

Sentiment review of coastal assessment using neural network and naïve Bayes

Oman Somantri¹, Santi Purwaningrum², Ratih HafSarah Maharrani¹

¹Department of Cyber Security Engineering, State Polytechnic of Cilacap, Cilacap, Indonesia

²Department of Informatics Engineering, State Polytechnic of Cilacap, Cilacap, Indonesia

Article Info

Article history:

Received Nov 14, 2022

Revised Mar 24, 2023

Accepted Apr 7, 2023

Keywords:

Coastal assessment

Feature weight

Naïve Bayes

Neural network

Sentiment review

ABSTRACT

An assessment of a place will provide an overview for other people whether the place is feasible to be visited or not. Assessment of coastal places will provide a separate assessment for potential visitors in considering visitation. This article proposes a model using the neural network (NN) and naïve Bayes (NB) methods to classify sentiment toward coastal assessments. The proposed NN and NB models are optimized using information gain (IG) and feature weights, namely particle swarm optimization (PSO) and genetic algorithm (GA) which are carried out to increase the level of classification accuracy. Based on the experimental results, the best level of accuracy for the classification of coastal assessments is 87.11% and is named the NB IG+PSO model. The best model obtained is a model that can be used as a decision support for potential beach visitors in deciding to visit the place.

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



Corresponding Author:

Oman Somantri

Department of Cyber Security Engineering, State Polytechnic of Cilacap

Dr. Soetomo street No.01 Sidakaya, Cilacap, Indonesia

Email: oman_mantri@yahoo.com

1. INTRODUCTION

The assessment of a place made by people who have visited or occupied the area is now often used as a benchmark for other potential visitors when considering whether to visit the area or not. The natural beauty of Indonesia, which is an archipelagic country, has a certain attraction for everyone. The charm of tourism which is currently the prima donna is the beauty of the coast. The beach has its charm, not only visitor can enjoy the beauty of the sea, but they can also play around the coast. The current beach review assessment is carried out via the internet, whether it's a website or social media, which is used as a benchmark by other visitors as reference material. The assessment of positive and negative sentiments provides input for decision support when assessing a place that has a coastline. The southern coast of the island of Java is currently being visited by many, this can provide many sources of information in the form of a review of coastal assessment sentiment, especially for the southern coast of Java Island, Indonesia.

Sentiment review is an intelligent-based technology that uses artificial intelligence to get the best model used in classifying positive, negative, and neutral sentiments. Sentiment review in its process uses machine learning (ML) as the main intelligence tool in analyzing the data used. Whether or not the machine learning method is used depends on the ML algorithm itself because the characteristics of each data differ from one another. Sentiment models are currently being developed from various fields such as sentiment product reviews [1], [2], movie and video reviews [3], sentiment comments on social media [4], [5], and reviews of tourist attractions and hotels [6], [7]. Currently, many sentiment reviews have also been carried out in the government sector, namely sentiment in the field of public services [8].

Neural network (NN) and naïve Bayes (NB) [9], [10] are machine learning algorithms that are currently widely used for sentiment analysis. It shows that these two methods are the best methods that can be applied to sentiment models [11]. Several studies that are currently being developed using NN for sentiment models include those carried out for the Chinese car review [12]. Other studies have proposed a graph neural network model for sentiment analysis [13]. There is research that applies NN to hurricane sentiment taken from social media tweets [14]. In other previous studies, naïve Bayes was widely used for product review sentiment [15], tourists, and hotel and restaurant services [16], [17]. However, sentiment reviews that use this method discuss more the sentiment side of tourist attractions, hotels, and other services and are not focused on reviewing the place itself. Based on the research that has been done, this article proposes a sentiment review that is directly focused on the place itself, namely a review of the coast.

The weakness of the method developed using NN and NB is that the accuracy level is still not optimal, so this article is proposed to optimize the model to increase the accuracy of sentiment classification. The determination of the classification of coastal sentiment review carried out in this study is to find a model that can provide the best level of accuracy. This article proposes an optimization of the model using feature weights [18], [19] to get the best model accuracy value and increase the accuracy value. Based on its advantages, the particle swarm optimization (PSO) [20] and genetic algorithm (GA) methods are proposed for the optimization model, and the information gain (IG) method is also applied as an effort to increase accuracy. The contribution proposed in this article is the application of an optimization model to get the best weight value for the model on the neural network and naïve Bayes to get the best accuracy level.

