A contest of sentiment analysis: k-nearest neighbor versus neural network

Fachrul Kurniawan[1], Triyo Supriyatno[2]
[1]Department of Informatics Engineering, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim, Malang, Indonesia
[2]Department of Islamic Education, Faculty of Tarbiyah and Teacher Training, Universitas Islam Negeri Maulana Malik Ibrahim, Malang, Indonesia

Article Info

Article history:

Received Mar 22, 2024
Revised Oct 21, 2024
Accepted Nov 14, 2024

Keywords:

Classification
Deep learning
K-nearest neighbor
Machine learning
Neural network

ABSTRACT

Discourse about public matters often encompasses sentences that address topics emerging within societal contexts, including issues related to Islamophobia. Debates surrounding this subject frequently evoke support and opposition within digital platforms and interpersonal interactions. Categorizing such dialogic expressions within online media facilitates an evaluation of their negative and positive implications. This study employs two distinct methodologies, specifically deep learning and machine learning techniques, to visualize the findings by implementing dual algorithms. According to the comparative analysis, deep learning achieves a higher accuracy rate of 78%, whereas machine learning achieves a rate of 71%. Thus, deep learning is a better method for textual data classification.

Corresponding Author:

Fachrul Kurniawan
Department of Informatics Engineering, Faculty of Science and Technology
Universitas Islam Negeri Maulana Malik Ibrahim
St. Gajayana 50, Malang City, East Java, Indonesia
Email: fachrulk@ti.uin-malang.ac.id

1. INTRODUCTION

Community discourse essentially shapes public perceptions and policy frameworks, especially in contentious topics like Islamophobia that evoke heated debates across social media platforms [1]. In order to devise proper policies for such issues, detailed methodology and analysis should be carried out with a view to evaluating and understanding the pulse of the public sentiment, which essentially requires the power of deep text-mining technologies for the precise identification and classification of discourse [2]. If applied, such technologies hold considerable potential for providing essential insights into the nature of public opinion.

Another valuable contribution of this study is the use of machine learning and deep learning algorithms for classifying conversations with positive and negative sentiments accordingly [3]. Deep learning is the subpart of machine learning that goes beyond the limitations of traditional approaches by modeling their operations just like the human brain and efficiently manages complex data types. Machine learning tries to improve system performance by acquiring knowledge from labeled training datasets [4], [5]. These developments offer a more complex understanding of community perceptions.

Deep learning, as introduced by Geoffrey Hinton in 2006, changed the main paradigm in machine learning by providing an automated feature engineering process and enhancing their performances in a very wide range of applications including speech analysis, text classification, and image recognition [6], [7]. The neural network (NN) architectures it covers, including NN, convolutional neural network (CNN), and recurrent neural network (RNN), prove their worth in both supervised and unsupervised learning scenarios [8]. The mere fact that it does so underlines the potential of deep learning to change how we analyze and comprehend digital communication.

On the other hand, traditional machine learning algorithms decision trees, random forests, support vector machines (SVM), naïve Bayes all require features to be manually predefined and also hardwire the programming to implement specific tasks [9]–[12]. In this review, two different algorithms are considered: NN under deep learning and k-nearest neighbor (K-NN) under traditional machine learning. NN models mimic human neural systems functions by interpreting stimuli to generate actionable responses [13], [14]. However, K-NN is appropriate for jobs of pattern identification and classification since it classifies data by the proximity of nearby data points to one another [15]–[19]. The insight about the comparisons of both the advantages and disadvantages, and their influence on sentiment analysis (SA) approaches.

Except for the underlying approaches, the research tries to compare and critically evaluate the classification performance of NN and K-NN algorithms using the same process phases. The study provided useful insight into their suitability for SA in discourse about Islamophobia through the structured evaluation of advantages and disadvantages of each and of measures of performance. The ultimate objective is to determine which of these techniques, deep learning-NN or traditional machine learning-K-NN, gives more accuracy and speed in the classification of sentiment in online discussions. Comparing these two aimed at improving our understanding of the technological foundation of SA and how it informs public and policy responses.

2. METHODS

The research method has been accordingly planned in such a way that it captures the aims of the research with step-by-step clarity. It describes a chronological sequence of steps involving the careful collection of data, processing of the same, and ending with the derivation of key research findings. Primarily, this is aimed at equipping the researcher with all that they may require in order to effectively undertake the research. Figure 1 shows a graphic that explains this organized research method and provides a clear guide for understanding the study's complex way of doing things.

