DOI: 10.14421/ijid.2020.09205 IJID (International Journal on Informatics for Development), *e-ISSN: 2549-7448*
Vol. 9, No. 2, 2020, Pp. 87-93

Public Sentiments Analysis about Indonesian Social Insurance Administration Organization on Twitter

Siti Rahmawati[1], Muhammad Habibi[2*]
Departement of Informatics, Universitas Jenderal Achmad Yani, Yogyakarta, Indonesia
[1]strhmwt24@gmail.com, [2]muhammadhabibi17@gmail.com

Article History
Received Oct 31, 2020
Revised Dec 11, 2020
Accepted Dec 18, 2020
Published Dec, 2020

***Abstract*— Insurance Administration Organization, which can be used by all people. However, this organization has received various criticisms from the public through social media, namely Twitter. This study aims to analyze public sentiment about the Indonesian Social Insurance Administration Organization on Twitter. The method used in this research is the Naive Bayes Classifier (NBC) method and uses the Support Vector Machine (SVM) method as a comparison. The amount of data used was 12,990 tweets with a data collection period from September 14, 2019 - February 18, 2020. The study compared the two classifier models built, namely the classifier model with two sentiment classes and four sentiment classes. The accuracy results show that the SVM method has a better accuracy value than the NBC method. SVM has an accuracy value of 63.60% and 82.77% for the two sentiment classes in the four sentiment classifier model. The tweet classification results show that the public's conversation about the Indonesian Social Insurance Administration Organization on Twitter has a negative polarity value tendency.**

Keywords—Naïve Bayes Classifier; Sentiment Analysis; Data Mining; Twitter; BPJS

1 INTRODUCTION

Health is the most important thing for society, and health is also the main thing in life's welfare. Through the Ministry of Health, the Indonesian government said that in 2014 it had promoted a health service called the Healthy Indonesia Program. One of the Healthy Indonesia Programs is the National Health Insurance/ Jaminan Kesehatan Nasional (JKN). As of April 2019, JKN has covered more than 82% of Indonesia's total population, so that it has improved the standard of health in Indonesia [1].

To support the JKN program, the government has formed a social security agency that is legally incorporated. This program was then legalized on October 29, 2011, and formulated in Law Number 24 the Year 2011 concerning the Indonesian Social Insurance Administration Organization. This organization's function is to organize health programs for all Indonesian people also has the objective of realizing the provision of proper health insurance for each participant or family member to fulfill the basic needs of life for the Indonesian population following Law No. 24 of 2011 Article 3. The Indonesian Social Insurance Administration Organization came into effect on January 1, 2014. However, this organization did not run as expected by the Indonesian people. The Indonesian Social Insurance Administration Organization reaps the pros and cons in terms of public services, increased contributions, etc.

Twitter is one of the social media that has the most users compared to other social media. Based on data from Oberlo, monthly active Twitter users in 2019 were 330 million users, and 145 million users used Twitter services every day [2]. Social media, especially Twitter, provides a place to express opinions to users. Public opinion regarding the pros and cons of Indonesian Social Insurance Administration Organization services is widely discussed on Twitter. Evaluation of public tweets related to the Indonesian Social Insurance Administration Organization has not been carried out so far. Tweets must be evaluated and analyzed to see that the tweet tends to positive or negative sentiment. The manual classification process for tweets usually requires extra effort and time. So we need a classification model that can automatically classify tweets into a sentiment class label.

Research related to text classification has been done a lot, including research related to student comments [3], thesis title classification [4], journal classification based on abstracts [5]. The research carried out has differences from existing research. In this study, the polarity of sentiment is described in more detail. Usually, sentiment polarity is only classified into two polarities, namely positive and negative. This study's Sentiment Polity is classified into four classes: happy, sad, satisfied, and disappointed, so that the classification of sentiment polarity in tweets becomes more detailed.

This study aims to develop a classification model for classifying the polarity of sentiments related to the Indonesian Social Insurance Administration Organization's tweet on Twitter. Apart from comparing the two sentiment polarity models' results, this study also uses the Naïve Bayes Classifier (NBC) method and the Support Vector Machine method as a comparison. The NBC method has been widely used in text classification. The use of the NBC method in this study is because this method is a classification method that is generally the most suitable to be applied with results that match expectations [6]; NBC can also work well, even in features that have a strong dependency on the dataset [7]. Again, this method is capable of producing higher classification accuracy with less complexity [8]. Besides, the SVM method is used to compare because this method has been proven to be the method that has the best accuracy in text classification [5], [9].