2. METHOD

2.1. Dataset and tools

The dataset used in this study comes from Indonesian language text data which contains words of review comments given by people who have visited or even seen the coasts in the southern region of Java Island, Indonesia. For now, the data obtained and used is a review of comments given for *Pantai Teluk Penyu* in the Cilacap Regency, Central Java Province, Indonesia which was taken between 2020 and 2022. Data taken from the dataset [21] contains the results of a review of comments on that place. Data labeling was carried out in this study by classifying data based on star ratings, where 1–3-star ratings were included in the "negative" category with 140 data comments and 5-star ratings in the "positive" category with 250 data. The process of data analysis in obtaining the best model in this study used RapidMiner studio software tools.

2.2. Proposed method and framework

This study carried out several stages of the experimental process to get the desired best model results. The experimental stage was carried out through several stages, namely, the data preprocessing stage, model application, model optimization stage, data validation, and the final stage model evaluation. This study classifies two categories of sentiment, namely "positive" and "negative". In the data pre-processing stage, the process of transform case, tokenization, and filtering of tokens is carried out. In the data pre-processing process, text document data is also transformed into weights using the TF-IDF [22] method using (1), where $W_{i,j}$ is the weight of i and j values, $tf_{i,j}$ is the number of occurrences of i this is j , df_i is the number of documents containing i , and N is the total number of documents.

$$W_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right) \quad (1)$$

For the stages of applying the model in this experiment, two main methods were used, namely neural networks and naïve Bayes. At this stage, various ways were carried out to determine the combination of the parameters of the two methods. The data validation stage uses cross-validation [23], at this stage, the data is split between the training data and the testing data; 90% of the training data is determined and 10% is the testing data. The model evaluation carried out in this article is done by doing a comparison of all the proposed experimental models to show the model with the best level of accuracy. Figure 1 shows the framework of the research method proposed in this article. The performance value of the model proposed in this study is obtained by using the evaluation value of the confusion matrix [24], [25] as in (2), where TP is True Positive, TN is True Negative, FP is False Positive, and FN is False Negative.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

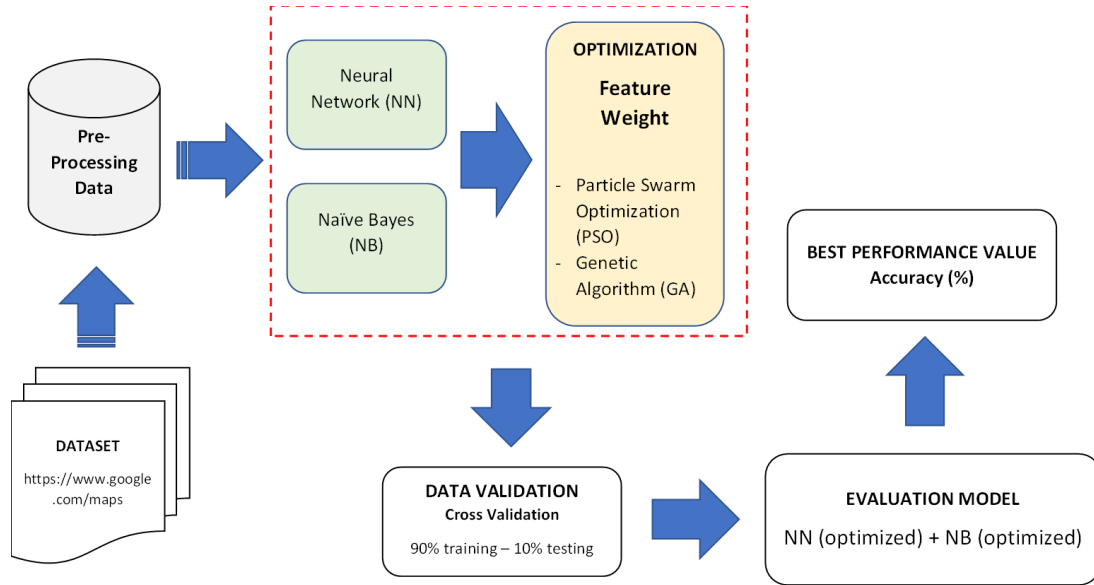


Figure 1. Proposed approach framework-based feature weights

2.3. Neural network and naïve Bayes

Neural network (NN) is a technique that is widely used in artificial intelligence, where NN is an algorithm whose way of working is to imitate the human brain with a computational element that mimics a neuron [26], [27]. The neural network is a method that simulates the workings of the human brain and tries to identify the underlying relationships in a data set. NN has a way of working where each layer consists of an input layer, hidden layer, and output layer that are interconnected to produce the desired mode.

