [8].

Moreover, many machine learning techniques are applied to classify or cluster the sentiments or opinions of text statements. Machine learning is a type of artificial intelligence that makes computers have the self-learning ability [9]. It can be categorized into 2 main types: supervised learning and unsupervised learning. Supervised learning is a learning of the input data in which the answers are already given, such as predicting the sentiment of a sentence by training examples of sentences with their opinion meaning, or the stock price at a particular time. This type of machine learning is prepared for the data prediction involving the problems like regression and classification. Unsupervised learning is a learning of the input data in which the answers are still unknown. The type of machine learning helps us get closer to the answers or understand more problems by arranging the data structure. The model will be prepared to use in the data structure in order to reduce duplication and categorize data into the same group, for example, the problem about clustering.

Furthermore, there are many researches about opinion mining with machine learning techniques in the last few years. For examples, the proposed method in Govindaraj and Gopalakrishnan [5] used acoustic and textual features to analyze opinions on customer product reviews from Amazon product reviews and YouTube. Customer feedback in the form of audio clips (.wave file) was proceeded by speech synthesis tool, speech recognizer and voice-to-text converter before feature selection using hand-coded rules. Acoustic and textual features were calculated and extracted to generate the training data sets for building three classification models with three different feature sets by the support vector machine (SVM) classifier. The opinions of a customer chatting were categorized into 5 levels of sentiment score: extremely positive, positive, neutral, negative and extremely negative. Valdivia *et al.* [10] anticipated an analysis for matching between users' sentiments and automatic sentiment-detection algorithms using TripAdvisor as a resource for sentiment analysis, including the challenges of sentiment analysis and TripAdvisor. The best-known sentiment analysis task aims to observe the sentiments at intervals documents, sentences, or words. This work is often separated into 3 steps: polarity detection (positive, negative, or neutral), aspect extraction (features for organizing the text) and classification (machine learning or lexicon approaches). There are various forms of texts, such as tweets, blog, and reviews. In addition, human language is complicated because of different grammatical rules, cultural variations and jargon in statements. This study obviously expressed the requirement of mining opinions beyond user ratings. Therefore, the implementation of sentiment analysis techniques to extract opinions is crucial to understanding the mind of a traveler and can influence quality improvement in tourism.

Pugsee *et al*. [11] implemented the sentiment analysis application to mine opinion on Twitter messages. Tweets about skin care (with "#skincare") was analyzed by combining word information with the machine learning techniques. SentiWordNet [8] was modified to improve the performance of application for skin care products and two machine learning techniques, i.e. Naïve Bayes and SVMs were implemented to identify the sentiments of messages. The user opinions on tweets were categorized into 5 levels of sentiments: very positive, positive, neutral, negative and very negative. The performance of result classification was evaluated by the accuracy, the precision, and the recall rate that all of their values are more than 75%.

Therefore, our research designed to use SentiWordNet with basic machine learning techniques like a decision tree and Naïve Bayes to implement the opinion mining tool for the laptop reviews. The reasons are that both methods are simple machine techniques that can be implemented and embedded in the software tools easier, including a not too long processing time. Moreover, the sentiment classification with Naïve Bayes in Pugsee *et al.* [11-12] has sufficient performance.

2.2 Opinion mining methodology

The proposed opinion mining tool can provide an organized summary of the product reviews for customers and assist them with the decision making when they want to buy laptop products. The overview of the proposed opinion mining tool following the framework for analyzing laptop reviews [2] with some modification is shown in Figure 1. This tool consists of four main processes: data preparation, subjective detection, aspect identification and sentiment classification.

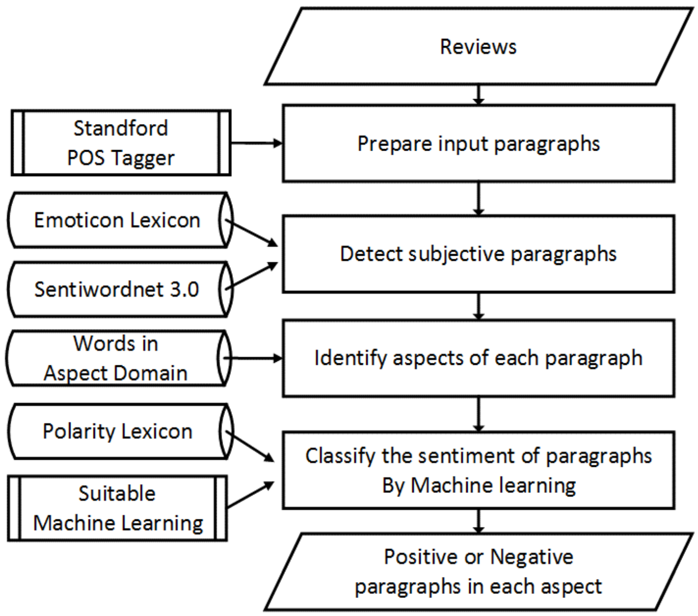

Figure 1. The overview of proposed opinion mining tool

According to Figure 1, the proposed tool is composed of four processes which are to prepare data for analysis, to detect subjective text paragraphs, to identify the aspect of each text paragraph and to classify the sentiments of each text paragraph. The objectives of the opinion mining tool are to categorize content paragraphs in subjective or objective paragraphs, to identify paragraphs' aspects into four aspects and to classify the sentiments of paragraph reviews. The inputs of this tool are the laptop reviews from online communities and the outputs are both groups of text paragraphs that are positive or negative text paragraphs in each aspect domain.

