Naive Bayes as opinion classifier to evaluate students satisfaction based on student sentiment in Twitter Social Media

Fahmi Candra Permana\textsuperscript{1}, Yusep Rosmansyah\textsuperscript{2}, and Atje Setiawan Abdullah\textsuperscript{1}

\textsuperscript{1} Department of Computer Science, Faculty of Mathematics and Natural Sciences Universitas Padjadjaran, Indonesia
\textsuperscript{2} School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia

E-mail: fahmicandrap@gmail.com

Abstract. Students activity on social media can provide implicit knowledge and new perspectives for an educational system. Sentiment analysis is a part of text mining that can help to analyze and classify the opinion data. This research uses text mining and naive Bayes method as opinion classifier, to be used as an alternative methods in the process of evaluating students satisfaction for educational institution. Based on test results, this system can determine the opinion classification in Bahasa Indonesia using naive Bayes as opinion classifier with accuracy level of 84\% correct, and the comparison between the existing system and the proposed system to evaluate students satisfaction in learning process, there is only a difference of 16.49\%.

1. Introduction

Social media sites like Twitter and Facebook provide great venues for students to share joy and struggle, vent emotion and stress, and seek social support\cite{1}. On social media sites, students discuss and share their everyday encounters in an informal and casual manner. Students activity on social media creates large amounts of implicit knowledge and a whole new perspective for teacher and educational researchers to evaluate and monitor students experiences outside the controlled classroom environment. Therefore, students satisfaction is one key to the success of teachers and academic staff in developing the existing education system. by analyzing a range of opinions given by the students, A teacher can know the response is positive, negative and questions from students the teaching materials or systems that have been done so that it can be a matter of evaluation of teachers and academic staff.

Opinion classification system development can be develop using text mining in sentiment analysis \cite{3}. Process classification of textual documents with sentiment analysis is done with dividing the document types into three categories, namely category sentiment positive, neutral, and negative \cite{3}. Many approaches are being made to perform sentiment analysis, of which the most famous is the Machine Learning Based (training-based machines), Naive Bayes as a method, Support Vector Machine and Maximum Entropy\cite{4}.

Twitter social media activity have huge development in the world, one of them is in Indonesia. Indonesia is a country with a Twitter user fifth largest in the world with 50 million registered
accounts [2], even when compared with the number of Internet users in a country, in Asia Pacific Indonesia ranks first [3].

![Twitter users in Asia Pacific](image)

**Figure 1.** Twitter users in Asia Pacific[7]

With the number of users in Indonesia and the characteristics of Twitter itself is only composed of a maximum of 140 letters in once post, Twitter can produce information that is simple and make sense, the use of Twitter at this time may be authorized to make an evaluation in various fields, including educational system. Therefore, at this time the need for a new system that could facilitate an educational institution in classifying the opinion of all the students who are in the process of education they do into the kind of positive opinion, negative or neutral that can be a useful tools for performance evaluation of educational institution. The purpose of this research is to create and review the results of the application of the system that can monitor and evaluate satisfaction level of student learning by students sentiment on social media, so it can be used as an alternative method for evaluating the performance of an educational institution.

2. Theoretical Background
To do research in Sentiment analysis and monitoring student satisfaction we doing literature reviews in related topics. Literature map is a graphic that helps to visualize the relationship between some of the papers that have between one mapping helps to identify issues such as connection and closeness of ideas and findings [2]. Literature map is useful because it helps divide the paper into sections or chapters.
Figure 2. Literature Map.

Figure 2 shows the research literature on the map Sentiment Analysis on social media and monitoring student satisfaction. Realizing the importance and issues of sentiment analysis can make better evaluate to teacher and academic staff who will guide and monitoring student satisfaction in social media.

