

Global recession sentiment analysis utilizing VADER and ensemble learning method with word embedding

Maylina Rahayu Ningsih¹, Kevyn Aalifian Hernanda Wibowo², Ahmad Ubai Dullah³, Jumanto⁴

^{1, 2, 3, 4}Department of Computer Science, Universitas Negeri Semarang, Indonesia

Article Info

Article history:

Received August 15, 2023

Revised September 28, 2023

Accepted September 28, 2023

Keywords:

Sentiment analysis

Global recession

VADER

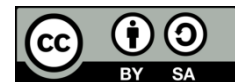
Word embedding

Ensemble learning

ABSTRACT

The issue of Global Recession is hitting various countries, including Indonesia. Many Indonesians have expressed their opinions on the issue of global recession in 2023, one of which is from Twitter. By understanding public sentiment, we can assess the impact felt by the public on the issue itself. Sentiment analysis in this research is a form of support to evaluate Indonesia's sustainability in dealing with the issue of Global Recession in accordance with the Sustainable Development Goals (SDGs). However, in previous research, it is still rare to find a model that has good performance in conducting Global Recession Sentiment Analysis. Therefore, the purpose of this research is to propose a machine learning model that is expected to provide good performance in sentiment analysis. The existing sentiment dataset is labeled with the Valence Aware Dictionary for Social Reasoning (VADER) algorithm, then an Ensemble Learning method is designed which is composed of Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine (SVM) algorithms. After that, the Countvectorizer feature extraction with N-Gram, Best Match 25 (BM25) and Word Embedding is carried out to convert sentences in the dataset into numerical vectors so as to improve model performance. The research results provide a more optimal accuracy performance of 95.02% in classifying sentiment. So that the proposed model successfully performs sentiment analysis better than previous research.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Maylina Rahayu Ningsih,
Department of Computer Science,
Universitas Negeri Semarang, Indonesia.
Email: maylinarahayuningsih@gmail.com
<https://doi.org/10.52465/joscecx.v4i3.193>

1. INTRODUCTION

The condition when the economies of most countries deteriorate as activity in the economic sector declines is called a global recession [1]. Recently, the IMF surprisingly published bad news for the world economy, namely the decline in world economic growth in 2023 to 2.9% [2]. The IMF also predicts that Indonesia's economy will still be hampered by global deceleration with inflation at 9.5% and economic growth at 3.9% [2], [3]. This issue invites Indonesian people to give their opinions on social media, especially twitter.

The unstructured sentiment of society is often overlooked, even though it is able to extract the viewpoints expressed [4], [5] and is the first stage to assess how far this recession has affected. So a sentiment

analysis strategy is needed that is able to generalize the condition of the community from the sentiment data they convey. So as to be able to evaluate the sustainability, prevention and handling of recessions for the government.

Therefore, many studies have developed sentiment analysis techniques from twitter or other datasets to be able to generalize and even predict from the processed sentiment data [6], [7]. Text mining is a machine learning method that aims to gather information and obtain high accuracy in classifying predictions on a class of data [8], [9]. Sentiment analysis is part of data mining, various optimizations with various algorithms are carried out in order to achieve high accuracy performance [10].

Various approaches are used for twitter sentiment analysis in various contexts such as sentiment regarding elections, government policies and so on [11], [12]. In twitter sentiment analysis, the use of ensemble learning can help in overcoming the diversity and variation of sentiments [13]. The ensemble learning method improves classification accuracy for multiclass segmentation [14]. Research [15], [16] confirmed that the selection of algorithms for ensemble learning affects the accuracy results and the need to use customized feature extraction to improve accuracy performance.

Related research has applied various methods for sentiment analysis by combining various data labeling methods, feature extraction, algorithms and pre-processing [17], [18]. Ensemble learning is one of the most popular classification models today. In other research [19], developed an ensemble learning model with Term Frequency-Inverse Document Frequency (TF-IDF) and Bag Of Word feature extraction from twitter datasets retrieved using API. This research was developed again by [20] However, the accuracy obtained is 77.15% which means it is still low so it is recommended to use feature extraction in deep learning. Other research that discusses sentiment analysis [18], [21], [22] contributes using various algorithms in the proposed ensemble learning method but not optimal in performance. So sentiment analysis is needed to detect and classify sentiments properly. From this study, several findings can be drawn as follows i) the majority of sentiment analysis uses existing datasets or is extracted manually using supporting tools ii) the use of ensemble learning needs to be supported by relevant algorithms with feature extraction and methods that can balance the data and minimize overfitting.

