International Journal of Frontier Technology and Engineering (IJFTE)
Vol. 01, No. 02, June 2023, pp. 95-103

Understanding People Opinion on Artificial Intelligence Ethics through Machine Learning-based Sentiment Analysis

Pramana Yoga Saputra[1], Arwin Datumaya Wahyudi Sumari[2,3], Yan Watequlis Syaifudin[1], Vian Satria Maulana Navalino[1], and Ekojono[1]

[1] *Department of Information Technology, Politeknik Negeri Malang, Malang, Indonesia*
[2] *Cognitive Artificial Intelligence Research Group (CAIRG), Department of Electrical Engineering, Politeknik Negeri Malang, Malang, Indonesia*
[3] *Department of Electrical Engineering, Faculty of Industrial Technology, Adisutjipto Institute of Aerospace Engineering, Yogyakarta, Indonesia*

Corresponding Author: Pramana Yoga Saputra, pramana.yoga@polinema.ac.id

Received Date: 27-03-2023 Revised Date: 09-04-2023 Accepted Date: 03-05-2023

Abstract

Artificial Intelligence (AI) ethics is so near to human existence. It contains a collection of beliefs, concepts, and methods that utilize generally recognized moral standards to govern moral behavior in developing and using AI technology. Besides providing many advantages to people, the development and use of AI also pose risks that may harm the humankind. At the crossroads between the need for AI and the risks, we want to know what people think about AI. For that purpose, we devised a sentiment analysis system powered by Naive Bayes Classifier and Term Frequency-Inverse Document Frequency (TF-IDF) methods. After analyzing 1.138 data taken from YouTube and Twitter which it is categorized into three opinion labels, namely positive, neutral, and negative, the system can achieve accuracy of 71% with average precision of 66%, average recall of 71%, and average F1-Score of 64% through K-Fold Cross-Validation.

Keywords : AI ethics, Naïve Bayes Classifier, sentiment analysis, TF-IDF

1 Introduction

The Industrial Revolution 4.0 (IR4) marked by the massive use of the Internet-of-Things (IoT), Artificial Intelligence (AI), and Big Data supported by advanced computing technology has brought many advantages to humankind such as the automation of numerous things. The change that is clearly visible due to the IR4 is the advancement of innovation caused by the AI technology's disruptive innovation. It is a condition in which such innovations create new markets, thereby disrupting existing markets, and eventually is able to replace older technologies [1]. Such innovations have influenced the employment absorption in the industries. In this case, a lot of human power have been replaced by automatic machines and robots [2]. Such description demonstrates the critical significance of AI in human existence. While AI, in one side, can be beneficial to people, in another side it also has the potential to be harmful to human existence. Consider the decision that self-driving vehicles would make when in a certain circumstance, confronted with the options of driving towards a child or into a wall in order to save the child's life but possibly may kill or wound the passengers [3]. Another example is the prejudice inherent in computer algorithms used to assess the likelihood that an accused would commit a subsequent crime in an unjust manner against the black race, with black defendants being classified as being at a greater risk than white defendants [4]. The discussion over the potential effect of AI has become a major issue in the international community, and is referred to as AI ethics. As a result of widespread concern about AI ethics, the United Nations Educational, Scientific and Cultural Organization (UNESCO) released the first global standard-setting instrument on AI ethics in the form of guidelines [5]. Müller says in an article that AI poses basic issues regarding what to do with the systems, what to deal with the systems themselves, and what risks the systems will face in the future [6]. As such, this research will use sentiment analysis to ascertain public opinion on the AI ethics.