2.2.1 Prepare input paragraphs

This process is implemented based on the technique in Chatchaithanawat and Pugsee [2]. The first step is to delete special characters and symbols in text paragraphs. There are more than 200 special characters and symbols were added from Chatchaithanawat and Pugsee [2], such as other characters not in English alphabets and symbols. In addition, photo and URL links will be deleted from the input reviews. When the photos from reviews in community website are saved in text format it will be saved as [IMG] tag. This process will delete [IMG] tag from the original reviews. Moreover, normal URL links will also be deleted from the reviews by detecting "http" and "www". Furthermore, picture links will be deleted from reviews by detecting the ".jpg", ".gif" and ".png".

The second step is to tag words with their parts of speech, after separating paragraphs and sentences by a tab character, a full stop, and a newline. Stanford POS Tagger [13] demonstrated to identify words' parts of speech, such as adjectives, adverbs, verbs and nouns. In the next step, the tagged words are changed into the basic forms using WordnetStemmer [14] to manage stemming method. Stemming is to transform into the base form of the focused words by removing the prefix and suffix of the words. The focused words are words in the adjective group, the adverb group, the verb group and the noun group. Figures 2 and 3 present an example of a review paragraph and a prepared paragraph, which are the input and output of this process.

```
The touchpad is able to recognise even the complex 3-finger gestures with great precision. During about 2 hours
of use I've only had 3 times when the mouse didn't do what I was expecting, mostly when trying to select text
(which is tricky business on touchpads anyway). I didn't have trouble with palm rejection either, though it might
be because my hands don't touch the touchpad while typing :p .
```

Figure 2. An example of a review paragraph

```
The_DT touchpad_NN is_VBZ able_JJ to_TO recognise_VB even_RB the_DT complex_NN 3-finger_NN
gestures_NNS with_IN great_JJ precision_NN ._.
During_IN about_RB 2_CD hours_NNS of_IN use_NN I_PRP 've_VBP only_RB had_VBN 3_CD times_NNS
when_WRB the_DT mouse_NN did_VBD n't_RB do_VB what_WP I_PRP was_VBD expecting_VBG ,_,
mostly_RB when_WRB trying_VBG to_TO select_VB text_NN -LRB-_-LRB- which_WDT is_VBZ tricky_JJ
business_NN on_IN touchpads_NNS anyway_RB -RRB-_-RRB- ._.
I_PRP did_VBD n't_RB have_VB trouble_NN with_IN palm_NN rejection_NN either_CC ,_, though_IN
it_PRP might_MD be_VB because_IN my_PRP$ hands_NNS do_VBP n't_RB touch_VB the_DT
touchpad_NN while_IN typing_NN :p_NN ._.
```

Figure 3. A prepared paragraph

2.2.2 Detect subjective paragraphs

The word information from SentiWordNet [8], which is categorized into the adjective group, and the adverb group, will be useful for identifying whether those words are subjective or objective. Both groups are interesting words in this research and the best information for analyzing subjective statements because most subjective words are in the adjective and the adverb groups. This process detects subjective paragraphs like algorithm described in Chatchaithanawat and Pugsee [2]. Every text paragraph, which has at least one subjective word or emoticon text, will be considered as a subjective paragraph. Additionally, emoticon texts will also be identified by comparing emoticon texts in the paragraph with data from an emoticon lexicon [15] including some emoticons found in experimental data. The steps of this process are shown in Figure 4.

The inputs of this process are prepared paragraphs from the previous process. Then, the emoticons in prepared paragraphs will be detected by comparing found emoticons of text paragraph to the emoticons in the lexicon. If emoticons found match with data in the emoticon lexicon at least one emoticon, those paragraphs will be collected as subjective paragraphs. If there are no detected emoticons in the prepared paragraphs, the subjective words will be detected in this process by comparing with words in SentiWordNet. If the subjective words are found at least one word, the paragraphs will be kept as subjective paragraphs. The emoticons in the lexicon and the subjective words in SentiWordNet are compared to sequential words in the paragraph using unicode matching and string matching, respectively. Consequently, only the subjective paragraphs are the outputs of this process. The detected subjective words and emoticon texts in the paragraph are exposed in Figure 5.