The ensemble learning method uses the idea of multiple algorithms to find the best result. However, we found research [23] using Ensemble Deep Learning (MVEDL) based on the Majority Voting technique which gives equal weight to each classifier. Which means that each classifier has the same ability and reliability, so the accuracy results are not optimal. Furthermore, the ensemble learning model is improved with bagging techniques and more optimal sentiment weighting. We propose an ensemble learning method with VADER algorithm with BM25 feature extraction and Word Embedding that successfully improves accuracy performance. Our research contributions are i) improving the accuracy model of 2023 global recession sentiment analysis using ensemble learning method and VADER algorithm [24], ii) optimizing the performance with BM25 feature extraction and Word Embedding.

2. METHOD

The proposed method uses ensemble learning method that combines several models, namely Logistic Regression, Decision Tree, Random Forest, and SVM. Before processing with the proposed model, data pre-processing is carried out starting from labeling, data cleaning, Countvectorizer feature extraction with N-gram, BM25, and word embedding. After going through the above process, the new data is processed by the proposed model and produces an evaluation matrix. The proposed method will be explained in the next section and shown in Figure 1.

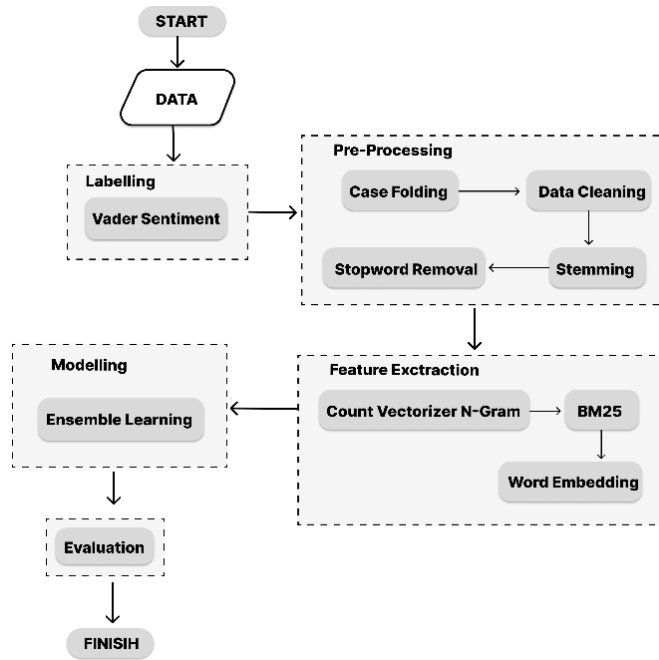


Figure 1. Flowchart of the proposed method

2.1 Dataset

The dataset contains the sentiment of Indonesian people regarding the issue of global recession on twitter. This data is sourced from Kaggle <https://www.kaggle.com/datasets/hafizhadinda/global-recession-in-2023-sentiment-indonesian> which was published in January 2023. The dataset contains 3074 sentiments with four columns: id, created at, username, and tweet. The dataset needs to be labeled first before processing, namely with the VADER algorithm. Sentiment labeling uses positive (1) and negative (0) labels.

2.2 VADER

The VADER algorithm [25] is used as a sentiment analysis model because it can measure the emotion of the data. In determining the polarity of the sentence, the compound in each word is given positive, negative or neutral criteria. In this study, the text label is based on the polarity score calculated by normalizing the sum of positive and negative scores. Equation (1) is the VADER calculation used in the labeling process.

$$Label \begin{cases} Positive & \text{if } score \geq 0.05 \\ Negative & \text{if } score \leq -0.05 \end{cases} \tag{1}$$

2.3 Preprocessing

At the beginning of data pre-processing, it is necessary to perform a data cleaning stage on the data that has been labeled. Because the data taken from twitter usually contains unnecessary words or characters such as non-alphanumeric punctuation marks, symbols, numbers, username (@), retweets and excess whitespace that can interfere with the data analysis process. After the data cleaning process, the sentiment is processed by going through several stages, namely, Case Folding, StopWord Removal, and stemming. Further process is needed to remove duplicate data and remove unnecessary links. Then the sentiments are merged back for further processing. The dataset containing sentiment is converted into a dataframe format containing six columns in the form of id, created at, username, tweet (before data pre-processing), label and tweet (after data pre-processing). The following examples of tweets before and after pre-processing can be seen in Table 1.