2 METHOD

This research has several steps, including retrieval of tweet data from the Twitter site, then preprocessing, after which the manual labeling process is carried out before entering the training process. After the classification model is formed in the training process, the classification process is carried out. The research stages can be seen in Figure 1.

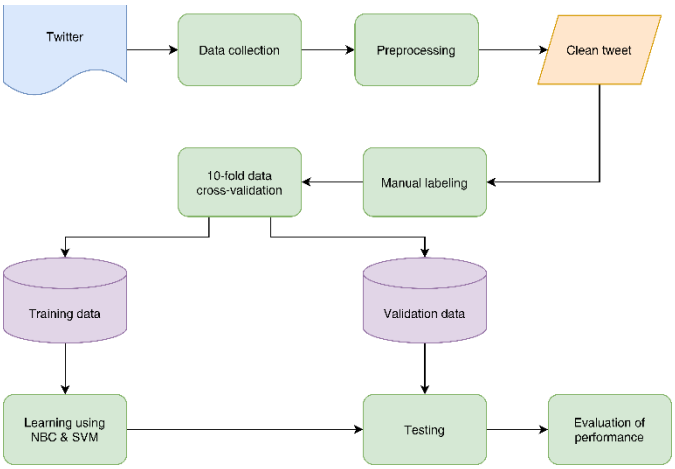

Figure 1. Research stages

2.1 Data Collection and Preprocessing

This study's data were obtained from the Twitter site data extraction, stored in a Microsoft Excel file. Data extraction is where data is analyzed and explored from data sources such as the web or database. The purpose of data extraction is to retrieve relevant information [10]. The tool used to extract data from Twitter is Twitter Scrapper. The data taken is a tweet or re-tweet in Indonesian using keyword #bpjs. The data used in this study were collected from September 14, 2019 - February 18, 2020, with 12,990 tweets.

The next stage is preprocessing; preprocessing is a method done before carrying out the data mining process to produce a more superficial meaning. The preprocessing step is divided into five stages: the first case folding process, which is changing capital letters to lowercase or lowercase letters. The second process of tokenization is splitting tweets into tokens separated by spaces. The third process is stopword removal, namely

removing conjunctions and words that have no meaning. The fourth process is correcting slang words, namely changing non-standard words into standard terms. Lastly, stemming is changing a word into a root word. After the tweet data is clean, the next step is the manual labeling process. Manual labeling is used as training data for the training process using the Naïve Bayes Classifier (NBC) and Support Vector Machine (SVM) method.

2.2 Sentiment Analysis

Sentiment analysis is a subtask of Natural Language Processing (NLP), which analyzes big data to detect people's opinions and emotions [11]. Sentiment analysis is also often referred to as opinion mining, and this analysis is used to help users and customers learn about other consumers' comments or opinions [12]. Also, sentiment analysis can serve as a considered tool for analyzing the products and services of e-commerce on Twitter [13].

2.3 Term Frequency-inverse document frequency

The feature used in this study is the Term Frequency-Inverse Document Frequency (TF-IDF). TF-IDF is a metric commonly used in the text categorization process [14]. TF-IDF is a statistical approach widely used to reflect the importance of a term in a particular corpus document [15]. The TF-IDF weighting scheme assigns weight to term t in document d, as shown by Equation (1) [16].

$$tf.idf_{t,d} = tf_{t,d} \times idf_t \quad (1)$$

The value of $tf_{t,d}$ is the weight of a term t in document d, while idf_t is the inverse document frequency of term t. Equation (2) is an equation for finding the value of idf_t. The value of idf_t is obtained from the result of the logarithm of N divided by df_t. is the total number of documents where df_t is the number of documents containing term t.

$$idf_t = \log \frac{N}{df_t} \quad (2)$$

2.4 Naïve Bayes Classifier (NBC)

The Naïve Bayes Classifier is a classification method rooted in the Bayes theorem. This classifier assumes that the presence of a feature in a class is not related to other features [17]. Equation (3) is the Bayes theorem equation.

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)} \quad (3)$$

Where P(X|Y) is the probability of occurring X if it is known Y. P(Y|X) is the chance of occurring Y if it is known X. P(X) is the probability of occurring X and P(Y) is the probability of occurring Y.

