Indian Journal of Science and Technology, *Vol 12(13), DOI: 10.17485/ijst/2019/v12i13/142833, April 2019*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Sentiment Analysis of Thai Sounds in Social Media Videos by using Support Vector Machine

Sumitra Nuanmeesri

Suan Sunandha Rajabhat University, Bangkok, Thailand; Sumitra.nu@ssru.ac.th

Abstract

Objective: To examine the sounds in Thai videos on YouTube, to analyze consumer opinion on beauty products. **Statistical Analysis:** The data was collected to analyze the sentiments of Thai sounds from 500 YouTube videos which were the reviews of beauty products; the length of each video being approximately 2-5 minutes. **Findings:** The accuracy rate of SVM (SVM) appears greater than those from the Naïve Bayes (NB) and K-Nearest Neighbor (KNN) techniques. The SVM used the RBF Kernel-typed Sequential Minimal Optimization (SMO) function, where c=50000 and gamma=0.1; the accuracy rate was 94.40%, when using K-fold Cross Validation, where K =10, that had 293 attributes. **Application/Improvement:** The SVM used the RBF Kernel-typed SMO function, where c=50000 and gamma=0.1*; it can be applied to* analysis of sentiments studies in social media derived from Thai videos which are evaluate processes need to be fast and able to provide negative, neutral, or positive results in a timely manner, in which it will become a purchase decision-making guide for consumers.

Keywords: Sentiments Analysis, Social Media, Support Vector Machine (SVM), Thai Sound, Video

1. Introduction

Social media is a prevalent communication channel for exchanging information, opinions, attitudes, and so on. Users can share information in different forms such as messages, pictures, sound audios and videos. Customers are also able to instantly share their impressions and experiences of a product or service via social media which serves as an intermediary to rapidly distribute information to others by using common electronic devices including smartphones, tablets, and computers, in the use of popular social media applications, including Facebook, Twitter, and YouTube[1]. It is common for people to provide an opinion, or product review, in the form of a video on YouTube with the video that customers share from their experiences being regarded as an important piece of information for potential customers. A product or service review video of a customer who has used the product or service before can become a highly influential piece of information in relation to consumer decision-making.

Sentiment analysis uses Natural Language Processing (NPL) to investigate and analyze opinion. Presently, there are a number of researchers that analyzed sentiments in communications (including sounds and images embedded in online videos). Most Thai sentiment analysis researchers however, mainly use texts written or posted about interesting topics such as products, services, and politics[2].

Analysis of social media sentiment relies on machine learning classifiers such as the Naïve Bayes, Decision Tree, Maximum Entropy, and Support Vector Machine (SVM). Lima and Castro (2013) used the Naïve Bayes (NB) technique to develop the sentiment analysis system using data gained from Twitter text messages about Brazilian TV channels, which were determined to have an accuracy rate of 91.94%[3]. In adopted the NB technique to analyze information regarding beauty and pharmaceutical products as obtained from Facebook and Twitter. Data was classified depending on whether it was positive or negative in nature. Research results indicated that the NB technique was 83% accurate when compared to the specialist classification technique[4] demonstrated the sentiment analysis process with the NB technique based on the audio information from YouTube videos. The accuracy rate was calculated to be as high as 87.70%[5].

