Sentiment analysis on feedback of higher education teaching conduct: An empirical evaluation of methods

Cite as: AIP Conference Proceedings **2470**, 050006 (2022); https://doi.org/10.1063/5.0080182
Published Online: 25 April 2022

Jimmy and Vincentius Riandaru Prasetyo

Sentiment Analysis on Feedback of Higher Education Teaching Conduct

An Empirical Evaluation of Methods

Jimmy[a)] and Vincentius Riandaru Prasetyo[b)]

Department of Informatics Engineering, Faculty of Engineering, University of Surabaya, Surabaya 60293, Indonesia

[a)] Corresponding author: jimmy@staff.ubaya.ac.id
[b)] vincent@staff.ubaya.ac.id

Abstract. Sentiment analysis aims to automatically identify and classify the tone or polarity of people's opinion in an unstructured text. This paper focuses on evaluating the accuracy of sentiment analysis methods to classify Indonesian higher education teaching conduct. In this context, we make the following contributions: (1) evaluation on the impact of text preparation methods in term of accuracy of sentiment analysis results, (2) evaluation the accuracy of three popular sentiment analysis methods (i.e., Naive Bayes, Support Vector Machine, Decision Tree) in classifying Indonesian text, and (3) proposal and evaluation on the effectiveness of combining the results from the three methods considered in this study with hope to improve the results' accuracy. Finally we analyzed cases where all classifiers suggested incorrect sentiment classification and highlighted areas for future works to improve the accuracy of sentiment analysis, in particular for Indonesian Text.

Keywords: Sentiment Analysis, Feedback, High Education Teaching Conduct

INTRODUCTION

Sentiment analysis is the task of clustering or classification of open ended and unstructured text which often contains opinions of people about a specific topic with objective to better understand people's opinion on the topic [1, 2]. For example, an organization performed sentiment analysis on their social media page (e.g., Twitter, Facebook) to understand how the general public perceived the organization and its products/services. In contrast to rating system in common online shopping platform such as Amazon, Shopee and Tokopedia, people expressed their opinion in social media as a qualitative open-ended text. While open ended question allows people to better express their opinion, there are times when the organization need to understand the polarization of an opinion (e.g., whether a person is expressing a negative or a positive experience). Understanding the polarization of thousands, millions or more opinions is time consuming and thus, a significant number of resources are needed if this task is to be performed manually. Hence, sentiment analysis has been the focus of many research in recent years to automatically cluster or classify many opinions of a particular topic.

In term of task type, sentiment analysis is commonly used to classify and/or to cluster a collection of unstructured text. The classification task focuses on grouping a set of text (e.g., opinions) into subsets of predetermined groups (often labeled as classes) [3]. In contrast, the clustering task focuses on grouping a set of text based on similarities or patterns within the text with no predetermined groups [3].

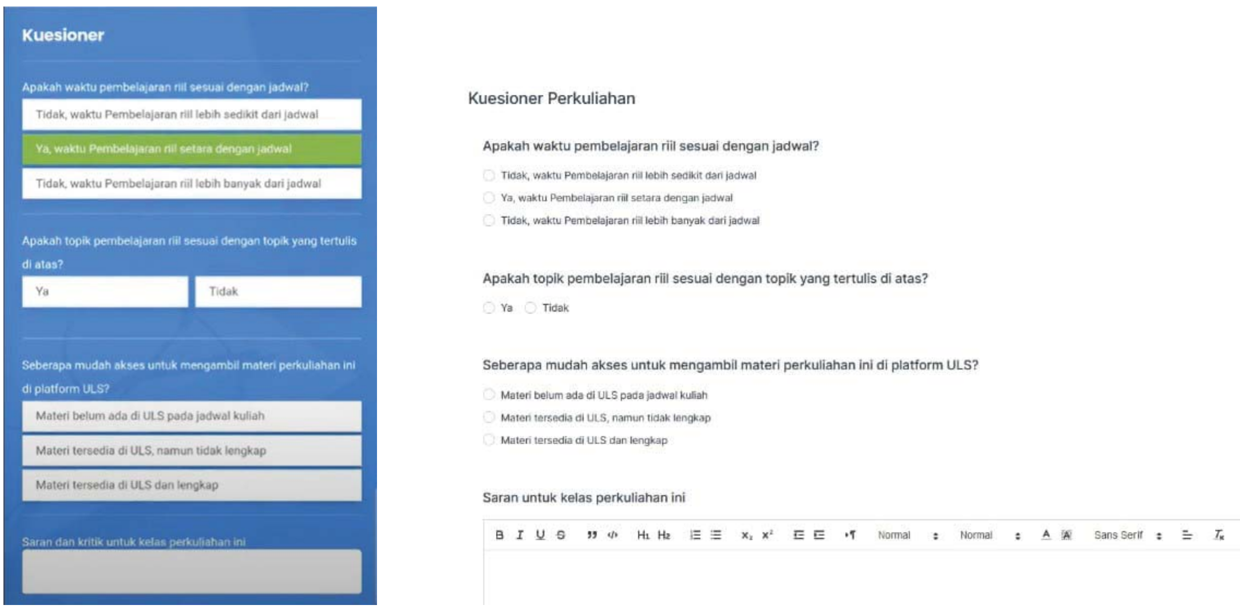
In this paper, we focus on applying sentiment analysis method to classify undergraduate student feedback on teaching conduct of their enrolled classes. The selected case study used in this paper is from the University of Surabaya, Indonesia. Hence, the data collection is in Indonesian. In early 2020, the COVID-19 pandemic forced each and every academic institution to radically and immediately switch their teaching methods from traditional here

International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)
AIP Conf. Proc. 2470, 050006-1–050006-9; https://doi.org/10.1063/5.0080182
Published by AIP Publishing. 978-0-7354-4180-4/$30.00

both the lecturer and the students physically attend classes to online learning. Such radical changes had raised many concerns and adjustments to do.