2.5 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a non-probabilistic binary linear classifier. For the training point set (x_i, y_i), where x is the feature vector y is the class. To determine the maximum limit of the hyperplane dividing the points by $x_i = 1$ and $y_i = 1$ [18]. To find the hyperplane Equation (4) is used.

$$w \cdot x + b = 0 \quad (4)$$

SVM is used to create a model classifier or prediction model on the test data by conducting a training process using data consisting of a collection of features and labels. Steps to determine the class in new data, first select the optimal hyperplane. The second step extends the first step for problems that cannot be separated nonlinearly. The final step is to map the data to a high-dimensional, easily accessible space.

2.6 Evaluation

To evaluate the classifier model that has been created, we use the k-fold cross-validation method. This method segments the data into k partitions of equal size. During the process, one of the divisions is selected for testing, while the rest is used for training. This procedure is repeated so k times that each partition is used for the test exactly once. The total errors are determined by adding up the errors for all k processes. To measure the accuracy of the classifier model made, we also use a Confusion matrix. A confusion matrix is an essential tool in the visualization method used in machine learning, which usually contains two or more categories [19].

3 RESULT AND DISCUSSION

In this study, we conducted a training process for making a classifier model using the Naïve Bayes Classifier method. The number of training data used for the training process was 1,442 tweets with manual labeling. The 1.442 tweet data consists of 362 tweets with sad sentiment class, 360 tweets with disappointed sentiment class, 360 tweets with satisfied sentiment class, and 360 tweets with happy sentiment class. Examples of tweets and sentiment classes can be seen in Table 1.

Table 1. Examples of tweets and sentiment classes

No	Tweet	class
1	do not you feel sorry for the spending money to be deducted to increase the cost of the bpjs increase	sad
2	in the past, my father controlled two polyclinics. if you use bpjs, you can't use publicly	sad
3	the root of the bpjs problem is not a small fee, but a messy management	disappointed
4	what are you doing with the bpjs list? you want to die; the liquid bpjs are already dying	disappointed
5	my grandmother used bpjs to stay on sunday, paying cheaply, even though the room was first class	happy
6	hooray, get treatment using bpjs for free eid homecoming anywhere	happy
7	the price of insulin is free, even with a bpjs of 80 thousand / month for the highest; thank god	satisfied
8	my brother used to be almost 8 million because his room was also vip, but yesterday my friend just had a free operation because he used bpjs	satisfied

3.1 Accuracy Score

At this stage, the four sentiment classes' classifier models, namely the disappointed, sad, happy, and satisfied that have been made, are evaluated using the k-fold cross-validation method. The classifier performance evaluation process uses the k-fold cross-validation approach. In this study, k = 10 was used, the 10-

fold cross-validation data will be divided into ten subsets with the same size and different data. In each iteration, one part is used for testing data, and the rest is used for training data.

This section compares the two methods used in making the classifier model—namely the Naïve Bayes Classifier (NBC) method and Support Vector Machine (SVM). The accuracy and F1-Score results on the 10-fold cross-validation of the NBC and SVM methods can be seen in Figure 2 and Figure 3.

