COMMENT ANALYSIS FOR PRODUCT AND SERVICE SATISFACTION FROM THAI CUSTOMERS' REVIEW IN SOCIAL NETWORK

Todsanai Chumwatana
College of Information and Communication Technology
Rangsit University, Pathumthani, Thailand

todsanai.c@rsu.ac.th

ABSTRACT

In the last decade, the amount of social media usage has rapidly increased exponentially in Thailand. A huge amount of Thai online reviews and comments are available on social network every second. Because of this fact, comment analysis, also called sentiment analysis, has then become an essential task to analyze people's emotions, opinion, attitudes and sentiments from the amount of these online posts. This paper proposed the technique for analyzing Thai customers' comments or opinions about the products and services by counting the polarity words of the product and service domains. To demonstrate the proposed technique, experimental studies on analyzing Thai customers' comments in the social media are presented in this paper. The comments are classified into neutral, positive or negative. The proposed technique benefits the business domain in guiding product improvement and quality of service. Hence, this paper also benefits the end-users in making a smart decision.

Keywords: Comment analysis, sentiment analysis, opinion mining, social network analysis.

INTRODUCTION

At present, a significant number of websites, blogs, forums and social media allow people to provide their opinions about the services and products (Bagheri, 2013). Such media has been regarded as a tool for people to express their comments in a positive or negative way. The survey shows that 510 comments are posted every 60 seconds on Facebook and 4.75 billion pieces of content are shared daily. These online reviews open a new era of business intelligence and online marketing in today's world. The expressed words contained in customers' reviews can be regarded as the main factor to judge the customers' satisfaction. Surprisingly, Thailand has been ranked in the top ten highest number of social media users in the world. With this rapidly increasing number of Thai online customers' reviews available in the social media and websites, sentiment or comment analysis technique, also called opinion mining, has become an important task in the past few years (Siersdorfer, Chelaru, Penro, Altingovde, & Nejdl, 2014). This technique aims to analyze people's emotions, opinion, attitudes and sentiments from a large number of customer reviews which are available on the social media. Sentiment analysis is one of the recent research fields in the area of information extraction and text processing that provides many opportunities to develop the performance of business work by using the extracted reviews on the social websites. For instance, the comments about products and services proposed by customers can be computed into account aspects such as positive, negative or neutral opinions about these products or services. These comments will then be used to improve the service and product quality. In addition, the sentiment analysis also provides significant benefits for customers to gain valuable information about the products and services so that they are able to make smart decisions. In order to perform sentiment analysis, the classical approaches for opinion mining represent the reviews as bag-of-words as many words can be used to identify positive or negative feedbacks. This makes these methods work well with European language reviews which are segmented texts where words are clearly defined by word delimiters such as white space or other special symbols. These texts are explicitly segmented into word tokens, and then the word tokens are used as a bag-of-words to be parameters for the sentiment analysis process. However, these bags-of-words-based methods face problems with Thai customers' reviews which are non-segmented texts, since Thai texts are formed as a long sequence of characters without word boundaries. When applying the sentiment analysis to Thai customers' reviews, the reviews need to be parsed into individual words using the word segmentation technique before sentiment analysis can be performed. Up to now, not much research has been conducted on sentiment analysis for Thai customer reviews on social network. This is because the syntax of the Thai language is highly ambiguous and Thai language is non-segmented

(i.e. a text document is written continuously as a sequence of characters without explicit word boundary delimiters). Because of this problem, this paper proposes a sentiment analysis technique for Thai customers' reviews. The proposed technique is based on the integration of Thai word extraction and sentiment analysis techniques for mining Thai customers' opinions. Before the proposed technique is described in more detail, the segmentation problems of the Thai language are first discussed in the next section to clarify the problem.

SEGMENTATION PROBLEMS OF THAI LANGUAGE

Thai language is considered as non-segmented texts where the structure of writing is a string of symbols without explicit word boundary delimiters. Words in these languages are not naturally separated by any word delimiting symbols such as white spaces. The spaces in the Thai language are usually used to interrupt an idea or to help the reader pay attention to the text, but they do not signify a split between words, phrases or sentences (Jaruskulchai, 2003). Additionally, the Thai language has no capital letters to identify proper nouns or starting points of sentences like the English language. Figure 1 shows an example of the Thai language comment.

กำลังช่างใจคิดว่าถอย Nokia Lumia 520 มาเพื่อใช้งานเปิดเมล์

เปิดไฟล์เอกสาร ราคาไม่แพงด้วยแต่ปัญหาคือเราต้องพกมือถือ 3 เครื่องกันเลย

Figure 1. Example of the Thai language comment.

Due to this reason, Thai word segmentation is an important process that is firstly needed to apply to Thai texts for segmenting the words or terms before other techniques can be performed. However, this process is regarded as a challenging task because there is usually more than one way to insert word separators, which can lead to segmentation ambiguity. From the literature review, there are two main problems for segmenting Thai texts: segmentation ambiguity and unknown words (Jaruskulchai, 1996; Aroonmanakun, 2002).