Figure 1. Research process

2.1. System design

In that respect, systemic architecture provides the roadmap in a planned manner to the flow of work right from the very beginning to the very end of results acquisition. The process starts with structured data gathering from Facebook. Data in the above form are sorted warily into different positive and negative groups, which would be useful in further steps of analysis. Figure 2 shows how this part of the system design provides for a stepped process of data flow management and the steps involved in the classification of data. Figure 2 shows useful information on various layers at which the system architecture and stages of progress occur. It depicts how integrated data collection and data categorization processes are in ensuring that information clarity is cardinal to both research output and analytics insight.

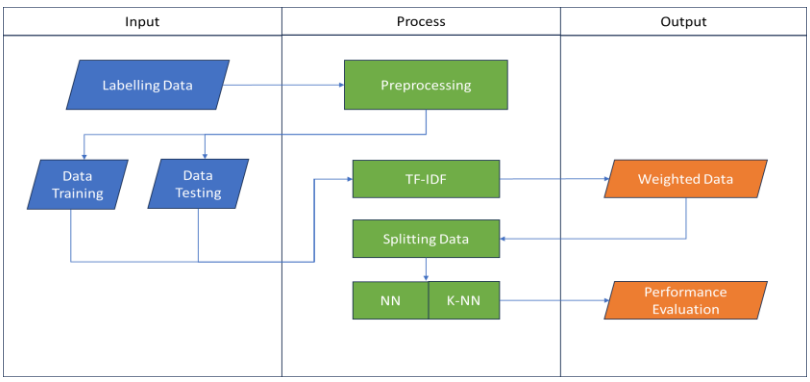

Figure 2. System design

2.2. Data collection

For this study, publicly available discussion data were used, collected from the social media platform Facebook. Due to the current restrictions Facebook puts on application programming interface (API) availability for non-user-owned accounts, this collection process required a purely manual approach by means of copy-and-paste methods. Such an exhaustive method was fundamental to capturing varied threads of discourse that encompass different topics relevant to the focus on Islamophobia. The dataset had 10,997 instances initially, represented majorly by chat messages and comments. After a deliberate elimination of duplicates, a database of 4,339 unique cases was retrieved to further the analysis. Following the refinement process, a language detection methodology was used to classify the dataset into the different linguistic varieties present. The most recurrent was English at 3,854 occurrences, which really pointed out that it is the premier language of communication in the dataset. Indonesian had 143 occurrences, while all the rest combined had only 31 instances. Above have meaningfully informed about the makeup and composition of the dataset and prepared it for further preparation and analysis steps.

Labeling data requires a very precise process in which operations should strictly follow the linguistic features and classification criteria developed by experts. The ten representative data points were randomly sampled from the dataset by the researchers, from which the language specialists rigorously analyzed and came up with comprehensive protocols for labeling. This model gave consistency and accuracy for the labeling techniques across the data, thus enabling the researchers to easily categorize and analyze sentiments and themes related to Islamophobia on social media channels. This labeled dataset came to be the backbone for further machine learning and deep learning analysis, trying to show certain trends in the dynamics of public conversation and assessment.

Preprocessing is an important step in any research process, where raw data is cleaned and processed into a refined format to make it understandable. In this research work, the preprocessing part involves four steps to improve data quality for further processing. The cleaning and preliminary preparation of data aim at removing unwanted items like numbers, punctuation, and emoticons [20]. While these factors are usually dominant in social media data, they only introduce noise and obscure meaningful patterns in the analysis. In removing them systematically, researchers ensure that the dataset focuses on relevant textual information, therefore enhancing the precision of further analyses. Case normalization is done after data cleansing; the text should be standardized by transforming all the characters of the text to their lowercase equivalents [21]. This uniformity at this stage is very important while retrieving and processing documents effectively since the elimination of incorrect phrases can result in duplicate or misleading results of searches. The words "I" and "shall" could have different meanings according to context; however, the language sustains an underlying meaning that consistency sometimes shows the key to effective gathering of data.

Tokenization is an important step in breaking down text information into individual words or tokens, which is essential in applying metrics such as term frequency-inverse document frequency (TF-IDF) [22]. By segmenting data into tokens, researchers enable quantitative analysis based on the frequency and distribution of specific terms within the dataset. This process lays the foundation for understanding word significance and their contextual relevance in relation to the greater body of collected data.

Stemming, the final step of preprocessing discussed here, is a process of bringing words to their root form through the removal of affixes: prefixes, infixes, suffixes, and confixes [23]. This technique aims at reducing words of similar meanings into a common word, hence reducing the dimensionality of the dataset and enhancing further computational analysis efficiency. By normalizing words into their root forms, stemming allows for a finer-tuned and coherent dataset; this enables better analysis of the usage of language and sentiment.