2 Literature Review

Numerous research has been performed to define AI, its applications, and its advantages as well as its disadvantages. According to the findings from our literature review, AI enables computers to learn and execute cognitive activities previously limited for humans [5]. The aim of developing an AI is to create a "post-human future" in which humanity's current issues must be resolved [7]. At the moment, AI has developed into seven distinct fields, namely expert systems, Natural Language Processing (NLP), speech understanding, robotics and sensory systems, computer vision and scene recognition, intelligent computer-assisted instruction, and neural computing [8]. AI is being developed in a variety of areas. With the growing applications of AI, a variety of good and bad consequences for human existence emerge. This is what eventually led to the development of AI ethics. To discover the public opinion about the development and the utilization of AI, we performed sentiment analysis on public opinion regarding the AI ethics in this research using AI method, especially NLP. We utilize the Naive Bayes Classifier (NBC) to generate the sentiment polarity. According to a study performed by Affandy [9], the average precision accuracy is 88.88%, recall achievement is 90%, and F-measure is 89% when NBC is used. Based on another study in the form of a comparative study conducted by Ipmawati, Kusrini, and Luthfi [10], the accuracy levels obtained by NBC, Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) with the results were 72%, 67.33%, and 56.83%, respectively. Meanwhile, the F-Measures obtained by NBC, SVM and KNN are 72%, 67.3%, 54.5%, respectively. According to another study conducted by Khurniawan [11], the SVM algorithm achieved the highest accuracy rate of 81.70%, followed by NBC with 80.90% and the Decision Tree with 74.55%, when examining the Indonesian people's sentiments and views on the 2019 Corruption Eradication Commission Law Revision.

2.1 Sentiment Analysis

Sentiment analysis is a technique for determining the text's contextual polarity. This method generates the text's positivity, negativity, or neutrality. [12] [13] Sentiment analysis is a subset of NLP because it contains the detection of emotions and opinions based on the textual components. User reviews on websites and comments on social media platforms are often utilized to elicit user feedback. The method of sentiment analysis is shown in Figure1 with the explanations as follows.

- the input of documents into the system is in the form of a corpus, such as Portable Data Format (PDF), Hypertext Markup Language (HTML), and Word.
- the corpus documents are then transformed into text and preprocessed using linguistic techniques such as stemming, tokenization, and part of speech tagging
- additionally, the system may make use of a vocabulary and linguistic resources. The system's primary component is the document analysis module, which makes use of linguistic
- annotations may be added to the whole documents (for document-based sentiments), to the individual phrases (for sentence-based feelings), or to the specific characteristics of a particular entity (for aspect-based sentiments). These annotations are generated by the system and may be seen by the user using a variety of tools [14].

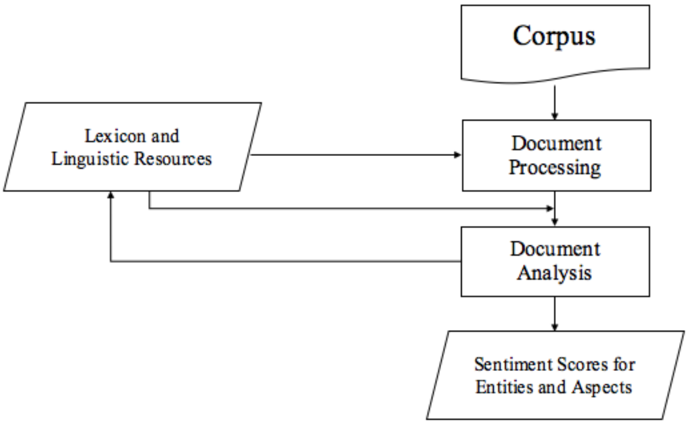

Figure 1: Sentiment Analysis Process

2.2 TF-IDF Algorithm

TF-IDF is a technique for integrating the Term Frequency (TF) and the Inverse Document Frequency (IDF). It serves as a representation of the value of each document in the training data set, forming a vector between documents and words [15] as shown in Equation (1) to Equation (3) [16].

$$W_{tf(pn)} = \frac{n_{pn}}{n} \tag{1}$$

$W_{tf(pn)}$ is the weight of TF, which is given to the p-th word in the n-th document. It's obtained from the fraction of the number of p-th word in the n-th document (n_{pn}) and the number of all words in the document (n).

$$idf_{(p)} = log\frac{N}{df_{(p)}} \tag{2}$$

We also calculate the IDF, which is symbolized as $idf_{(p)}$. IDF is defined as the inverse document frequency weighting on the n-th word. Words that rarely appear in documents have a high inverse document frequency weight. It's obtained from the total number of documents in the collection (N) divided by the number of documents that contain the p-th word ($df_{(p)}$).

$$w_{(pn)} = w_{tf(pn)} x idf_p \tag{3}$$

After we get the TF and IDF, we can calculate the TF-IDF by multiplying $w_{tf(pn)}$ and idf_p.