Among the many adjustments to do was the lecturers and students' attendances system. Traditionally, prior starting a class, the lecturer(s) need to get an empty set of attendance form, fill in the date and time of the class, sign in, write the topic of the day. Sequentially, students who attend the class should also put their signatures in the attendance form. Such mechanism was no longer feasible when the teaching and learning activities went online.

Therefore, in early 2021, the university implemented a new attendance system which is expected to be applicable for any method of learning activities (i.e., online or offline). To confirm his/her attendance, the new attendance system requires each student to provide feedback using browser using their mobile device (accessible through https://m.ubaya.ac.id) or using device with larger screens such as laptop (https://neo.ubaya.ac.id). Figure 1 shows the user interface for both mobile and desktop version. In general, students were asked to provide feedback on the duration, conformance to the topic, and the availability of teaching materials in the University's online learning space. In addition to the four closed questions, students were asked to provide comments as an open-ended answer. The number of questions was limited to minimize the amount of student's effort in confirming his/her attendance.



(a) Mobile version (b) Desktop version

FIGURE 1: Student's questionnaire form for mobile. The mobile version (left) was designed for devices with small screen (e.g., mobile phones and tablets). The desktop version (right) was designed for devices with large screen such as laptop and personal computer.

As noted in Fig. 1, there was no option for students to rate how well or bad a class was being conducted by the lecturer. The comment was made available to capture students' expressions on the class conduct. While such feedback is important for the lecturer to further improve the teaching and learning process, the university's management need an aggregated view to quickly understand how well a class was conducted. Such view will also enable the management to evaluate the quality of teaching and learning over a period of time from the students' perspective.

Hence, the problem that this paper aims to solve is how to better understand students' sentiments on teaching conducts based on their feedback. To solve the problem, we empirically evaluated the effectiveness of sentiment analysis methods in classifying the students' open-ended comments. In this context, we make the following contribution:

1. We evaluated the impact of text preparation methods on the accuracy of sentiment analysis methods. The considered text preparation methods include stop words removal, stemming, and tokenization.
2. We evaluated three general sentiment analysis methods adapted to the problem of classifying feedback of higher education teaching conduct: (1) Naive Bayes, (2) SVM, and (3) Decision Tree.

3. We evaluate the effectiveness of combining results of the three sentiment analysis methods using a vote mechanism.

RELATED WORKS

In this paper, we performed empirical evaluation following the five steps to analyze data as suggested by Alessia et al. [4]: (1) data collection, (2) text preparation, (3) sentiment detection, (4) sentiment classification, and (5) presentation of output. Here, we provided discussion on alternatives to perform each of the five steps as found in the literature.

Data Collection

Data collection is the first step with purpose to collect data for training and evaluation purposes. Data may be collected specifically to fit a specific research purpose (i.e., primary data) or collected from other systems that collected data for other purposes (i.e., secondary data) [5]. This research uses secondary data from an attendance system in the University of Surabaya containing students' feedback with purpose to improve teaching and learning activities in the university.

Data can be either quantitative that uses numbers to represent phenomenon or qualitative that describes phenomenon using text, audio, or other visual medium. Depending on its structure, data can be classified as simple or complex [5]. In this paper, the data used is qualitative with simple data structure.

Text Preparation

After data is collected, data is then transformed to enable better data analysis results. Data preparation includes removing special characters (e.g., $, %, &), removing numbers, lower-casing, stemming, and removing stop-words [6]. In this paper, we performed lower-casing and, following the work of, Pradana and Hayaty [7], evaluated the impact of stemming and stop-words removal on sentiment analysis.

We particularly interested on evaluating the impact of stemming and stop-words removal as experiments on the use of the two data preparation methods show contradictory results. Sentiment analysis studies on Indonesian text suggest that stemming and stop-words removal degrade the accuracy of sentiment analysis results. A study on sentiment analysis on Indonesian text by Pradana and Hayaty [7] suggests that stemming and stop-words removal actually reduced the accuracy of sentiment analysis as evaluated in their study. Similarly, a study on sentiment analysis on Indonesian tweets by Hidayatullah, Ratnasari, and Wisnugroho [8] shows that stemming does not improve the sentiment classifiers evaluated in their study. In contrast, a sentiment analysis study by Angiani et al. [9] suggests that stemming and stop-words removal does, in general, improve sentiment analysis accuracy.

Sentiment Classification

Sentiment classification aims to detect the subjectivity in text and/or classifying opinions and sentiment inherent in text [4, 10]. Sentiment classification commonly aims to detect the polarity of sentiment in a text into contrasting groups such as negative or positive (or neutral), good or bad, pros or cons, and support or oppose. In this context, we aim to classify student feedback into two classes: negative or positive.

In term of approaches, Alessia et al. [4] suggest that sentiment classification approaches can be classified into three main categories: machine learning, lexicon based, and hybrid (that combines the benefit of both machine learning and lexicon-based approaches). Machine learning approaches utilize term presence and frequency, and negations to classify a collection of text. Lexicon base approaches classify text by detecting the presence of set of terms (i.e., sentiment dictionary) in a text. In this paper, we investigate the accuracy of three supervised machine learning approaches, Naive Bayes, Support Vector Machine (SVM), and Decision Tree in classifying the polarity of student feedback. Our study is similar to study by Laksana and Purwarianti [11] that compare the use of Naive Bayes, Support Vector Machine (SVM), and Decision Tree to classify Indonesian tweeets.

Naive Bayes is known to be an accurate and yet, simple (i.e., computationally efficient) in solving text classification problem. The Naive Bayes Classifier has been found as accurate in classifying text into polar categories: neutral, positive, and negative [12, 13]. Fiarni, Maharani, and Pratama [12] study used Indonesian text (as in our study) and shows that their Naive Bayes classifier implementation is able to classify customers feedback with 89.21% accuracy. Thus, we seek to replicate their success in this study.