The effectiveness of different sentiment analytic techniques compared classification parameters of the Naïve Bayes, Decision Tree, Maximum Entropy and SVM methods. In used a model to analyze customer satisfaction based on recommendation messages obtained from opinion mining by comparing the Decision Tree and NB models. These findings revealed that the Decision Tree technique was more accurate at 95.50%[6]. In also did a comparison of sentiment classifications in the Tunisian language as seen on Facebook when using the NB and SVM techniques. The researchers used the same set of data to experiment in which 60% of it was classified into information for trial, and the remaining 40% was separated for testing. The results indicated that the efficiency of the SVM technique was more accurate when compared to the NB technique[7]. In the same manner, classified English language sentiments by extracting each sound from YouTube videos. They also developed the simulative sentiment model and the automatic speech recognition system by using opinion-based data such as the reviews of products, movies, and social issues from both male and female reviewers. The researchers used the Maximum Entropy technique to create the simulative sentiment model. Additionally, they also investigated the system with the NB, SVM, and Maximum Entropy techniques. When comparing the accuracy, between the simulative sentiment model and the standard technique; the simulative model was more accurate at 80%[8]. In developed the sentiment estimation process based on Thai comments, where 6,000 comments from news, entertainment, and product review websites. The specialists classified the comments into 6 groups of feelings (love, joy, surprise, sadness, fear, and anger; each group consisting of 1,000 opinions). Then, they compared the technical performances between the Naïve Bayes, SVM, and Decision tree. The findings revealed that the SVM technique had the highest accuracy rate with 69.15%[9] classified emotions towards the Hurricane Sandy hashtag on Twitter into 4 types consisting of positive, anger, fear, and other. The SVM technique was 75.90% accurate, while the NB technique was determined to be 69.10% accurate[10]. In classified emotions towards video clip sharing on the YouTube web board. The information consisted of 2,771 Thai Music Videos (MV) and 3,077 Thai commercial Advertisement Videos (AD). This information was classified according to Ekman's six basic emotions theory (anger, disgust, fear, happiness, sadness, and surprise), to compare the Multinomial Naïve Bayes (MNB), Decision Tree, and SVM techniques. Sentiment classification of the comments in the AD group using SVM had the highest accuracy with 76.14%, while the sentiment classification in the MV group using the MNB technique had the highest accuracy score with 84.48%[11].

According to the previous studies about the sentiment analysis of the texts and videos on social media, it can be concluded that the most effective sentiment analysis was via the SVM. The SVM is the process to Supervise Learning to create a classifier for categorizing an unknown dataset. It is used for finding a decision plane to separate data into two groups, by building a midline between two groups for widening the scopes so they are as distant as possible. The SVM is usually applied for high-dimensional data. It is an artificial neural network that copies neuron characteristics to classify an input space in the form of a high-dimensional dataset in the feature space by using the Kernel Function to adjust the form of the data. The above-mentioned ability will be used for facilitating the development of a data classifier with the quadratic equation in the feature space and verifying the data classification processes. The appropriate midline is called the optimal separating hyper plane.

To illustrate, (x_i, y_i)....,(x_n, y_n) is set to be to be a sample in the teaching demonstration. 'n' is the number of samples. m is the number of input space dimensions[12]. 'y' is the result which is set to be +1 or -1, as shown in formulae.

$$(x_i,y_i),\ldots,(x_n,y_n) \text{ when } x \in R^m,\ y \in \{+1, -1\} \tag{1}$$

In terms of the linear problem, the high-dimensional data is separated into two groups[12]. The decision plane is calculated according to formulae (2).

$$(w{*}x)+b=0 \tag{2}$$

When w is weight value, and b is bias value[12], it will be used for the calculation according to formulae (3) and (4) in order to classify data.

$$(w{*}x)+b>0 \ \text{ if } \ y_i=+1 \tag{3}$$

$$(w{*}x)+b<0 \ \text{ if } \ y_i=-1 \tag{4}$$

Using the SVM technique in measuring sentiment classified Indonesian social media sentiment into 4 groups, (positive, negative, neutral and questions). Findings based on the SVM technique were reported to be 83.50% accurate[13]. In applied the SVM technique when classifying basic sentiments with a 96.43% accuracy[14]. In categorized YouTube Spanish video sentiments regarding general topics, where SVM was used to classify texts, snds, and videos with a 75% accuracy[15].

It is evident that there is a gap in sentiment analysis applications of the contemporary Thai social media. A positive or negative comment can directly influence a consumer's purchase decision. As such, this research aims to examine the sounds in Thai videos on YouTube, to analyze consumer opinion on beauty products. In this study, the researcher used the audio extraction, Automatic Speech Recognition (ASR), word segmentation, database comparison, sentiment analysis, word adjustment, document representation, and modeling methods. The processed data was classified by the SVM technique to analyze negative, neutral, or positive opinion, for the development and implementation of purchase decision-making guidenes.