2.3 Naïve Bayes Classifier

Naive Bayes Classifier (NBC) is a probabilistic classifier that is based on the Bayes' Theorem. The benefit of NBC is that it needs minimal training data to estimate the classification parameters [17]. It is also called as conditional probability model that provides practical learning methods and allows for the combination of previous knowledge and observed data. The NBC technique's basic aim is to determine the likelihood of a particular category in a text document by combining the probabilities of words and categories that is based on the assumption of word independence. Even the NBC algorithm careless with the connection between individual words, it is still often used to categorize text [18]. The NBC's benefits are simple, quick, and accurate [19]. To train NBC, we have to group together all documents that belong to the same category, use the frequency of the terms in the group, and do the computation to obtain the probability with the highest likelihood [20] using Equation (4).

$$P(x_i|y) = \frac{count(x_i, y) + 1}{\Sigma x \in V(count(x_i, y) + 1) + |V|} \tag{4}$$

$P(x_i|y)$ is the probability of occurrence of the word x in the bag of words, which is calculated as a fraction of occurrences of x_i in all documents with topic y. All documents with category y are combined, so that it becomes a combined text with category y . Next we use the frequency of x_i in this combined text to get the maximum likelihood estimate of the probability. V here consists of unions of all types of words in all classes, not only words that are in category y. Furthermore, to avoid the probability of being zero, we use Laplace smoothing.

2.4 K-Fold Cross Validation

K-Fold Cross Validation is a model validation technique that uses k-fold testing to estimate the average performance of a model [21]. K-Fold Cross Validation splits the dataset into k-parts, with one part serving as the test data and the remaining k-1 portions serving as the training data. Figure 2 depicts the example of 5-fold cross-validation in practice.

2.5 K-Fold Cross Validation

K-Fold Cross Validation is a model validation technique that uses k-fold testing to estimate the average performance of a model [21]. K-Fold Cross Validation splits the dataset into k-parts, with one part serving as the test data and the remaining k-1 portions serving as the training data. Figure 2 depicts the example of 5-fold cross-validation in practice.

2.6 Confusion Matrix for Multi-Class Classification

The confusion matrix is used to evaluate a machine learning performance [22]. Precision, re-call, F1-Score, and accuracy will be used to assess performance of the model in this research. According to the confusion matrix, performance assessment requires six variables [23], which include:

- True Positives (TP), positive prediction results and the actual value is also positive.
- True Negatives (TNg), the prediction result is negative and the actual value is also negative.
- True Neutral (TNt), the prediction result is neutral and the actual value is also neutral.
- False Positives (FP), the predicted result is positive but the actual value is not positive, can be neutral or negative.
- False Negatives (FNg), the predicted result is negative but the actual value is not negative, can be neutral or positive.
- False Neutral (FNt), the predicted result is neutral but the actual value is not neutral, can be positive or negative.

Because we have three classes, namely positive, negative and neutral, so we use confusion matrix for multi-class classification. In the multi-class classification, the performance calculation is different from the binary calculation, in the multi-class classification performance calculation, the true positive and true negative locations will be explained in the following table 1.

Table 1: The Results of The Precision of Each Fold and The Average from k=4 to k=10

Actual	Prediction Negative	Prediction Neutral	Prediction Positive
Negative	True Negative (TNg)	False Neutral 2 (FNt2)	False Positive 1 (FP1)
Neutral	False Negative 2 (FNg2)	True Neutral (TNt)	False Postive 2 (FP2)
Positive	False Negative 1 (FNg1)	False Neutral 1 (FNt1)	True Positive (TP)

Based on table 1, we can measure accuracy, precision and recall, as shown in Equation (6) to Equation (12) [24].

$$accuracy = \frac{TP + TNg + FNt}{TP + FNg1 + FNg2 + TNt}.100 \tag{5}$$

$$precision.positive = \frac{TP}{TP + FP1 + FP2}.100 \tag{6}$$

$$precision.negative = \frac{TNg}{TNg + FNg1 + FNg2}.100 \tag{7}$$

$$precision.neutral = \frac{TNt}{TNt + FNt1 + FNt2}.100 \tag{8}$$

$$recall.positive = \frac{TP}{TP + FNg1 + FNt1}.100 \tag{9}$$

$$recall.negative = \frac{TNg}{FP1 + TNg + FNt2}.100 \tag{10}$$

$$recall.neutral = \frac{TNt}{FP2 + FNg2 + TNt}.100 \tag{11}$$

Figure 2: Sentiment Analysis Model

3 Method

In this section, we will explain the methodology to get the sentiment analysis of the prepared dataset. We started from collecting the data from YouTube comments and Twitter tweets. At the end, we calculated the performance of the sentiment analysis system.