Table 1. Pre-processing stages

Tweets before Pre-Processing	Tweets after Pre-Processing
------------------------------	-----------------------------

After going through the plenary session process, continue...	After the plenary session process, continue with the points...
@papa_loren Really reckless even though it's 2023...	How reckless is the world's economic condition...

After cleaning the dataset, determine the tweet columns for the x variable and y variable. Then, the data is divided into two with a ratio of 80: So 80% (2426 data) for training data and 20% (607 data) for test data.

2.4 Count Vectorizer Feature Extraction with N-Gram

In machine learning analysis, to process unstructured text data, it needs to be converted into numerical form so that it can be processed. One method that can be used is Count vectorizer feature extraction with N-grams. In this method, the text is divided into a sequence of n consecutive words and then the frequency with which each sequence of words appears is calculated. The result is a feature vector representing the text, with each element of the vector indicating the frequency of occurrence of a particular word sequence.

2.5 BM25 Vectorizer

The Best Matching 25 Vectorizer (BM25) [26] is a major improvement over the classic TF-IDF-based algorithm. The BM25 ranking system sorts document match results based on the searched keywords. This feature is used to estimate the relevance between the query and the text during information indexing. Equations (2) and (3) are BM25 calculations, and BM25 Vector Visualization with Principal Component Analysis (PCA) is shown in Figure 2.

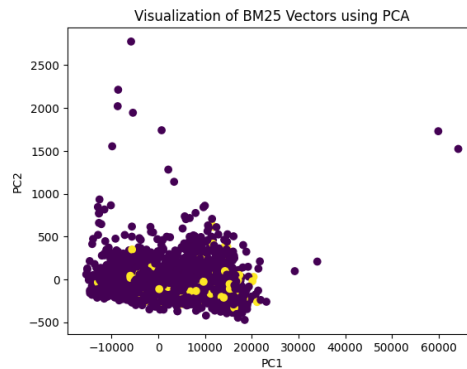


Figure 5. BM25 Vectors using PCA

$$BM25 = \sum_{t \in Q} \log_{10} \frac{(N - nt)}{nt} \cdot \frac{(k1 + 1)fd,t}{K + fd,t} \cdot \frac{(k3 + 1)fq,t}{k3 + fq,t} \tag{2}$$

Where Q is the Query used as input, N is the total sentences in the document, nt is the total terms containing the Query, fd,t as the total term frequency, fq,t as the total query frequency, and k1,k3,b is a constant or parameter.

$$K = k1((1 - b) + \frac{b \cdot dl_d}{avl}) \tag{3}$$

Where k1 and b are constants or parameters, $[[dl]]_d$ is the total sentences in the document, and avl is the average document length.

2.6 Word Embedding

The word embedding [27] used in this research is the word2vec model with a dimension of 300. Each word is a vector that represents a point in a certain dimension in space. Words that have certain characteristics, such as being in the same context or having the same semantic meaning, are connected to each other in a certain

space. Using the one-hot encoding method, a vector with a value of 1 indicates the position of the word, and a value of 0 indicates the position of other words.

Word2vec consists of CBOW and Skip-gram, Given a sequence of words $W = \{w_1, w_2, \dots, w_N\}$ as input, which can maximize the log-likelihood probabilities used in CBOW and Skip-gram written in equations (4) and (5) [28].

$$L_{CBOW}(W) = \frac{1}{N} \sum_{i=1}^N \log \Pr(w_i | w_{cxt}) \quad (4)$$

$$L_{Skip-gram}(W) = \frac{1}{N} \sum_{i=1}^N \sum_{-c \leq j \leq c, j \neq 0} \log \Pr(w_{i+j} | w_i) \quad (5)$$

Where, w_{cxt} represents the context word of the current word w_i , c is the window size of the context, and conditional probability $\Pr(w_i | w_j)$ is calculated by equation (6).

$$\Pr(w_i | w_j) = \frac{\exp(\mathbf{w}_i \cdot \mathbf{w}_j)}{\sum_{w \in W} \exp(\mathbf{w} \cdot \mathbf{w}_j)} \quad (6)$$

Where \mathbf{w} , \mathbf{w}_i and \mathbf{w}_j are the represent of word w and W is the word vocabulary Word2vec assumes that a word has semantic similarity with the surrounding words. However, the surrounding words may not have the same semantics as the current word [29].