Segmentation Ambiguity

This ambiguity happens when two or more phrases or sentences are spelled alike but have different meanings or pronunciation. Many Thai words are

compound words consisting of multi-syllables and two or more simple words, making it possible to have different segmentations for the same compound word. Figure 2 shows different ways of inserting word separators in terms of computer systems.

Thai and English texts	Different ways to insert word separators (Thai and English)	
'คนตากลมนอนตากลม'	คน|ตาก|ลม|นอน|ตาก|ลม	Hu-
	คน|ตาก|ลม|นอน|ตา|กลม	man|expose|wind|sleep|expose|wind
	คน|ตา|กลม|นอน|ตาก|ลม	Human|expose|wind|sleep|round|eye
'A round eye human sleep and expose wind'		Human|eye|round|sleep|expose|wind
	คน|ตา|กลม|นอน|ตา|กลม	**(Correct segmentation)**
		Human|eye|round|sleep|round|eye

Figure 2. Different ways to insert word separators in Thai texts.

Another problem is that most Thai words have a variety of meanings, which is known as polysemy. Polysemy refers to a set of words that are spelled alike but each belongs to a different grammatical category or meaning. For example, the word 'ตา' as shown in Figure 2 can mean 'eyes', 'a grandfather' or 'a turn'.

Unknown Word Problem

Another difficulty of segmentation is the existence of unknown words in the texts. This is a serious problem that occurs when the dictionary-based approach is used to perform segmentation (Jaruskulchai, 1996). Thai texts usually contain unknown words such as person or place names. This makes Thai word segmentation more complicated because the Thai language does not use capital letters to indicate proper nouns. The unknown words in the Thai language can be classified into six categories: proper noun, loan word, acronym, foreign word, mistype and official places. These words cause mistakes in segmenting Thai texts.

RELATED WORKS

Since the last decade, the competition in business industry has been very intense. Most businesses seek good strategies to gain more customers and

retain the customer base. This is because customer retention is the key to business success and growth. Businesses need to make customers satisfied with their products or services. A number of studies have shown that satisfaction is one of the critical indicators affecting repeat purchase behavior. To gain customers's satisfaction, businesses have to provide better service to match customers' needs. Therefore, knowing customers' needs and understanding what drives customers' satisfaction are the primary steps to conduct research. Many methodologies are used to collect customers' needs and analyze customer satisfaction. The majority of the traditional analysis methods use questionnaires and interviews. KANO's model is one of the well-known models applied to analyze customer needs in order to improve product and service quality in many business areas such as logistics customer service (Huiskonen & Pirttilä, 1998), ceramic product design (Tama, Azlia, & Hardiningtyas, 2015), and automotive design (Xu et al., 2009). However, using questionnaires and interviews is not suitable for sentiment analysis nowadays since most people tend to express their opinions, emotions, satisfaction and dissatisfaction via the social media. This makes text and data mining become an important task to analyse amounts of texts and data from the social media server and data warehouse (Chayanukro, Mahmuddin, & Husni, 2014; Samsudin, Puteh, & Hamdan, 2013) One of the essential techniques in this area is the sentiment analysis. This technique is used to analyses people is emotions and sentiment which has spread widely in many countries and languages. Many researches have proposed the techniques for analyzing comments, reviews or posts on the social media and communities. The results of these researches provide useful information for businesses and users to make the right decision. Opinion mining and sentiment analysis approaches have been reviewed and surveyed in the last half decade in order to find the best strategy (Ravi & Ravi, 2015).

To study sentiment analysis from the web service, Serrano-Guerrero presented sentiment analysis, a review and comparative analysis of web services (Serrano, 2015). This work provided reviews and compared some free access web services, analyzed their capabilities to classify and score different pieces of text with respect to the sentiments contained. In addition, a survey on sentiment detection of reviews was revealed by Cheng (Cheng, 2009). This survey discussed document sentiment classification and opinion extraction and some main approaches in this area.

In 2014, the application was implemented by applying the semantic information of words and decision tree classifier to mine customer's opinions about airline services (Pugsee, Chongvisuit & Na-Nakorn, 2014). The input data was collected from the opinions and comments from many contents on

Twitter. The outcome of this application helps both businesses and customers to make decisions to choose the best airline service among different airline brands.

Sentiment classification is another interesting research area that developed a lexicon-enhanced method to generate a set of sentiment words contained in corpus by using the word information from a sentiment lexicon (Dang, Zhang, & Chen, 2010). This corpus is categorized into five sets of online product reviews which will be used as classification models.

Furthermore, not only the English language, but the sentiment technique is also employed to non-English language. Some researches that work for other languages have also been proposed. In 2008, an empirical study of sentiment analysis for Chinese documents was proposed by Tan (Tan, 2008). This technique presents an empirical study of sentiment categorization on Chinese documents. The experimental results indicate that IG performs the best for sentimental terms selection. In the Czech social media, a supervised sentiment analysis was presented by Ivan Habernal (Habernal, 2014). This article described in-depth research on machine learning methods for sentiment analysis of the Czech social media. The author also established a common ground for further research by providing a large human-annotated Czech social media corpus. In this research, Facebook posts were used as input dataset. The results of this research were divided into three-class classifications: positive, negative and neutral. The system's measure was very low (less than 0.5) for negative posts, around 0.54 for neutral and around 0.75 for positive posts. The average score for Facebook posts was 0.69 which meant most of the posts just shared general discussion and positive comments.