2.1. Mining Twitter Data
Researchers from diverse fields have analyzed Twitter content to generate specific knowledge for their respective subject domains. For example, someone analyzes tweets with mentioned @edulabindonesia using histograms, user networks, and frequencies of top keywords to quantify online activism. In this research, we built a classification model based on inductive content analysis. This model was applied and validated on a brand new dataset. Therefore, we emphasize not only the insights gained from one dataset, but also the application of the classification algorithm to other datasets for detecting student problems. The human effort is thus augmented with large-scale data analysis.

2.2. Machine Learning Technique
Machine Learning techniques use a training set and a test set for classification. Training set contains input feature vectors and their corresponding class labels. Using this training set, a classification model is developed which tries to classify the input feature vectors into corresponding class labels. Then a test set is used to validate the model by predicting the class labels of unseen feature vectors. A number of machine learning techniques like Naive Bayes (NB), Maximum Entropy (ME), and Support Vector Machines (SVM) are used to classify reviews [3].

2.3. Classification Method with Naive Bayes
Opinion classification using Naive Bayes method proved to have a fairly good accuracy, especially in the English-speaking opinion. Based on research on some English-speaking opinion data obtained through the website of a company, the level of accuracy of the method Naive Bayes classification opinion in doing quite well with the percentage of accuracy 81.43322% [4]. The existence of many differences in the structure of language, the term non-standard, and the expression in English and Indonesian prompted the authors to conduct research building
classification system in Indonesian language opinion obtained via the social media Twitter by using Naive Bayes.

2.4. Purposive Sampling
Purposive sampling (e.g. judgment and quota) is an approach where members conform to certain criteria for selection. In judgment sampling a researcher may only want to survey those who meet a certain criteria. Judgment sampling is also appropriate in early stages of research where selection is made based on screening criteria[8]. Recruiting survey respondents through Twitter was highly effective for the research conducted but has several limitations. The survey was conducted with purposive sampling which is a non-probability sampling technique.

3. Methodology
This research separated in three step, the first step is mine the student sentiment in Twitter with text mining, the next steps is classify and analyze the data from twitter data to evaluate and monitor student satisfaction every mid semester, and the last step is compare the result from this system with existing system at academic consultant.

3.1. Text Mining
There are two basic methodologies to detect sentiments from text. They are Symbolic techniques and Machine Learning techniques [3]. This Research used the Machine Learning techniques with supervised learning as a output the technique because Machine Learning techniques are simpler and efficient than Symbolic techniques and these techniques can be applied for twitter sentiment analysis[3] and Naive Bayes as the classifier method, due to the accuracy of the most well among other classification process[3].

3.2. Sentiment Analysis
In Sentiment Analysis an overview of steps and techniques commonly used in sentiment classification approaches, as shown in Figure 3. Part of speech model in which a document is represented as a vector, whose entries correspond to individual terms of a vocabulary. Part-of-speech information is supposed to be a significant indicator of sentiment expression. The work on subjectivity detection [4] reveals a high correlation between the presence of adjectives and sentence subjectivity.

After get the database from twitter with text mining using the NodeXL Excel Template. In the rst phase preprocessing is done. Then a feature vector is created using relevant features. Finally using Naive Bayes as classifiers, tweets are classified into positive, negative or neutrals classes. Based on the number of tweets in each class, the nal sentiment is derived. And this research develop a system to evaluate and monitoring student satisfaction based on sentiment analysis classification.

3.3. Compare This Proposed System with Existing System
After we get the data from twitter, uses sentiment analysis and have a data who was classified we develop a system that will be used every mid semester to evaluate and monitor student satisfaction and compared with manual survey to student which common use in the consultant academic.

4. Implementation
This section describe about the result of our research and discuss it to make sure this research can well implemented, this is part of our plan result after mine the twitter data, using Naive Bayes as Classifier. And can be compare with existing system at consultant academic.
4.1. Opinion Chronology Classifier System
Sentiment analysis system developed consists of a system capable of classifying opinions into positive and negative types using Naive Bayes. In this system, the raw data obtained in advance is separated into two types of data that training data and test data.