2.7 Ensemble Learning

The proposed model in this research uses an Ensemble Learning algorithm [30]–[32] that uses a voting system to get the best results. There are several methods combined into this model, namely Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), and Support Vector Machine (SVM). To determine the best results, voting will be carried out, which will later be taken as the final result in the final evaluation. For more details can be seen in Figure 3.

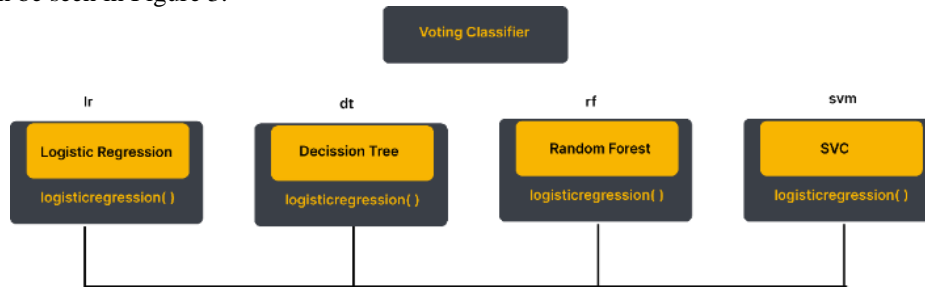


Figure 3. Ensemble model

Equations (7), (8), (9) and (10) show the calculation of Logistic Regression, Decision Tree, Random Forest and SVM respectively [33]–[38].

$$P = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)}} \quad (7)$$

$$Entropy(D) = \sum_{i=1}^c p_i \log_2(p_i) \quad (8)$$

$$f(x) = \frac{1}{s} \sum_{i=1}^s T_i(x) \quad (9)$$

$$H(x) = w^T x + b = 0 \quad (10)$$

2.8 Evaluation

After analyzing the experiments, the ensemble learning model needs to be evaluated to see the resulting performance using Confusion Matrix. Confusion Matrix can display the condition of actual data and predictions made by machine learning algorithms. True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) are the number of true positive classes, the number of false positive, true negative, and false negative data in the dataset. Equations (13), (14) and (15) are calculations to determine Accuracy, Recall, and f1 score values, respectively.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (13)$$

$$Recall = \frac{TP}{TP + FN} \quad (14)$$

$$f1\ score = \frac{2 \times precision \times recall}{precision + recall} \quad (15)$$

3. RESULTS AND DISCUSSIONS

In the experimentation and testing process, this research involved the Ensemble Learning method as the best performing result. Previously, the team had conducted experiments on other experiments with several parameters.RAM.

3.1 Results of SVM and TF-IDF algorithm approach with Manual Labeling

Initial experiments were conducted using a dataset labeled with label 0 as negative and label 1 as positive manually by the team. TF-IDF feature extraction was used in assessing the frequency of a word and then Word Embedding was used to represent the word. The results of this approach are shown in Table 2.

SVM + TF-IDF + Word Embedding	
Accuracy	61.76%
Precision	54.12%
Recall	61.76%
F1-Score	50.13%

3.2 Results of SVM and BM25 algorithm approach with manual labeling

The experiments were then continued using the manually labeled dataset. Table 3. shows the results of pre-processing, using BM25 and word embedding in SVM.

SVM + BM25 + Word Embedding	
Accuracy	65.79%
Precision	59.22%
Recall	65.79%
F1-Score	60%

3.3 Ensemble Learning Approach Results

Evaluating previous experiments, this approach represents the use of multiple algorithms with the Ensemble Learning method. In addition, since it was judged that the manually calculated labeling based on different preferences made the data unbalanced and resulted in poor performance. So the dataset was updated with labeling done using VADER automatically. This model is used for sentiment analysis and can determine the diversity of data through the intensity of emotional strength.