RESEARCH METHODOLOGY

In this section, the proposed method that combines the Thai word extraction technique and the sentiment analysis technique is described. In this paper, the proposed method aims to analyses Thai customers' opinions from the comments on the social media using Thai word segmentation to extract Thai words and sentiment analysis to check customers' opinion, in order to classify customers' comments into three categories: neutral, positive and negative. Therefore, the proposed method consists of the following steps: collecting customers' comments from the social media, Thai word extraction, detecting polarity words, calculating polarity word score from the comment and classifying the comments into groups. The overview of the research process is shown in Figure 3.
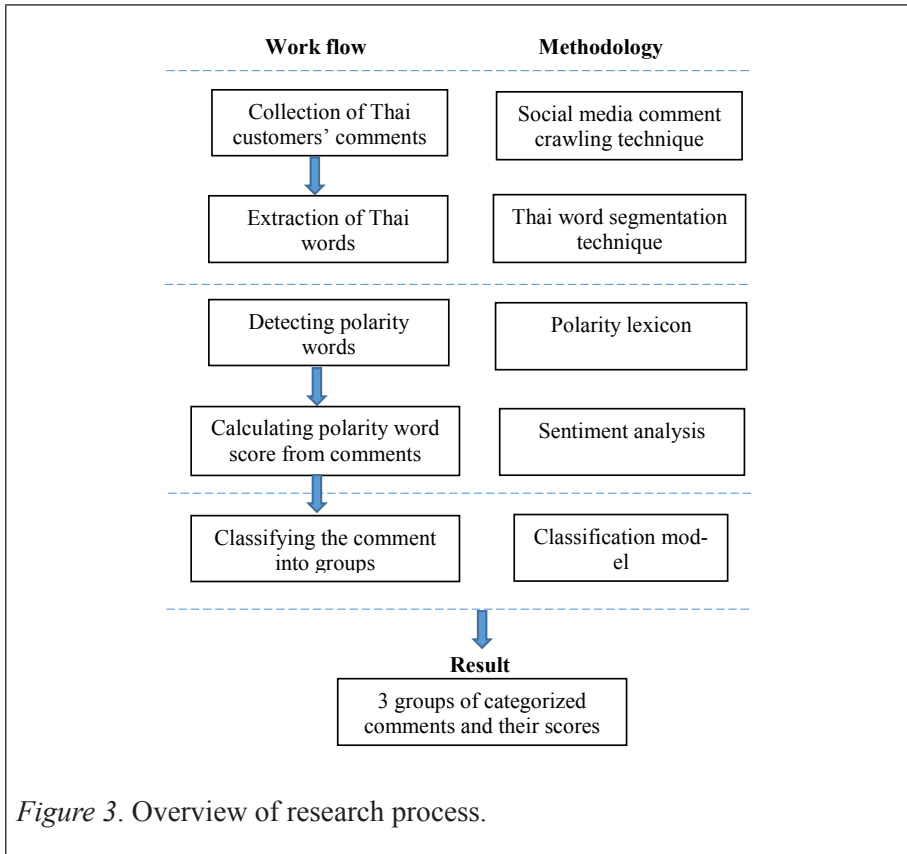
Figure 3. Overview of research process.

The overview of the research process of the proposed technique which consists of many processes is described below.

Collection of Thai Customers' Comments

The collection of Thai customers' comments is regarded as the first process for the proposed technique. To perform comment analysis, customers' comments or posts need to be collected from the social media websites. However, collecting a huge amount number of comments from the social media is very time-consuming and labor intensive if it is done manually. Technically, these comments can be collected by the RapidMiner tool which is on the Internet to collect the posts or comments as shown in Figure 4.

The Thai customers' comments collected from the social media are then used as an input to extract Thai words in the next step.
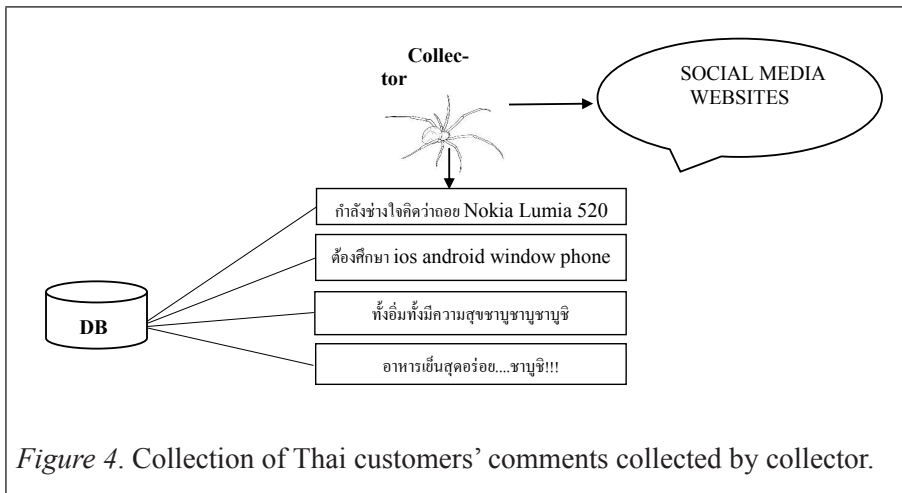
Figure 4. Collection of Thai customers' comments collected by collector.