3.1 Data Collection Technique

The data for this research acquired from the YouTube comments section and Twitter tweets. The data was gathered using the YouTube Application Programming Interface (API) and the Twitter API. All collected data, as many as 1,033 comments and tweets, was, subsequently stored in the form of an .xlsx file to facilitate the sentiment analysis. The data for this research came from scraping the videos on the YouTube platform named "How Far is Too Far the Age of A.I." from March 6 to 8, 2021, as well as a Twitter tweet with the keyword "ai" on February 16, 2021, and a Twitter account @SteveStuWill on July 2, 2021.

3.2 Data Processing Technique

After the scraped data is gathered, the following step is its labeling. Three annotators are responsible for the labeling process. If two annotators label their comments differently, the third annotator is the determinant of the comments' polarity. Furthermore, the processed data will be tagged as positive, negative, and neutral in the context of sentiment, for the preprocessing step. The preprocessing steps include the case folding, which converts all letters to lowercase, and the lemmatization, which converts the words into their basic words. The next process was to do the stop-words removal to remove meaningless words such as conjunctions, and the last step is the filtering that removes the punctuation marks. After the labeling and the preprocessing steps, the TF-IDF calculation will be performed using Equation (1) to Equation (3), while the last step will be the classification process by using Equation (4). The flowchart on Figure 3 illustrates the general description of the system design process as follows.

1. Taking the data from the video comments column about the AI ethics on the YouTube platform, and the Twitter tweets about AI
2. Data preprocessing including case folding, lemmatization, stop-words removal, and filtering. The preprocessing step does not require stemming because the lemmatization process has been carried out and produced more relevant results [25].
3. Dividing the data using K-Fold Cross Validation into k-times which one part is used as the test data while the rest are used as the training data. Furthermore, the data will be calculated to obtain the TF-IDF values as the basis for to creating the sentiment analysis model using the NBC algorithm. After the model has sufficient knowledge, it then is tested with the test data.
4. The K-Fold process continues to the next fold and uses the second 250 data as test data, and another 1000 training data, and also performs weighting as well as other processes until the last iteration.
5. The last step is to calculate the model performance to obtain the values of the accuracy, precision, recall, and F1-Score values of the NBC-based sentiment analysis model. The final process is to compile and average all the model performances to get the final measurement.