Support Vector Machine (SVM) is a supervised machine learning approach that are commonly used to solve classification problem [14]. SVM relies on a set of feature vector x_i with its corresponding value y_i indicating whether the feature vector is within or outside a particular class. An empirical study on the use of SVM to classify sentiment within tweets on a transportation service in Indonesia by Syahputra, Basyar, and Tamba [15] suggest that their SVM classifier implementation was able to accurately classify 91.8% of their testing data.

The last classification method used as a comparison is Decision Tree. Decision tree is used as a predictive model in decision tree learning, mapping observations about an item to conclusions about the item's target value. The root, internal, and leaf nodes of a decision tree structure. It's a tree structure resembling a flow chart, with each internal node representing a test condition on an attribute, each branch represents the test condition's outcome, and each leaf node (or terminal node) is labeled with a class label. The root node is the highest node in the hierarchy. The decision tree is built in a divide-and-conquer manner. Each decision rule is formed by a path in the decision tree. From top to bottom, it generally employs a greedy approach. The decision tree classification technique has two stages: tree construction and tree pruning. The top-down approach is used when building trees. The tree is recursively partitioned during this phase until all data items have the same class label. Because the training dataset is traversed repeatedly, it is extremely computationally intensive [16]. An empirical study by Bayhaqy et al. [17] shows that the accuracy in performing sentiment analysis of decision tree approach outperformed the accuracy of K-NN, and Naive Bayes approaches. Hence, we consider decision tree as a promising contender to the Naive Bayes and SVM approaches in performing sentiment analysis.

Evaluation and Presentation of Output

The last, not the least important, step is to evaluate the accuracy of the selected sentiment classifier approaches and present it in a meaningful and easy to understand manner [4]. As suggested by Yadav and Shukla [18], in this study, we focus on maximising accuracy as it is considered as the most critical measure in classification task.

In terms of data validation, Yadav and Shukla [18] compares the use of k-fold cross-validation and hold-out validation on large data sets. Their empirical results suggest that k-fold cross-validation can be as an alternative to hold-out validation with comparable results. In this study, we use k-fold cross-validation with k=5 following the suggestion from the work of Laksana and Purwarianti [11] and Rodriguez, Perez, and Lozano [19].

METHOD

This section discusses the first four steps to analyze data [4]: data collection, text preparation, sentiment detection, and sentiment classification. Discussion on the presentation of output is left for the following Section Results.

Data Collection

We used student's feedback from the University Surabaya between 30 April 2021 and 11 May 2021. Within the data collection period, we gathered 60,615 student feedbacks. We removed feedback with comments less than 2 characters long and obtained 9,076 comments to be considered. The 9,076 comments were submitted by 1,701 undergraduate students for 469 distinct courses across eight faculties. The 469 courses were delivered by 362 distinct lecturers. Then, we removed duplicate feedback to acquire 3,601 unique feedbacks. Finally, we removed feedback that was not related to the teaching and learning activity such as feedback to the attendance system. In the end, we had 3,570 feedbacks.

We (i.e., two authors) manually labeled the remaining 3,570 feedback to either negative or positive. 198 of the 3,570 feedbacks were labeled by both authors to measure our agreement level. Results shows that our agreement level is 97% of the 198 feedbacks. Nevertheless, when disagreements occurred, we used label as provided by the first author over the second.

Text Preparation

Text preparation aims to transform the collected data to a representation to fit the sentiment analysis processes. This includes reducing and simplifying features within the data. As previously mentioned in literature review, in this study, we performed stop-words removal and stemming. These two text preparation methods are commonly use in text related studies such as in [20].

A stop word is a word that appears frequently and therefore, has little contribution in classifying the intention of a text. Stop-words includes conjunctions (e.g., “dan”, “atau”, “dari”, “adalah”) and words with a single character that are likely available in documents regardless the intent. In this study we used Indonesian stop-words list as provided in Sastrawi project (i.e., an Indonesian open-source library to work with Indonesian text available in https://github.com/sastrawi/sastrawi).

Stemming aims to transform a word to its basic form. Such transformation allows a variety of a single word to be considered as one. For example, stemming transform “run”, “running”, and “ran” to the same basic form “run”. In Indonesian text documents, the stemming process is based on affixes, which can take the form of a prefix, a suffix, or a combination of prefixes and suffixes between the root words. For example, “lari”, “berlari”, and “melarikan” are of the same root word: “lari”. In this study we used Indonesian stemmer as provided by the Sastrawi project (Sastrawi version 1.0.1).

Sentiment Classification

Sentiment classification aims to detect and group intents or tone intrinsic in a text to a set of predefined classes. Prior performing sentiment classification, we performed feature weighting based on the importance of a token in our data collection. Here, we considered a token as a uni-gram word. Token with higher importance (i.e., keywords) will be weighted higher than token with less importance. In this paper, the importance of a token was measured by its $tf.idf$ (i.e., the multiplication products of term frequency and inverse term frequency). Token with higher $tf.idf$ score is commonly considered as more important than token with lower $tf.idf$ score.

The weight of each token was then considered in classifying the sentiment of a text in our data collection. In this study, we compared the accuracy of three supervised learning methods: Naive Bayes, Support Vector Machine (SVM), and Decision Tree. A brief discussion on each algorithm used in this paper is available on the previous section: related works.

In addition to contrasting the performance of Naive Bayes, SVM, and decision tree approach, in this study, we also investigated the use of a vote mechanism that combine the results of the three approaches. Our vote mechanism determined the class of a text (i.e., negative or positive) based on the most agreement in the three classifier results. If any two of the three classifiers consider a text as negative then vote mechanism output will be negative, and vice-versa.

EMPIRICAL RESULTS

This section presents and discusses findings from our experiment. Table 1 list sample data with every possible combination of judgement from the two reviewers. The average word count in the feedback is 9,13 words with standard deviation 8,37 words. The relatively high standard deviation suggests that feedback in our collection varies in length.