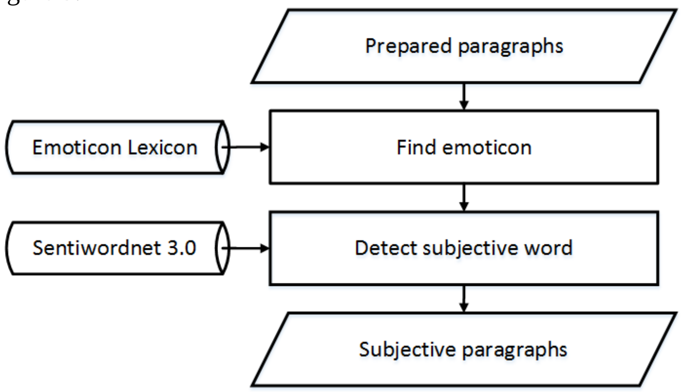

Figure 4. The steps of detecting subjective paragraphs process

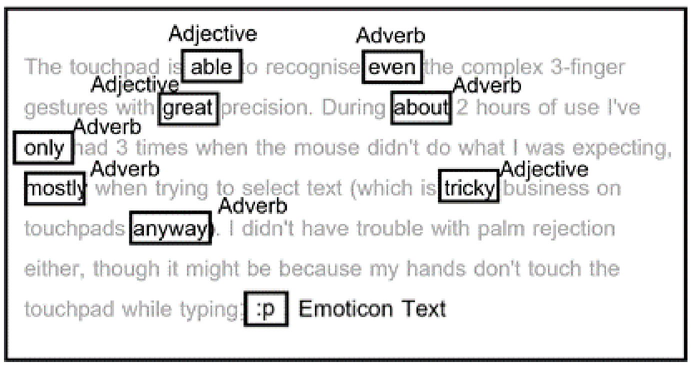

Figure 5. The detected subjective words and emoticon texts in the paragraph

2.2.3 Identify aspects of each paragraph

In this process, the subjective paragraphs from the previous process will be divided into four different aspects (“Performance”, “Design”, “Feature” and “Other”) by comparing words in review paragraphs with words in word lists of three aspect domains. Individual subjective paragraphs can match more than one aspect. Otherwise, some subjective paragraphs that cannot be recognized in previous groups will be identified into “Other” aspect. The words of each aspect domain are listed by analyzing the popular words found in laptop reviews. The steps of this process are shown in Figure 3. The challenge of this process is creating the word lists in each aspect that are useful to categorize the aspect of paragraph correctly. Finding the frequency of all content words in laptop reviews and determining the threshold of the word frequency to count as words in each aspect were proceeded to generate the word lists.

According to Figure 6, this process will detect words in aspect domains for identifying types of review aspects. The examples of words in each aspect domain are shown in Table 1, and the detected words of a paragraph are shown in Figure 7. These words will be collected from all review paragraphs by using AntConc [16], which helps to find the frequency of words in each paragraph. Then, the aspect words with high frequency from all reviews will be categorized into each aspect domain by the researcher's judgment to classify the aspect of paragraphs.

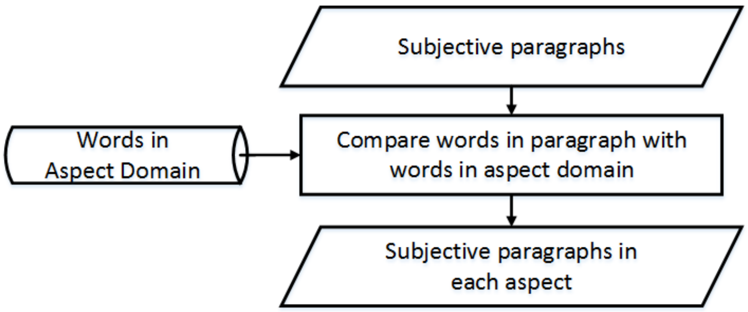

Figure 6. The step of identifying aspects of each paragraph process

Table 1. Examples of words in three aspects

Performance	Design	Feature
battery	display	Bluetooth
CPU	height	camera
GPU	materials	DVI
memory	screen	HDMI
processor	size	touchpad
ram	weight	USB
resolution	width	wireless

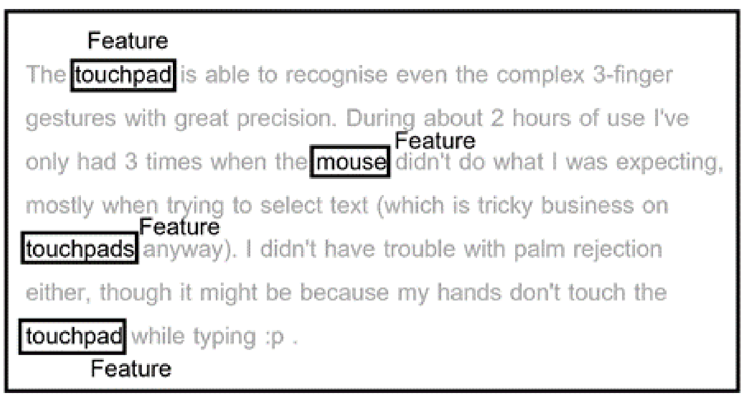

Figure 7. The detected words in the feature aspect

2.2.4 Classify the sentiments of paragraphs

The subjective paragraphs in individual aspect will be classified into the sentiment types of paragraphs by using the selected machine learning technique. The results of this process are two groups of text paragraphs (positive or negative paragraphs). There are 6,234 text paragraphs in experiments and these texts are categorized into 2,534 positive paragraphs and 3,700 negative paragraphs. The selected features of the classification model are all adjectives, adverbs, their parts of speech, and their polarity score will be learned and classified by WEKA (Waikato Environment for Knowledge Analysis) [17] to choose the suitable feature set and a classifier. WEKA is one of the popular machine learning software implemented in JAVA programming language by the University of Waikato, New Zealand. This software tool is free to use under the General Public License (GPL). It is a collection of machine learning algorithms for data mining tasks, such as data pre-processing, classification, regression, clustering, association rules, and visualization. Our research tried to test on three different feature sets with two machine learning techniques (Naïve Bayes and Decision Tree). Then, Decision Tree (J48) and Naïve Bayes classifier of WEKA are executed to generate the classification models of sentiment analysis. To test the performance of classification, the confusion matrix is applied with labeled the positive and negative paragraphs by manual annotation in order to assess the performance of the classification model with evaluation values. The confusion matrix and three evaluation values are displayed in Table 2.