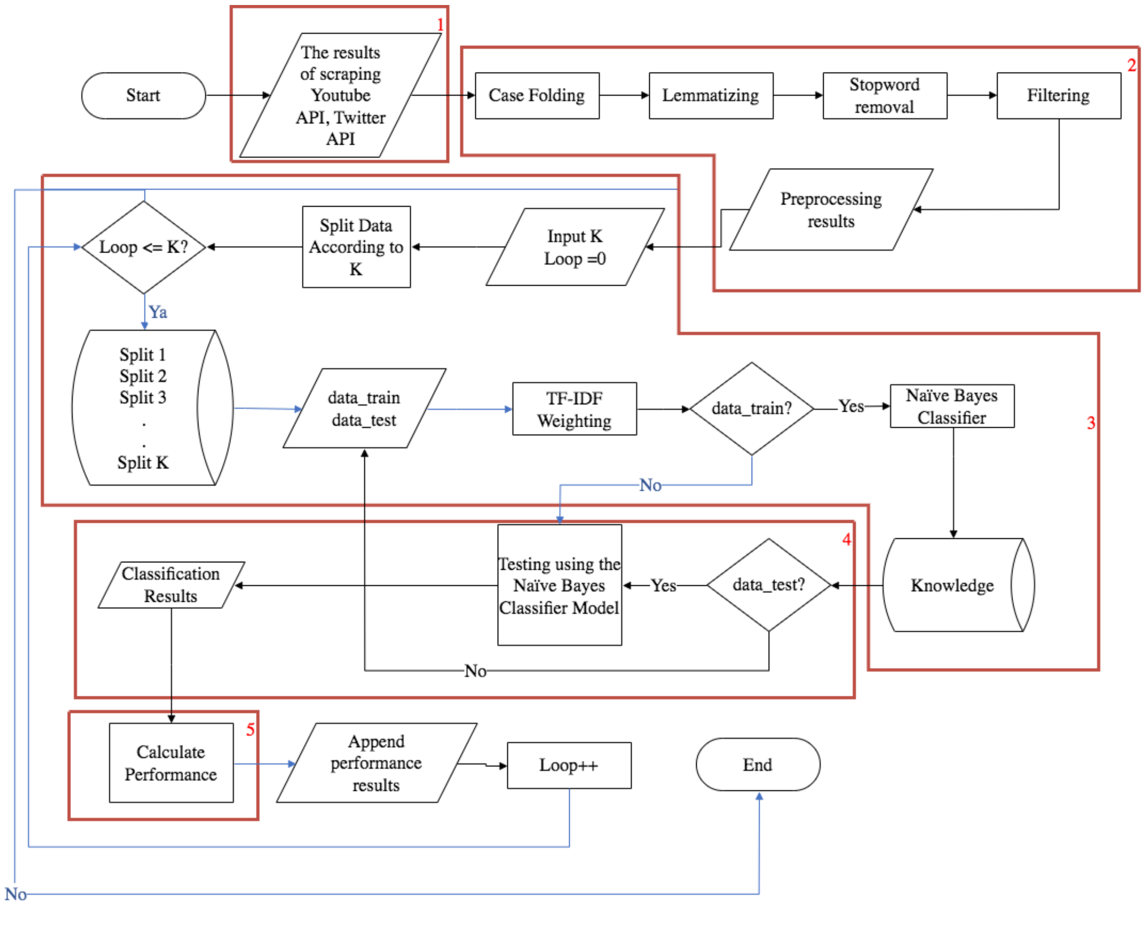

Figure 3: Sentiment Analysis Model

4 Results and Discussion

We are using 1.033 data, taken from tweets on Twitter and YouTube comments. Three linguists have labeled the data, to determine the polarity namely, positive, neutral, and negative. So, the 1.033 data consist of 262 positives, 126 neutrals, and 645 negatives. To test the system's performance, we are using a confusion matrix, that was already been explained before. We utilized K-fold Cross Validation to split the training and test data. From Table 2 until Table 5, respectively, shows the precision, recall, F-1 score, and accuracy results. The following is a discussion of the research findings based on the outcomes of tests on research data performed by researchers.

1. The TF-IDF computation may be used to assign a weight value to each word in the public comments and tweets about AI ethics.
2. The NBC technique may be used to classify comments and tweets on artificial intelligence's ethics.
3. The K-Fold Cross Validation method may be used to evaluate the NBC algorithm's performance.

Table 2: The Results of The Precision of Each Fold and The Average from k=4 to k=10

	K=4	K=5	K=6	K=7	K=8	K=9	K=10
Fold 1	0,72	0,70	0,73	0,72	0,71	0,72	0,72
Fold 2	0,60	0,68	0,73	0,82	0,76	0,71	0,68
Fold 3	0,59	0,59	0,59	0,60	0,65	0,66	0,80
Fold 4	0,68	0,58	0,59	0,54	0,58	0,65	0,56
Fold 5		0,72	0,64	0,54	0,67	0,53	0,54
Fold 6			0,70	0,74	0,58	0,57	0,68
Fold 7				0,66	0,77	0,61	0,50
Fold 8					0,63	0,78	0,62
Fold 9						0,64	0,84
Fold 10							0,65
Average	0,65	0,65	0,66	0,66	0,67	0,65	0,66

Table 3: The Results of The Recall of Each Fold and The Average from k=4 to k=10

	K=4	K=5	K=6	K=7	K=8	K=9	K=10
Fold 1	0,75	0,73	0,75	0,76	0,75	0,76	0,76
Fold 2	0,63	0,72	0,76	0,80	0,74	0,73	0,70
Fold 3	0,70	0,70	0,59	0,63	0,71	0,71	0,80
Fold 4	0,73	0,66	0,71	0,69	0,58	0,60	0,62
Fold 5		0,74	0,76	0,62	0,79	0,69	0,57
Fold 6			0,72	0,83	0,65	0,67	0,82
Fold 7				0,66	0,84	0,73	0,57
Fold 8					0,60	0,86	0,74
Fold 9						0,60	0,89
Fold 10							0,58
Average	0,70	0,71	0,71	0,71	0,71	0,70	0,71