Naïve Bayes is an algorithm that belongs to the machine learning category that can be used for classification in sentiment analysis models and has proven that this method can provide models with the best level of accuracy. In addition, NB is very efficient for use in a classification model for multivariate analysis [28]. The equation of the naive Bayes method can be seen in (3). In the naive Bayes equation, a probability of c being true given that x is true is sought.

$$P(c|x) = \frac{P(x|c) P(c)}{P(x)} \quad (3)$$

2.4. Genetic algorithm and particle swarm optimization

Genetic algorithm (GA) and particle swarm optimization (PSO) are optimization algorithms that allow obtaining the best weights for the model we want [29]–[31]. In this research, NN and NB are optimized so that they can have an effect, namely increasing the level of classification accuracy for coastal sentiment review. The PSO calculation process is shown in (4).

$$V_i^{t+1} = \omega V_i^t + c_1 r_1 (P_{best(i)}^t - P_i^t) + c_2 r_2 (P_{bestglobal}^t - P_i^t) \quad (4)$$

3. RESULTS AND DISCUSSION

The experimental results have produced several models with different levels of accuracy according to the method produced. In this study, the search for a model used the Core i7 computer specifications and 8GB of memory, in addition to data analysis, the RapidMiner studio 9.0 software based on the Windows 7 operating system was used for data analysis. This section of the chapter shows some experimental results from the proposed models according to predetermined scenarios, namely neural networks, naïve Bayes, and optimization models from the two methods. In this article, each model either has different or similar results depending on the experiments conducted.

3.1. Neural network experiment

The application of the neural network algorithm in this experiment is set by several predetermined parameters, to produce the best accuracy value, this is done by conducting experiments. The first stage of the process is to apply a classical neural network (NN), in this model optimization is carried out using only the

information gain (IG) method. The results of the NN implementation experiment are shown in Table 1 and Table 2, in this experiment the cross-fold validation parameter=10. The results shown in Table 1 and Table 2 show that the average level of accuracy produced is in the range of 74% to 76%. The model with the highest level of accuracy is 76.15% using the training cycle parameter=700 and momentum 0.9 using stratified sampling as shown in Table 1. This accuracy value is the same as the model generated using NN+IG by applying the learning parameter. rate=0.04, momentum=0.9, and training cycle=900 using stratified sampling as shown in Table 2. The experiments that have been carried out with the application of NN+IG using fold=10 are still not good enough because they are not in line with expectations. The desired level of accuracy is still below 80%, so it is deemed necessary to improve again. The next experiment was carried out by applying NN and IG using cross-validation fold=5, at this stage, an experiment was carried out to find out the difference when changing the value of different fold parameters. Broadly speaking, the results of this model produce accuracy values as shown in Table 3 and Figure 2.

Table 1. The accuracy of neural network with IG using momentum parameters 0.9

Training cycle	Stratified Sampling	Shuffled Sampling	Linear Sampling
100	75.64%	74.36%	74.10%
200	75.64%	75.38%	72.31%
300	75.38%	75.38%	72.05%
400	75.90%	75.13%	71.28%
500	75.90%	74.87%	71.03%
600	75.90%	74.87%	71.03%
700	76.15%	75.38%	71.03%
800	75.90%	75.13%	70.77%
900	75.64%	74.87%	70.51%
1000	75.38%	74.62%	70.51%

Table 2. The accuracy of neural network+IG with momentum 0.9 and training cycle 900

Learning rate	Stratified Sampling	Shuffled Sampling	Linear Sampling
0.01	76.67%	74.36%	72.31%
0.02	76.15%	75.38%	71.28%
0.03	75.90%	74.36%	71.03%
0.04	76.15%	74.36%	71.03%
0.05	75.64%	74.87%	71.03%
0.06	74.62%	75.13%	70.77%
0.07	75.13%	74.36%	70.77%
0.08	75.64%	74.36%	70.77%
0.09	75.38%	74.62%	70.77%
0.1	75.38%	74.87%	71.03%