Table 2. A confusion matrix and evaluation values

Actual class	**Predicted class**		**Accuracy**	**Precision**	**Recall**
	Positive	Negative			
Positive	True positive (TP)	False negative (FN)	TP+TN/ (TP+FN+FP+TN)	TP/ (TP+FP)	TP/ (TP+FN)
Negative	False positive (FP)	True negative (TN)		TN/ (TN+FN	TN/ (TN+FP)

According to Table 2, there are three evaluation values of the performance of classification that are accuracy, precision and recall. The accuracy is calculated from the number of data with the correct prediction comparing to the total number of data. The precision is counted using the number of data with the correct prediction comparing to the number of predicted data in each class, while the recall is calculated by comparing to the number of actual data in each class.

In the first experiment, all adjective and adverb words with their parts of speech will be used in the training data. Figure 8 displays the adjective and adverb words in the paragraph. The percent of accuracy, precision and recall rates will be calculated by confusion matrices. The confusion matrices of the first experiment and the percentage of accuracy, precision, and recall rates are shown in Table 3 and Table 4, respectively.

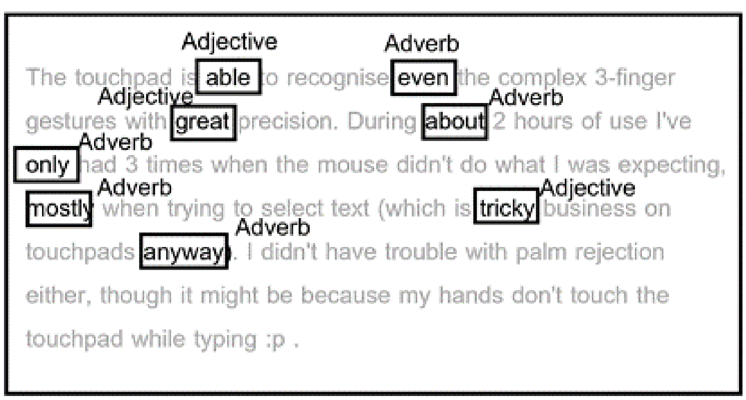

Figure 8. The adjective and adverb words in the paragraph

Table 3. The confusion matrices of results in Experiment I

Actual class		Predicted class			
		Naïve Bayes		J48	
		Positive	Negative	Positive	Negative
Positive	2,534	1,499	1,035	1,695	839
Negative	3,700	922	2,778	821	2,879
Total	6,234	2,421	3,813	2,516	3,718

Table 4. The percentage of accuracy, precision and recall of Experiment I

Classifier		Accuracy	Precision	Recall
Naïve Bayes	Positive	68.61%	61.91%	59.16%
	Negative		72.86%	75.08%
J48	Positive	73.37%	67.37%	66.89%
	Negative		77.43%	77.81%

In the second experiment, the word information from SentiWordNet [8] was modified to create the polarity lexicon which consisted of words and their polarity levels ("strong positive", "positive", "neutral", "negative" and "strong negative"). Some concatenated adjective words with their polarity level are added into the polarity lexicon to enhance the tagged polarity, e.g., high-end, full-colored, and industry-standard. The polarity levels of all adjectives and adverbs are included into the training data with words and their parts of speech from the first experiment. Figure 9 presents the polarity levels of adjectives and adverbs in the paragraph. The confusion matrices of the second experiment and the percentage of accuracy, precision and recall rates are shown in Tables 5 and 6, respectively.

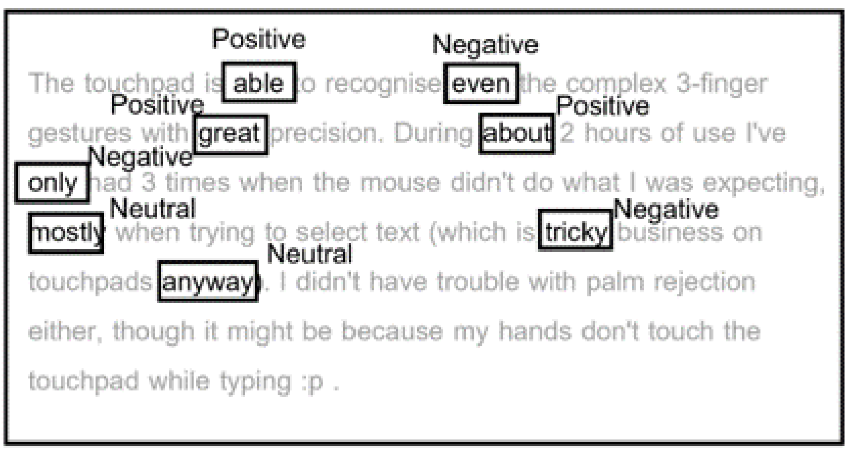

Figure 9. The polarity levels of adjectives and adverbs in the paragraph

Table 5. The confusion matrices of results in Experiment II

Actual class		Predicted class			
		Naïve Bayes		**J48**	
		Positive	Negative	Positive	Negative
Positive	2,534	1,676	858	958	1,576
Negative	3,700	860	2,840	368	3,332
Total	6,234	2,536	3,698	1,326	4,908