Table 4: The Results of The F1 Score of Each Fold and The Average from k=4 to k=10

	K=4	K=5	K=6	K=7	K=8	K=9	K=10
Fold 1	0,69	0,67	0,69	0,69	0,68	0,69	0,69
Fold 2	0,54	0,66	0,71	0,77	0,69	0,67	0,64
Fold 3	0,62	0,60	0,48	0,54	0,64	0,66	0,76
Fold 4	0,67	0,58	0,62	0,59	0,48	0,49	0,52
Fold 5		0,68	0,69	0,52	0,72	0,58	0,47
Fold 6			0,66	0,78	0,57	0,57	0,74
Fold 7				0,60	0,80	0,66	0,47
Fold 8					0,53	0,81	0,67
Fold 9						0,54	0,86
Fold 10							0,53
Average	0,63	0,64	0,64	0,64	0,64	0,63	0,64

Table 5: The Results of The Accuracy of Each Fold and The Average from k=4 to k=10

	K=4	K=5	K=6	K=7	K=8	K=9	K=10
Fold 1	0,75	0,73	0,75	0,76	0,75	0,76	0,76
Fold 2	0,63	0,72	0,76	0,80	0,74	0,73	0,70
Fold 3	0,70	0,70	0,59	0,63	0,71	0,71	0,80
Fold 4	0,73	0,66	0,71	0,69	0,58	0,60	0,62
Fold 5		0,74	0,76	0,62	0,79	0,69	0,57
Fold 6			0,72	0,83	0,65	0,67	0,82
Fold 7				0,66	0,84	0,73	0,57
Fold 8					0,60	0,86	0,74
Fold 9						0,60	0,89
Fold 10							0,58
Average	0,70	0,71	0,71	0,71	0,71	0,70	0,71

5 Conclusion

The following findings were drawn from the test results acquired from the deployment of the NBC to determine the sentiment of public reactions to AI ethics. There were 645 data with a negative classification, 126 with a neutral classification, and 262 with a positive classification among the 1,033 data collected. The average for precision is 66%, recall 71%, F1-Score 64%, and accuracy 71%. It means that the NBC and TF-IDF techniques is prospective to be used to predict the sentiment of public reactions to the AI ethics.

Acknowledgements

We would like to thank to the State Polytechnic of Malang and the Research Center Office (P2M) who have supported this research project.

References

[1] A. A. Rahman, U. Z. A. Hamid, and T. A. Chin, “Emerging Technologies With Disruptive Effects: A Review”, *PERINTIS eJournal*, vol. 7, no. 2, 2021, pp. 111–128.

[2] M. F. Kamaruzaman, R. Hamid, A. A. Mutalib, and M. S. Rasul, “Comparison of Engineering Skills with IR 4.0 Skills”, *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 15, no. 10, 2019, pp. 15–28.

[3] M. Coeckelbergh, “AI Ethics”, *MIT Press*, 2020.

[4] J. L. Julia Angwin, “Machine Bias”, *ProPublica*, 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

[5] The World Commission on the Ethics of Scientific Knowledge and Technology COMEST, “Preliminary Study on the Ethics of Artificial Intelligence”, *IRCAI*, 2017, https://ircai.org/project/preliminary-study-on-the-ethics-of-ai.

[6] V. C. Müller, “Ethics of Artificial Intelligence and Robotics”, *Stanford Encyclopedia of Philosophy*, 2020, https://plato.stanford.edu/entries/ethics-ai.

[7] A. C. Neubauer, “The future of intelligence research in the coming age of artificial intelligence – With a spe-cial consideration of the philosophical movements of trans- and posthumanism”, *Intelligence*, vol. 87, 2021.

[8] A. Pannu, “Artificial Intelligence and its Application in Different Areas”, *International Journal of Engineer-ing and Innovative Technology (IJEIT)*, vol. 4, no. 10, 2015, pp. 79–84.

[9] A. Affandy and O. Nandiyati, “Sentiment Analysis Berbasis Algoritma Naïve Bayes Classsifier untuk Identifikasi Persepsi Masyarakat Terhadap Produk / Layanan Perusahaan”, *JOINS (Journal of Information System)*, vol. 5, no. 1, 2020, pp. 126–135.

