BAREKENG: Jurnal Ilmu Matematika dan Terapan
March 2022 Volume 16 Issue 1 Page 007–014
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol16iss1pp007-014

COMBINATION OF KNN AND PARTICLE SWARM OPTIMIZATION (PSO) ON AIR QUALITY PREDICTION

Sugandi Yahdin[1], Anita Desiani[2*], Shania Putri Andini[3], Dian Cahyawati[4], Rifkie Primartha[5], Muhammad Arhami[6], Ditia Fitri Arinda[7]

[1,2,3,4]Mathematics Department, Mathematics and Natural Science Faculty, Universitas Sriwijaya
[5]Informatics Engineering, Computer Science Faculty, Universitas Sriwijaya
[7]Nutrition Department, Public Health Faculty, Universitas Sriwijaya
Jl Palembang-Prabumulih, Inderalaya, Ogan Ilir, Sumatera Selatan, 30862, Indonesia
[6]Informatics Engineering, Polteknik Negeri Lhokseumawe
Jl Banda-Medan, Blang Mangat, Lhokseumawe Aceh, 24301, Indonesia

Corresponding author e-mail: [2] anita_desiani@unsri.ac.id*

Abstract. *The increase in the use of energy sources causes air pollution. The Air Pollutant Index (API) is information about the air quality of a place and at a certain time. API has several parameters, namely SO_2, PM_{10}, NO_2, O_3, and CO. In this study, the KNN method was used to assist categorize air quality. However, all training data were used during the classification process with KNN causes a long prediction process. Another problem with KNN is difficult to determine the optimal value of the K parameter in KNN. The Particle Swarm Optimization (PSO) method can be used for problems on KNN. Therefore, the aim of this study is to predict air quality based on the API by combining the KNN-PSO method. The dataset used is the API dataset for the DKI Jakarta area 2017-2019 totaling 1075 data. The results showed the accuracy for the KNN-PSO method was 98.42% with a precision value of 97.75% and a recall value of 98.13%. To further analyze the results on the combined method, the results of this study were compared with the KNN method only. The results obtained from the KNN method are lower than the KNN-PSO method. So it can be concluded that the KNN-PSO method is great and robust in air quality classification or prediction.*

Keywords: *air quality, air pollutant index, prediction, k-nearest neighbors, particle swarm optimization*

Article info:
Submitted: 15th July 2021 *Accepted: 07th January 2022*

How to cite this article:
S. Yahdin, A. Desiani, S. P. Andini, D. Cahyawati, M. Arhami and N. Eliyati, "COMBINATION OF KNN AND PARTICLE SWARM OPTIMIZATION (PSO) ON AIR QUALITY PREDICTION", *BAREKENG: J. Il. Mat. & Ter.*, vol.16, iss.1, pp. 007-014, Mar, 2022.

1. INTRODUCTION

The development of industrialization and urbanization increases the use of high energy sources and produces residual energy combustion which can cause air pollution [1]. Air pollution has a negative impact, such as human health problems, especially children, women or the elderly who suffer from respiratory problems [2]. To monitor air quality, the government sets the National Ambient Air Quality Standard which is used as the maximum limit for the levels of elements present in ambient air [3]. Air pollution is assessed from the concentration of several parameters compiled in the Air Pollution Standard Index (ISPU). In the Decree of the State Minister for the Environment Number KEP45/MENLH/1997, ISPU is used as information on air quality in a place at a certain time. ISPU has parameters according to Decree of the Head of the Environmental Impact Management Agency Attachment Number 107 of 1997, which are SO_2, PM_{10}, NO_2, O_3, and CO. From the five parameter values obtained, those will be classified based on the highest value [3]. To assist in determining the category in the ISPU, it can predict or classify the air quality [1].

One of the popular and simple methods for classification problems is KNN. KNN has the advantage that the calculation is simple, easy to implement and can be used on data with multiple labels or classes and gives significant results. Some researchers agree that KNN is the most popular classification method in data science [4]–[6]. Several studies have shown good performance from KNN including Munter et. al [7] compared KNN, Kmeans and Supervised Neural Network for network flow classification. The results of this study indicate that KNN provides the best performance with the highest accuracy value. Zing & Bay uses KNN for medical health classification [8] with great accuracy. Another study was conducted by Xiong and Yaou [9] who used KNN to classify adaptive thermal comfort. Imandoust [10] have used KNN to predict economic events.