4.2. Dataset
The document level sentiment classification dataset in sample courses at this academic consultant, processed with this system and get the data, this research focus only on positive, negative and neutrals opinion, and after that this result will be use every mid semester to evaluate student satisfaction and will be compare with manual surveys of students which is commons used in the consultant academic.
4.3. Implementation
The system of evaluating students' level of satisfaction that we develop and designed to classify students' opinion on social media in several categories: positive, negative and neutral, then the system can provide the value of the students' level of satisfaction. In the implementation of the system is very dependent on the database, because the database connection is used to store and add training data that serves as the core classification system that opinion. This database connection serves to connect the program or application that we made to the database "twittersentimentanalyzer" in which there is a table "table-tweettraining".

4.4. Validation Test System
After getting the results of the system that we made, the next step is to test the system validation to determine the level of accuracy and performance of the system that has the we made, before they were used and compared with the already running previously. The system that the we created consisting of 300 data option is divided into two pieces of data, training data and test data, with the ratio of training data: test data = 70: 30, consisting 90 of 210 training data and test data. validation of test results and test the accuracy of the classification system based on opinion can be seen in Figure 5.

| SENTIMENT DOMINANT = 77.0% POSITIVE |
|--------------------------------------|
| TINGKAT KEPUASAN SISWA EDULAB = 3.85 |

| TWEET POSITIVE          | POSITIVE |
|-------------------------|----------|
| [1]cape matematika @EdulabIndonesia | F 0 |
| [2]@EdulabIndonesia tapi sekarang dating sama kang Haroki yeyeye | T 1 |
| [3]@edulabindonesia edulab enak belajarnya. feel like home :bd | T 1 |
| [4]TO edulab tadi siang susah tapi semoga hasilnya bagus, semangat to2 berikuuyaa @Edulatlnonesia | T 1 |
| [5]Ga keraa to @EdulabIndonesia tinggal 2x lagi..semangat buat to selanjutnya :) | T 1 |
| [6]Daripada meminang gada kerjaan meing belajar @EdulabIndonesia , semangat SBMPTN, hehe demi masa depan :) | T 1 |
| [7]@EdulabIndonesia kesaaku belajar di edu itu pokoknya ga bozenin, seneng boleh tambahan kpn saja, bisa dating- asik laah. | T la |
| [8]@EdulabIndonesia mungkin tempat belajar yang bisa ngajay saya malem hanya guru2 di edulab , Thanks :) | T 1 |
| [9]cuma di @EdulabIndonesia yang gurunya bisa diajak belajar tambahan sampai jan segin, always feel like home :) | T 1 |
| [10]Gasabar pengen cepet les di @EdulabIndonesia , susah kalau sendiri | T 1 |
| [11]Suasana belajar di @EdulabIndonesia enak banget, bikin nyaman belajar | T 1 |
| [12]RT @HasnaSittiH: Gasabar pengen cepet les di @EdulabIndonesia , susah kalau sendiri | T 1 |
| [13]RT @HasnaSittiH: Suasana belajar di @EdulabIndonesia enak banget, bikin nyaman belajar | T 1 |

**Figure 5.** Data validation test results.

The test results in Figure 5 can be used to calculate the level of accuracy of the system.

\[
\text{Level of accuracy} = \frac{\text{the number of tests is true}}{\text{amount of data validation}} \times 100\% = \frac{75}{90} \times 100\% = 84\%.
\]
Based on the calculation of the percentage of accuracy, the accuracy of the values obtained are quite high, with an accuracy of 84%. Implementation of the proposed system in an environment there will always be comparisons with the old system and has been existence, therefore we compare the test results of satisfaction level assessment of student learning in the Edulab using the old system can be seen in Figure 6.

**Figure 6.** Student satisfaction questionnaire.

In the Figure 6 there are 54 aspects of assessment, and conducted a survey to all students in the branch Edulab K7, with a total of 326 student respondents and divided into two categories namely CSI ”Customer Satisfied Indicator” and CEI ”Costumer Enggagement Indicator”, while the results of the survey can be seen in Figure 7.