As previously mentioned in Related Words, in this work we validated our findings using K-Fold cross validation with k = 5 (i.e., 80% of data is used for training and the remaining is used for testing). Furthermore, here, we evaluated performance based on accuracy where accurate means that the classifier judgment is the same as the manual judgement. Consequently, results are presented in percentage of accuracy for each fold, and we averaged the fivefold results at the end to provide a performance summary.

First, we evaluated the impact of text preparation on classification accuracy that includes stop-words removal and stemming. Table 2 shows impact of each text preparation method on the accuracy of each sentiment classifier results. Our experiment results indicate that, in general, the use of stop-word removal and stemming produce lower accuracy than without using them in all classifiers considered in this study. The average accuracy further highlights that all three classifiers are more accurate when the data is used without any data preparation (NP) compared to when the data is first treated with stop-words removal (SW), stemming (ST), or both stop-words removal and stemming (SW+ST). These findings are in line with findings by Pradana and Hayaty [7] and Hidayatullah, Ratnasari, and Wisnugroho [8] who also found that performing stop-word removal and stemming degraded the accuracy of sentiment analysis on Indonesian text.

TABLE 1. Sample Data with every possible judgement from the two reviewers

Student Feedback	Reviewer 1	Reviewer 2
materi nya sesuai dan metode pembelajaran nya cukup baik	Positive	Positive
Perlu tambahan video materi di setiap pertemuan melalui ULS	Negative	Negative
Dosen memasukin ruangan zoom ketika banyak mhs yang belum masuk. Materi yang disampaikan sesuai, dan memang tidak ada di ULS karena memakai UU	Positive	Negative
Saran saya,ibu tetap mempertahankan cara mengajarnya yang menerapkan sistem tanya jawab yang membuat mahasiswa untuk bisa memberanikan diri menjawab pertanyaan yang diberi sehingga cepat memahami materi.	Negative	Positive

TABLE 2. Impact of text preparation on sentiment classifier accuracy. NP: results without preparation. SW: results with stop words removal only; ST: results with stemming only; SW+ST: results with stop words removal and stemming. All results are in % of accuracy. Numbers in bold indicate the best among results per method and iteration or average.

	Naïve Bayes				**SVM**				**Decision Tree**			
Iteration	**NP**	**SW**	**ST**	**SW+ ST**	**NP**	**SW**	**ST**	**SW+ ST**	**NP**	**SW**	**ST**	**SW+ ST**
1	**83.31**	80.92	82.47	81.49	**90.32**	87.93	89.90	87.80	**84.15**	82.89	83.73	82.33
2	80.95	79.97	**81.23**	79.97	**89.07**	88.51	88.65	88.09	83.19	81.37	**83.33**	82.07
3	**82.49**	80.67	**82.49**	80.81	88.09	86.13	**88.65**	86.41	83.47	83.19	**84.73**	84.17
4	**89.35**	88.93	88.51	88.93	**89.49**	89.07	88.93	87.25	84.31	**84.87**	84.45	83.05
5	**85.43**	86.13	85.29	85.99	**87.81**	86.55	86.55	85.01	81.93	**82.49**	79.80	80.11
AVG	**84.31**	83.32	84.00	83.44	**88.96**	87.64	88.54	86.91	**83.41**	82.96	83.21	82.35

Our results contrast findings from Angiani et al. [9] which argues that stemming and stop-words removal generally improve sentiment analysis accuracy. One of the main differences between our findings is language and therefore, the stop-words dictionary and the stemming algorithm used are also different. This implies the need of a better stop-words and stemming algorithm for Indonesian text. To date, we found no study that contrast or propose better stop-words dictionary and stemming algorithm than the Sastrawi as used in this study.

Next, we evaluated the accuracy of the three sentiment classifiers considered in this paper. Additionally, we evaluated the impact of combining results from the three classifiers through a vote mechanism. For this purpose, we considered results using both stop-words removal and stemming. In Table III shows the average accuracy of the three classifiers plus vote mechanism for all k iteration.

We analyzed the variation amongst the four groups of results as shown in Table III using ANOVA Single Factor. As noted in the last column of Table 3, the four groups of results from all iterations are significantly varied. We then performed post-hoc test using t-test to understand variability between every pair of group results. In Table 3, we reported statistically significant differences ($\alpha < 0.05$) using case superscripts.

Table 3 indicates that, in general, SVM performs better in term of accuracy compared to Naive Bayes and Decision Tree. Exceptions were found in iteration four and five where Naive Bayes performed better accuracy than SVM. Nevertheless, the accuracy differences between the two algorithms in this iteration were not statistically different. In contrast, the decision tree method performed worst accuracy across all methods considered in this paper.

Next, we considered to combine results from all methods using a vote mechanism. If at least two methods classify a comment as “Positive”, the comment will also be labeled as “Positive”. On the other end, if at least two methods classify a comment as “Negative”, the comment will also be labeled as “Negative”. In general, Table 3 shows that the vote mechanism performed less accuracy than SVM as the best among the three considered sentiment classifiers. Nevertheless, the difference was found to be statistically insignificant.

TABLE 3. Percentage of sentiment classifier accuracy. Numbers in bold indicate the best among results per method and iteration or average. P-Value indicates variations between the four groups of results measured using ANOVA Single Factor. The superscripts denote statistical significance difference using t-test paired two sample for means with $\alpha < 0.05$.