The rest of this article is structured as follows: Section 2 illustrates the research methodology of the sentiment analysis of the sounds in Thai social media videos by using the support vector machine. Section 3 presents the experimental results compared by the NB learning technique, K-Nearest Neighbor (KNN) learning technique, and SVM (SVM). This part also explains the parameter setting and non-adjusted documents and adjusted documents. Finally, Section 4 consists of the discussion, conclusions and recommendations for future work.

2. Materials and Methods

To analyze the Thai sounds gained from social media videos by using the sentiment analysis which applies the Support Vector Machine, the researcher had adopted the following research methodology.

2.1 Data Collection

Data was collected to analyze the sentiments of Thai sounds from 500 YouTube videos which were the reviews of beauty products; the length of each video being approximately 2-5 minutes.

2.2 The Audio Extraction and Automatic Speech Recognition (ASR)

The 500 Thai review videos were extracted by Web speech API which worked with online and recorded videos. The recorded videos can assist in filtering the dataset for better quality. These videos were checked for correctness in the Thai language, extracted as texts, and then converted into .txt files. Fifty Thai language specialists (who had Thai listening, speaking, reading, and writing skills at a high level) were employed in this process. A 10-fold cross validation was used where each specialist investigated 10 different videos, as shown in Figure 1. If errors were found, the texts were rectified prior to the word segmentation stage.

Figure 1. Checking text from using web speech API.

2.3 Word Segmentation

The researcher segmented Thai words from the extracted text were saved in the .txt file by using ThSplitLib, a dictionary-based word segmentation program. The Longest Word Pattern Matching technique for was used to increase vocabularies, to analyze the sentiments of the words that may have not occurred in the dictionary. The .txt-file extracted texts were then converted into PHP language, and applied to the word segmentation function. The ThSplitLib supports the normalization of words in social media, for example, the phrase 'It is Goooood!' will be changed into 'It is good!' as shown in Figure 2 before moving onto the step of comparing databases.

2.4 Database Comparison

To substitute words with the Vector Space Model (VSM) technique, the Term Frequency (TF) gained from the data retrieval for word substitution was used. This research uses data retrieval to search for emotional adjectives by comparing the results of word segmentation as collected

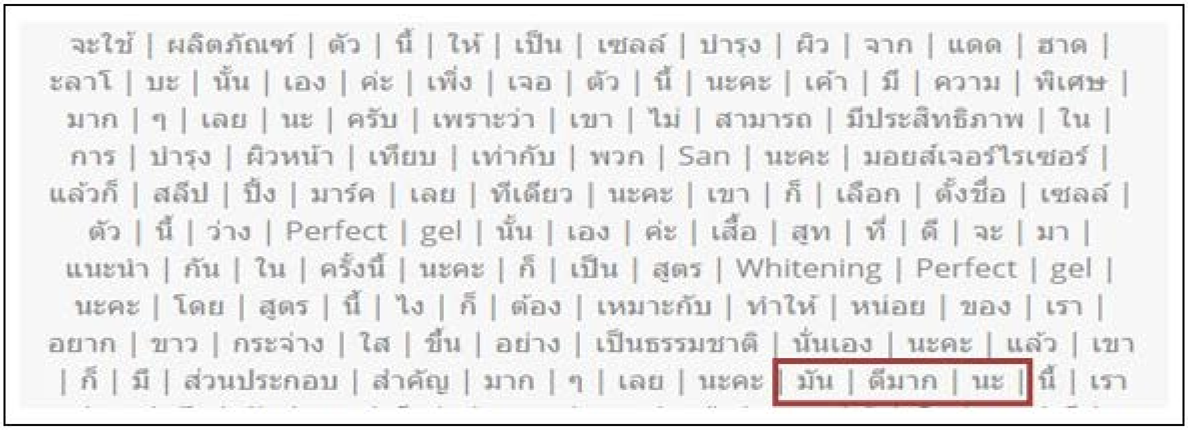

Figure 2. A result of word segmentation.

in the database. the results from the compared database where data was classified into four main groups, being: 1. Positive group showing the total frequency of retrieved positive adjectives with word frequency; 2. Negative group showing the total frequency of retrieved negative adjectives with word frequency; 3. Neutral group with the total frequency of retrieved neutral adjectives with word frequency; and 4. Redundant words or words that do not belong to any group of negative, neutral, and positive sentiments.