TF-IDF is the statistical method for measuring word importance in specific documents with respect to the whole corpus [24]. This scheme thus assigns weights to the words based on their frequency of occurrence in the document and focuses on terms that best represent the matter of the document. Larger values of TF-IDF reflect terms that are frequently appearing within the given document but relatively less frequent in the entire dataset; these indicate the defining feature concerning the thematic focus or the sentiment of the document. In other words, preprocessing is quite crucial in preparing a dataset for full analysis, enhancement of quality, and standardization relevant to study purposes. These procedures will ensure that subsequent analytical methodologies, such as machine learning or statistical models, act upon the thus-prepared data to provide useful insights that assist in driving well-informed decisions in the context of the study.

Categorization in SA employs several different algorithms, all with different ways of mining meaningful knowledge from texts. One of the big artilleries in SA is the NN, initially developed by Warren McCulloch and Walter Pitts in 1943 [25]. In developing the NN, it was designed to mimic the human nervous system and thus relied upon a network of interconnected neurons for information processing to resolve complex problems related to data mining, text analysis, and image recognition, among others. Central to NN operations is the backpropagation algorithm, which is pivotal for adjusting weights iteratively to minimize errors during training [26]. This iterative adjustment enhances the algorithm's ability to provide accurate outputs aligned with expected values, which is crucial for optimizing classification outcomes in SA tasks [27].

On the other hand, the K-NN is a proximity-based learning classification. The algorithm classified new observations through their similarities in the previously labeled examples of observations [28]. In this regard, K-NN is considered a supervising learning technique, where it learns from the prelabelled training datasets through taking the type of new instances on the nearest neighbors with respect to a set of training observations. The main challenge with 'K-NN' classification involves choosing 'k', which balances the number of neighbors to be considered [29]. A large value of k smooths out the noise, though on the other side, it contributes to an over-smoothing effect. On the other side, a small 'k' gives less smoothed predictions with vulnerability to outliers.

However, their effectiveness in respect to both NN and K-NN depends on the quality of the training data and properties of the data. Due to the hierarchical structure of NN, combined with learning processes, NN become very good at recognizing complex patterns and relations inside the data. On the other hand, the openness and reliance on local similarity make K-NN very useful for some classification tasks. It is particularly good in those cases in which the points constitute discrete or discontinuous clusters. This is because one can only tell, by experience, about the adequacy of various algorithms to several specific needs or attributes sought within the analyzed dataset, having deep knowledge of their respective merits and demerits. In addition, NNs and K-NN are among a broad range of algorithms for classification that could be used in SA, and each will have its relative merits based on the different factors involved in a dataset, such as complexity, size, and noisiness. These techniques surely keep evolving with each step-in machine learning and artificial intelligence, probably hybrid models or new ways to obtain better performance and efficiency in the broad applications and contexts of SA.

2.3. Evaluation

A confusion matrix serves as an essential instrument for evaluating the effectiveness of classification algorithms or models when ground truth data is accessible [30]. To systematically illustrate the model's predicted ability by distinguishing between correct and incorrect classifications. Table 1 illustrates the core concepts within the confusion matrix. From Table 1, true positives (TP) and true negatives (TN) represent correct predictions, where the model accurately identifies positive and negative instances, respectively. Conversely, false positives (FP) and false negatives (FN) represent inaccurate predictions where the model misclassifies positive instances as negative and vice versa.

Table 1. Confusion matrix

Actual class	Prediction class	
	Positive	Negative
Positive	True positive	False negative
Negative	False negative	True negative

Classification performance metrics from the confusion matrix encompass accuracy, precision, recall, and the F1-score, all articulated as percentages to quantify the model's effectiveness in distinguishing between classes. Precision (1) quantifies the ratio of accurately predicted positive cases to the total instances anticipated as positive, emphasizing the model's accuracy in optimistic predictions instances predicted as positive, highlighting the model's exactness in positive predictions. Recall (2) gauges the fraction of correctly predicted positive instances out of all actual positive instances, emphasizing the model's ability to capture all positives. Accuracy (3) calculates the ratio of correctly predicted instances (both positive and negative) to the total number of instances, providing an overall assessment of the model's correctness. The F1-score (4) harmonizes precision and recall, offering a balanced measure that considers both metrics' contributions to the model's performance.

$$Precision = \frac{TP+TN}{(TP+FP)} \times 100\% \tag{1}$$

$$Recall = \frac{TP}{(TP+FN)} \times 100\% \tag{2}$$

$$Accuracy = \frac{TP+TN}{(TP+FN+FP+TN)} \times 100\% \tag{3}$$

$$F\ Measure = 2 \times \frac{Precision\ x\ Recall}{(Precision+Recall)} \times 100\% \tag{4}$$

These metrics collectively evaluate how well a classification model performs, informing researchers and practitioner about its strengths and areas needing improvement in various real-world applications of SA and beyond.