Iteration	Naïve Bayes^{n}	SVM^{s}	Decision Tree^{d}	Vote^{v}	P-Value
1	81.49^{sv}	$\mathbf{87.80}^{ndv}$	82.33^{sv}	85.69^{nsd}	0.003
2	79.97^{sv}	$\mathbf{88.09}^{ndv}$	82.07^{sd}	86.41^{nsd}	0.000
3	80.81^{sdv}	86.41^{n}	84.17^{nv}	$\mathbf{86.69}^{nd}$	0.007
4	$\mathbf{88.93}^{d}$	87.25^{d}	83.05^{nsv}	88.65^{d}	0.003
5	$\mathbf{85.99}^{d}$	85.01^{d}	80.11^{nsv}	$\mathbf{85.99}^{d}$	0.005
Average	83.44	**86.91**	82.35	86.69	

Finally, we evaluated in which cases most classifiers went wrong. For this, we selected feedbacks that were incorrectly classified by all three classifiers. We found 220 out of 3,570 feedbacks met this criterion. Table 4 shows some of the feedback and the problem that likely caused the automated classifiers to produce inaccurate classification. In general, we found inaccurate classification were caused by the following three problems:

- Contradicting Statements: some feedback contained positive statements that were actually not available during class and therefore should be classified as negative feedback. For example: "lecturers can explain clearly". The sentiment of such feedback is hard to classify even for human.
- Multiple sentiments: some feedback contained both positive and negative statements. Human may be able to accurately classify the overall sentiment by contrasting the weight of the negative and the positive statements. However, such task is considered hard to be accurately performed automatically.
- Inaccurate manual label: on rare occasions, the automated sentiment classifiers were found to be more accurate than the initial human reviewer label.

TABLE 4. Sample of feedback inaccurately classified by all three automated classifiers.

Manual	Automated	Problem Category	Sample Feedback
Negative	Positive	Positive Statements	"ada ringkasan materi" "dosen dapat menjelaskan dengan jelas" "lebih baik lagi"
Negative	Positive	Multiple Sentiments	"baik dan perlu ditingkatkan" "dosen menjelaskan sudah baik, namun mungkin lebih baik dibuatkan contoh soal, karena bila sekali jalan akan susah bagi kami mahasiswa memahami"
Positive	Negative	Negative Statements, Multiple Sentiments	"kelas sudah efektif dengan lock meeting apabila waktu sudah melebih jadwal"
Positive	Negative	Inaccurate Manual Label	"Materi di sediakan di uls dan jam perkuliahan lebih tepat waktu"

Problems as shown in Table 4 indicate that classifying the sentiment within text in our collection was not easy, even for human. This highlights the many rooms for improvements in future works.

CONCLUSION

In this paper, we aimed to answer the question how to better understand students' sentiments on teaching conducts based on their feedback? Our approach to answer this question is by using sentiment analysis classifier. We empirically evaluated the accuracy of three sentiment analysis classifier: Naive Bayes, Support Vector Machine (SVM), and Decision Tree. We used student feedback in Indonesian as data. First, we investigated the impact of applying stop-words removal and stemming prior classifying feedback. Our empirical results suggest that applying stop-words removal and stemming using the Sastrawi libary (a popular Indonesian stop-words and stemming library) degrades the accuracy of all classifiers evaluated in this study.

With regards to evaluation on the accuracy of the three-sentiment classifier, we found that the SVM classifier performed best compared to the other two (i.e., on average 89% of sentiments were accurately classified using

SVM). In contrast, the decision tree classifier was found to be, in general, statistically worst among the three considered classifiers.

Sequentially, we experimented with aim to combine and leverage results from the tree classifiers using a vote mechanism. Our experiment results suggest that the accuracy of the Vote mechanism was comparable to the SVM accuracy (i.e., not statistically different).

Finally, we evaluated cases where all three classifiers failed to correctly classify feedback. We identify three main root problems: contradicting statements, multiple sentiments, and inaccurate manual label. Apart from the last problem, problems found in our study indicate that automating sentiment analysis to our data collection is a challenging task even for human annotator. Thus, there are many rooms for improvement for future works.

REFERENCES

1. R. Feldman, "Techniques and applications for sentiment analysis," Communications of the ACM 56, 82–89 (2013).
2. P. Gonçalves, M. Araújo, F. Benevenuto, and M. Cha, "Comparing and combining sentiment analysis methods," in Proceedings of the first ACM conference on Online social networks (2013) pp. 27–38.
3. H.-L. Nguyen, Y.-K. Woon, and W.-K. Ng, "A survey on data stream clustering and classification," Knowledge and information systems 45, 535–569 (2015).
4. D. Alessia, F. Ferri, P. Grifoni, and T. Guzzo, "Approaches, tools and applications for sentiment analysis implementation," International Journal of Computer Applications 125 (2015).
5. J. J. Hox and H. R. Boeije, "Data collection, primary versus secondary," (2005).
6. M. Denny and A. Spirling, "Text preprocessing for unsupervised learning: Why it matters, when it misleads, and what to do about it," When It Misleads, and What to Do about It (September 27, 2017) (2017).
7. A. W. Pradana and M. Hayaty, "The effect of stemming and removal of stopwords on the accuracy of sentiment analysis on indonesian-language texts," Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control , 375–380 (2019).
8. A. F. Hidayatullah, C. Ratnasari, and S. Wisnugroho, "The influence of stemming on indonesian tweet sentiment analysis," in Proceeding of International Conference on Electrical Engineering, Computer Science and Informatics (EECSI 2015) (2015) pp. 127–132.
9. G. Angiani, L. Ferrari, T. Fontanini, P. Fornacciari, E. Iotti, F. Magliani, and S. Manicardi, "A comparison between preprocessing techniques for sentiment analysis in twitter." in KDWeb (2016).
10. T. Pranckevicius and V. Marcinkeviˇ cius, "Comparison of naive bayes, random forest, decision tree, support vector machines, and logisticˇ regression classifiers for text reviews classification," Baltic Journal of Modern Computing 5, 221 (2017).
11. J. Laksana and A. Purwarianti, "Indonesian twitter text authority classification for government in bandung," in 2014 International Conference of Advanced Informatics: Concept, Theory and Application (ICAICTA) (IEEE, 2014) pp. 129–134.
12. C. Fiarni, H. Maharani, and R. Pratama, "Sentiment analysis system for indonesia online retail shop review using hierarchy naive bayes technique," in 2016 4th international conference on information and communication technology (ICoICT) (IEEE, 2016) pp. 1–6.
13. P. Gamallo and M. Garcia, "Citius: A naivebayes strategy for sentiment analysis on english tweets," in Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014 (Citeseer, 2014).
14. W. H. Press, H. William, S. A. Teukolsky, W. T. Vetterling, A. Saul, and B. P. Flannery, Numerical recipes 3rd edition: The art of scientific computing (Cambridge university press, 2007).
15. H. Syahputra, L. Basyar, and A. Tamba, "Setiment analysis of public opinion on the go-jek indonesia through twitter using algorithm support vector machine," in Journal of Physics: Conference Series, Vol. 1462 (IOP Publishing, 2020) p. 012063.
16. S. D. Jadhav and H. Channe, "Comparative study of k-nn, naive bayes and decision tree classification techniques," International Journal of Science and Research (IJSR) 5, 1842–1845 (2016).
17. A. Bayhaqy, S. Sfenrianto, K. Nainggolan, and E. R. Kaburuan, "Sentiment analysis about e-commerce from tweets using decision tree, k-nearest neighbor, and naïve bayes," in 2018 international conference on orange technologies (ICOT) (IEEE, 2018) pp. 1–6.
18. S. Yadav and S. Shukla, "Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification," in 2016 IEEE 6th International conference on advanced computing (IACC) (IEEE, 2016) pp. 78–83.