2.5 Sentiment Analysis

The positive-negative-neutral sentiment analysis process is a step to illustrating the results after comparing the words with the adjectives database by determining the most frequent terms from negative, neutral, and positive groups. It also used the weight of each word to calculate TF and IDF values, to determine term frequency in the documents. The positive sentiment analysis results can be calculated by using the total frequency. For example, if the positive frequency is equal to 8 words which are higher than the negative frequency, while the negative frequency is equal to 0 words, the result in the Class column is P (Positive).

2.6 Word Adjustment

Presently, Thai words have been developed and increased for fulfilling utilization purposes and providing sufficient words to communicate on social media. For an accurate sentiment analysis, it is necessary to apply word adjustment by combining the most used adjectives from the results with other words in the texts. The words were combined in this pattern: one for a preceding word and another for a following word. The researchers could decide whether the result of a combined word was an adjective in negative, neutral, positive, redundant, or unidentifiable groups. The results from the word adjustment process also improved the sentiment text database. The combined words that were selected as the adjectives of negative, neutral or positive groups will be added into the sentiment text database. This was expected to increase the effectiveness of sentiment analysis.

The word adjustment process from the positive sentiment analysis utilized term frequency with 8 words consisting of 'compact', 'strength', 'slow', 'safe', 'restored', 'precious', and 'valuable'. The word 'slow' had the highest term frequency (taking up two words), so combined with a preceding word and a following word, or 4 Thai words in total; the phrases consisted of 'which slow…', 'slow … (something)', 'help to slow', and 'slow down aging'. Then, the users chose the positive adjectives from these combined words. Therefore, 'help to slow' and 'slow down aging' were selected to be positive adjectives, and added into the sentiment text database.

2.7 Document Representation

To build the matrix (n x m) which consisted of the Attribute (m) from the sentiment text database, there are (n) comments from the 500 opinions. The documents were separated into two trial sets below:

Set 1: A word was valued based on the accuracy value using the TF method for word substitution because it was not too complex. The 293 attributes of adjectives expressing emotions were separated into two groups which were a group of positive adjectives (166 attributes) and a group of negative adjectives (127 attributes).

Set 2: The size of the matrix increased due to the sentiment text database adjustment. There were 2,379 attributes of adjectives expressing emotions. These were separated into two groups, including a group of positive adjectives (1,812 attributes) and a group of negative adjectives (567 attributes).

The data was then converted into an ARFF file. This is a text file describing data that consists of attributes. The program has the @HEADER function which identifies the relations, attributes, and types of attribute data. Also, @DATA function identifies the detail of each attribute, in which the last attribute is the class label which has been classified into the negative, neutral or positive group to be used for both the trial session and then testing with the Weka Program.

2.8 Machine Learning Modeling

The document representatives were prepared into two trial sets, which were derived from 500 comments in the sample videos. The Weka Program was used to build a model and to compare several learning machine techniques, including NB, KNN, and SVM (SVM), by using RBF Kernel-typed Sequential Minimal Optimization (SMO), Poly Kernel and Puk, Lib SVM with Polynomial Kernel and Radial Basis function, and Linear alongside with parameter adjustment.

K-Fold Cross Validation and Percentage Split techniques were used to build the sentiment analysis model for the two documents. K-Fold Cross Validation is a popular research method to test model effectiveness because the outcome is expected to be reliable. It separates data into k parts and each k part is equal. In this research, k is set as 10; meaning each of the datasets included 500 opinions which were separated into 10 equal parts. Subsequently, a part of the data will be used as the subject for testing model efficacy. This process was repeated to completion.

