Indonesian online travel agent sentiment analysis using machine learning methods

Abimanyu D Poernomo, Suharjito
Computer Science Department, BINUS Graduate Program - Master of Computer Science,
Bina Nusantara University Jakarta, Indonesia

Article Info

ABSTRACT

Many companies use social media to support their business activities. Three leading online travel agent such as Traveloka, Tiket, and Agoda use Facebook for supporting their business as customer service tool. This study is to measure customer satisfaction of Traveloka, Tiket, and Agoda by analyzing Facebook posts and comments data from their fan pages. That data will be analyzed with three machine learning algorithms such as K-Nearest Neighbors (KNN), Naïve Bayes, and Support Vector Machine (SVM) to determine the sentiment. From the classification results, data will be selected with the highest f-score to be used to calculate the Net Sentiment Score used to measure customer satisfaction. The result shows that KNN result better than Naive Bayes and SVM based on f-score. Based on Net Sentiment Score shows companies that get the highest satisfaction value of Traveloka followed by Tiket.com and Agoda.

Corresponding Author:

Suharjito,
Computer Science Department, BINUS Graduate Program,
Master of Computer Science,
Bina Nusantara University Jakarta, Indonesia.
Email: suharjito@binus.edu

1. INTRODUCTION

High penetration of the internet and the high demand of the people of Indonesia for the needs of the holiday makes the new opportunity for the travel business. Many emerging new travel agencies serving ticket purchases and accommodation through the website. Social Media Marketing is a new way to promote products and reach new customer with less budget and resources. Social media marketing gives benefit for business by reducing cost and staff time, and increasing revenue generation [1]. Since social media have become popular for people to share their opinions, companies have opportunity to use the existing data for business intelligence applications for enterprise marketing services and customer relationship management [1]. Product review and feedback from product on social media is the best source to gain public opinion [2]. The public opinion is obtained using sentiment analysis [3]. Sentiment analysis is a kind of text classification that groups text based on the orientation of opinion, determining whether the given text is positive, negative or neutral [4], [5]. Sentiments found in comments, feedback or criticism provide useful indicators for various purposes. Sentiment analysis gives tool to companies to estimate product acceptance levels and determine strategies to improve the product quality [6]-[8].

Many companies use social media to support their business activities. Three leading online travel agent such as Traveloka, Tiket and Agoda use Facebook for supporting their business as customer service tool. Facebook has a Fan Page feature that online travel agencies use for promotions and get feedback from

posts or comments. Facebook comments and posts can become data sources to understand customers feeling, opinions and satisfactions.

This study is to measure customer satisfaction of Traveloka, Tiket and Agoda by analyzing Facebook posts and comments data from their fan pages. That data will be analyzed with three machine learning algorithms such as K-Nearest Neighbors (KNN), Naïve Bayes, and Support Vector Machine (SVM) to determine the sentiment.

From the classification results, data will be selected with the highest accuracy to be used to calculate the Net Sentiment Score used to measure customer satisfaction [9]. NSS score range is between -100 and 100, higher score indicates better customer satisfaction. The result of this study can benefit Traveloka, Tiket and Agoda to improve their service quality to satisfy the customers.

2.   RELATED WORKS

Sentiment analysis has been used in various industries for like telco [5], [10], transportation [9], automotive [11] and bank [12]. Social media as a data source that can be processed into sentiment that can be used to measure the satisfaction of a company's customers. Research on sentiment analysis as a benchmark of this satisfaction has been done by several researchers.

Indonesian researchers analyze tweet written in Indonesia to get Net Brand Reputation of Mobile Provider [10]. They use tweet from 3 operators as a data source. They classified and compared with Naïve Bayes, Support Vector Machine, and Decision Tree classifier. The Testing of the classification algorithm is based on cross validation technique. The results showed that SVM gives better performance than other two classifiers (Naïve Bayes and Decision Tree) in accuracy.

Another research about sentiment for online transportation provider [9]. They use tweet from 2 operator as a data source. Researchers do manual labeling before applying to classification model. They classify and compare with Naïve Bayes, Support Vector Machine, and Decision Tree Classifier. The Testing of the classification algorithm is based on cross validation technique. The results showed that SVM and Decision Tree give better performance than Naive Bayes.