From Figure 7 can be seen that the category CSI ”Customer Satisfied Indicator” get results with a 3.40 rating, while the category of CEI ”Costumer Enggagement Indicator” get assessment results 3.25, so the combined Edulab student learning satisfaction level can be averaged averaged as follows:

\[
Edulab \text{ student satisfaction} = \frac{CSI + CEI}{2} = \frac{3.40 + 3.21}{2} = 3.305
\]
| No | Cabang Program | K7 |
|----|----------------|----|
|    |                | D  | S  |
| 42 | Customer Service Edulab selalu memberikan informasi kepada saya dengan tepat dan akurat | 3.69 | 3.59 |
| 43 | Edulab memberikan pelayanan peranganan keluhan saya dengan cepat dan responsif | 3.23 | 3.10 |
|    | CSI | 3.40 | 3.37 |

**Figure 7.** Result of level student satisfaction in edulab.

The Systems that we develope is a system that has the same goal with a system that is already running in Edulab, that the system evaluator satisfaction levels of student learning based on student sentiment on social media, the system uses the Naive Bayess method as a classifier student opinion in social media. Sentiment analysis system that has been built based on the data that is included @edulabindonesia tweet on his tweet, in the period October 2014 to May 2015, and after the process of training and testing data in the system result in that student satisfaction rates reached 3.85, from the results obtained both types of measurement methods of evaluation and monitoring of student satisfaction levels can be calculated accuracy edulab both types of measurement methods in the following:

\[
\text{Improvement} = \frac{\text{Proposed Method} - \text{Old Method}}{\text{Old Method}} \times 100% = 3.85 - 3.305 \times 100% = 16.49% 
\]

5. Conclusion

Based on test results, this system can determine the opinion classification in Bahasa Indonesia using naive Bayes as opinion classifier with accuracy level of 84% correct, and the comparison between the existing system and the proposed system to evaluate students satisfaction in learning process, there is only a difference of 16.49%. So the proposed system can be used as an alternative system to evaluate student’s satisfaction on educational system.
Acknowledgments
The work reported in this paper is supported by the Academic Leadership Grant Unpad 2016, so the authors thank to Rector Universitas Padjadjaran and Directorate Research and Community Services Unpad for supporting funding of research and paper dissemination.

References
[1] Konopnicki D, Shmueli-Scheuer M, Cohen D, Sznajder B, Herzig J, Raviv A, Zwerling N, Roitman H, Mass Y 2013 A statistical approach to mining customers’ conversational data from social media IBM Journal of Research and Development 57(3/4) pp 14:1 –14:13
[2] Chen X, Vovoreanu M, Madhavan K Mining Social Media Data for Understanding Students' Learning Experiences Learning Technologies, IEEE Transactions on PP(99) p 1
[3] Pang B, Lee L, Vithyanathan S 2002 Thumbs Up ? SentimentClassification Using Machine Learning Techniques Proceedings of TheACL-02 conference on Empirical methods in natural language processing(pp. 79–86). Stroudsburg: Association for Computational Linguistic
[4] Neethu MS, Rajasree R 2013 Sentiment analysis in twitter using machine learning techniques Computing, Communications and Networking Technologies (ICCCNT), 2013 Fourth International Conference p 1,5
[5] Bahrainian, Seyed-Ali, Dengel, Andreas 2013 Sentiment Analysis and Summarization of Twitter Data Computational Science and Engineering (CSE), 2013 IEEE 16th International Conference pp 227 –234
[6] Omar Y, Ashaari NS 2010 Futuristic model for school’s content management systems: A beginning Information Technology (ITSim), 2010 International Symposium 3 pp 1387 –1392
[7] http://www.statista.com/statistics/254803/twitter-penetration-in-selected-apac-countries/ (accessed at 25 May 2015)
[8] Sibona, Christopher, Walczak, Steve 2012 Purposive Sampling on Twitter: A Case Study 45th Hawaii International Conference on System Sciences