The Percentage Split technique separated data into two groups with random percentages. For example, 80:20% split, by which 80% of the data was used for the trial, and 20% of the data was used for testing model efficacy. In this research, the split value was arbitrarily 66%, with 34% of data randomly selected for the trial set, and 66% of data for the real experiment. However, this method used the randomized data for one time only. If the random data for the experiment session had the same characteristics as those in the trial session, the performance results were assumed to be effective. Alternatively, if the random data for the experiment was significantly different from the trial session data, the performance results were assumed to be ineffective. Hence, the data was randomized several times. The advantage of this method is to build a model in a short time, being appropriate for a large-scale dataset.

2.9 The Comparison of Techniques

A comparison of learning machine performance comparing the NB technique, KNN, and SVM (SVM) with the accuracy measurement, can be calculated according to formulae (5)[16].

$$Accuracy = \frac{TP - TN}{TP + FP + FN + TN} \tag{5}$$

TP = Positive Opinion tend to have True Positive Opinion.

FP = Positive Opinion tend to have False Positive Opinion.

TN = Negative Opinion tend to have True Negative Opinion.

FN = Negative Opinion tend to have False Negative Opinion.

3. Results

Research findings of the sentiment analysis for Thai sounds in social media videos by using the SVM were obtained using specialists who experimented with 500 beauty product review videos on YouTube using the machine learning techniques consisting of NB, KNN, and SVM (SVM). Technique performance was compared, based on the concept of information retrieval accuracy by using two documents. Document 1 used the Term Frequency (TF) method for word substitution which gave out 293 attributes of adjectives. Document 2 used the word adjustment method which provided 2,379 attributes of adjectives. The techniques used for the comparison were K-Fold Cross Validation, where K is equal to 10, and Percentage Split, where a split is equal to 66% (after experimenting with the data, this was determined to be the most accurate rate of data separation by K-Fold Cross Validation and Percentage Split techniques). The experiment results of the accuracy rate are shown in Table 1.

Table 1. Comparison of effectiveness in terms of accuracy in each technique and the parameters

Technique and Parameter value	Accuracy			
	k=10		split=66%	
	293 (series1)	2379 (series2)	293 (series3)	2379 (series4)
1. NB	86.20	72.80	82.35	72.94
2. KNN	89.40	48.80	87.64	46.47
3. SVM (SMO), kernel=poly, c=1	85.20	74.40	80.00	74.70
4. SVM (SMO), kernel=poly, c=50000	91.20	60.40	89.41	54.12
5. SVM (SMO), kernel=rbf, c=1, gamma=0.01	48.20	74.20	47.06	74.70
6. SVM (SMO), kernel=rbf, c=50000, gamma=0.01	92.00	69.20	90.00	65.29
7. SVM (SMO), kernel=rbf, c=50000, gamma=0.1	94.40	81.60	93.88	76.47
8. SVM (SMO), kernel=puk, c=1, omega=1, sigma=1	92.40	83.00	93.53	76.47
9. SVM (SMO), kernel=puk, c=50000, omega=1, sigma=1	92.60	83.00	92.94	76.47
10. Lib SVM, kernel=linear, cost=1000, gamma=10	92.40	62.00	90.00	56.47
11. Lib SVM, kernel=polynomial, cost=1000, gamma=10	90.60	48.80	88.82	45.88
12. Lib SVM, kernel=radial, cost=1000, gamma=10	94.39	91.40	92.35	91.12
13. Lib SVM, kernel=sigmoid, cost=1000, gamma=10	91.00	73.80	87.05	71.18

Table 1, when considering the 13 models derived from the parameter adjustments for each technique, the experiment results can be concluded as follows:

Comparison between Document 1 (Series 1) which had 293 attributes of adjectives, and Document 2 (Series 2) which had adjusted words and 2,379 attributes of adjectives by using K-Fold Cross Validation and Percentage Split techniques, it was determined that the average accuracy rate of K-Fold Cross Validation, where K =10, was higher than the percentage split (at 66%). When the number of attributes increased, the processing time for the K-Fold Cross Validation technique also increased. Comparison data indicated that data separation using the K-Fold Cross Validation, where K=10, affected the accuracy of the SMO function, where Kernel=ref, c=50000, and gamma=0.1, in which the accuracy rate was 94.40% in the document (Series 1) which had 293 attributes of adjectives, as shown in Figure 3.