19. J. D. Rodriguez, A. Perez, and J. A. Lozano, "Sensitivity analysis of k-fold cross validation in prediction error estimation," IEEE transactions on pattern analysis and machine intelligence 32, 569–575 (2009).
20. Jimmy, G. Zuccon, B. Koopman, et al., "Payoffs and pitfalls in using knowledge-bases for consumer health search," Information Retrieval Journal 22, 350–394 (2019).
21. V. R. Prasetyo, "Searching cheapest product on three different e-commerce using k-means algorithm," in 2018 International Seminar on Intelligent Technology and Its Applications (ISITIA) (IEEE, 2018) pp. 239–244.
22. V. R. Prasetyo, B. Hartanto, and A. A. Mulyono, "Penentuan pembimbing tugas akhir mahasiswa jurusan teknik informatika universitas surabaya dengan metode dice coefficient," Teknika 8, 44–51 (2019).
23. S. Styawati and K. Mustofa, "A support vector machine-firefly algorithm for movie opinion data classification," IJCCS (Indonesian Journal of Computing and Cybernetics Systems) 13, 219–230 (2019).
24. S. Tan, X. Cheng, Y. Wang, and H. Xu, "Adapting naive bayes to domain adaptation for sentiment analysis," in European Conference on Information Retrieval (Springer, 2009) pp. 337–349.
25. H. Tang, S. Tan, and X. Cheng, "A survey on sentiment detection of reviews," Expert Systems with Applications 36, 10760–10773 (2009).
26. A. Wibawa, A. Kurniawan, D. Murti, R. P. Adiperkasa, S. Putra, S. Kurniawan, and Y. Nugraha, "Naïve bayes classifier for journal quartile classification," International Journal of Recent Contributions from Engineering, Science IT (iJES) 7, 91 (2019).
27. M. Adnan, R. Sarno, and K. R. Sungkono, "Sentiment analysis of restaurant review with classification approach in the decision tree-j48 algorithm," in 2019 International Seminar on Application for Technology of Information and Communication (iSemantic) (IEEE, 2019) pp. 121–126.
28. T. Astuti, H. A. Nugroho, and T. B. Adji, "The impact of different fold for cross validation of missing values imputation method on hepatitis dataset," in 2015 International Conference on Quality in Research (QiR) (2015) pp. 51–55.

AIP Conference Proceedings

scitation.org/journal/apc

Volume 2470

International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)

Leveraging Smart Engineering

Surabaya, Indonesia • 25–26 August 2021

Editors • Joniarto Parung, Nemuel Daniel Pah, Putu Doddy Sutrisna and Marcellinus Ferdinand Suciadi

InCITE 2021

Leveraging Smart Engineering

Faculty of Engineering Universitas Surabaya

UBAYA UNIVERSITAS SURABAYA

AIP Publishing

AIP Conference Proceedings

Table of Contents

INTERNATIONAL CONFERENCE ON INFORMATICS, TECHNOLOGY, AND ENGINEERING 2021 (InCITE 2021): Leveraging Smart Engineering

Conference date: 25–26 August 2021
Location: Surabaya, Indonesia
ISBN: 978-0-7354-4180-4
Editors: Joniarto Parung, Nemuel Daniel Pah, Putu Doddy Sutrisna and Marcellinus Ferdinand Suciadi Suciadi
Volume number: 2470
Published: Apr 25, 2022

PRELIMINARY

Free . April 2022

Preface: International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)

AIP Conference Proceedings **2470**, 010001 (2022); https://doi.org/10.1063/12.0009021

Free . April 2022

Committees: International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)

AIP Conference Proceedings **2470**, 010002 (2022); https://doi.org/10.1063/12.0009708

SMART DESIGN AND INNOVATION

Free . April 2022

How energy efficiency, smart factory, and mass personalization affect companies in industry 4.0

Yu-Chung Tsao, Felix Arril Simbara Barus and Chien-Wei Ho

AIP Conference Proceedings **2470**, 020001 (2022); https://doi.org/10.1063/5.0080526

SHOW ABSTRACT

Free . April 2022

Kansei engineering and product-service systems (KEPSS) integrative framework for customer-centered experience

Markus Hartono

AIP Conference Proceedings **2470**, 020002 (2022); https://doi.org/10.1063/5.0080177