Table 3. The accuracy of neural network with IG using momentum parameters 0.9 with fold=5

Training cycle	Stratified Sampling	Shuffled Sampling	Linear Sampling
100	75.13%	74.36%	55.64%
200	75.38%	73.85%	56.92%
300	75.13%	73.85%	67.95%
400	75.64%	73.85%	67.95%
500	75.38%	73.85%	67.95%
600	74.87%	73.59%	67.95%
700	74.62%	74.10%	67.69%
800	74.36%	74.10%	68.21%
900	74.36%	74.10%	67.95%
1000	73.85%	74.10%	67.95%

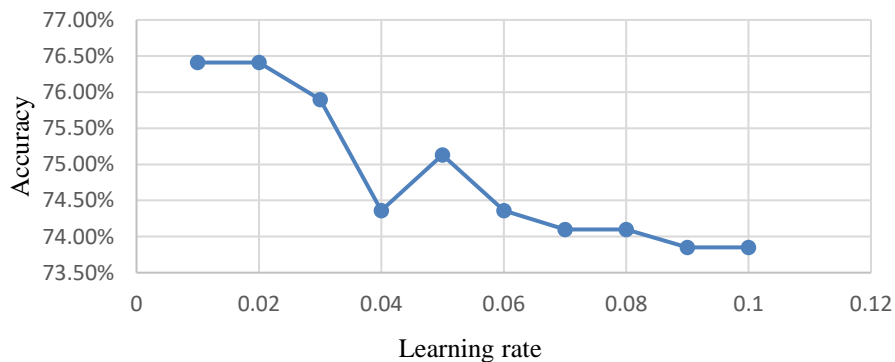


Figure 2. The performance model using NN+IG with fold=5, momentum=0.3, training cycle=900

Figure 3 shows the best model of NN+IG, obtained by using the momentum parameter value=0.3 and training cycle=900 resulting in the best accuracy rate of 76.41%. This experiment uses a data validation method with stratified sampling. The results of experiments carried out using fold=5 are almost the same as

previous studies, namely using fold=10, which is around 76%. The accuracy level equation of the models that have been obtained gives the impression that the model applied is a classical neural network is not enough if it is only optimized using information gain, therefore other efforts by using the application of other optimization methods need to be pursued.

3.2. Feature weights in neural network

Efforts to increase the accuracy of the NN model are to optimize using feature weights. At this stage, the NN is optimized using the PSO particle swarm optimization and genetic algorithm (GA) methods. The performance value of the experimental results at this stage is shown in Table 4.

Table 4. Performance model NN using genetic algorithm

Model	Population	Selection scheme	crossover type	Accuracy	Time (minute)
Model 1	5	tournament	uniform	76.67%	13.25
Model 2	10	tournament	uniform	77.44%	31.06
Model 3	5	roulette wheel	uniform	76.41%	27.12

In Table 4, the parameters specified in NN are learning rate=0.01, momentum=3, and training cycle=900. Data validation using cross-validation was set to fold=10 through the stratified sampling method. It can be seen in this experiment that the highest accuracy rate is 77.44% with the GA parameter set population=10, selection scheme=tournament, crossover type=uniform, with an analysis processing time of 31.06 minutes. The results in Table 4 on average are still below 80% so efforts are still needed to increase the level of accuracy. In addition, it appears that the best model produced is the model that requires more time to find.

The next experimental result is to optimize the model using PSO. The PSO method is applied to optimize the weight values of the neural network algorithm. The parameters specified in this PSO are using the parameter learning rate 0.01, momentum=0.3, and training cycle=700. The results of the accuracy value for the NN+IG+PSO model can be seen in Table 5, the highest accuracy value is 76.92%. Model evaluation using feature weights optimization is done by comparing the models that have been obtained based on the highest level of accuracy produced. Table 6 and Figure 3 show the performance differences of all optimized models, it can be seen that the highest level of accuracy is as follow.