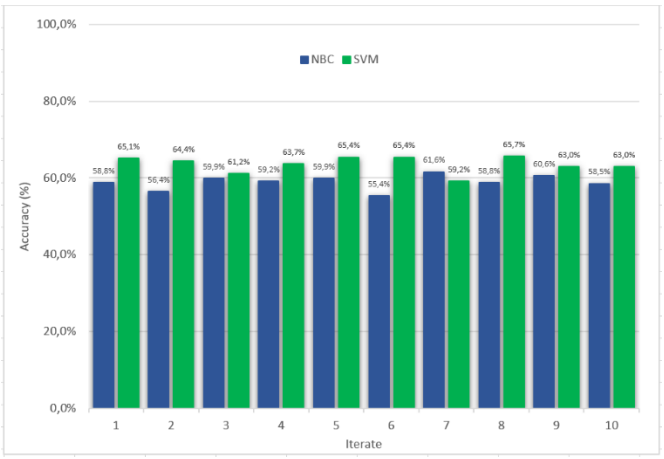

Figure 2. Iteration accuracy for four sentiment classes

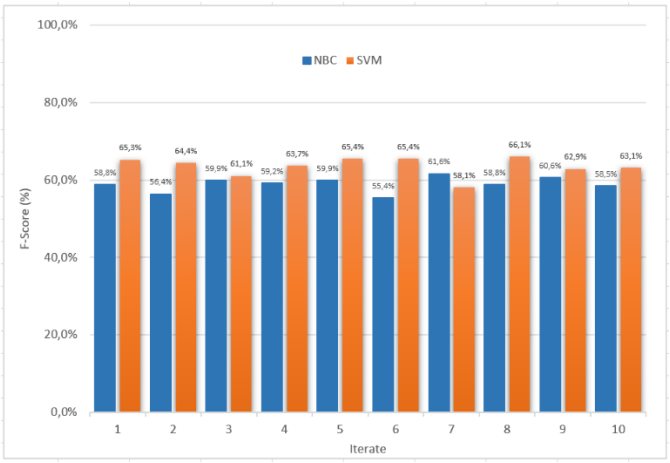

Figure 3. Iteration F1-Score for four sentiment classes

The average confusion matrix value is obtained based on the 10-fold cross-validation algorithm of the naïve Bayes classifier, as shown in Table 1.

Table 2 Confussion Matrix NBC for four sentiment classes

		Predicted			
		Disappointed	**Sad**	**Happy**	**Satisfied**
Actual	**Disappointed**	43	15,8	5,1	9
	Sad	15,9	35,6	7,4	12,2
	Happy	7,3	7,1	36,7	21,4
	Satisfied	5,3	4,1	8,2	54,9

We can see the average confusion matrix value in the 10-fold cross-validation of the support vector machine method for the four sentiment classes in Table 2.

Table 3 Confussion Matrix SVM for four sentiment classes

		Predicted			
		Disappointed	**Sad**	**Happy**	**Satisfied**
Actual	**Disappointed**	46,5	19,1	4,6	5,5
	Sad	15	40,3	7,9	6,9
	Happy	6,7	10,2	44,5	9,6
	Satisfied	5,6	6,2	7,9	52,5

Accuracy and F1-Score are methods that are often used to see classifier performance [20]. The F1-Score determines the predictive power; the higher the F1-Score, the better [21]. Based on the results of the average confusion matrix in Table 1 and Table 2, it is found that the average accuracy value of the NBC method for the four sentiment classes is 58.89%. Meanwhile, the SVM method has a higher average accuracy value of 63.60%. Also, the F1-Score for NBC is lower, namely, 58.53% compared to the F1-Score for SVM, 63.65%.

As a comparison, we made a classifier model of two sentiment classes, namely, positive and negative classes. We try to combine them into two sentiment classes of the four sentiment classes, namely being disappointed and sad, entering the negative sentiment class. Meanwhile, a happy and satisfied sentiment class is included in the positive sentiment class. The two sentiment class classifier models are also created using two methods, namely NBC and SVM. We can see the accuracy F1-Score results for the 10-fold cross-validation of the NBC and SVM methods in Figure 4 and Figure 5.