Extraction of Thai Words

All Thai comments collected from the first step will be an input for word extraction. This process is necessary for the Thai language because of its non-segmented nature. Word extraction can be regarded as an important task that has to be performed before comment analysis can be performed. A number of techniques have been proposed to extract the words from Thai texts. One of the existing techniques is Thai word segmentation. The Thai word segmentation technique is the way to pass human knowledge to computers, so that computers can understand the Thai language like humans do. As a result, the computer can use this technique to extract Thai words before the next process. The Thai word segmentation technique can be mainly categorized into three approaches: Rule-based, Dictionar- based and Machine-learning-based approaches (Chumwatana, 2013). In the proposed technique, the machine-learning technique is employed to segment Thai customer's reviews as this technique is well known and most efficient (Haruechaiyasak, 2004; Kruengkrai, 2006). The machine-learning-based approach can be divided into four main algorithms: Naïve Bayes (Lewis, 1998), decision tree (Quinlan, 1986), Support Vector Machine (Vapnik, 1995), and Conditional Random Field (Haruechaiyasak, Kongyoung & Damrongrat, 2004; Kruengkrai & Isahara, 2006). Naïve Bayes is based on Bayes' Theorem in the probabilistic framework. The basic idea is to use the joint probabilities of words and categories to estimate probabilities of the categories that best describe the document. The Naïve Bayes algorithm makes the assumption of word independence, (i.e. the conditional probability of a word given a category, is assumed to be independent from the conditional probabilities of other words given that category. The decision tree technique is a simple structure where non-leaf nodes represent the conditional tests

278

of attributes or features and leaf nodes containing the class label. Another technique is the Support Vector Machine, also called SVM, which is based on machine-learning algorithm. This algorithm is based on the Structural Risk Minimization principle with the error-bound analysis. The method is defined over a vector space where the problem is to find a decision surface. In this paper, the Conditional Random Field algorithm was applied to Thai texts, which were shown to perform better than other machine-learning algorithms for the task of labeling and segmenting Thai texts. This technique uses the machine-learning technique to learn from a Thai text corpus. As learning is the essential part of this approach, it is necessary to have an appropriate word-segmented Thai text corpus, called the ORCHID corpus. The ORCHID corpus has consisted of 153,404 manually tagged words since 1996 (Haruechaiyasak & Kongyoung, 2008). Using a tagged corpus in which word boundaries are explicitly marked with special annotations, the machine-learning algorithm builds statistical models based on the features of the characters surrounding the boundaries. The most common features used for Thai word segmentation models are the identities and categories of characters within a defined n-gram of characters surrounding a word boundary candidate. Character types can be quite diagnostic when used for word segmentation. For example, certain leading vowels often appear at the beginning of a word, whereas tone-marking characters can never begin a word. In machin-learning approaches, a binary classification task is used to assign each character in the string. Each character in the Thai language can be tagged into one of two classes. The first class is the character that is placed at the beginning of the word, called word-beginning is be labeled as class 'B'. The second class is the character that is placed inside the word, called intra-word character that is labeled as class 'I'.

Table 1 shows an example of a string where each character is tagged with 'B' or 'I'. By using the tagged corpus in which word boundaries are explicitly marked with a special character, the Conditional Random Fields algorithm can be applied to train a model based on the features surrounding these boundaries (Haruechaiyasak, Kongyoung, & Damrongrat, 2004). For the Thai language, characters can be distinguished for segmentation tasks into ten different types as shown in Table 2.

The set of character types is designed based on Thai linguistic knowledge. Machine-learning algorithms can build classification models by extracting patterns of character types. For example, the Conditional Random Field algorithm could learn that the character which is a vowel by type is most likely to begin a word (i.e. class B). The final feature set for constructing a model is the n-gram of characters preceding and following the word boundary with their character types.

Table 1

Example of String of Characters Tagged as Word-Beginning (B) or Intra-word (I) Characters

Character position		Tag	Character position		Tag
ห	n	B	ย	c	I
ม	c	I	า	v	I
อ	c	I	ค	c	B
ร	c	B	น	c	I
้	t	I	ไ	w	I
ก	c	I	ข	c	I
			้	t	I

Table 2

Character Types for Building a Feature Set Learned by Conditional Random Field Algorithm

Tag	Type	Example
c	Characters that can be the final consonant in a word	ก ข ค ม ง จ ช ซ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป พ ฟ ภ ม ย ร ล ว ศ ษ ส ฬ อ
n	Characters that cannot be the final consonant in a word	ค ฉ ผ ฝ ฌ ห ฮ
w	Vowels that can begin a word	เ แ โ ใ ไ ฤ ฦ
v	Vowels that cannot begin a word	ะ า ิ ี ึ ื ุ ู ็ ำ
t	Tonal characters	่ ้ ๊ ๋
s	Symbols	์ ๆ ฯ
d	Digit characters	0-9
q	Quote characters	_
p	Space character inside a word	'-' '-'
o	Other characters	A-Z

The following table shows the example of segmenting customers' reviews by using the Conditional Random Field algorithm described above.