Following data classification, Document 1 (Series 1) had 293 attributes of adjectives and Document 2 (Series 2) had adjusted words and 2,379 attributes of adjectives by using K-Fold Cross Validation, where K =10, it showed that the average accuracy rate of Document 1 was higher than Document 2, as illustrated in Figure 4.

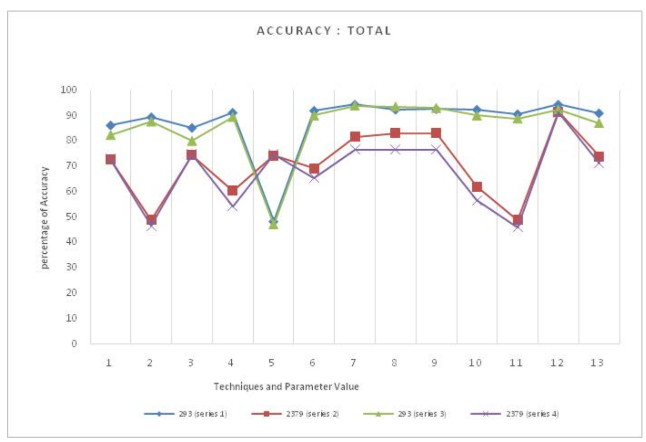

Figure 3. Comparison of all models accuracy

By categorizing the data for Document 3 (Series 3) which had 293 attributes of adjectives, and Document 4 (Series 4) which had adjusted words and 2,379 attributes of adjectives by using a percentage split at 66%, it showed that the average accuracy rate of Document 3 was higher than that of Document 4, as shown in Figure 5.

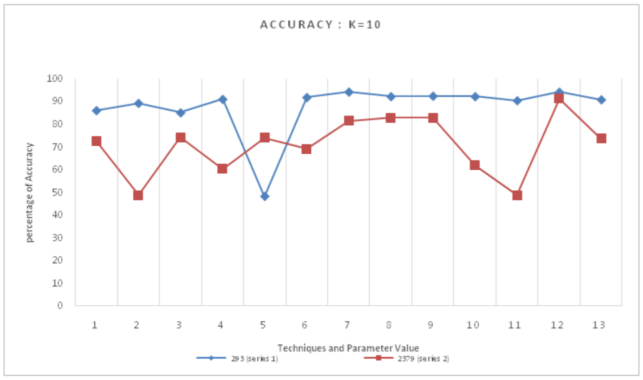

Figure 4. Accuracy rate comparison between document 1 and document 2 by k-fold cross validation, where K =10.

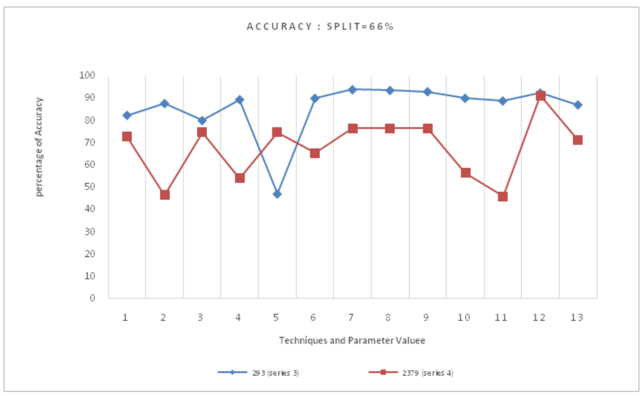

Figure 5. Accuracy rate comparison between document 3 and document 4 by percentage split at 66%.