K-Nearest Neighbors or KNN is a machine learning supervised algorithm used for classification or prediction. The KNN manipulates the training data and classifies the new test data based on distance by finding the k-th neighbor closest to the test data, and classified according to the majority class label. KNN does not work optimally for predictions on datasets that have many attributes and the need for selecting the right parameter to get good accuracy results [11]. Choosing the optimal K value to achieve maximum model accuracy has always been a challenge in the application of KNN [12], [13]. Mismatch in choosing the value of K can cause the model to be over/under fit. If K value is too small, it can cause noise in the dataset to have a high impact on predictions, but if K value is too large it can make the computational cost expensive. The main disadvantage of KNN is computationally inefficient and difficult to select the "correct" value of K, but the advantages of KNN are versatile for different proximity calculations and very intuitive [14], [15]. Selection of K is very important for the model, if the wrong choice can cause the model to be over or under fit. K value that is too small can cause noise in the data to have a high influence on the prediction, but a K value that is too large can make computationally expensive. Its main drawbacks are computationally inefficient and difficult to choose the correct value of K, but the advantage of this algorithm is versatile for different proximity calculations and is very intuitive

Particle Swarm Optimization (PSO) is an optimization method in computing that iteratively works to improve candidate solutions based on a given quality measure [16]. PSO generates a population of candidate solutions called particles, and moves these particles around the search space based on their position and velocity using mathematical equations. Each candidate solution is updated when a better position is found by another particle. It aims to get the best solution on a problem. PSO is a heuristic method that does not have assumptions on an optimized problem so it can generate a large candidate solution space[17], [18]. In KNN the selection of the best K parameter values is chosen by trial and test. Trial and test generation of k parameter values resulted in the chosen k not necessarily being the best k value in the problem of air quality classification and prediction. To overcome the shortcomings of KNN, this study combines KNN and PSO techniques, namely KNN-PSO. The expectation of the combination of these two methods is hoped the result of classification or prediction process on air quality can be more accurate by optimizing the value of the k parameter on the KNN with the PSO optimization algorithm.

The previous research on air quality prediction was Jun ma et al [1] using the Transferred bi-directional long short term memory method in deep learning. The other research was using Cloud Model Granulation method in predicting air quality in Wuhan [19]. Air quality prediction was also used for the case in Shenyang using the Random Forest method [20]. Unfortunately, this research only measures the accuracy value and does not measure the other performances. This study was proposed new method the K-Nearest Neighbor and

Particle Swarm Optimization (KNN-PSO) method to predict air quality. To evaluate the performance of KNN-PSO, it would be measured the other performances, namely precision and recall.

2. RESEARCH METHODS

2.1 Data Collecting

This study used the data from the DKI Jakarta Provincial Environmental Service in 2017-2019 which can be accessed on the website https://data.jakarta.go.id/dataset. The dataset consisted of 10 attributes, which were five attributes for parameters according to the Decree of the Head of the Environmental Impact Management Agency Number 107 of 1997, 4 information attributes, and 1 category attribute containing the status of air quality on the day when the ISPU value was recorded. Four information attributes including max which is the highest value for the five ISPU parameters, critical is the name of the parameter that has the highest value, location is the location of the Air Quality Monitoring Station (SPKU) which gets the highest parameter value, and category is the status of air quality. The five ISPU parameters, which are PM_{10} *(Particulates)*, SO_2 *(Sulfur Dioxide)*, *CO (Carbon Monoxide)*, O_3 *(ozone)*, and NO_2 *(Nitrogen Dioxide)* are monitored and recorded by the SPKU. ISPU dataset can be seen in Table 1.

Table 1. Air Pollutant Standard Index Dataset in 2018-2019 (DKI Jakarta Environmental Service)