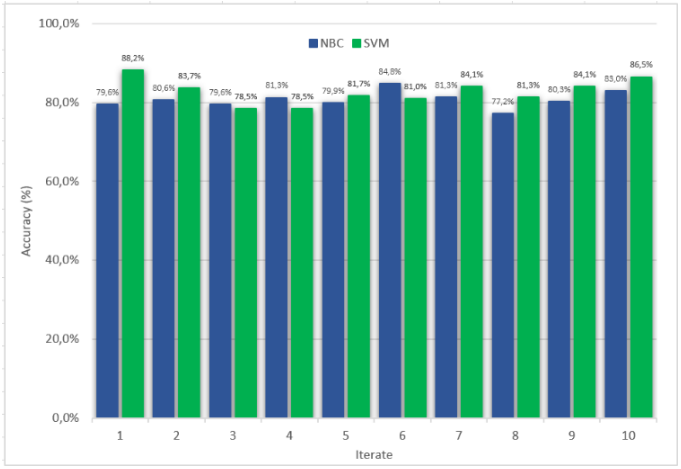

Figure 4. Iteration accuracy for two sentiment classes

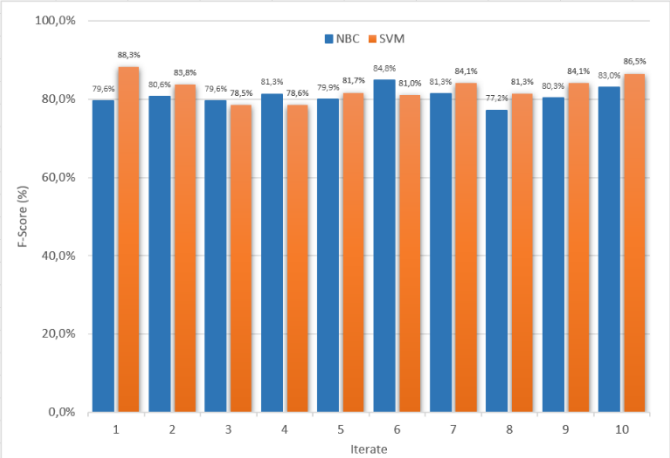

Figure 5. Iteration F1-Score for two sentiment classes

The average confusion matrix value is obtained based on the 10-fold cross-validation of the naïve Bayes classifier method for the two sentiment classes, as shown in Table 3.

Table 4 Confussion Matrix NBC for two sentiment classes

		Predicted	
		Negative	**Positive**
Actual	**Negative**	117,5	31,4
	Positive	24,2	115,9

Meanwhile, we can see the average confusion matrix value in the 10-fold cross-validation of the support vector machine method for the two sentiment classes in Table 4.

Table 5 Confussion Matrix SVM for two sentiment classes

		Predicted	
		Negative	**Positive**
Actual	**Negative**	118,9	25,2
	Positive	24,6	120,3

Based on the average confusion matrix results in Table 3 and Table 4, we get the precision, recall, F1-Score, and accuracy values for the NBC and SVM methods as in Table 5.

Table 6 Performance measures

Parameter	**NBC**	**SVM**
Precision	78,68%	82,68%
Recall	82,73%	83,02%
F1-Score	80,65%	82,85%
Accuracy	80,76%	82,77%

The performance measure of the two sentiment classes' NBC method has a precision value of 78.68%, a recall value of 82.73%, an F1-score value of 80.65%, and an average accuracy value of 80.76%. The SVM algorithm's performance measure has a precision value of 82.68%, a recall value of 83.02%, an F1-score value of 82.85%, and an average accuracy value of 82.77%. Based on the results of the performance measure, it is known that the SVM method has a higher average accuracy value than the NBC method.

The classifier of four sentiment classes and two sentiment classes' average accuracy value shows that the SVM method's average accuracy value has a higher value than the NBC method. This shows that the SVM method performs better than the NBC method[9], [22].

Besides that, the two sentiment classifiers have better average accuracy values because the amount of training data used is deemed insufficient for the four sentiment classes. The more sentiment classes used in the classification, the amount of training data must also adjust. Determining the amount of training data can affect a classifier's accuracy because the training data pattern will be used as a rule to determine the class in the test data.

3.2 Sentiment Analysis Result

This section will discuss the sentiment analysis results regarding the Indonesian Social Insurance Administration Organization data on Twitter. Of the 12,990 tweets with the keyword #bpjs, classification has been carried out using the NBC and SVM algorithms that have been made. We can see the classification results of the tweet data of the four sentiment classes in Figure 6.

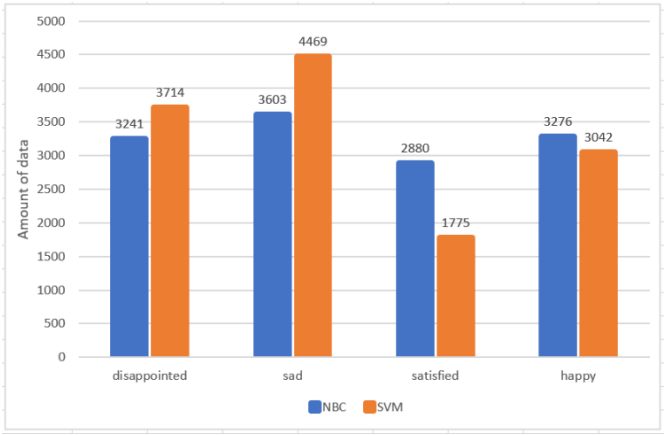

Figure 6. Classification result for 4 sentiment class

Based on Figure 6, the tweets' results using the NBC method show that 3241 tweets belong to the sentiment class disappointed, 3603 tweets sad, 2880 tweets satisfied, and 3276 tweets happy. Meanwhile, the SVM method classification shows that there are 3741 disappointed tweets, 4469 sad tweets, 1775 satisfied tweets, and 3042 happy tweets. The classification of tweets using the SVM shows more data on the disappointed and sad sentiment classes. Meanwhile, in the satisfied and happy class, the SVM tweet classification results show a smaller

amount of data compared to NBC. We can see the results of the classification of the two sentiment classes in Figure 7.