SHOW ABSTRACT

The role of supply chain performance factors to determine the firm performance

Yoni Kristiawan, Edi Purwanto and Rustono Farady Marta

AIP Conference Proceedings **2470**, 020003 (2022); https://doi.org/10.1063/5.0080235

SHOW ABSTRACT

Free . April 2022

Real-time monitoring design for make-to-order industry

Senapati C. Adisasmito, Pradipta Deffinika Pamungkas and Anas Ma'ruf

AIP Conference Proceedings **2470**, 020004 (2022); https://doi.org/10.1063/5.0080747

SHOW ABSTRACT

Free . April 2022

Reducing production process lead time using Value Stream Mapping and Kaizen approaches: A case study in

the musical instrument industry

Indra Setiawan and Hernadewita

AIP Conference Proceedings **2470**, 020005 (2022); https://doi.org/10.1063/5.0080159

SHOW ABSTRACT

Free . April 2022

Simple heuristics for scheduling apache airflow: A case study at PT. X

Hans Tanto, Indriati N. Bisono and Hanijanto Soewandi

AIP Conference Proceedings **2470**, 020006 (2022); https://doi.org/10.1063/5.0081042

SHOW ABSTRACT

Free . April 2022

Quality improvement with PDCA approach and design of experiment method in single socks industry in Indonesia

Purba

AIP Conference Proceedings **2470**, 020007 (2022); https://doi.org/10.1063/5.0080179

SHOW ABSTRACT

Free . April 2022

Integration of corporate social responsibility for improved company performance: Evidence from the Indonesian manufacturing industry

Esti Dwi Rinawiyanti, Xueli Huang and Sharif As-Saber

AIP Conference Proceedings **2470**, 020008 (2022); https://doi.org/10.1063/5.0080731

SHOW ABSTRACT

Free . April 2022

Community mobility during Covid-19 pandemic and tourism performance: Data mining approach

Gunawan

AIP Conference Proceedings **2470**, 020009 (2022);
https://doi.org/10.1063/5.0080170

SHOW ABSTRACT

Free . April 2022

Inventory system improvement for poultry

Noverta Brilly Leksana Putra, Indri Hapsari and Dina Natalia Prayogo

AIP Conference Proceedings **2470**, 020010 (2022);
https://doi.org/10.1063/5.0080160

SHOW ABSTRACT

Free . April 2022

Users' perception of digital prototypes in Indonesian fashion industry: A qualitative study

Christabel Parung and Prayogo Waluyo

AIP Conference Proceedings **2470**, 020011 (2022); https://doi.org/10.1063/5.0080178

SHOW ABSTRACT

SMART MANUFACTURING AND PROCESSES

Free . April 2022

Dynamic study of batch milk cooling process at KUD SAE Pujon

Rudy Agustriyanto, Puguh Setyopratomo and Endang Srihari

AIP Conference Proceedings **2470**, 030001 (2022); https://doi.org/10.1063/5.0080176

SHOW ABSTRACT

Free . April 2022

Recent advances and application of Selective Laser Melting (SLM) technology in the aerospace industry

Anel Yerubayeva, Essam Shehab and Md. Hazrat Ali

AIP Conference Proceedings **2470**, 030002 (2022); https://doi.org/10.1063/5.0080173

SHOW ABSTRACT

POWER SYSTEM AND SMART ENERGY MANAGEMENT

Free . April 2022

Immobilization of xylanase on acid pretreatment bentonite as green biocatalyst

Lieke Riadi, Ruth Chrisnasari, Joshua Kristanto, Cahaya Caesar Bigravida and Meyta Sanoe

AIP Conference Proceedings **2470**, 040001 (2022); https://doi.org/10.1063/5.0080318

SHOW ABSTRACT

Free . April 2022

Rheological behavior and antioxidant activity of carrageenan extracted from Green seaweed (*Eucheuma cottonii*) using alkaline solution at low temperature

Puguh Setyopratomo and Lanny Sapei

AIP Conference Proceedings **2470**, 040002 (2022); https://doi.org/10.1063/5.0080286

SHOW ABSTRACT

Free . April 2022

Simulation of impact azimuth angle on specific energy output of a fixed mounting rooftop PV system in Jakarta, Indonesia

Elieser Tarigan

AIP Conference Proceedings **2470**, 040003 (2022);
https://doi.org/10.1063/5.0080228

SHOW ABSTRACT

Free . April 2022

Rice husk ash for the stabilization of the outer interfacial layer of W/O/W double emulsion

Lanny Sapei, Rudy Agustriyanto, Endang Wahyu Fitriani, Zerravym Levy and Cindy Sumampouw

AIP Conference Proceedings **2470**, 040004 (2022);
https://doi.org/10.1063/5.0080285

SHOW ABSTRACT

The effects of material to solvent ratio on the performances of natural dyes extraction

Putu Doddy Sutrisna, Meyta Sanoe, Rifando Gogo Adiyaksa, Kristina Wahyu Agustine, Hadiyatni Rita Priyantini, Prayogo Widyastoto Waluyo and I. Made Ronyastra

AIP Conference Proceedings **2470**, 040005 (2022); https://doi.org/10.1063/5.0080171

SHOW ABSTRACT

Free . April 2022

Carbon based sulfonated catalyst as an environment friendly material: A review

Putu Padmareka Deandra, Herry Santoso and Judy Retti B. Witono

AIP Conference Proceedings **2470**, 040006 (2022); https://doi.org/10.1063/5.0080728

SHOW ABSTRACT

Free . April 2022

A performance study of

magnetite-lignin composites as photothermal materials in solar steam generation system

Marta Devega Yuharma, W. Widiyastuti, Ni Made Intan Putri Suari and Heru Setyawan

AIP Conference Proceedings **2470**, 040007 (2022);
https://doi.org/10.1063/5.0080989