Date	PM_{10}	...	NO_2	Max	Critical	Location	Category
1/1/17	76	…	9	76	PM_{10}	Rg5	MEDIUM
2/1/17	23	…	14	39	O_3	Reg5	GOOD
3/1/17	53	…	23	101	O_3	Reg5	UNHEALTHY
4/1/17	53	…	15	57	O_3	Reg3	MEDIUM
5/1/17	44	…	10	44	PM_{10}	Reg3	GOOD
6/1/17	20	…	5	31	O_3	Reg3	GOOD
7/1/17	26	…	5	46	O_3	Reg5	GOOD
8/1/17	62	…	17	73	O_3	Reg5	MEDIUM
9/1/17	45	…	14	63	O_3	Reg3	MEDIUM
10/1/17	42	…	17	95	O_3	Reg3	MEDIUM
11/1/17	40	…	19	81	O_3	Reg2	MEDIUM
12/1/17	37	…	17	83	O_3	Reg5	MEDIUM
13/1/17	51	…	21	93	O_3	Reg5	MEDIUM
14/1/17	33	…	14	73	O_3	Reg5	MEDIUM
15/1/17	30	…	12	72	SO_2	Reg4	MEDIUM
16/1/18	25	…	10	70	SO_2	Reg4	MEDIUM
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
31/12/19	52	…	9	55	O_3	Reg2	MEDIUM

Determination of the air quality category is based on the value of those five parameters. The range of category values for the ISPU value is 0-50, which is a good category, meaning the level of air quality has no effect on all life, 51-100 is a medium category, meaning the level of air quality only affects sensitive plants, 101-199 is an unhealthy category, meaning the air quality of this level can be harmful to humans and sensitive animals, 200-299 categories are very unhealthy and can cause health problems in a number of exposed population segments, and more than 300 is hazardous categories, meaning the air quality of this level can seriously harm the health. However, in the ISPU data in this study, only 4 categories were recorded, which are good, medium, unhealthy, and very unhealthy. Based on Table 1, the amount of data for each label was 142 data for good labels, 509 data for medium labels, 424 data for unhealthy labels, and 35 data for very unhealthy labels. The total data recorded was 1075 data. DKI Jakarta had 5 SPKU which spread over 5 regions. The SPKU location in DKI Jakarta had its own code, such as Reg1 in Central Jakarta, Reg2 in North Jakarta, Reg3 in South Jakarta, Reg4 in East Jakarta, and Reg5 in Central Jakarta. The dataset composition of the five SPKU consisted of 384 data from Reg1, 103 data from Reg2, 172 data from Reg3, 51 data from Reg4, and 365 data from Reg5 which was collected into the ISPU dataset in Table 1.

2.2 Data Pre-Processing

In this study, pre-processing was carried out to select the attributes used during prediction. The selection of attributes was based on the Decree of the Head of Bapedal Number 17/1997 about parameters used in the air quality category. Before prediction process, the data would be divided using the n-fold cross validation method. N-fold cross validation begins by dividing the data as many as n partitions equally and the testing and training process is also carried out n times [21].

2.3 Air Quality Prediction Using KNN-PSO

The steps in the KNN-PSO algorithm are as follows:

1. Determine the best *k* parameters by performing the KNN algorithm. Determine the value of *k* based on the best accuracy value obtained by the KNN method. The algorithm steps in the KNN method are as follows [22]:
   a. Specify the *k* parameters that will be tested.
   b. Calculate the distance between the new data and all training data using Euclidean Distance on equation (1)[13].
   $$d_i = \sqrt{\sum_{i=1}^{n}(x_{2i} - x_{1i})^2} \quad (1)$$
   where d_i is defined as the distance between x_1 and x_2, x_1 is the sample data and x_2 is the test data, *i* is the data variable, and *p* is the data dimension.
   c. Sort the results of the calculation of the distance from the smallest and determine the nearest neighbor based on the *k*-th minimum distance.
   d. The data will be classified to the class that has the most values in the same number of classes.
2. Initialize the population and parameters in PSO. One particle in PSO has a particle position that is $P_i(t) = p_{i1}(t), p_{i2}(t), \dots, p_{id}(t)$ and has a randomly initialized particle velocity between 0 and 1 written in $V_i(t) = v_{i1}(t), v_{i2}(t), \dots, v_{iN}(t)$, where *i*=1,2,3,..., n is the particle index, d is the dimension of the particle and *t* is the iteration.
3. Determine the fitness value based on the accuracy value obtained through the KNN process as in step 1.
4. Determine the local best and global best. The local best is determined from a particle that has the best fitness value in all iterations. The global best is determined from all particles that have the best fitness value in all iterations.
5. Determine the velocity for each particle with equation (2):
   $$V_{i,d}(t+1) = (W)(V_{i,d}(t)) + c_1 r_1 \left(P_{i,d}^{L}(t) - P_{i,d}(t)\right), + c_2 r_2 (P_d^G(t) - P_{id}(t)) \quad (2)$$
   where $P_{i,d}^{L}$ represents the local best of the *i*-th particle in the d-dimensional, P_d^G represents the global best on the d dimension, W is the weight of inertia, c_1, c_2 is a positive constant called the learning factor, and r_1, r_2 is a random number with a value between 0 to 1.
6. Update the particle with the following equation:
   $$P_{i,d}(t+1) = V_{i,d}(t) + P_{i,d}(t+1) \quad (3)$$
7. Do the steps 3-5 until the specified iteration is complete.
8. Do all the steps to fix the particles to a predetermined iteration. The best particles in the last iteration will be used during the training process.