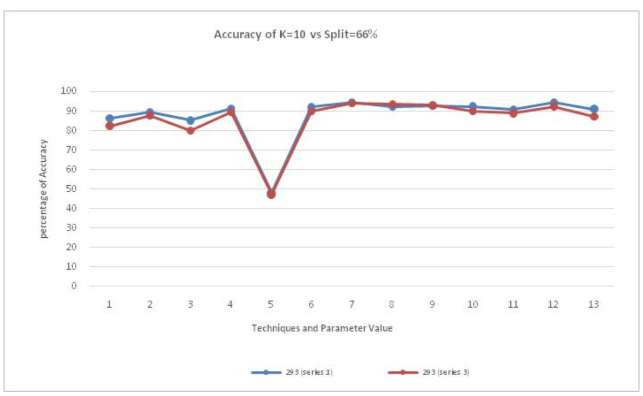

Figure 6. Accuracy rate comparison between series1 and series 3 by k=10 and percentage split at 66%.

Determining the data from Document 1 (Series 1) and Document 3 (Series 3), that had 293 attributes of adjectives by using the K-fold Cross Validation, where K =10, and a 66% percentage split, the accuracy rate of the K-fold Cross Validation technique, where K =10, was greater than the 66%percentage split, as presented in Figure 6.

The average accuracy rate of SVM appears greater than those from the NB and KNN techniques. The SVM used the RBF Kernel-typed SMO function, where c=50000 and gamma=0.1; the accuracy rate was 94.40%, when using K-fold Cross Validation, where K =10, that had 293 attributes, as shown in Figure 7.

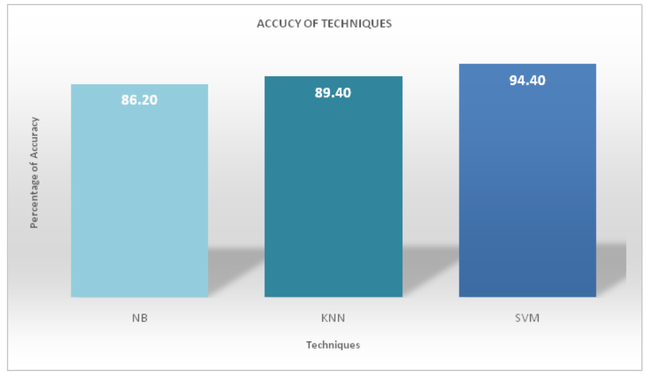

Figure 7. The accuracy rate comparison between NB, KNN, and SVM learning technique.

4. Conclusions

This study allowed for the development of sentiment analysis for Thai sounds in social media videos, by using a SVM to analyze 500 sample videos which were beauty product reviews on YouTube. Texts from YouTube videos were used to prepare the data for further classification. The data was then processed using audio extraction, speech recognition (ASR), word segmentation, database comparison, sentiment analysis, word adjustment, document representation, and modeling. The findings revealed that the effectiveness of the machine learning techniques, using the Term Frequency (TF) approach for word substitution, SVM technique with SMO function for adjectives by categorizing the data from tests, and K-Fold Cross Validation for data learning, where K is equal to 10, is 94.40% accurate when compared to the learning techniques of the NB and the KNN learning machine techniques. Moreover, sentiment analysis using the SVM technique was accurate and conformed with previous studies.

Analysis of sentiments studies in social media derived from texts, images, sounds, and videos continues. Users of social media can express their opinions and reviews to others with ease; with these reviews directly and indirectly affecting consumer choice. Information dissemination and subsequent market trends occurs at a fast rate. As such, the process of sentiment analysis continues as new technologies and analysis tools are developed. Data evaluation processes need to be fast and able to provide negative, neutral, or positive results in a timely manner, in which it will become a purchase decision-making guide for consumers.