Table 3

Segmentation of: "มือถือยี่ห้อนี้ค้างบ่อยและก็แย่ด้วย" by using Conditional Random Field Algorithm

| Customer's review	มือถือยี่ห้อนี้ค้างบ่อยและก็ไม่ดีด้วย
Translate Customer's review	This mobile phone brand is always hanged and it is terrible as well.
Thai word segmentation	มือถือ-ยี่ห้อ-นี้-ค้าง-บ่อย-และ-ก็-แย่-ด้วย

Detecting Polarity Words

After the Thai words are segmented and extracted from the comments, the next process is detecting polarity words. To detect polarity words, this paper applied the polarity lexicon based on the SentiWordNet which was reviewed by the experts to identify polarity scores (positive, negative or neutral) (Esuli & Sebastiani, 2006). These polarity words were then assigned to each segmented word from the customers' comments in order to detect and analyse their emotion. This was because the polarity words can reveal customers' emotions, opinions, attitudes and sentiments as well as their satisfaction. In this research, polarity words in lexicon were be represented in 3 levels: (a) Positive represented by +1, (b) Neutral represented by 0, and (c) Negative represented by -1 as shown in Table 4 below.

Table 4

Polarity Lexicon

Positive (+1)	Neutral (0)	Negative (-1)
เจ๋ง (Cool)	เฉย ๆ (so so)	แย่ (Terrible)
เยี่ยม (Great)	ธรรมดา (Normal)	น่าเกลียด (Awful)
มหัศจรรย์ (Amazing)	โอเค (OK)	เลวร้าย (Bad)
สวย (Beautiful)	ปกติ (Usual)	น่าผิดหวัง (Disappoint)
ดีที่สุด (Best)		เกลียด (Hate)
ดีกว่า (Better)		ปัญหา (Problem)
ยอด (Awesome)		ค้าง (Hang)

To make it easier to understand, an example of first three processes of the proposed technique is revealed in Figure 5. The words in **Bold** are indicated as polarity words.

Figure 5. Example of the first three processes of the proposed technique.
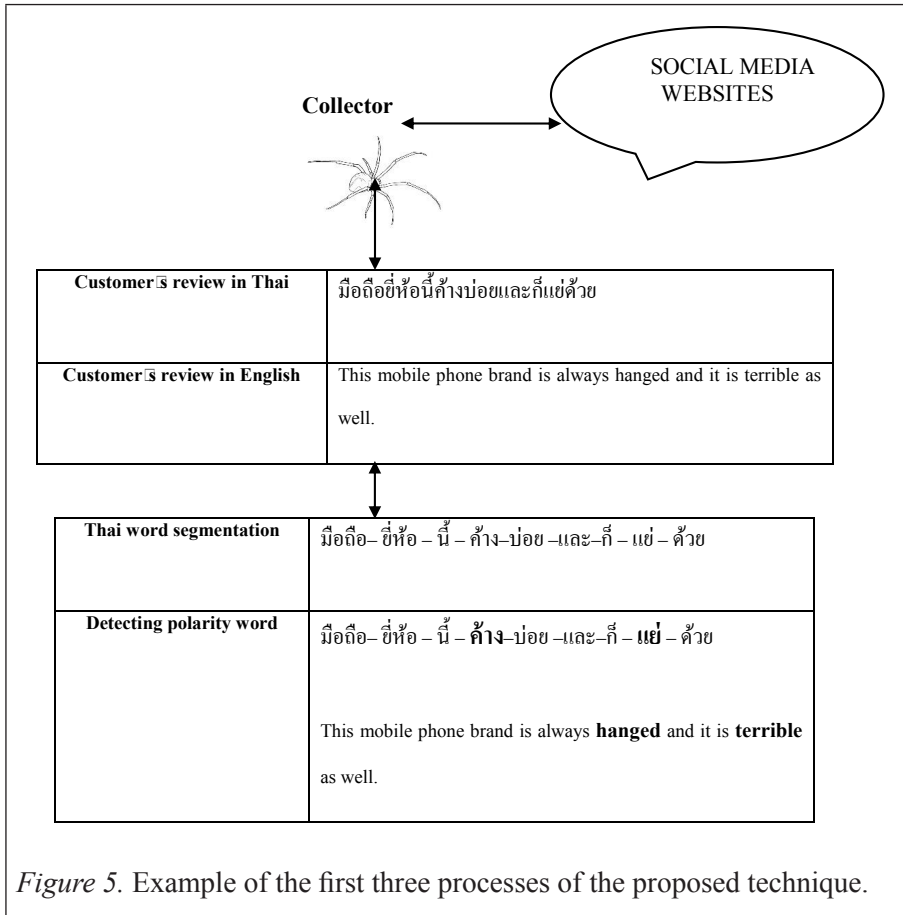
Calculating Polarity Word Score from Comments

In calculating the polarity scores of the comment, the summation of all polarity word scores in each comment is calculated. If the polarity score of the comment is more than zero, this comment is classified to be a positive comment. If the polarity score of the comment is zero, the comment is classified to be a neutral comment. On the other hand, the comment will be classified as negative if the summation of the polarity scores is less than zero. To illustrate the process of calculating the polarity word score, Table 5 shows the summation of the polarity scores for the comments. In Table 5, the words in **Bold** stand for negative words, the words in *Italic*

stand for neutral words while the words <u>Underlined</u> stand for positive words attached with the summation of the polarity score for the comments.