Table 6. The percentage of accuracy, precision and recall of Experiment II

Classifier		Accuracy	Precision	Recall
Naïve Bayes	Positive	72.44%	66.09%	66.14%
	Negative		76.80%	76.76%
J48	Positive	68.82%	72.25%	37.81%
	Negative		67.89%	90.05%

The classification results in Experiment I and Experiment II are different. The performance of the classification model of decision tree technique is higher than those by the Naive Bayes classifier in the first experiment. The reason is that there are various adjective and adverb words found in reviews, so only the probability of words is not sufficient to classify the sentiment, while the decision tree has bias in the majority of data. On the other hand, the classification model of Naive Bayes classifier has capacity more than the classification model of decision tree technique. It is found that the polarity level of words can help improve the sentiment classification performance, but the decision tree is overfitted to the data with bias in the majority class. Therefore, there is a test on the third feature set that there is only the polarity level.

In the third experiment, only polarity levels of words in adjective and adverbs which are strong positive, very positive, positive, neutral, negative, very negative and strong negative will be used in the training data. The confusion matrices and the percentage of accuracy, precision and recall are shown in Tables 7 and 8.

Table 7. The confusion matrices of results in Experiment III

Actual class		Predicted class			
		Naïve Bayes		**J48**	
		Positive	Negative	Positive	Negative
Positive	2,534	1,907	627	1,811	723
Negative	3,700	781	2,919	830	2,870
Total	6,234	2,688	3,546	2,641	3,593

Table 8. The percentage of accuracy, precision and recall of Experiment III

Classifier		Accuracy	Precision	Recall
Naïve Bayes	Positive	77.41%	70.94%	75.26%
	Negative		82.32%	78.89%
J48	Positive	75.09%	68.57%	71.47%
	Negative		79.88%	77.57%

From three experiments, the results showed that the performance of Naïve Bayes classification (accuracy, precision and recall) is higher than those of J48 classification for the second and the third experiment. Moreover, the performance of Naïve Bayes classification with the feature set in the third experiment is the highest performance. Furthermore, all evaluation values of the Naïve Bayes classification are higher than those of decision tree technique. Therefore, this research selects the feature set in the third experiment and Naïve Bayes classification to generate classification models and to implement the opinion mining tool.

3. Results and Discussion

3.1 Implementation

This implemented software tool is an easy way to apply the opinion mining methodology for analyzing laptop reviews. The developers have designed the layout of the user interface for this tool as one page to make the software easier to use. The main screen consists of three areas: the menu bar, a middle text area and four bottom text areas as shown in Figure 10. The menu bar includes “Single Review” button for analyzing a review, “Multiple Review” button for analyzing reviews, text box for inputting a filtered word and drop-down list for selecting output types. The middle text area shows only subjective paragraphs in the review and four bottom text areas show the subjective paragraphs in each aspect domain as displayed in Figures 11 and 12.