[10] J. Ipmawati, Kusrini, and E. T. Luthfi, “Komparasi Teknik Klasifikasi Teks Mining Pada Analisis Sentimen”, *IJNS - Indonesian Journal on Networking and Security*, vol. 6, no. 1, 2017, pp. 28–36.

[11] F. S. Khurniawan and Y. Ruldeviyani, “Twitter Sentiment Analysis: Case Study on the Revision of the Indonesia’s Corruption Eradication Commission (KPK) Law 2019”, *2020 International Conference on Data Science and Its Applications (ICoDSA)*, 2020.

[12] M. Devika, C. Sunitha, and A. Ganesh, “Sentiment Analysis: A Comparative Study on Different Ap-proaches”, Procedia Computer Science, vol. 87, 2016, pp. 44–49.

[13] S. Qaiser, N. Yusoff, F. K. Ahmad, and R. Ali, “Sentiment Analysis of Impact of Technology on Em-ployment from Text on Twitter”, *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 14, no. 07, 2020, p. 88.

[14] R. Feldman, "Techniques and Applications for Sentiment Analysis", *Communications of the ACM*, vol. 56, no. 4, 2013, pp. 82–89.

[15] A. Aninditya, M. A. Hasibuan, and E. Sutoyo, "Text Mining Approach Using TF-IDF and Naive Bayes for Classification of Exam Questions Based on Cognitive Level of Blooms Taxonomy", *2019 IEEE In-ternational Conference on Internet of Things and Intelligence System (IoTaIS)*, 2019.

[16] M. Riefky and W. Pramesti, "Sentiment Analysis of Southeast Asian Games (SEA Games) in Philippines 2019 Based on Opinion of Internet User of Social Media Twitter with K-Nearest Neighbor and Support Vector Machine", *Jurnal Matematika, Statistika dan Komputasi*, vol. 17, no. 1, 2020, pp. 26–41.

[17] A. P. Wibawa, A. C. Kurniawan, D. M. P. Murti, R. P. Adiperkasa, S. M. Putra, S. A. Kurniawan, and Y. R. Nugraha, "Naïve Bayes Classifier for Journal Quartile Classification", *International Journal of Recent Contributions from Engineering, Science & IT (iJES)*, vol. 7, no. 2, 2019, p. 91.

[18] A. Setiawan, H. Praherdhiono, and S. Suthoni, "Penggunaan Game Edukasi Digital Sebagai Sarana Pembelajaran Anak Usia Dini", *JINOTEP (Jurnal Inov. dan Teknol. Pembelajaran) Kaji. dan Ris. dalam Teknol. Pembelajaran*, vol. 6, no. 1, 2019, pp. 39–44, doi: 10.17977/um031v6i12019p039

[19] D. A. Putri, D. A. Kristiyanti, E. Indrayuni, A. Nurhadi, and D. R. Hadinata, "Comparison of Naive Bayes Algorithm and Support Vector Machine using PSO Feature Selection for Sentiment Analysis on E-Wallet Review", *Journal of Physics: Conference Series*, vol. 1641, 2020, p. 012085.

[20] D. Jurafsky and J. H. Martin, "CHAPTER Naive Bayes and Sentiment Classification", *Speech and Language Processing*, 2020, https://web.stanford.edu/ jurafsky/slp3/4.pdf

[21] S. Raschka, "Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning", *arXiv.org*, 2018, https://arxiv.org/abs/1811.12808v2.

[22] J. Xu, Y. Zhang, and D. Miao, "Three-way confusion matrix for classification: A measure driven view", *Information Sciences*, vol. 507, 2020, pp. 772–794.

[23] S. J. Putra, Y. Sugiarti, G. Dimas, M. N. Gunawan, T. Sutabri, and A. Suryatno, "Document Classification using Naïve Bayes for Indonesian Translation of the Quran", *2019 7th International Conference on Cyber and IT Service Management (CITSM)*, 2019.

[24] A. E. S. Saputro, K. A. Notodiputro, Indahwati, "Study of Sentiment of Governor's Election Opinion in 2018", *International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET)*, ISSN: 2456-3307, vol. 4, no. 11, 2018, pp. 231-238

[25] V. Balakrishnan and L.-Y. Ethel, "Stemming and Lemmatization: A Comparison of Retrieval Performances", *Lecture Notes on Software Engineering*, vol. 2, no. 3, 2014, pp. 262–267.