Table 5. NN optimization model performance using PSO

Population	Sampling Type	Fold	Accuracy
5	Stratified	10	76.67%
6	Stratified	5	76.92%

Table 6. Comparison NN with feature weight

Fold	Model	Accuracy
10	NN_IG	76.67%
5	NN_IG+PSO	76.92%
10	NN_IG+GA	77.44%

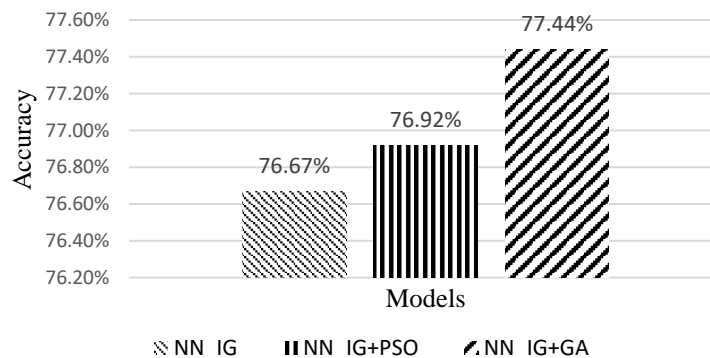


Figure 3. Graphic comparison NN+FW

3.3. Naïve Bayes experiment

The next research stage is to apply the naïve Bayes (NB) method to the sentiment review model for coastal assessments. The NB model at this stage produces a level of sentiment classification accuracy that is still not optimal and by what is desired. The feature weights method on the NB model is applied using the PSO and GA algorithms. The experimental results obtained in the FS-based NB model are shown in Table 7 and Table 8. In Table 7, an optimization model using the genetic algorithm (GA), the parameter value is determined by using population=5. In addition, the highest accuracy value obtained is 86.61% where the GA parameters set are using selection scheme=tournament, crossover type=uniform, in addition to model validation using fold=9 with sampling type=linear. The same results were obtained in model number 7 with the highest accuracy of 86.81% using fold=9 with sampling=linear.

Other models that are slightly different provide the highest accuracy rate of 85.08% where the PSO model is assigned the parameters population=5, fold=9, and sampling=linear. The other experiments shown in Table 9 have determined the PSO model using the population parameter = 10 and there is a difference in accuracy, which is 87.11%, slightly higher than the previous model in Table 8. These results show that the difference in the value of the population parameter gives a performance that has a different accuracy value.

3.4. Evaluation model and recommendation

The search that was carried out to get the best model for the classification of sentiment review assessment of the coast of the southern region of Java in Indonesia with the best level of accuracy is the focus of this article. Various experiments are conducted to get the best model, including optimizing the model using feature weights (FS). FS methods such as genetic algorithms and particle swarm optimization are applied to the main algorithms, namely neural networks and naïve Bayes. Based on the results obtained, it was found that the model with the highest accuracy rate was 87.11%, which we named the NB_IG+PSO model.

The best model in Figure 4 that produces the highest level of classification accuracy is the one that use the naïve Bayes method based on information gain (IG) with PSO algorithm feature weights optimization. There is a slight difference between the models that use NN as the main method, but the results are not much different, the accuracy rate is only 0.5%. The difference occurs because apart from the different methods used, the parameter values set in each method are slightly different as well as the parameters set.

The experiments carried out and those that have been proposed in this article have not yet reached the perfect model. It is because several other efforts can still be made, especially during the data preprocessing stage. The process that has not been optimally carried out in this research is at the stage of the steaming process, because the data used is in the form of text in Indonesian, so there is a slight difference in the method with the existing one, which is in English. Future research can be a challenge as more in-depth research is needed to get the best model with a better level of accuracy.

Table 7. Results of model experiments using GA-based NB

No_model	Selection scheme	crossover type	Fold	Sampling	Accuracy (%)
1	tournament	uniform	9	linear	86.61
2	tournament	uniform	10	linear	80.51
3	tournament	uniform	9	shuffled	76.15
4	tournament	uniform	10	shuffled	76.41
5	tournament	uniform	9	Stratified	76.71
6	tournament	uniform	10	Stratified	77.18
7	tournament	shuffled	9	linear	86.61
8	tournament	shuffled	10	linear	80.51
9	tournament	shuffled	9	shuffled	76.15
10	tournament	shuffled	10	shuffled	76.41
11	tournament	shuffled	9	Stratified	76.71
12	tournament	shuffled	10	Stratified	77.18
13	roulette wheel	uniform	9	linear	86.34
14	roulette wheel	uniform	10	linear	81.28
15	roulette wheel	uniform	9	shuffled	76.17
16	roulette wheel	uniform	10	shuffled	76.67
17	roulette wheel	uniform	9	Stratified	76.41
18	roulette wheel	uniform	10	Stratified	76.92
19	roulette wheel	shuffled	9	linear	86.34
20	roulette wheel	shuffled	10	linear	81.28
21	roulette wheel	shuffled	9	shuffled	76.17
22	roulette wheel	shuffled	10	shuffled	76.67
23	roulette wheel	shuffled	9	Stratified	76.41
24	roulette wheel	shuffled	10	Stratified	76.92