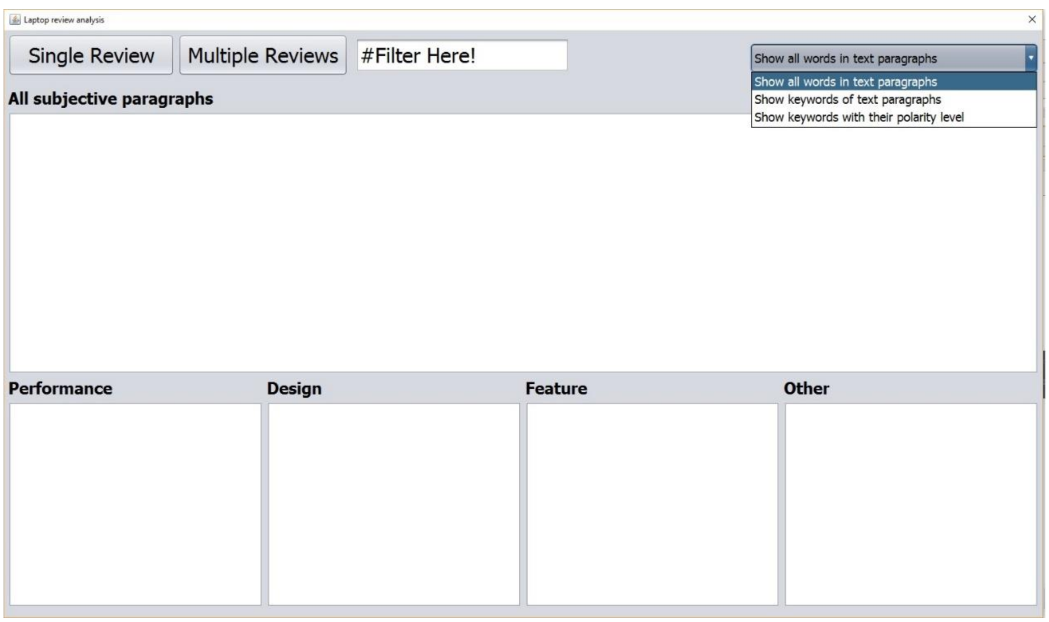

Figure 10. The main user interface of the proposed opinion mining tool

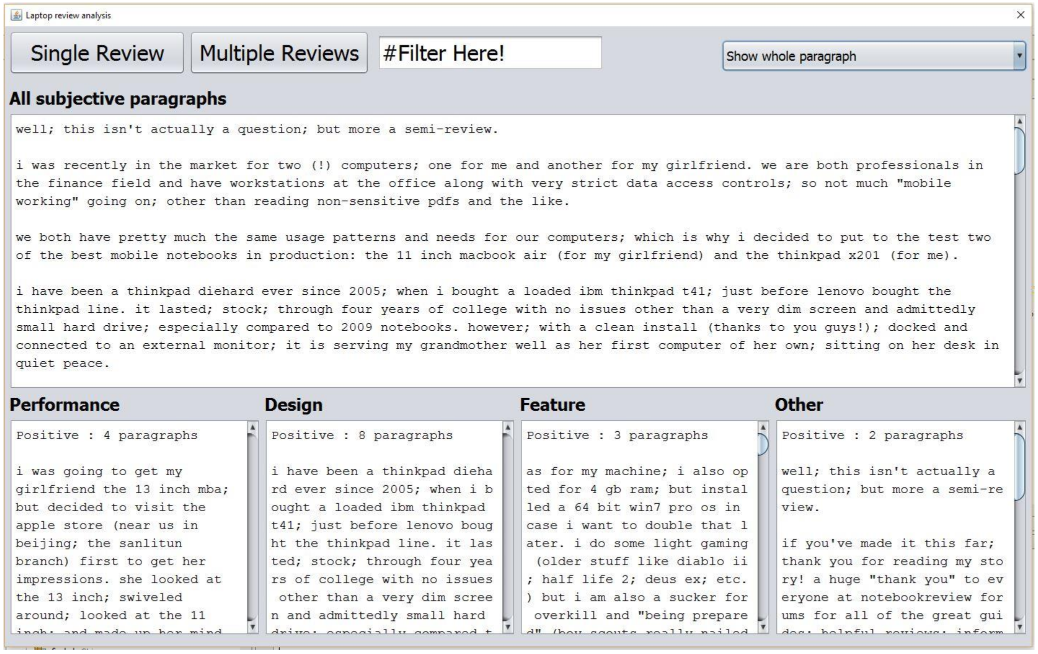

Figure 11. The output interface of the single review

Figure 12. The output interface of the multiple reviews

The main user interface is the first page that the users will see when they start the opinion mining tool. This page can be divided into 2 main functional parts. Single review, the single review button, is used to analyze only one text review in one text file (the output result is shown in Figure 11). Multiple reviews, the multiple reviews button is used to analyze more than one text reviews in one text file (the output result is shown in Figure 12).

Additionally, the output on the screen of this software implementing the opinion mining methodology can be divided into 3 main types: show all words in text paragraphs, show only the keywords of text paragraphs and show only the keywords with their polarity level of text paragraphs. The default output results show all words in paragraphs (Figures 11 and 12). The bottom text area shows all positive paragraphs and all negative paragraphs in the review separated by the aspect domains. For each aspect of text area, there are the total number of positive paragraphs, all positive text paragraphs, the total number of negative paragraphs and all negative text paragraphs, respectively.

The next output results (as shown in Figure 13) display the keywords of the text paragraphs. All text areas show only adjective words, adverb words and some words displaying the aspect of paragraphs instead of all words in paragraphs. These words are the keywords for detecting aspect and the keywords to generate feature sets for classifying sentiments. The last output results (as shown in Figure 14) demonstrate the keywords with their polarity levels. All adjective and adverb words with their polarity levels generated from their polarity scores are displayed, including words in each aspect domain. The keywords' polarity levels are the feature set of the sentiment classification by the machine learning.