The algorithms used in Ensemble Learning are Logistic Regression, Decision Tree, Random Forest, and SVM. This method assesses by voting on the results shown by the model. In generating model performance we chose BM25 to represent words because from experimental results the BM25 method is more sophisticated for calculating relevance scores than simple methods such as TF-IDF. In the overall experimental results the model is superior to both Accuracy, Precision, Recall and F1 Score. The best results of Ensemble Learning performed well in the evaluation metrics shown in Table 4.

Table 4. Ensemble Learning Results

VADER + Ensemble + BM25 + Word Embedding	
Accuracy	95.02%
Precision	95.27%
Recall	95.02%
F1-Score	92.89%

The use of BM25 and Word Embedding greatly supports the model in producing good performance. This can be seen from the visualization of Word Embedding with a comparison using PCA and t-Distributed Stochastic Neighbor Embedding (t-SNE). PCA is used to identify and reduce word dimensionality. The visualization in Figure shows the patterns and relationships between words, where each point distribution in the image represents a word vector representation. A comparison is made using t-SNE where this visualization method pays more attention to the similarity of the distribution between points. This pattern helps to illustrate the hidden patterns of word relationships. The visualization is shown in Figure 5.

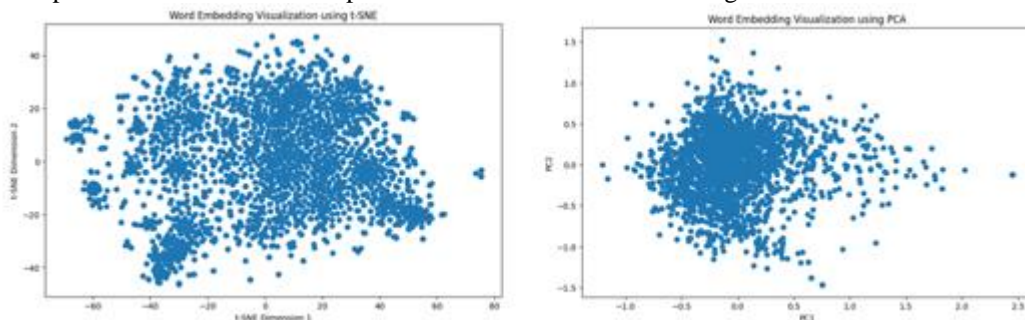


Figure 4. Word Embedding dengan t-SNE (Left), Word Embedding dengan PCA (Right)

The t-SNE visualization (left) results explore complex semantic relationships between words. This can help in finding groups of words that are related in a way that is not always visible linearly. Meanwhile, PCA visualization (right) results provide a more linear and orderly general picture of the relationships between words.

After considering the results from testing the SVM algorithm, the research team decided to use the Ensemble Learning method as the main approach in experimentation and testing. As a result, the ensemble model was able to overcome the uncertainty problem and obtain an accuracy of 95.02%. Thus, from the results of the experiments and tests conducted, it can be concluded that the Ensemble Learning approach provides the best performance in data classification. To compare the performance of the proposed model, it is shown in Table 5.

Table 5. Comparison model

Author	Algorithm Machine Learning	Objective	Result
Bengesi et al. [39]	KNN, MLP, Random Forest, Logistic Regression, MLP, Naïve Bayes, XGBoost, SVM	Analisis Sentimen pada Wabah Cacar Monyet dari Opini Publik dari Kicauan di Twitter	Highest accuracy of SVM-93.48%
Obiedat et al. [40]	SVM, Logistic Regression, Random Forest, Decision Tree, KNN, XGBoost, SVM-PSO	Sentiment Analysis of Customer Reviews	Highest accuracy of SVM-PSO -89.99%

Wazrah et al. [27]	LSTM, SVM, AraBERT, SGRU, SBi-GRU	Sentiment Analysis for Arabic Tweets	Accuracy of the ensemble model AraBERT, SGRU, SBi-GRU =90.21%
Proposed Method	Vader + Ensemble Model Voting + BM25+ Word Embedding	Sentiment analysis for recession tweets for Indonesians	95.02%

Based on table 5, it can be compared that the proposed model is superior to models from previous studies. The proposed model has more optimal accuracy than other models, with an accuracy value of 95.02%. Therefore, it can be said that this proposed model is more optimal than previous research models.