Table 5

The Summation of Polarity Scores for the Comments

Comment	Score	Source
จะ-ว่า-ไป-พัก-หลังๆ-โนเกีย-ทำ-สี-ตัว-เครื่อง-*เฉย ๆ* -แล้ว-อ่ะ -แถม-ยัง-**ค้าง**-บ่อย-และ-ก็-**แย่**ด้วย	-2	Facebook
กำลัง-ช่าง-ใจ- คิด-ว่า-ถอย-Nokia- Lumia –520-มา-เพื่อ-ใช้-งาน-เปิด-เมล์-เปิด-ไฟล์-เอกสาร-แต่-ปัญหา-คือ-เรา-ต้อง-พก-มือถือ-3-เครื่อง-กัน-เลย !!	0	Twitter
เพื่อน-เรา-ใช้- Android- มา-นาน-มาก-หลาย-ปี - บอก-ว่า-<u>ดี</u>-นะ - ใช้-ง่าย -ราคา-<u>ถูก</u>-แต่ พอ-เปลี่ยน-มา-ใช้- Lumia - 530 - ได้-แปปเดียว-ต้อง-กลับ-ไป-ซื้อ –Android -ใหม่-มัน บอก-**ห่วย**-มาก	+1	Twitter

Table 5 shows that the first comment is negative because it consists of one neutral word represented by 0, and two negative words represented by -1. Consequently, the summation of the polarity score of the first comment is equal to 0 - 1 - 1 which is -2, while the second comment does not consist any polarity words so this comment is neutral. The last comment consists of one negative word, and two positive words so its polarity score is equal to + 1 + 1 - 1 which is +1. As a result, the last comment can be expected to be positive.

Classifying the Comment into the Group

In the final process, all comments identified by their scores from the previous step will be separated into three groups based on their levels by using the classification technique. In this research, the sentiment analysis was employed to label each comment to the different groups of similar comments. The groups of comments can be divided into three: positive, negative and neutral. The positive group consists all comments which have a polarity score of more than 0. Meanwhile, the negative group contains all negative comments which have a score of less than 0, and the neutral group has only comments scored 0. In Figure 6, classifying the comments into the three groups using sentiment analysis is depicted.
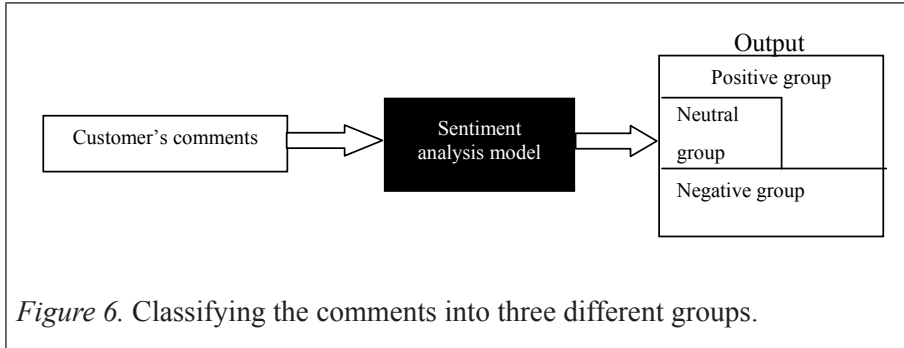
Figure 6. Classifying the comments into three different groups.

Based on all the above processes, the results of the proposed technique can be used to educate people involved in business. The numerous feedbacks from customers together with the polarity level can help businesses to improve their product quality and services in order to be able to compete with business competitors. In addition, the results of this research also provide significant benefits for users to gain information and then make smart decisions to pick the right product or services among competitive brands.

EXPERIMENTAL RESULTS

In this section, an experiment for analyzing customers' reviews based on the proposed method is presented. 3348 reviews were used as input in the experimental study. All reviews used were collected from Facebook and Twitter websites that consisted of 7 different categories: Nokia Thailand, Nokia X, Overall Mobile, Nescafe Red Cup Machine, MK Restaurant, Shabushi Restaurant and JBL Speaker as shown in Table 6.