SHOW ABSTRACT

Free . April 2022

Organosolv lignin from coconut coir as potential biomaterials for sunscreen

Diana Novita Sari, Mahardika F. Rois, W. Widiyastuti and Heru Setyawan

AIP Conference Proceedings **2470**, 040008 (2022);
https://doi.org/10.1063/5.0080768

SHOW ABSTRACT

Free . April 2022

One-step electrochemical synthesis of silica-coated magnetite nanofluids

Delyana Ratnasari, W. Widiyastuti and Heru Setyawan

AIP Conference Proceedings **2470**, 040009 (2022); https://doi.org/10.1063/5.0080982

SHOW ABSTRACT

Free . April 2022

Synthesis and characterization of chitosan-Allium sativum film

Emma Savitri, Maria Irine Tjahayani and Anastasia Apriyani Ngene Say

AIP Conference Proceedings **2470**, 040010 (2022); https://doi.org/10.1063/5.0080203

SHOW ABSTRACT

Free . April 2022

The effect of sodium lauryl sulfate on silica nanofluid stabilization using microbubble method

Ratri Sekaringgalih, I. Made Joni, W. Widiyastuti and Heru Setyawan

AIP Conference Proceedings **2470**, 040011 (2022);

https://doi.org/10.1063/5.0080985

SHOW ABSTRACT

Free . April 2022

Textile wastewater purification using corona discharge method

Rezha Carina Rachmady, Ari Rahman and Wahyu Kunto Wibowo

AIP Conference Proceedings **2470**, 040012 (2022); https://doi.org/10.1063/5.0080740

SHOW ABSTRACT

THE ROLE OF IT IN INNOVATION ENHANCEMENT

Free . April 2022

Hate speech content detection system on Twitter using K-nearest neighbor method

Vincentius Riandaru Prasetyo and Anton Hendrik Samudra

AIP Conference Proceedings **2470**, 050001 (2022);
https://doi.org/10.1063/5.0080185

SHOW ABSTRACT

Free . April 2022

School finder, intelligent recommendation system for elementary school selection

Susana Limanto, Endah Asmawati and Yusuf Wira Kencana Putra

AIP Conference Proceedings **2470**, 050002 (2022);
https://doi.org/10.1063/5.0080461

SHOW ABSTRACT

Free . April 2022

Designing a recommender system based on the application of decision tree algorithm in data mining with KNIME (for recommending the topic of undergraduate's thesis)

Liyansah

AIP Conference Proceedings **2470**, 050003 (2022);
https://doi.org/10.1063/5.0081214

SHOW ABSTRACT

Free . April 2022

Weather image classification using convolutional neural network with transfer learning

Mohammad Farid Naufal and Selvia Ferdiana Kusuma

AIP Conference Proceedings **2470**, 050004 (2022);
https://doi.org/10.1063/5.0080195

SHOW ABSTRACT

Free . April 2022

Student performance prediction in higher education: A comprehensive review

Ellysa Tjandra, Sri Suning Kusumawardani and Ridi Ferdiana

AIP Conference Proceedings **2470**, 050005 (2022);
https://doi.org/10.1063/5.0080187

SHOW ABSTRACT

Free . April 2022

Sentiment analysis on feedback of higher education teaching conduct: An empirical evaluation of methods

Jimmy and Vincentius Riandaru Prasetyo

AIP Conference Proceedings **2470**, 050006 (2022); https://doi.org/10.1063/5.0080182

SHOW ABSTRACT

Free . April 2022

Implementing directed pairwise judgement approach in web-based AHP survey application to reduce inconsistency ratio

Daniel Hary Prasetyo

AIP Conference Proceedings **2470**, 050007 (2022);
https://doi.org/10.1063/5.0080169

SHOW ABSTRACT

Free . April 2022

Leaf geometric properties measurement using computer vision system based on camera parameters

Joko Siswantoro, Ida Bagus Made Artadana and Muhammad Z. F. N. Siswantoro

AIP Conference Proceedings **2470**, 050008 (2022);
https://doi.org/10.1063/5.0080190

SHOW ABSTRACT

Free . April 2022

A design of secure supply chain management system with blockchain technology

Alexander Yohan, Nai-Wei Lo and Kevin Valentino

AIP Conference Proceedings **2470**, 050009 (2022);
https://doi.org/10.1063/5.0080266

SHOW ABSTRACT

Free . April 2022

Implementation of behavior tree for creating an in-game cut-scene

Delta Ardy Prima

AIP Conference Proceedings **2470**, 050010 (2022);
https://doi.org/10.1063/5.0080233

SHOW ABSTRACT

Free . April 2022

The online attendance system models for educational institutions

Andre and Marcellinus Ferdinand Suciadi

AIP Conference Proceedings **2470**, 050011 (2022); https://doi.org/10.1063/5.0080180

SHOW ABSTRACT

Free . April 2022

Genetic algorithm with adaptive diversification and intensification for the vehicle routing problem

Eric Wibisono, Iris Martin and Dina Natalia Prayogo

AIP Conference Proceedings **2470**, 050012 (2022); https://doi.org/10.1063/5.0080197

SHOW ABSTRACT

Resources

AUTHOR

LIBRARIAN

General Information

ABOUT

CONTACT

PRIVACY POLICY

TERMS OF USE

scopus.com/sourceid/26916?origin=recordpage

Scopus Preview

Author Search Sources Creat

Source details

Feedback > Compare sources >

AIP Conference Proceedings

Scopus coverage years: from 1973 to 1978, from 1983 to 1984, 1993, from 2000 to 2001, from 2003 to Present

ISSN: 0094-243X E-ISSN: 1551-7616

Subject area: Physics and Astronomy: General Physics and Astronomy

Source type: Conference Proceeding

View all documents > Set document alert Save to source list

CiteScore 2020
0.7

SJR 2020
0.177

SNIP 2020
0.314