4. CONCLUSION

Based on the results of experiments and testing of the proposed new model in analyzing Indonesian sentiment about the 2023 global recession on Twitter using the Ensemble Learning model by combining several Machine Learning methods, it can be concluded that in terms of accuracy, this model shows good performance in analyzing Indonesian sentiment about the 2023 global recession with an accuracy of 95.02%. So this model can be one of the effective solutions in distinguishing Indonesian sentiment, positive or negative. However, it should be noted that the proposed model has language limitations in performing sentiment analysis. So it is hoped that future research can develop models that can perform sentiment analysis with foreign languages.

REFERENCES

- [1] B. Erten and J. Antonio, "The future of commodity prices and the pandemic-driven global recession : Evidence from 150 years of data," *World Dev.*, vol. 137, p. 105164, 2021, doi: 10.1016/j.worlddev.2020.105164.
- [2] S. M. S. Yobel Rayfinando Tua Hutagaol, Ronaldo Putra Pratama Sinurat, "Strategi penguatan keuangan negara dalam menghadapi ancaman resesi global 2023 melalui green economy," *J. pajak dan Keuang. negara*, pp. 378–385, 2023.
- [3] K. Ahmed *et al.*, "Breaking down linguistic complexities : A structured approach to aspect- based sentiment analysis," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 35, no. 8, p. 101651, 2023, doi: 10.1016/j.jksuci.2023.101651.
- [4] A. Gutub, M. Khaled, and M. A. Abu-hashem, "Coronavirus impact on human feelings during 2021 Hajj season via deep learning critical Twitter analysis," *J. Eng. Res.*, vol. 11, no. 1, p. 100001, 2023, doi: 10.1016/j.jer.2023.100001.
- [5] F. Altuntas, S. Altuntas, and T. Dereli, "International Journal of Information Management Data Insights Social network analysis of tourism data : A case study of quarantine decisions in COVID-19 pandemic," *Int. J. Inf. Manag. Data Insights*, vol. 2, no. 2, p. 100108, 2022, doi: 10.1016/j.jjime.2022.100108.
- [6] F. Shatnawi, "Detecting Epidemic Diseases Using Sentiment Analysis of Arabic Tweets," vol. 26, no. 1, pp. 50–70, 2020.
- [7] K. Garcia and L. Berton, "Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA," *Appl. Soft Comput. J.*, vol. 101, p. 107057, 2021, doi: 10.1016/j.asoc.2020.107057.
- [8] H. Yan, M. Ma, Y. Wu, H. Fan, and C. Dong, "Heliyon Overview and analysis of the text mining applications in the construction industry," *Heliyon*, vol. 8, no. September, p. e12088, 2022, doi: 10.1016/j.heliyon.2022.e12088.
- [9] R. Hillan, "LAPORAN TEK NIK KOMPILASI SUNDANESE LANGUAGE," 2021.
- [10] E. Rosenberg *et al.*, "Results in Engineering Sentiment analysis on Twitter data towards climate action," *Results Eng.*, vol. 19, no. July, p. 101287, 2023, doi: 10.1016/j.rineng.2023.101287.
- [11] N. Leelawat, S. Jariyapongpaiboon, A. Promjun, and S. Boonyarak, "Heliyon Twitter data sentiment analysis of tourism in Thailand during the COVID-19 pandemic using machine learning," *Heliyon*, vol. 8, no. September, p. e10894, 2022, doi: 10.1016/j.heliyon.2022.e10894.
- [12] H. Xu, R. Liu, Z. Luo, and M. Xu, "Telematics and Informatics Reports COVID-19 vaccine sensing : Sentiment analysis and subject distillation from twitter data," vol. 8, no. July, 2022, doi: 10.1016/j.teler.2022.100016.
- [13] N. Saleena, "ScienceDirect An Ensemble Classification System for Twitter Sentiment Analysis An Ensemble Classification System for Twitter Sentiment Analysis," *Procedia Comput. Sci.*, vol. 132, no.