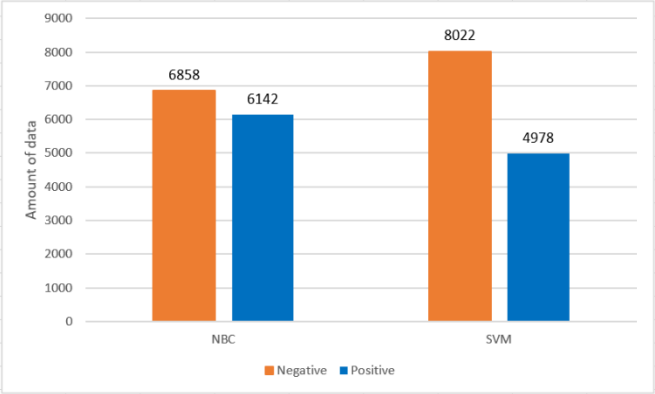

Figure 7. Classification result for 2 sentiment class

If we compare the sentiment classification results, four sentiment classes and two sentiment classes have similarities in the number of tweets' classifications. Both the NBC and SVM methods classify more tweet data into negative classes, namely for NBC as many as 6858 and SVM as many as 8022. Negative sentiment class means the sentiment class, which contains disappointed and sad classes. The sentiment classification results show that the public's conversation about the Indonesian Social Insurance Administration Organization on Twitter negatively predicts sentiment polarity. The classification results identify that the Indonesian people are not satisfied with the Indonesian Social Insurance Administration Organization's services and policies.

4 CONCLUSION

This study succeeded in analyzing public sentiment about the Indonesian Social Insurance Administration Organization on Twitter. In this study, two classifier models were made using the Naive Bayes Classifier (NBC) and Support Vector Machine (SVM) methods. The first classifier model with four sentiment classes, the NBC method has an accuracy value of 59.89% and the SVM method of 63.60%. Meanwhile, the second classifier model with two sentiment classes, the NBC method has an accuracy value of 80.76% and the SVM method of 82.77%. Based on the results of the performance measurement, it is known that the SVM method has a better accuracy value than the NBC method in classifying sentiment on tweets. The tweet classification results show that the public's conversation about the Indonesian Social Insurance Administration Organization on Twitter has a negative polarity value tendency.

ACKNOWLEDGMENT

The research team would like to thank the Department of Informatics and the Study Center and Data Analytic Services of Universitas Jenderal Achmad Yani Yogyakarta, who have supported the research team to add insight and knowledge through this research. Hopefully this research can bring benefits to the progress of the Indonesian nation.

REFERENCES

[1] Widyawati, "Program Indonesia Sehat untuk Capai Tingkat Kesehatan Tertinggi - Sehat Negeriku," *Biro Komunikasi dan Pelayanan Masyarakat*, 2019. [Online]. Available: http://sehatnegeriku.kemkes.go.id/baca/rilis-media/20190521/5530314/program-indonesia-sehat-capai-tingkat-kesehatan-tertinggi/. [Accessed: 23-Sep-2020].

[2] Y. Lin, "10 Twitter Statistics Every Marketer Should Know in 2020 [Infographic]." [Online]. Available: https://id.oberlo.com/blog/twitter-statistics. [Accessed: 23-Oct-2020].

[3] M. Habibi and Sumarsono, "Implementation of Cosine Similarity in an automatic classifier for comments," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 3, no. 2, pp. 38–46, 2018.

[4] A. F. Hidayatullah and M. R. Maarif, "Penerapan Text Mining dalam Klasifikasi Judul Skripsi," in *Seminar Nasional Aplikasi Teknologi Informasi (SNATi) Agustus*, 2016, pp. 1907–5022.

[5] M. Habibi and P. W. Cahyo, "Journal Classification Based on Abstract Using Cosine Similarity and Support Vector Machine," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 4, no. 3, pp. 48–55, 2020.

[6] F. J. Yang, "An implementation of naive bayes classifier," in *Proceedings - 2018 International Conference on Computational Science and Computational Intelligence, CSCI 2018*, 2018, pp. 301–306.