3. RESULTS AND DISCUSSION

This study investigated the effects of SA algorithms on public health discourse extracted from social media platforms, focusing on the Indonesian Ministry of Health's Facebook account. While earlier studies have explored SA in various contexts, few have explicitly addressed its application in assessing public health perceptions through social media platforms like Facebook. The execution of the two algorithms was conducted utilizing the Python programming language, adhering to the procedural stages delineated in the preceding discourse. A total of 2,000 data points, originating from conversational sentences on the Indonesian Ministry of Health's Facebook account, were employed to evaluate the system. Subsequently, the data was bifurcated, with 80% allocated to training and 20% to testing. The distinction during the implementation phase lies solely in applying the two algorithms. The subsequent sections delineate the implementation procedures for the NN and K-NN algorithms.

In implementing NN, the initial determinations involve the number of nodes in the hidden layer, the epoch value, and the learning rate. The following initializations were employed in this study: the number of nodes is 10, and the learning rate is 0.1. The resultant accuracy, precision, recall, and F1-score are elucidated in Table 2, referring to the initializations above.

Table 2. Measurement results of NN implementation

Initial determinations	Accuracy	Precision	Recall	F1-score
Number of nodes: 10 learning rate: 0.1	0.78	0.80	0.79	0.80
Number of nodes: 10 learning rate: 0.2	0.77	0.79	0.77	0.79

Table 2 shows that the maximum accuracy attained is 0.78, corresponding to a configuration consisting of 10 nodes and a learning rate of 0.1. Subsequently, the chosen test outcomes will be assessed using a confusion matrix classification method. The subsequent findings delineate the confusion matrix's results, which were obtained utilizing a 0.1 learning rate, ten nodes, 80% training data, and 20% testing data partition. Figure 3 shows the results of NN.

```
[[180  37]
 [ 50 133]]
              precision    recall  f1-score   support

     negatif       0.78      0.83      0.81       217
     positif       0.78      0.73      0.75       183

    accuracy                           0.78       400
   macro avg       0.78      0.78      0.78       400
weighted avg       0.78      0.78      0.78       400
```

Figure 3. The results of NN measurement

Upon examining the calculations using the confusion matrix, 180 negative class predictions are discerned. This result indicates the algorithm's proficiency in accurately predicting positive and negative classes. The subsequent section presents a manual computation of the metrics, including accuracy, precision, recall, and F1-score as in (5)-(8).

$$Precision = 180/(180 + 55) \times 100\% = 0.765 \times 100\% \approx 76.5\% \tag{5}$$

$$Recall = 180/(180 + 37) \times 100\% = 0.829 \times 100\% \approx 82.9\% \tag{6}$$

$$Accuracy = (180 + 133)/(180 + 133 + 37 + 50) \times 100\% = 0.776 \times 100\% \approx 77.6\% \tag{7}$$

$$F1 - score = 2 \times (0.765 \times 0.829)/(0.765 + 0.829) \times 100\% = 0.796 \times 100\% \approx 79.6\% \tag{8}$$

Classification employing the K-NN algorithm is achieved by identifying the predominant occurrence of data within a specified number of proximal neighbors. In this investigation, the chosen number of neighbors is three. Initially, data weighting is necessitated, serving as the foundation for determining the nearest neighbor values. Subsequently, the performance of the implemented methodology is evaluated utilizing a confusion matrix. This evaluation aims to ascertain the precise measurement values within the system. The conducted measurements on K-NN algorithms yield the subsequent values in Figure 4.

```
[[161  55]
 [ 60 124]]
              precision    recall  f1-score   support

           0       0.73      0.75      0.74       216
           1       0.69      0.67      0.68       184

    accuracy                           0.71       400
   macro avg       0.71      0.71      0.71       400
weighted avg       0.71      0.71      0.71       400
```

Figure 4. The results of the K-NN measurement

Figure 4 presents the measures using the K-NN algorithm. Such measures enable the extraction of a confusion matrix applied to metrics calculation of accuracy, precision, recall, and F1-score. These are obtained in the use of formulae, as presented in the previous chapter. Thereby, the values of accuracy, precision, recall, and F1-score will be obtained based on the confusion matrix values as in (9)-(12).