- Iccids, pp. 937–946, 2018, doi: 10.1016/j.procs.2018.05.109.
- [14] M. Rahil, B. N. Anoop, G. N. Girish, A. R. Kothari, S. G. Koolagudi, and J. Rajan, “A Deep Ensemble Learning-Based CNN Architecture for Multiclass Retinal Fluid Segmentation in OCT Images,” *IEEE Access*, vol. 11, no. December 2022, pp. 17241–17251, 2023, doi: 10.1109/ACCESS.2023.3244922.
- [15] E. H. Y. Wahyu Fadli Satryaa, Ria Aprilliyani, “Sentiment analysis of Indonesian police chief using multi-level Sentiment analysis of Indonesian police chief using multi-level ensemble model ensemble model,” *Procedia Comput. Sci.*, vol. 216, no. 2022, pp. 620–629, 2023, doi: 10.1016/j.procs.2022.12.177.
- [16] D. Sunitha, R. Kumar, N. V Babu, A. Suresh, and S. Chand, “Twitter sentiment analysis using ensemble based deep learning model towards COVID-19 in India and European countries,” *Pattern Recognit. Lett.*, vol. 158, pp. 164–170, 2022, doi: 10.1016/j.patrec.2022.04.027.
- [17] A. S. Talaat, “Sentiment analysis classification system using hybrid BERT models,” *J. Big Data*, 2023, doi: 10.1186/s40537-023-00781-w.
- [18] O. Abiola, A. A. Alli, O. A. Tale, and S. Misra, “Sentiment analysis of COVID - 19 tweets from selected hashtags in Nigeria using VADER and Text Blob analyser,” *J. Electr. Syst. Inf. Technol.*, vol. 9, 2023, doi: 10.1186/s43067-023-00070-9.
- [19] A. Mohi, U. Din, S. Tanzeel, and Q. Rayees, “Detecting twitter hate speech in COVID-19 era using machine learning and ensemble learning techniques,” *Int. J. Inf. Manag. Data Insights*, vol. 2, no. 2, p. 100120, 2022, doi: 10.1016/j.jjimei.2022.100120.
- [20] A. Mohi *et al.*, “NNPCov19 : Artificial Neural Network-Based Propaganda Identification on Social Media in COVID-19 Era,” vol. 2022, 2022.
- [21] L. Corti, M. Zanetti, G. Tricella, and M. Bonati, “Social media analysis of Twitter tweets related to ASD in 2019 – 2020 , with particular attention to COVID - 19 : topic modelling and sentiment analysis,” *J. Big Data*, 2022, doi: 10.1186/s40537-022-00666-4.
- [22] N. Parveen, P. Chakrabarti, B. T. Hung, and A. Shaik, “Twitter sentiment analysis using hybrid gated attention recurrent network,” *J. Big Data*, 2023, doi: 10.1186/s40537-023-00726-3.
- [23] S. Malla and P. J. A. Alphonse, “COVID-19 outbreak : An ensemble pre-trained deep learning model for detecting informative tweets,” *Appl. Soft Comput.*, vol. 107, p. 107495, 2021, doi: 10.1016/j.asoc.2021.107495.
- [24] J. Jumanto, M. A. Muslim, Y. Dasril, and T. Mustaqim, “Accuracy of Malaysia Public Response to Economic Factors During the Covid-19 Pandemic Using Vader and Random Forest,” *J. Inf. Syst. Explor. Res.*, vol. 1, no. 1, pp. 49–70, Dec. 2022, doi: 10.52465/joiser.v1i1.104.
- [25] C. J. Hutto and E. Gilbert, “VADER: A Parsimonious Rule-based Model for,” *Eighth Int. AAAI Conf. Weblogs Soc. Media*, pp. 216–225, 2014.
- [26] S. Robertson and H. Zaragoza, “The probabilistic relevance framework: BM25 and beyond,” *Found. Trends Inf. Retr.*, vol. 3, no. 4, pp. 333–389, 2009, doi: 10.1561/15000000019.
- [27] A. A. L. Wazrah and S. Alhumoud, “Sentiment Analysis Using Stacked Gated Recurrent Unit for Arabic Tweets,” *IEEE Access*, vol. 9, pp. 137176–137187, 2021, doi: 10.1109/ACCESS.2021.3114313.
- [28] T. Huanling *et al.*, “Representation of Semantic Word Embeddings Based on SLDA and Word2vec Model,” *Chinese J. Electron.*, vol. 32, no. 3, pp. 647–654, 2023, doi: 10.23919/cje.2021.00.113.
- [29] S. Y. Ihm, J. H. Lee, and Y. H. Park, “Skip-gram-KR: Korean word embedding for semantic clustering,” *IEEE Access*, vol. 7, pp. 39948–39961, 2019, doi: 10.1109/ACCESS.2019.2905252.
- [30] R. Jayapermana, A. Aradea, and N. I. Kurniati, “Implementation of Stacking Ensemble Classifier for Multi-class Classification of COVID-19 Vaccines Topics on Twitter,” *Sci. J. Informatics*, vol. 9, no. 1, pp. 8–15, 2022, doi: 10.15294/sji.v9i1.31648.
- [31] M. A. Muslim *et al.*, “New model combination meta-learner to improve accuracy prediction P2P lending with stacking ensemble learning,” *Intell. Syst. with Appl.*, vol. 18, no. December 2022, p. 200204, 2023, doi: 10.1016/j.iswa.2023.200204.
- [32] S. M. Abd-Elsalam, M. M. Ezz, S. Gamalel-Din, G. Esmat, A. Salama, and M. ElHefnawi, “Early diagnosis of esophageal varices using Boosted-Naïve Bayes Tree: A multicenter cross-sectional study on chronic hepatitis C patients,” *Informatics Med. Unlocked*, vol. 20, p. 100421, 2020, doi: 10.1016/j.imu.2020.100421.
- [33] S. K. Arts and S. K. Arts, “Performance Evaluation of Machine Learning Algorithms for Email Spam Detection,” pp. 1–4, 2020.