Table 6

Experimental Categories

Category	No. of reviews	Channel
Nokia Thailand	1939	Twitter
Nokia X	91	Twitter
Overall Mobile	7	Twitter
Nescafe Red Cup Machine	112	Twitter
MK Restaurant	668	Twitter
Shabushi Restaurant	199	Facebook
JBL Speaker	332	Facebook

Table 7

Group of Customers' Comments Based on Polarity Scores

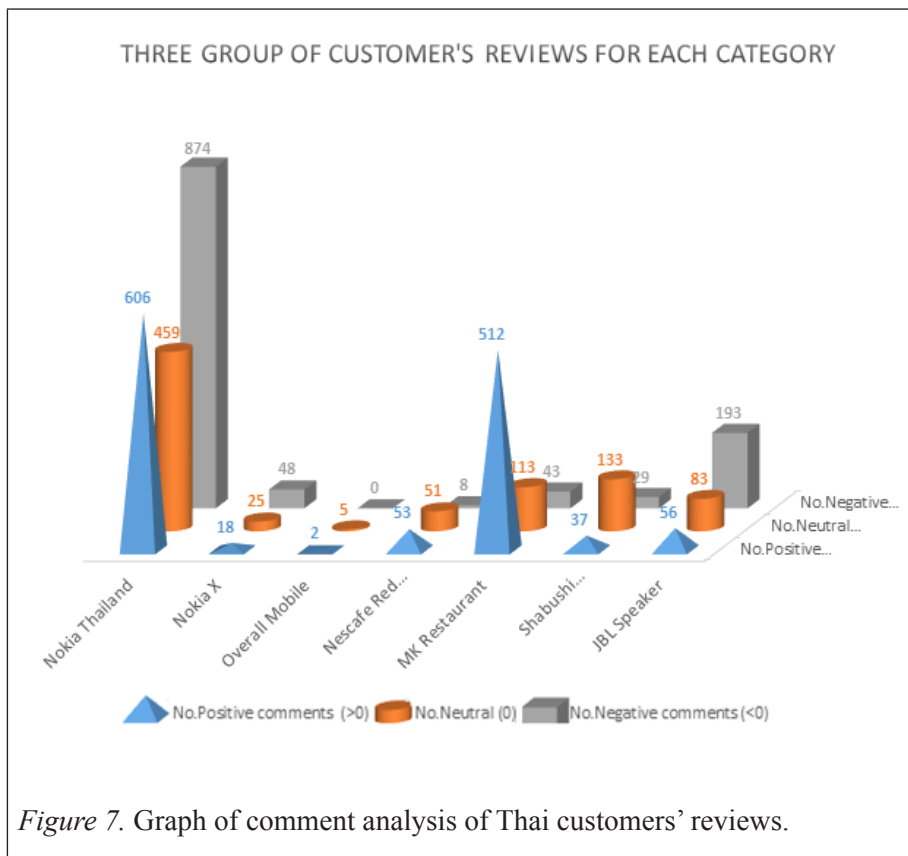| Category	Positive opinion (>0)	No opinion (0)	Negative opinion (<0)
Nokia Thailand	606	459	874
Nokia X	18	25	48
Overall Mobile	2	5	0
Nescafe Red Cup Machine	53	51	8
MK Restaurant	512	113	43
Shabushi Restaurant	37	133	29
JBL Speaker	56	83	193

Figure 7. Graph of comment analysis of Thai customers' reviews.

In order to determine whether the proposed method was appropriate for comment analysing Thai customers' reviews, the experimental study was needed. In this experiment, the comments were first segmented and extracted using the Thai word segmentation technique. Next, detecting polarity words and calculating the polarity word score from the comments were then applied to the resultant set of extracted words. Finally, sentiment analysis was employed to label each comment to three groups of similar comments: positive, negative and neutral based on their polarity scores. Based on the experimental study, customers' negative, positive and neutral opinions are revealed in Table 7 and the resultant graph shown in Figure 7.

The graph of the experiment result in Figure 7 shows that each category provides different results depending on their polarity scores. As shown in Table 7, more Thai customers have negative opinions of Nokia Thailand, Nokia X and JBL speaker compared to the number of customers who have positive opinions while most Thai people just talk about general things for the Overall Mobile category. In contrast, many people provided good reviews for 3 categories: Nescafe Red Cup Machine, MK and Shabushi Restaurants. These results would help the company to know about the strengths and weaknesses of their products and services in order to improve their qualities. Furthermore, this proposed comment analysis method is an advantage in decision-making for customers to pick quality products and services.

CONCLUSION

Today, numerous customer reviews are available on the Internet. Websites, blogs, forums and the social media are the main channels which people use to provide their opinions about the services and products. However, it is a challenging task to manually monitor such a huge number of comments as it is time-consuming and labor-intensive. This makes sentiment analysis an important task to analyse customers' satisfaction. However, unlike the English language, applying the sentiment analysis technique to Thai reviews requires pre-processing because the language is of non-segmented texts. Due to this problem, this paper proposed the comment analysis technique for Thai customers' reviews. The proposed method consists of five main steps: **(1)** collecting customer's comments from the social media, **(2)** word extraction, **(3)** detecting polarity words, **(4)** calculating polarity word score and **(5)** classifying the comment. The experiment results show that different categories provide different results depending on the brand and quality. All customer comments of each category are classified into three different groups: positive, negative and neutral. In the groups of Nokia Thailand, Nokia X and

JBL speaker, more Thai customers have negative opinions than the number of customers who have positive opinions. Meanwhile, most people just share a general discussion without any positive or negative opinion for the Overall Mobile category. As for the Nescafe Red Cup Machine, MK and Shabushi Restaurants groups, many people provided good opinions on these three categories. This reveals the customers' opinions about the products and services. Consequently, the results of the proposed technique would be of significant benefit to business to open the way for improvement of their products and service quality. The technique also provides opportunity for users to gain information in order to make smart decisions to choose the best product or service from among competitive brands.