[7] P. Domingos and M. Pazzani, "On the Optimality of the Simple Bayesian Classifier under Zero-One Loss," *Mach. Learn.*, vol. 29, no. 2–3, pp. 103–130, 1997.

[8] N. Salmi and Z. Rustam, "Naïve Bayes Classifier Models for Predicting the Colon Cancer," in *IOP Conference Series: Materials Science and Engineering*, 2019, vol. 546, no. 5, p. 052068.

[9] M. Habibi and E. Winarko, "Analisis Sentimen dan Klasifikasi Komentar Mahasiswa pada Sistem Evaluasi Pembelajaran Menggunakan Kombinasi KNN Berbasis Cosine Similarity dan Supervised Model," Universitas Gadjah Mada, 2017.

[10] I. A. Ahmad Sabri, M. Man, W. A. W. Abu Bakar, and A. N. Mohd Rose, "Web Data Extraction Approach for Deep Web using WEIDJ," in *Procedia Computer Science*, 2019, vol. 163, pp. 417–426.

[11] A. B. Nassif, A. Elnagar, I. Shahin, and S. Henno, "Deep learning for Arabic subjective sentiment analysis: Challenges and research opportunities," *Applied Soft Computing*. Elsevier Ltd, p. 106836, 02-Nov-2020.

[12] H. T. Phan, V. C. Tran, N. T. Nguyen, and D. Hwang, "Improving the Performance of Sentiment Analysis of Tweets Containing Fuzzy Sentiment Using the Feature Ensemble Model," *IEEE Access*, vol. 8, pp. 14630–14641, 2020.

[13] A. Perti, M. C. Trivedi, and A. Sinha, "Development of intelligent

model for twitter sentiment analysis," *Mater. Today Proc.*, Sep. 2020.

[14] N. N. Amir Sjarif, N. F. Mohd Azmi, S. Chuprat, H. M. Sarkan, Y. Yahya, and S. M. Sam, "SMS spam message detection using term frequency-inverse document frequency and random forest algorithm," in *Procedia Computer Science*, 2019, vol. 161, pp. 509–515.

[15] A. Thakkar and K. Chaudhari, "Predicting stock trend using an integrated term frequency–inverse document frequency-based feature weight matrix with neural networks," *Appl. Soft Comput. J.*, vol. 96, p. 106684, Nov. 2020.

[16] A. A. Putri Ratna, A. Kaltsum, L. Santiar, H. Khairunissa, I. Ibrahim, and P. D. Purnamasari, "Term Frequency-Inverse Document Frequency Answer Categorization with Support Vector Machine on Automatic Short Essay Grading System with Latent Semantic Analysis for Japanese Language," in *ICECOS 2019 - 3rd International Conference on Electrical Engineering and Computer Science, Proceeding*, 2019, pp. 293–298.

[17] K. L. Priya, M. S. Charan Reddy Kypa, M. M. Sudhan Reddy, and G. R. Mohan Reddy, "A Novel Approach to Predict Diabetes by Using Naive Bayes Classifier," in *Proceedings of the 4th International Conference on Trends in Electronics and Informatics, ICOEI 2020*, 2020, pp. 603–607.

[18] S. G. Kanakaraddi, A. K. Chikaraddi, K. C. Gull, and P. S. Hiremath, "Comparison Study of Sentiment Analysis of Tweets using Various Machine Learning Algorithms," in *Proceedings of the 5th International Conference on Inventive Computation Technologies, ICICT 2020*, 2020, pp. 287–292.

[19] C. D. Manning, P. Raghavan, and H. Schutze, *An Introduction to Information Retrieval*. Cambridge, England: Cambridge University Press, 2009.

[20] A. A. Anees, H. Prakash Gupta, A. P. Dalvi, S. Gopinath, and B. R. Mohan, "Performance analysis of multiple classifiers using different term weighting schemes for sentiment analysis," in *2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019*, 2019, pp. 637–641.

[21] A. Yadav, C. K. Jha, A. Sharan, and V. Vaish, "Sentiment analysis of financial news using unsupervised approach," in *Procedia Computer Science*, 2020, vol. 167, pp. 589–598.

[22] S. Dey, S. Wasif, D. S. Tonmoy, S. Sultana, J. Sarkar, and M. Dey, "A Comparative Study of Support Vector Machine and Naive Bayes Classifier for Sentiment Analysis on Amazon Product Reviews," in *2020 International Conference on Contemporary Computing and Applications, IC3A 2020*, 2020, pp. 217–220.