$$Precision = (161)/(161+60) \times 100\% = 0.728 \times 100\% = 73\% \tag{9}$$

$$Recall = (161)/(161+55) \times 100\% = 0.745 \times 100\% = 75\% \tag{10}$$

$$Accuracy = (161+124)/(161+124+55+60) \times 100\% = 0.712 \times 100\% = 71\% \tag{11}$$

$$F1-score = 2 \times (0.73 \times 0.75)/(0.73+0.75) \times 100\% = 0.739 \times 100\% = 74\% \tag{12}$$

Our results show that NN and K-NN can classify the sentiment in Indonesian public health discussions well. The best result for the NN model can reach 78% with ten nodes and 0.1 as the learning rate. It can be seen that the K-NN algorithm showed quite a good performance, at an accuracy of 71% with three nearest neighbors. This implies that though NN performs best concerning precision and recall, K-NN is doing great at proximity-based classification tasks.

Table 3 shows the classification results of both methods NN and K-NN. Each of the methods has been run using 10 different sample data in order to see various classifications. Table 3 also presents that the classification results obtained from the test of each algorithm are in concert with the prelabeled training data provided by the expert. Even as the metrics from the experiments conducted do not show very high values, the results of classification are always in line with the training data, which means both algorithms had effective implementations.

Our findings concur with the available literature that machine learning algorithms, especially, play a very crucial role in SA. It is illustrated from the study that each of these approaches provides distinct advantages in the fact that both NN and K-NN algorithms attain respectable accuracies of classification 78% for NN and 71% for K-NN. The NN model is adaptable and can change weights based on the subtleties of the training data; hence, it is befitting for intricate sentiment patterns of health-related discussions [31]. Contrarily, the simplicity of K-NN and reliance on nearest neighbors make a much more straightforward approach to initial sentiment assessments so valuable in a much faster analysis of big volumes of data [32].

The comparison of the obtained results with previous studies gives complex views about the approaches applied in the field of SA. The improved performance in real applications may not indicate an increase in the accuracy of the classification metrics. Since the NN model can handle complex patterns, this promises more deep SA processes, especially in those fields that put high demands on contextual understanding. On the other hand, the efficiency of K-NN concerning proximity-based classification underlines its application for rapid sentiment estimations, though it is less sensitive to subtle shifts in sentiment.

The present study developed a comprehensive dataset of one social media platform for a fixed period. It is, nevertheless, a very specific dataset on Indonesian health discourse on Facebook, which has its generalisability across platforms or geographies rather circumscribed. Future research should expand these datasets to a diverse set of social media sources and languages in order to validate findings across broader contexts.

These results indicate the necessity to continue the search for other methods of SA and data extension to multilingual and multicultural contributions. Advanced natural language processing (NLP) techniques may be integrated into future studies to further enhance sentiment classification accuracy between platforms. It could also describe temporal trends and dynamic shifts in the ways public view public health.

Thus, our study shows that machine learning algorithms such as NN and K-NN represent a feasible tool for performing SA in public health discourse on social media. The findings give an insight into the application of the algorithms in understanding the public perception and sentiment, thus serving as an input for targeted health communication strategies. More diversified datasets and further advances in analysis techniques will provide a greater robustness of SA on health-related social media content for further research in the future.