Table 8. Performance of model NN_IG+PSO using population=5

Population	Sampling	Fold	Accuracy
5	linear	9	85.08%
5	linear	10	81.54%
5	shuffled	9	77.20%
5	shuffled	10	77.18%
5	Stratified	9	77.17%
5	Stratified	10	77.44%

Table 9. Performance of model NN_IG+PSO by using population=10

Population	Sampling	Fold	Accuracy
10	linear	9	87.11%
10	linear	10	83.33%
10	shuffled	9	77.69%
10	shuffled	10	77.17%
10	Stratified	9	77.95%
10	Stratified	10	77.44%

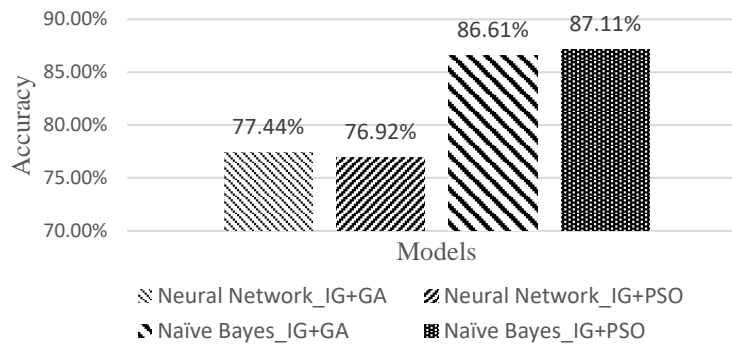


Figure 4. Evaluation performance model

4. CONCLUSION

A sentiment review model for assessing the state of the coast of the southern region of Java in Indonesia has been produced, namely by using the naive Bayes hybrid method based on information gain and particle swarm optimization and we call it the NB_IG+PSO model. The highest best model proposed in this article is 87.11%. The model obtained still has deficiencies in the process of finding the best parameter values, therefore efforts are needed to determine more precise parameter values. In addition, other methods need to be tested on this model to find out whether the proposed model can be optimized and whether there is an increasing level of accuracy. Other methods that can be applied for future research are by implementing support vector machines (SVM) and k-NN, as well as other methods that can be recommended according to the data type of the dataset since several methods do not match the characteristics of the data used.

ACKNOWLEDGEMENTS

The author would like to thank the Directorate General of Vocational Education, Ministry of Education, Culture, Research, and Technology, which has provided research funding through the Beginner Lecturer Research (PDP) scheme in 2022.




REFERENCES

- [1] N. Punetha and G. Jain, "Bayesian game model based unsupervised sentiment analysis of product reviews," *Expert Systems with Applications*, vol. 214, p. 119128, Mar. 2023, doi: 10.1016/j.eswa.2022.119128.
- [2] M. Zhang, L. Sun, G. A. Wang, Y. Li, and S. He, "Using neutral sentiment reviews to improve customer requirement identification and product design strategies," *International Journal of Production Economics*, vol. 254, p. 108641, Dec. 2022, doi: 10.1016/j.ijpe.2022.108641.
- [3] C. V. D., "Hybrid approach: naive bayes and sentiment VADER for analyzing sentiment of mobile unboxing video comments," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 5, p. 4452, Oct. 2019, doi: 10.11591/ijece.v9i5.pp4452-4459.
- [4] L. K. Ramasamy, S. Kadry, Y. Nam, and M. N. Meqdad, "Performance analysis of sentiments in Twitter dataset using SVM models," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 3, p. 2275, Jun. 2021, doi: 10.11591/ijece.v11i3.pp2275-2284.
- [5] N. Kewsuwun and S. Kajornkasirat, "A sentiment analysis model of Agritech startup on Facebook comments using naive Bayes classifier," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 3, p. 2829, Jun. 2022, doi: 10.11591/ijece.v12i3.pp2829-2838.
- [6] D. Apriliani, T. Abidin, E. Sutanta, A. Hamzah, and O. Somantri, "SentiHotel: a sentiment analysis application of hotel services using an optimized neural network," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 3, Jun. 2021, doi: 10.11591/eei.v10i3.3040.
- [7] Y. H. Hu, Y. L. Chen, and H. L. Chou, "Opinion mining from online hotel reviews – A text summarization approach," *Information Processing and Management*, vol. 53, no. 2, pp. 436–449, 2017, doi: 10.1016/j.ipm.2016.12.002.