From both studies, the similarities of the studies [9], [10] are on the objective of measuring the level of customer satisfaction with the firm. Then, the similarities in the study [9], [10] are the use of naive bayes and SVM algorithms. In addition to KNN as an algorithm to be tested along with Naive Bayes and SVM. The data used differently by using data derived from facebook that has a character length longer than 140 characters. The process of labeling training data is not done manually [9], [10] but by using a dictionary [13] [14]. The data is then tested with split validation. Then calculate Net Sentiment Score based on classification results from experiment with the best performance. The result shows which company offers better service quality and the best algorithm to classify data.

3.   RESEARCH METHOD

Figure 1 shows a framework that proposed in this research. Data collection from Facebook will be cleaned before made into features. After data is collected, it will be classified and evaluated. There are more explanations about the phases below.
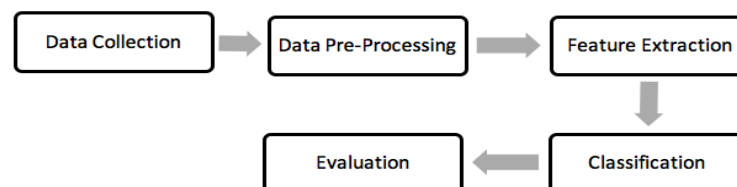

Figure 1. Research Framework

3.1. Data Collection

The Facebook posts and comments are collected using a custom script written in ASP.NET to contact Facebook Graph API and captured the data. The Crawler only crawls posts and comments from Indonesia online travel agencies such as Traveloka, Agoda and Tiket. The length of the data is started from January 2017 until December 2017. The data from crawling process can be stored in SQL Server Database.

3.2. Data Preprocessing

The goal in this step is to prepare the data that has been taken by eliminating the parts that are not needed later be used as a feature that will be used in the process of making a model of sentiment analysis. The processes undertaken in this stage are as follows:

1. Case Folding

   This process changes the existing word into lowercase. The purpose of this process is to avoid case sensitive validation when words comparing with dictionaries. Example: "Kerja Bakti" -> "kerja bakti"

2. Remove Link

   Datas are taken from the website link. Website links such as *http://facebook.com* are not required in the feature extraction process so that it is removed by performing a replacement with "".

3. Emoticon Conversion

   Emoticons contained in the data can be converted to words that will be obtained into a feature that can be used in sentiment analysis. Converting Emoticons to Word will be based on research from [15].

3.3. Feature Extraction

After preprocessing process, data will be transformed into unigram feature. Unigram feature will be transformed to vector model with Term Frequency – Inverse Document Frequency (TF-IDF). Feature extraction processed with stages as follows:

1. Remove Stopwords

   Stopwords is a word that does not mean much in the sentence and can be omitted. The dictionary to be used in removing stopwords in Bahasa Indonesia will be based on research from Tala [16].

2. Tokenization

   The tokenization processed by separating the sentence into words per word which can be performed for the analysis of sentiment.

3. Stemming

   Stemming converts the word to be a basic word. Stemming rules using Bahasa Indonesia use the Nazief and Adriani algorithm based on research from Asian [17], [18].

3.4. Classification

The research is done with the Rapidminer software for the prediction, Rapidminer is a data mining open source software that is used in many researches. The Feature will be applied with several machine learning method like K-Nearest Neighbor, Naïve Bayes and SVM algorithm. For data validation, model apply split validation with 70% training data and 30% testing data. Training data is provided by automatic labeling with opinion word dictionary [13], [14].

3.5. Evaluation

Evaluation process by comparison the results of sentiment analysis model using confusion matrix. Confusion matrix is a useful tool for analyzing how well a classifier model can predict the classification results correctly on a large number of classes [19].