Table 3. Classification results in 10 different data samples

No		Text	Expert	NN	K-NN
1	In Indonesian	*Marilah kita berdoa agar seluruh rakyat Indonesia positif Islamophobia aamiin.*	positive	positive	positive
	In English	Let us pray for all Indonesians to be positively impacted by Islamophobia. Amen.			
2	In Indonesian	*Ya Alloh lindungilah rakyat Indonesia dari Islamophobia.*	positive	positive	positive
	In English	Oh Allah, protect the people of Indonesia from Islamophobia.			
3	In Indonesian	*Rintangan terbesar tangani Islamofobia adalah perbedaan definisi dan konteks.*	positive	positive	Positive
	In English	The biggest challenge in addressing Islamophobia is the difference in definitions and contexts.			
4	In Indonesian	*Pembohongan publik, udah pake masker udah jaga jarak, tapi tetep aja nambah, ini konspirasi Islamophobia.*	positive	positive	positive
	In English	Public deception: even after wearing masks and maintaining distance, cases still increase. This is an Islamophobia conspiracy.			
5	In Indonesian	*Alhamdulillah peningkatan pemahaman Islamophobia semakin banyak.*	positive	positive	positive
	In English	Praise be to Allah, understanding of Islamophobia is increasingly improving.			
6	In Indonesian	*Obat belum ketemu yg sembuh bnyak.. aneh.*	positive	positive	positive
	In English	No medicine has been found, yet many cases have been cured... suspicious.			
7	In Indonesian	*Di negeri wkwkland pendapat pakar tak akan didengar. Jangankan pakar, Tuhannya saja tidak. yg didengar hanya #buzzerp yg selalu kreatif dlm memancing emosi ummat Islam melalui Ulamafobia dan Islamofobia meski bermodalkan hoax #kamisetiabersamaibhrs.*	positive	positive	positive
	In English	In this country of "clownland," expert opinions are never heard. Not even God's opinion is considered. What's heard is only the buzzers, who are always creative in provoking the emotions of Muslims through anti-scholar and Islamophobia narratives, even if it's all based on hoaxes. #westandwithscholars.			
8	In Indonesian	*Maka, dari sini sudah bisa ditebak bahwa mereka yang benci Islam dan menjadi penjaga ideologi kapitalisme adalah muslim yang berkarakter islam moderat. mereka anti-syariah kafah dan mengidap Islamofobia.*	positive	positive	positive
	In English	From this, it's clear that those who hate Islam and become defenders of capitalist ideology are Muslims with a moderate Islamic character. They are anti-Sharia, ambiguous, and foster Islamophobia.			
9	In Indonesian	*Slogan Islamofobia disembur dengan cat berhampiran pusat Islam brixton utara. #fmtnews.*	positive	positive	positive
	In English	Islamophobia slogans were erased with paint near the center of Brixton North. #fmtnews.			
10	In Indonesian	*Pemerintah kota London mengecam slogan Islamofobia.*	positive	positive	positive
	In English	The London city government erased Islamophobia slogans.			

4. CONCLUSION

The conclusion drawn from this study asserts that, in the context of classifying public discourse about Islamophobia, the accuracy achieved by employing machine learning techniques is surpassed by that of deep learning approaches. Experimental results indicate a 71% accuracy rate for machine learning, while deep learning demonstrates a superior performance with a 78% accuracy rate. This discrepancy in accuracy amounts to a 6% difference, with the deep learning method displaying a higher percentage. As expounded in preceding chapters, deep learning constitutes an advanced development of machine learning techniques. This investigation substantiates the notion that deep learning implementation yields more precise outcomes than machine learning, as evidenced by the accuracy rates associated with applying both models. Consequently, it is advisable to utilize deep learning methodologies for text-based classification tasks that necessitate elevated levels of accuracy, as opposed to alternative approaches.

ACKNOWLEDGMENTS

Universitas Islam Negeri Maulana Malik Ibrahim Malang who has supported this work.

REFERENCES

[1] Y. A. Ahmed, M. N. Ahmad, N. Ahmad, and N. H. Zakaria, "Social media for knowledge-sharing: a systematic literature review," *Telematics and Informatics*, vol. 37, pp. 72–112, Apr. 2019, doi: 10.1016/j.tele.2018.01.015.

[2] Z.-H. Zhou, *Machine learning*. Springer Nature, 2021, doi: 10.1007/978-981-15-1967-3.

[3] J. Huyan, W. Li, S. Tighe, Z. Xu, and J. Zhai, "CrackU-net: a novel deep convolutional neural network for pixelwise pavement crack detection," *Structural Control and Health Monitoring*, vol. 27, no. 8, Aug. 2020, doi: 10.1002/stc.2551.

[4] E. H. Houssein, A. Hammad, and A. A. Ali, "Human emotion recognition from EEG-based brain–computer interface using machine learning: a comprehensive review," *Neural Computing and Applications*, vol. 34, no. 15, pp. 12527–12557, Aug. 2022, doi: 10.1007/s00521-022-07292-4.

[5] S. Nurmaini *et al.*, "An automated ECG beat classification system using deep neural networks with an unsupervised feature extraction technique," *Applied Sciences*, vol. 9, no. 14, Jul. 2019, doi: 10.3390/app9142921.

[6] G. Kocher and G. Kumar, "Machine learning and deep learning methods for intrusion detection systems: recent developments and challenges," *Soft Computing*, vol. 25, no. 15, pp. 9731–9763, Aug. 2021, doi: 10.1007/s00500-021-05893-0.

[7] P. Sharma, S. Jain, S. Gupta, and V. Chamola, "Role of machine learning and deep learning in securing 5G-driven industrial IoT applications," *Ad Hoc Networks*, vol. 123, Dec. 2021, doi: 10.1016/j.adhoc.2021.102685.