- [34] M. J. Nayeem, S. Rana, F. Alam, and M. A. Rahman, "Prediction of Hepatitis Disease Using K-Nearest Neighbors, Naive Bayes, Support Vector Machine, Multi-Layer Perceptron and Random Forest," *2021 Int. Conf. Inf. Commun. Technol. Sustain. Dev. ICICT4SD 2021 - Proc.*, pp. 280–284, 2021, doi: 10.1109/ICICT4SD50815.2021.9397013.
- [35] H. Chen, N. Wang, X. Du, K. Mei, Y. Zhou, and G. Cai, "Classification Prediction of Breast Cancer Based on Machine Learning," *Comput. Intell. Neurosci.*, vol. 2023, pp. 1–9, 2023, doi: 10.1155/2023/6530719.
- [36] F. Gorgan, M. Taher, R. Mohammad, and Z. Kermani, "Decision tree models in predicting water quality parameters of dissolved oxygen and phosphorus in lake water," *Sustain. Water Resour. Manag.*, vol. 9, no. 1, pp. 1–13, 2023, doi: 10.1007/s40899-022-00776-0.
- [37] A. Alcántara, I. M. Galván, and R. Aler, "Deep neural networks for the quantile estimation of regional renewable energy production," *Appl. Intell.*, pp. 8318–8353, 2022, doi: 10.1007/s10489-022-03958-7.
- [38] C. S. Eke, E. Jammeh, X. Li, C. Carroll, S. Pearson, and E. Ifeakor, "Early Detection of Alzheimer's Disease with Blood Plasma Proteins Using Support Vector Machines," *IEEE J. Biomed. Heal. Informatics*, vol. 25, no. 1, pp. 218–226, 2021, doi: 10.1109/JBHI.2020.2984355.
- [39] S. Bengesi, T. Oladunni, R. Olusegun, and H. Audu, "A Machine Learning-Sentiment Analysis on Monkeypox Outbreak: An Extensive Dataset to Show the Polarity of Public Opinion From Twitter Tweets," *IEEE Access*, vol. 11, no. January, pp. 11811–11826, 2023, doi: 10.1109/ACCESS.2023.3242290.
- [40] R. Obiedat, R. Qaddoura, A. L. A. M. Al-zoubi, and L. Al-qaisi, "Sentiment Analysis of Customers' Reviews Using a Hybrid Evolutionary SVM-Based Approach in an Imbalanced Data Distribution," *IEEE Access*, vol. 10, pp. 22260–22273, 2022, doi: 10.1109/ACCESS.2022.3149482.