- [8] S. Verma, "Sentiment analysis of public services for smart society: Literature review and future research directions," *Government Information Quarterly*, vol. 39, no. 3, p. 101708, Jul. 2022, doi: 10.1016/j.giq.2022.101708.
- [9] C. Zhang, D. Jia, L. Wang, W. Wang, F. Liu, and A. Yang, "Comparative research on network intrusion detection methods based on machine learning," *Computers & Security*, vol. 121, p. 102861, Oct. 2022, doi: 10.1016/j.cose.2022.102861.
- [10] F. M. Delpino, A. K. Costa, S. R. Farias, A. D. P. Chiavegatto Filho, R. A. Arcêncio, and B. P. Nunes, "Machine learning for predicting chronic diseases: a systematic review," *Public Health*, vol. 205, pp. 14–25, Apr. 2022, doi: 10.1016/j.puhe.2022.01.007.
- [11] M. E. Shaik, M. M. Islam, and Q. S. Hossain, "A review on neural network techniques for the prediction of road traffic accident severity," *Asian Transport Studies*, vol. 7, p. 100040, 2021, doi: 10.1016/j.eastsj.2021.100040.
- [12] F. Chen and Y. Huang, "Knowledge-enhanced neural networks for sentiment analysis of Chinese reviews," *Neurocomputing*, no. xxxx, 2019, doi: 10.1016/j.neucom.2019.08.054.
- [13] W. Liao, B. Zeng, J. Liu, P. Wei, X. Cheng, and W. Zhang, "Multi-level graph neural network for text sentiment analysis," *Computers and Electrical Engineering*, vol. 92, p. 107096, Jun. 2021, doi: 10.1016/j.compeleceng.2021.107096.
- [14] F. Yao and Y. Wang, "Domain-specific sentiment analysis for tweets during hurricanes (DSSA-H): A domain-adversarial neural-network-based approach," *Computers, Environment and Urban Systems*, vol. 83, p. 101522, Sep. 2020, doi: 10.1016/j.compenvurbysys.2020.101522.
- [15] M. J. Sánchez-Franco, A. Navarro-García, and F. J. Rondán-Cataluña, "A naive Bayes strategy for classifying customer satisfaction: A study based on online reviews of hospitality services," *Journal of Business Research*, vol. 101, pp. 499–506, Aug. 2019, doi: 10.1016/j.jbusres.2018.12.051.
- [16] H. Kang, S. J. Yoo, and D. Han, "Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews," *Expert Systems with Applications*, vol. 39, no. 5, pp. 6000–6010, 2012, doi: 10.1016/j.eswa.2011.11.107.
- [17] D. Apriliani, T. Abidin, E. Sutanta, A. Hamzah, and O. Somantri, "Sentiment analysis for Indonesia hotel services review using optimized neural network," in *Journal of Physics: Conference Series*, Jun. 2020, vol. 1538, no. 1, p. 12060. doi: 10.1088/1742-6596/1538/1/012060.
- [18] W. Qian, Q. Ye, Y. Li, and S. Dai, "Label distribution feature selection with feature weights fusion and local label correlations," *Knowledge-Based Systems*, vol. 256, p. 109778, Nov. 2022, doi: 10.1016/j.knosys.2022.109778.
- [19] R. Shang *et al.*, "Unsupervised feature selection via discrete spectral clustering and feature weights," *Neurocomputing*, vol. 517, pp. 106–117, Jan. 2023, doi: 10.1016/j.neucom.2022.10.053.
- [20] Q. Li, S. Liu, and M. Gao, "Multiobjective optimization of roadheader shovel-plate parameters using gray weight and particle swarm optimization," *IEEE Access*, vol. 10, pp. 104555–104566, 2022, doi: 10.1109/ACCESS.2022.3210165.
- [21] O. Somantri, "Dataset of coastal assessment-Teluk Penyau beach," Zenodo, 2022. Accessed: Jan. 14, 2022. [Online], Available: <https://zenodo.org/records/10004643> (
- [22] A. Thakkar and K. Chaudhari, "Predicting stock trend using an integrated term frequency-inverse document frequency-based feature weight matrix with neural networks," *Applied Soft Computing*, vol. 96, Art. no. 106684, Nov. 2020, doi: 10.1016/j.asoc.2020.106684.