Table 1. Confusion Matrix

	Actual Positive Class	Actual Negative Class
Predictive Positive Class	True Positive(tp)	False Negative (fn)
Predictive Negative Class	False Positive(fp)	True Negative (tn)

Based on Table 1, confusion matrix reports the number of false positive, false negative, true positive and true negative. The processed data is entered into the matrix for each of each classification method. From the confusion matrix then we calculate for the accuracy, precision and recall. Accuracy measures the ratio of correct prediction over total number of items [20].

$$\frac{tp + tn}{tp + fp + tn + fn}$$

Precision measures positive pattern that correctly predicted from total predicted items in positive class [20].

$$\frac{tp}{tp + fp}$$

Recall measure fraction of positive patterns that classified as correct [20].

$$\frac{tp}{tp + tn}$$

F-Score is harmonic mean of precision and recall [21] that show the performance of model. A high value of F-Score indicates that the model performs better on the positive class.

$$\frac{2 \; x \; Precision \; x \; Recall}{Precision + Recall}$$

From the classification results, data will be selected with the highest accuracy to be used to calculate the Net Sentiment Score used to measure customer satisfaction [9]. NSS score range is between -100 and 100, higher score indicates better customer satisfaction.

$$NSS = \frac{Positive \; Mentions - Negative \; Mentions}{Positive \; Mentions + Negative \; Mentions} \; x \; 100$$

4. RESULTS AND ANALYSIS
4.1. Model Evaluation

In testing the sentiment model, all data from the three companies were analyzed using the algorithm. The results of the matrix confusion calculation for each algorithm are accuracy, precision, recall and f-score. Table 2 shows about accuracy, precision, recall for every algorithm.

Table 2. Customer Satisfaction with NSS

	Accuracy	Precision	Recall	F-Score
Naïve Bayes	47.60%	12.87%	76.83%	22,04%
KNN	96.32%	58.22%	21.05%	30,92%
SVM	93.34%	98.48%	8.61%	15,84%

Data in Table 2 shows that KNN has the best accuracy score with 96.32% better than others. Judging best model from accuracy is not fair enough to see the overall performance it will be calculated f-score with based on precision and recall. From Figure 2, KNN has the highest f-score compared with Naïve Bayes and SVM so it can be concluded that KNN is the best model compared to Naïve Bayes and SVM.
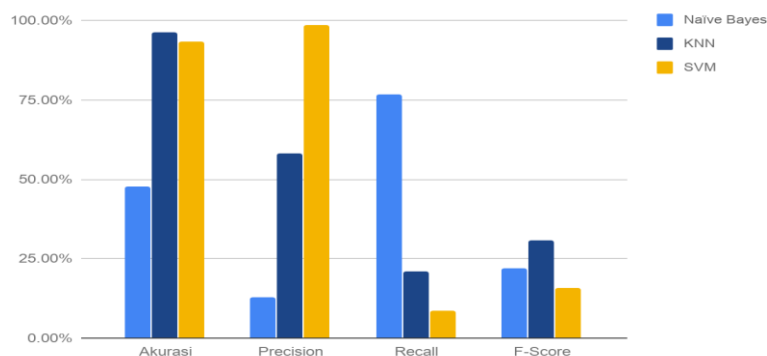

Figure 2. F-Score Comparison

4.2. Customer Satisfaction Evaluation

Customer satisfaction is measured using Net Sentiment Score calculated using positive and negative mention. This research will use the model with the highest f-score value based on the calculation result. Based on Table 3, it shows that Traveloka has Net Sentiment Score 97.53, followed by Tiket.com has 91.55

and Agoda has -96.03. Data shows that Agoda has a negative value of NSS which indicate low customer satisfaction different with Traveloka and Tiket.com.Traveloka get best score that means customer very satisfy with their services than others.

Table 3. Customer Satisfaction with NSS

	Accuracy	Precision	Recall	F-Score	Positive	Negative	NSS
Traveloka	95.44%	88.67%	13.47%	23.38	15441	193	97.53
Tiket.com	95.50%	53.73%	21.43%	30.64	11656	514	91.55
Agoda	97.71%	35.71%	11.90%	17.85	136	6722	-96.03

5. CONCLUSION

Companies that get the highest satisfaction value of Traveloka with a value of 97.53 followed by Tiket with a value of 91.55 and less satisfactory in the can by Agoda with a value of -96.03. Classification with good f-score in this study was KNN compared with Naïve Bayes and SVM. For future research, we recommend classification of services from online travel such as hotel services and ticket services to get a specific satisfaction data per service and improvements dictionary for Indonesian language with additional non-standard language that is often used on the Internet.