Figure 13. The output interface showing only keywords

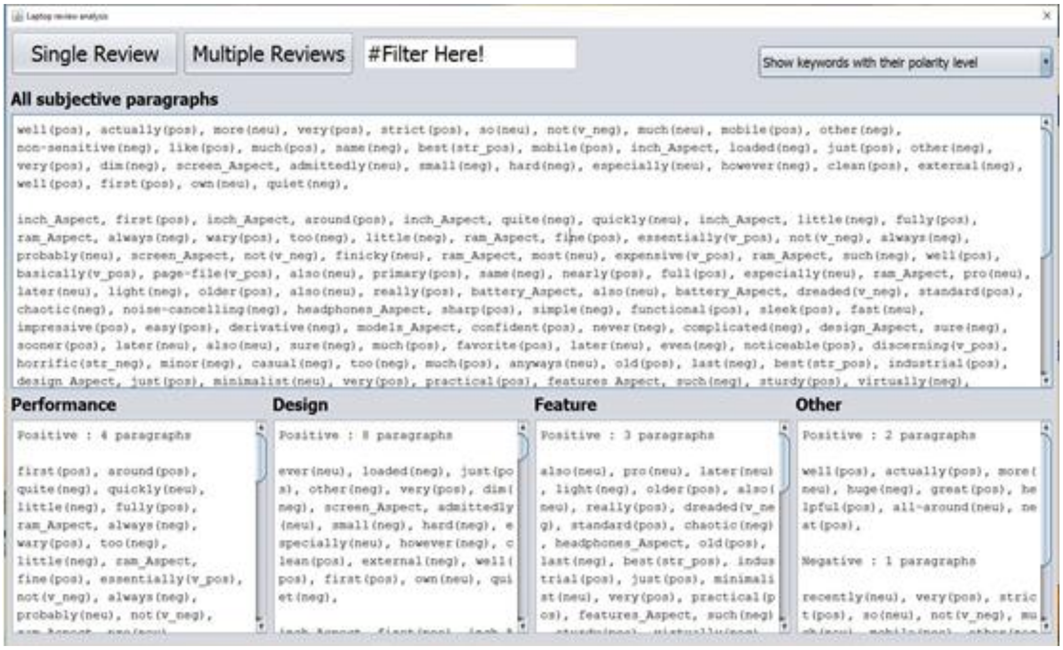

Figure 14. The output interface showing only keywords and their polarity level

3.2 Results and evaluation

In the results and evaluation section, there are 3 parts of the experiment in this research which are subjective detection, aspect identification, and sentiment classification. The results are explained in the form of confusion matrices, and all evaluations of three parts will be measured by calculating accuracy, precision and recall values from the confusion matrices.

3.2.1 Subjective detection

All selected review topics were separated into 15,384 paragraphs by detecting a new line. They can be separated into 6,399 subjective paragraphs and 8,985 objective paragraphs. Words in text paragraphs of the experimental data were compared with words in SentiWordNet and the modified emoticon lexicon to detect subjectivity. The result of the subjective detection is classified into 6,992 subjective paragraphs and 8,392 objective paragraphs. The confusion matrix of the result in the subjective detection is shown in Table 9 and the evaluation values are expressed in Table 10.

Table 9. The confusion matrix of subjective detection

Actual class		Predicted class	
		Subjective	Objective
Subjective	6,399	6,367	32
Objective	8,985	625	8,360
Total	15,384	6,992	8,392

Table 10. The percentage of accuracy, precision and recall of subjective detection

Actual class	Accuracy	Precision	Recall
Subjective	95.73%	91.06%	99.50%
Objective		99.62%	93.04%

Referring to Table 10, the accuracy, precision and recall rates are more than 90% for subjective paragraph detection. This means that the opinion mining tool can detect subjective review paragraphs effectively. The reason is that if there is at least one emoticon text or one subjective word in the paragraphs, these paragraphs will correctly be detected subjective paragraphs.

3.2.2 Aspect identification

Subjective paragraphs were identified into Performance, Design, Feature and Other aspects. The subjective paragraphs (6,992 paragraphs) can be divided into 1,347 performance paragraphs, 1,796 design paragraphs and 1,658 feature paragraphs by researchers reading categorizing manually. The confusion matrix of the result in the aspect identification is shown in Table 11.

Table 11. The confusion matrix of aspect identification

Aspect	Actual class	Predicted class	True positive	False positive	True negative	False negative
Performance	1,347	1,489	1,334	155	5,490	13
Design	1,796	2,150	1,737	413	4,783	59
Feature	1,658	1,811	1,614	197	5,137	44

The percentage of three evaluation values for the aspect identification is also shown in Table 12. The percentage of accuracy and recall rate of all aspect domains is greater than 90%, and the precision rate is about 80% or more. As a result, the aspect identification has high accuracy to identify the aspect of collected reviews in this research. This means that the generated aspect word lists are very useful for aspect identification.

Table 12. The percentage of accuracy, precision and recall of aspect identification

Actual class	Accuracy	Precision	Recall
Performance	97.60%	89.59%	99.03%
Design	93.25%	80.79%	96.71%
Feature	96.55%	89.12%	97.35%

3.2.3 Sentiment classification

All paragraphs in each aspect were classified the sentiments of contents by using the Naïve Bayes classifier. Only polarity levels of adjective and adverb words are training data in the sentiment classification. In this case, the objective paragraphs (625 of 6,992 paragraphs) and the neutral paragraphs (133 of 6,992 paragraphs) are removed from the training data, so there are only 6,234 paragraphs in the experiment. The confusion matrix and the percent of accuracy, precision and recall rate of the sentiment classification are displayed in Tables 13 and 14, respectively.