- [23] A. L. Pomerantsev and O. Y. Rodionova, "New trends in qualitative analysis: Performance, optimization, and validation of multi-class and soft models," *TrAC Trends in Analytical Chemistry*, vol. 143, p. 116372, Oct. 2021, doi: 10.1016/j.trac.2021.116372.
- [24] X. Deng, Q. Liu, Y. Deng, and S. Mahadevan, "An improved method to construct basic probability assignment based on the confusion matrix for classification problem," *Information Sciences*, vol. 340–341, pp. 250–261, May 2016, doi: 10.1016/j.ins.2016.01.033.
- [25] J. Xu, Y. Zhang, and D. Miao, "Three-way confusion matrix for classification: A measure driven view," *Information Sciences*, vol. 507, pp. 772–794, Jan. 2020, doi: 10.1016/j.ins.2019.06.064.
- [26] N. Portillo Juan and V. Negro Valdecantos, "Review of the application of Artificial Neural Networks in ocean engineering," *Ocean Engineering*, vol. 259, p. 111947, Sep. 2022, doi: 10.1016/j.oceaneng.2022.111947.
- [27] I. Casas, "Networks, Neural," in *International Encyclopedia of Human Geography*, Elsevier, 2020, pp. 381–385. doi: 10.1016/B978-0-08-102295-5.10410-X.
- [28] R. Blanquero, E. Carrizosa, P. Ramírez-Cobo, and M. R. Sillero-Denamiel, "Variable selection for Naïve Bayes classification," *Computers and Operations Research*, vol. 135, p. 105456, Nov. 2021, doi: 10.1016/j.cor.2021.105456.
- [29] Z. U. Haq, H. Ullah, M. N. A. Khan, S. Raza Naqvi, A. Ahad, and N. A. S. Amin, "Comparative study of machine learning methods integrated with genetic algorithm and particle swarm optimization for bio-char yield prediction," *Bioresource Technology*, vol. 363, p. 128008, Nov. 2022, doi: 10.1016/j.biortech.2022.128008.
- [30] R. P. Cherian, N. Thomas, and S. Venkitachalam, "Weight optimized neural network for heart disease prediction using hybrid lion plus particle swarm algorithm," *Journal of Biomedical Informatics*, vol. 110, p. 103543, Oct. 2020, doi: 10.1016/j.jbi.2020.103543.
- [31] T. Taher, H. Chantar, J. Too, M. Mafarja, H. Turabieh, and E. H. Houssein, "Boolean Particle Swarm Optimization with various Evolutionary Population Dynamics approaches for feature selection problems," *Expert Systems with Applications*, vol. 195, p. 116550, Jun. 2022, doi: 10.1016/j.eswa.2022.116550.

BIOGRAPHIES OF AUTHORS






Oman Somantri    received the S.Kom. degree in informatics engineering from STMIK Sumedang, Indonesia, in 2011 and the M.Kom degree in informatics engineering from Dian Nuwanto University, Semarang Indonesia. Currently, he is a lecturer at the Department of Cyber Security Engineering, Cilacap State Polytechnic, Indonesia. His research interests include data mining, text mining, sentiment analysis, and natural language processing. He can be contacted at email: oman_mantri@yahoo.com.



Santi Purwaningrum    received the S.Kom. and M.Kom. degrees in informatics engineering from Dian Nuswantoro University, Semarang, Indonesia, in 2016 and 2018. She is currently the lecturer of informatics engineering, Cilacap State Polytechnic, Indonesia. His research interests include the natural language processing, and text mining. She can be contacted at email: santi.purwaningrum@pnc.ac.id.



Ratih Hafsarah Maharrani    received the S.Kom. and M.Kom. degrees in informatics engineering from Dian Nuswantoro University, Semarang, Indonesia, in 2007 and 2010. She is currently the lecturer of cyber security engineering, Cilacap State Polytechnic, Indonesia. His research interests include the decision support system, and data mining. She can be contacted at email: ratih.hafsarah@pnc.ac.id.