[8] K. Korfmann, O. E. Gaggiotti, and M. Fumagalli, "Deep learning in population genetics," *Genome Biology and Evolution*, vol. 15, no. 2, Feb. 2023, doi: 10.1093/gbe/evad008.

[9] G. Battineni, N. Chintalapudi, and F. Amenta, "Machine learning in medicine: performance calculation of dementia prediction by support vector machines (SVM)," *Informatics in Medicine Unlocked*, vol. 16, 2019, doi: 10.1016/j.imu.2019.100200.

[10] A. All Tanvir, E. M. Mahir, S. Akhter, and M. R. Huq, "Detecting fake news using machine learning and deep learning algorithms," in *2019 7th International Conference on Smart Computing & Communications (ICSCC)*, Jun. 2019, pp. 1–5, doi: 10.1109/ICSCC.2019.8843612.

[11] B. He, D. J. Armaghani, and S. H. Lai, "Assessment of tunnel blasting-induced overbreak: a novel metaheuristic-based random forest approach," *Tunnelling and Underground Space Technology*, vol. 133, Mar. 2023, doi: 10.1016/j.tust.2022.104979.

[12] A. Coatrini-Soares *et al.*, "Microfluidic e-tongue to diagnose bovine mastitis with milk samples using machine learning with decision tree models," *Chemical Engineering Journal*, vol. 451, Jan. 2023, doi: 10.1016/j.cej.2022.138523.

[13] J. Mei, E. Muller, and S. Ramaswamy, "Informing deep neural networks by multiscale principles of neuromodulatory systems," *Trends in Neurosciences*, vol. 45, no. 3, pp. 237–250, Mar. 2022, doi: 10.1016/j.tins.2021.12.008.

[14] C. Sun, X. Liu, Q. Jiang, X. Ye, X. Zhu, and R.-W. Li, "Emerging electrolyte-gated transistors for neuromorphic perception," *Science and Technology of Advanced Materials*, vol. 24, no. 1, Dec. 2023, doi: 10.1080/14686996.2022.2162325.

[15] S. M. Ayyad, A. I. Saleh, and L. M. Labib, "Gene expression cancer classification using modified k-nearest neighbors technique," *Biosystems*, vol. 176, pp. 41–51, Feb. 2019, doi: 10.1016/j.biosystems.2018.12.009.

[16] H. Saadatfar, S. Khosravi, J. H. Joloudari, A. Mosavi, and S. Shamshirband, "A new k-nearest neighbors classifier for big data based on efficient data pruning," *Mathematics*, vol. 8, no. 2, Feb. 2020, doi: 10.3390/math8020286.

[17] D. A. Anggoro and D. Novitaningrum, "Comparison of accuracy level of support vector machine (SVM) and artificial neural network (ANN) algorithms in predicting diabetes mellitus disease," *ICIC Express Letters*, vol. 15, no. 1, pp. 9–18, 2021, doi: 10.24507/icicel.15.01.9.

[18] S. Uddin, I. Haque, H. Lu, M. A. Moni, and E. Gide, "Comparative performance analysis of k-nearest neighbour (KNN) algorithm and its different variants for disease prediction," *Scientific Reports*, vol. 12, no. 1, Apr. 2022, doi: 10.1038/s41598-022-10358-x.

[19] S. R. S. Chakravarthy, N. Bharanidharan, and H. Rajaguru, "Deep learning-based metaheuristic weighted k-nearest neighbor algorithm for the severity classification of breast cancer," *IRBM*, vol. 44, no. 3, Jun. 2023, doi: 10.1016/j.irbm.2022.100749.

[20] B. A. H. Murshed, S. Mallappa, O. A. M. Ghaleb, and H. D. E. Al-ariki, "Efficient twitter data cleansing model for data analysis of the pandemic tweets," *Studies in Systems, Decision and Control*, vol. 348, pp. 93–114, 2021, doi: 10.1007/978-3-030-67716-9_7.

[21] B. Zhong *et al.*, "Deep learning-based extraction of construction procedural constraints from construction regulations," *Advanced Engineering Informatics*, vol. 43, Jan. 2020, doi: 10.1016/j.aei.2019.101003.

[22] R. K. Dey and A. K. Das, "Modified term frequency-inverse document frequency based deep hybrid framework for sentiment analysis," *Multimedia Tools and Applications*, vol. 82, no. 21, pp. 32967–32990, Sep. 2023, doi: 10.1007/s11042-023-14653-1.