Table 13. The confusion matrix of sentiment classification

Actual class		**Predicted class**	
		Positive	Negative
Positive	2,534	1,907	627
Negative	3,700	781	2,919
Total	6,234	2,688	3,546

Table 14. The percentage of accuracy, precision and recall of sentiment classification

Actual class	**Accuracy**	**Precision**	**Recall**
Positive	77.41%	70.94%	75.26%
Negative		82.32%	78.89%

Referring to Table 14, the accuracy, precision and recall rates are more than 70% for sentiment classification. This means that the proposed methodology can classify the sentiment of subjective review paragraphs acceptably on collected reviews.

A major limitation of this tool is unseen words which are not included in the polarity lexicon, but may be found on laptop reviews. The reason is that these review analysis methods focus on words and the polarity of words to identify the aspect and to classify the sentiment of review paragraphs. If there are some missing input words from our lexicons, such as words with wrong spelling or technical words, the polarity level finding in the sentiment classification process cannot give the correct polarity level of words. Therefore, the performance of the sentiment classification process will be reduced by this error.

4. Conclusions

Nowadays, many laptops are manufactured with various features. When consumers decide to purchase a laptop, they normally search for laptop reviews in order to get the information first. Moreover, many reviews are created to let the consumers know more about each laptop. For those reasons, this research developed an opinion mining tool which helps users to know what is mentioned in the laptop reviews. The tool is separated into four main processes: to prepare data for analysis, to detect subjective text paragraphs, to identify the aspect of each text paragraph and to classify the sentiments of each text paragraph. The results of performance evaluation show that the subjective detection and the aspect identification has high accuracy and precision, including acceptably accurate and precise sentiment classification.

In conclusion, this opinion mining tool is useful for developing the review analysis system of laptops in order to help consumers gain information before purchasing a laptop. However, the user interface and feature of this tool will be improved in the future, such as data visualization and selected aspect comparison.

5. Acknowledgements

The first author was supported in part by a grant from the Thailand Research Fund and Office of the Higher Education Commission (Grant No. MRG6180250).

References

[1] Karamibekr, M., 2015. *A Sentiment Analysis Framework for Social Issues*. Ph.D., University of New Brunswick, Canada.

[2] Chatchaithanawat, T. and Pugsee, P., 2015. A framework for laptop review analysis. *Proceeding of the 2nd International Conference on Advanced Informatics: Concepts, Theory and Applications*, Chonburi, Thailand, August 19-22, 2015, 1-5.

[3] Pang, B., Lee, L. and Vaithyanathan, S., 2002. Thumbs up? Sentiment classification using machine learning techniques. *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing*, Philadelphia Pennsylvania, USA, July 6-7, 2002, 79-86.

[4] Turney, P.D., 2002. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, Philadelphia, Pennsylvania, USA, July 6-12, 2002, 417-424.

[5] Govindaraj, S. and Gopalakrishnan, K., 2016. Intensified sentiment analysis of customer product reviews using acoustic and textual features. *ETRI Journal*, 38(3), 494-501.

[6] Multi-perspective Question Answering, University of Pittsburgh, 2005. *The MPQA Opinion Corpus*. [online] Available at: http://mpqa.cs.pitt.edu/corpora/mpqa_corpus/

[7] Text Learning Group, 2010. *The SentiWordNet SentimentLlexicon.* [online] Available at: http://sentiwordnet.isti.cnr.it

[8] Baccianella, S., Esuli, A. and Sebastian F., 2010. SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. *Proceedings of the* 7th *International Conference on Language Resources and Evaluation*, Valletta, Malta, May 19-21, 2010, 2200-2204.

[9] Mitchell, T., 1997. *Machine Learning*, USA: McGraw-Hill Education -Europe.

[10] Valdivia, A., Luzon, M.V. and Herrera, F., 2017. Sentiment analysis in TripAdvisor. *IEEE Intelligent System*, 32(4), 72-77.

[11] Pugsee, P., Nussiri, V. and Kittirungruang, W., 2019. Opinion mining for skin care products on Twitter. *Communications in Computer and Information Science*, 937, 261-271.

[12] Pugsee, P., Sombatsri, P. and Juntiwakul, R., 2017. Satisfactory analysis for cosmetic product review comments. *Proceeding of 2017 International Conference on Data Mining, Communications and Information Technology*, Phuket, Thailand, May 25-27, 2017, 1-6.

[13] Stanford Natural Language Processing Group, Stanford University, 2008. *Basic English Stanford Tagger*. [online] Available at: https://nlp.stanford.edu/software/tagger.shtml

[14] Massachusetts Institute of Technology and Finlayson, M.A., 2013. *WordnetStemmer*. [online] Available at: https://projects.csail.mit.edu/jwi/api/edu/mit/jwi/morph/WordnetStemmer.html

[15] Yamamoto, Y., Kumamoto, T. and Nadamoto, A., 2014. Role of emoticons for multidimensional sentiment analysis of twitter. *Proceeding of the 16th International Conference on Information Integration and Web-based Applications & Services*, Hanoi, Viet Nam, December 4-6, 2014, 107-115.

[16] Anthony, L., 2019. *AntConc: A Freeware Corpus Analysis Toolkit for Concordance and Text Analysis*. [online] Available at: https://www.laurenceanthony.net/software

[17] Machine Learning Group, University of Waikato, 2019. *Weka: The Workbench for Machine Learning*. [online] Available at: https://www.cs.waikato.ac.nz/ml/weka/