REFERENCES

[1]   Bhardwaj N, Shukla A, Swarnakar P. Users' Sentiment Analysis in Social Media Context using Natural Language Processing. *The International Conference on Digital Information, Networking, and Wireless Communications.* 2014; 3: 103-111.
[2]   Sharma DM, Baig MM. Sentiment Analysis on Social Networking: A Literature Review. *International Journal on Recent Innovation Trends in Computing and Communication (IJRITCC). 2015; 3(2): 22-27.*
[3]   Jotheeswaran J, Koteeswaran S. Sentiment Analysis: A Survey of Current Research and Techniques. *International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE).* 2015; 3(5): 3749-3757.
[4]   Devika MD, Sunitha C, Ganesh A. *Sentiment Analysis: A Comparative Study on Different Approaches.* Fourth International Conference on Recent Trends in Computer Science & Engineering. Chennai. 2016; 87: 44-49.
[5]   Susanti AR, Djatna T, Kusuma WA. Twitter's Sentiment Analysis on Gsm Service using Multinomial Naive Bayes. *TELKOMNIKA .* 2017; 15(3):1354-1361.
[6]   Prabowo R, Thelwall M. Sentiment Analysis: A Combined Approach. Journal of Informetrics. 2009; 3(2): 143-157.
[7]   Akkineni H, Lakshmi P.V.S., Babu B.V. Online Crowds Opinion-Mining it to Analyze Current Trend: A Review. *International Journal of Electrical and Computer Engineering (IJECE).* 2015; 5(5): 1180-1187.
[8]   Kumar PK, Nandagopalan S. Insight to Problem, Research Trend and Progress in Techniques of Sentiment Analysis. *International Journal of Electrical and Computer Engineering (IJECE).* 2017; 7(5): 2818-2822.
[9]   Anastasia S, Budi I. Twitter Sentiment Analysis of Online Transportation Service Providers. International Conference on Advanced Computer Science and Information Systems (ICACSIS) Malang. 2016; Page(s): 359-365.
[10] Wijaya H, Erwin A, Soetomo A, Galinium M. Twitter Sentiment Analysis and Insight for Indonesian Mobile Operators. *Information Systems International Conference (ISICO),* 2013; 1:367.
[11] Shukri SE, Yaghi RI, Aljarah I, Alsawalqah H. Twitter Sentiment Analysis: A Case Study in Automotive Industry. *IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)* 2015: 1-5.
[12] Alamsyah A, Indraswari AA. Social Network and Sentiment Analysis for Social Customer Relationship Management in Indonesia Banking Sector. *Advanced Science Letters.* 2017;23: 3808-3812
[13] Wahid DH, Azhari SN. Peringkasan Sentimen Ekstraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity. *Indonesian Journal of Computing and Cybernetics Systems (IJCCS).* 2016; 10(2): 207-218.
[14] Liu B, Hu M, Cheng J. *Opinion Observer: Analyzing and Comparing Opinions on the Web.* Proceedings of the 14th international World Wide Web Conference (ICTS). Chiba. 2005;
[15] Vashist G, Thakur S. Facebook as a Corpus for Emoticon-Based Sentiment Analysis. *International Journal of Emerging Technology and Advanced Engineering (IJETAE).* 2014; 4(5): 904-908.
[16] Tala FZ. A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia. M.Sc Thesis. University of Amsterdam; 2003.
[17] Asian J. Effective Techniques for Indonesian Text Retrieval. PhD Thesis. School of Computer Science and Information Technology RMIT University; 2007.
[18] Arifin AZ, Mahendra IPAK, Ciptaningtyas HT. *Enhanced Confix Stripping Stemmer and Ants Algorithm for Classifying News Document in Indonesian Language.* Proceedings of the 5th International Conference on Information and Communication Technology & Systems (ICTS). Surabaya. 2009;
[19] Han J, Kamber M, Pei J. Data Mining Concepts and Techniques. 3rd ed. London: Morgan Kaufmann; 2012: 386.
[20] Hossin M, Sulaiman MN. A Review on Evaluation Metrics for Data Classification Evaluation. International Journal of Data Mining & Knowledge Management Process (IJDKP). 2015;5(2): 1-11.
[21]   Bekkar M, Djemaa HK, Alitouche TA. Evaluation measures for models assessment over imbalanced data sets. Journal of Information Engineering and Applications. 2013;3(10): 27-38.