[23] K. Wanjale, A. Chitre, P. Patil, R. Parmar, S. Raka, and S. Ghattuwar, "A comprehensive survey of stemming methods in information retrieval," *International Journal of Multidisciplinary Engineering in Current Research*, vol. 7, no. 10, pp. 15–22, 2022.

[24] A. Xiong, D. Liu, H. Tian, Z. Liu, P. Yu, and M. Kadoch, "News keyword extraction algorithm based on semantic clustering and word graph model," *Tsinghua Science and Technology*, vol. 26, no. 6, pp. 886–893, Dec. 2021, doi: 10.26599/TST.2020.9010051.

[25] J. Yang *et al.*, "Neuromorphic engineering: from biological to spike-based hardware nervous systems," *Advanced Materials*, vol. 32, no. 52, Dec. 2020, doi: 10.1002/adma.202003610.

[26] A. Goldstein *et al.*, "Shared computational principles for language processing in humans and deep language models," *Nature Neuroscience*, vol. 25, no. 3, pp. 369–380, Mar. 2022, doi: 10.1038/s41593-022-01026-4.

[27] H. Dagdougui, F. Bagheri, H. Le, and L. Dessaint, "Neural network model for short-term and very-short-term load forecasting in district buildings," *Energy and Buildings*, vol. 203, Nov. 2019, doi: 10.1016/j.enbuild.2019.109408.

[28] L. Jiao, X. Geng, and Q. Pan, "BP kNN: k-nearest neighbor classifier with pairwise distance metrics and belief function theory," *IEEE Access*, vol. 7, pp. 48935–48947, 2019, doi: 10.1109/ACCESS.2019.2909752.

[29] M. R. Romadhon and F. Kurniawan, "A comparison of Naive Bayes methods, logistic regression and K-NN for predicting healing of COVID-19 patients in Indonesia," in *2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT)*, pp. 41–44, Apr. 2021, doi: 10.1109/EIConCIT50028.2021.9431845.

[30] D. Chicco, N. Tötsch, and G. Jurman, "The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation," *BioData Mining*, vol. 14, no. 1, Feb. 2021, doi: 10.1186/s13040-021-00244-z.

[31] S. Thangamayan, S. N. Jagdale, T. R. K. Lakshmi, J. G. Thatipudi, P. K, and B. Khan, "Artificial intelligence oriented user sentiment evaluation system on social networks using modified deep learning principles," in *2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI)*, pp. 272–279, Jan. 2024, doi: 10.1109/ICMCSI61536.2024.00046.

[32] Rizkiansyah, A. Herliana, D. P. Alamsyah, and T. F. Tjoe, "Comparison of the k-nearest neighbor and decision tree algorithm to the sentiment analysis of investment applications users in Indonesia," in *2022 Seventh International Conference on Informatics and Computing (ICIC)*, pp. 01–06, Dec. 2022, doi: 10.1109/ICIC56845.2022.10006970.

BIOGRAPHIES OF AUTHORS

Fachrul Kurniawan received a B.S. in the Department of Informatics Engineering from the University of 17 Agustus 1945 Surabaya, master of information management technology, and doctor degrees in the Department of Electrical Engineering from Institut Teknologi Sepuluh November, Surabaya, Indonesia. Now he is a lecturer and researcher in the Department of Informatics Engineering, Master of Informatics, Faculty of Sains and Technology from Universitas Islam Negeri Maulana Malik Ibrahim Malang. Research interests include smart cities, information technology, big data, serious games, data mining, and intelligent systems-IAENG members. He can be contacted at email: fachrulk@ti.uin-malang.ac.id.

Triyo Supriyatno a Professor in Faculty of Tarbiyah and Teacher Training Universitas Islam Negeri Maulana Malik Ibrahim, Malang, East Java Indonesia. He received his doctor degree in the Faculty of Education from University of Malaya (UM) Malaysia, and his master degree in Islamic Education from Universitas Muhammadiyah Jakarta Indonesia and his bachelor degree from IKIP Negeri, Malang, East Java, Indonesia. In 2000 he joined the Department of Islamic Education in Faculty of Tarbiyah and Teacher Training of Universitas Islam Negeri Maulana Malik Ibrahim, Malang, East Java Indonesia as lecturer and 2021 as a Professor on Education. He has written several papers in the areas of islamic education and training, evaluation of learning, learning media, and ICT for learning. His research interests also include ICT and vocational education and training (TVET), standardization in educational, teaching and learning, pedagogical innovations in TVET, skills and personal development, and innovations in IT and education. He can be contacted at email: triyo@pai.uin-malang.ac.id.