These research results in data classification using trial and test sets showed that the appropriate parameter setting also affects the accuracy of sentiment analysis. This indicates that the data from a vast number of trials and tests can be influenced by sentiment analysis processes. Future research should focus on developing an automatic Thai vocabulary adjustment technique for sentiment analysis. They should also develop the sentiment analysis technique for Thai sounds that have a drawl to express one's feelings, improving sentiment analysis accuracy and its use in the market place.

5. Acknowledgements

We would like to express our gratitude to the Faculty of Science and Technology and Institute for Research and Development, Suan Sunandha Rajabhat University who offered us the opportunities to conduct this research.

6. References

1. Viriyavisuthisakul S. Thai Social Media Alert System for Business by Support Vector Machine, Panyapiwat Journal. 2016; 8:223–33.
2. Neethu MS, Rajasree R. Sentiment analysis in twitter using machine learning techniques, International Journal for Research in Applied Science and Engineering. 2016; 4(4):488–94.
3. Lima ACES, Castro deLN. Automatic sentiment analysis of Twitter messages. Forth International Conference on Computational Aspects of Social Networks; 2013. p. 52–57.
4. Isah H, Neagu D, Trundle P. Social Media Analysis for Product Safety Using Text Mining and Sentiment Analysis. 14th UK Workshop on Computational Intelligence; 2014. https://doi.org/10.1109/UKCI.2014.6930158.
5. Kadmateekarun P, Nuanmeesri S. Automatic sentiment analysis from opinion of Thais speech audio. Inter-national Conference on Science and Technology; 2015. p. 288–91. https://doi.org/10.1109/TICST.2015.7369372.
6. Pongthanoo P. Customer Satisfaction Analysis from Comment Text using Text using Opinion Mining, International Conference on Knowledge and Smart Technologies; 2012. p. 53–60.
7. Akaichi J, Dhouioui Z, Perez MJLH. Text mining Facebook status updates for sentiment classification. In-ternational Conference on System Theory Control and Computing; 2013. p. 640–45. https://doi.org/10.1109/ICSTCC.2013.6689032.
8. Kaushik L, Sangwan A, Hansen JHL. Sentiment extraction from natural audio streams, International Conference on Acoustics. Speech and Signal Processing; 2013. p. 8485–89.
9. Chaisal A, Sukhahut, R. Emotion Prediction from Thai Comments using Machine Learning Technique, The 9th National Conference on Computing and Information Technology; 2013. p. 260–66.
10. Brynielsson J, Johansson Carl F, Jonsson C, Westling A. Emotion classification of social media posts for estimating people's reactions to communicated alert messages during crises, Security Informatics. 2014; 3(1). https://doi.org/10.1186/s13388-014-0007-3.
11. Sarakit P, Theeramunkong T, Haruechaiyasak C, Okumura M. Classifying emotion in Thai Youtube comments, International Conference of Information and Communication Technology for Embedded Systems; 2015. p. 1–5.
12. Chirawichitchai N, Sanguansat P, Meesad P. Developing the effectiveness of automatic Thai document categorization, NIDA Development Journal. 2011; 51(3):188–205.
13. Naradhipa AR, Purwarianti A. Sentiment classification for Indonesian message in social media. In: Proceedings of the 2011 International Conference on Electrical Engineering and Informatics; 2011. p. 1–4. https://doi.org/10.1109/ICEEI.2011.6021696.
14. Binali HH, Wu C, Potdar Wu. A New Significant area: Emotion Detection in E-learning using Opinion Mining Techniques. International Conference on Digital Ecosystems and Technologies; 2009. p. 259–64. https://doi.org/10.1109/DEST.2009.5276726.
15. Rosas VP, Mihalcea R, Morency LP. Multimodal Sentiment Analysis of Spanish Online Videos. IEEE Intelligent Systems; 2013. p. 1–8.
16. Nuanmeesri S, Sriurai W. Development of guidelines for the academic support of students by using the a priori algorithm, Indian Journal of Science and Technology. 2018; 11(39):1–7. https://doi.org/10.17485/ijst/2018/v11i39/131268.