REFERENCES

Aroonmanakun, W. (2002). Collocation and Thai word segmentation. *Proceedings of the 5th SNLP & 5ᵗʰ Oriental COCOSDA Workshop*, 68-75.

Bagheri, A., Saraee, M., Jong, F.D. (2013). Care more about customers: Unsupervised domain-independent aspect detection for sentiment analysis of customer reviews. *Knowledge-Based Systems Journal, 52,* 201-213.

Chayanukro, S., Mahmuddin, M., & Husni, H. (2014). A generalized e-learning usage behaviour model by data mining technique. *Journal of Information and Communication Technology, 13,* 37-53.

Cheng, X. (2009). A survey on sentiment detection of reviews. *Expert Systems with Applications*, *36*(7).

Chumwatana, T. (2013). A survey of automatic indexing techniques for thai text documents. *Information Technology Journal, 17.* King Mongkut's University of Technology North Bangkok. Thailand,

Chumwatana, T. (2013). A survey of automatic word segmentation techniques for thai text documents. *Rangsit Journal of Information Technology (RJIT), 1*(2), 36 – 43.

Dang, Y., Zhang, Y., Chen, H. (2010). A lexicon-enhanced method for sentiment classification: An experiment on online product reviews. *IEEE Intelligent System, 25(4)*, 46-53.

Esuli, A., Sebastiani, F. (2006). SENTIWORDNET: A publicly available lexical resource for opinion mining. *Proceeding of LREC'06*, 417-422.

Habernal, I., Ptáček, T., Steinberger, J. (2014). Supervised sentiment analysis in Czech social media. *Information Processing & Management 50693-707*. Online publication.

Haruechaiyasak, C., Kongyoung, S., Dailey, M.N. (2008). A comparative study on thai word segmentation approaches. *Proceedings of Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology*.

Haruechaiyasak, S., Kongyoung, S., Damrongrat, C. (2004). LearnLexTo: A machine-learning based word segmentation for indexing Thai texts. *ACM 17th Conference on Information and Knowledge Management.*

Huiskonen, J., & Pirttilä, T. (1998). Sharpening logistics customer service strategy planning by applying Kano's quality element classification. *International Journal of Production Economics, 56–57*, 253-260.

Jaruskulchai, C. (1996). *Thai text segmentation: problems and potential solutions*. The Sixth Annual Workshop on Science and Technology Exchange between Thai Professionals in North America and Thailand. Edmonton, Alberta, Canada.

Jaruskulchai, C., Kruengkrai C. (2003). A practical text summarizer by paragraph extraction for Thai. *Proceedings of the Sixth International Workshop on Information Retrieval with Asian Languages*. Sapporo. Japan, 9-16.

Kruengkrai, C., Isahara, H. (2006). A conditional random field framework for Thai morphological analysis. *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC-2006)*.

Lewis, D. (1998). Naive (Bayes) at forty: The independence assumption in information retrieval. *Proceeding of 10th European Conference on Machine Learning*. 4-15.

Pugsee, P., Chongvisuit, T., Na Nakorn, K. (2014). Subjectivity analysis for airline services from twitter. *Proceeding of ITC-CSCC*, 944-947.

Quinlan, J.R. (1986). Induction of decision trees. *Machine Learning*, *1*(1), 81-106.

Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. Knowledge-Based Systems, *89*, 14-46.

Samsudin, N., Puteh, M., Hamdan, R. A. (2013). Normalization of noisy texts in Malaysian online reviews. *Journal of Information and Communication Technology*, *12*, 147-159.

Siersdorfer, S., Chelaru, S. Penro, J., Altingovde, I. S., Nejdl, W. (2014). Analyzing and mining comments and comment rating on the social web. *ACM Trans. Web., 5(10),* 17-39.

Serrano, G, J., Olivas, J. A., Romero P., Herrera-Viedma, E. (2015). Sentiment analysis: A review and comparative analysis of web service. *Information Science, 311,* 18-38.

Tama, I. P., Azlia, W., & Hardiningtyas, D. (2015). Development of customer oriented product design using Kansei engineering and Kano model: Case study of ceramic souvenir. *Procedia Manufacturing, 4*, 328-335.

Tan, S., Zhang, J. (2008). An empirical study of sentiment analysis for Chinese documents. *Expert Systems with Applications, 34(4),* 2622-2629.

Vapnik, V. (1995). *The nature of statistical learning theory.* New York: Springer.

Xu, Q., Jiao, R. J., Yang, X., Helander, M., Khalid, H. M., & Opperud, A. (2009). An analytical Kano model for customer need analysis. *Design Studies, 30*(1), 87-110.