2.4 Evaluation of KNN-PSO Method

The results are evaluated based on the confusion matrix which generated from the confusion matrix method. The matrix was generated to show the accuracy of the classification of the dataset on the active and inactive classes of the algorithm used [23]. The confusion matrix was used to calculate the values of accuracy, precision, and recall. The confusion matrix consists of two classes, which are positive class and negative class. In the positive class, the one which predicted correctly by the machine is called by true positive (TP), while the wrong prediction is called by false positive (FP). In the negative class, the one which predicted correctly by the machine is called by true negative (TN) and the wrong prediction is called by false negative (FN).

The steps above are described on the flowchart which can be seen in Figure 1.

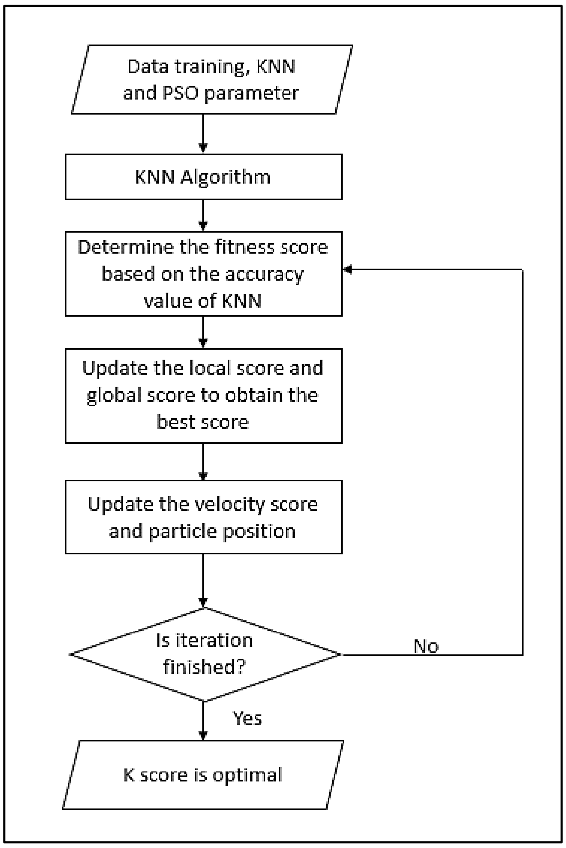

Figure 1. KNN-PSO Method on Proposed Method

Accuracy is used to measure the accuracy of the classification results correctly using equation (4).

$$\text{Acccuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}} \times 100\% \tag{4}$$

Precision in equation (5) is used to calculate the accuracy between the correct prediction results on the requested info.

$$\text{Precision} = \frac{TP}{(FP+TP)} \times 100\% \tag{5}$$

Recall is used to calculate the ratio of the selected relevant items to the total number of available items using equation (6).

$$\text{Recall} = \frac{\text{TP}}{(\text{TP+FN})} \times 100\% \tag{6}$$

3. RESULTS AND DISCUSSION

3.1. Data Pre-Processing

The attributes selected for air prediction were based on the Decree of the Head of Bapedal No. 17 in 1977. The attributes were *PM_{10} (Particulates), SO_2 (Sulfur Dioxide), CO (Carbon Monoxide), O_3 (ozone)*, and *NO_2 (Nitrogen Dioxide)*. The attributes which issued were date, max, critical and location, because these attributes were only additional attributes that contain information related to ISPU parameters. The data used in the study consisted of 5 parameter attributes and 1 attribute as a label. In the n-fold cross validation process, it used 10-fold that the data would be divided into 10 equal partitions.

3.2. Air Quality Prediction Using KNN-PSO

At the testing process, it used 10-fold cross validation. The best K was searched by testing several times with different K values to take the best accuracy value. To compare the prediction results, it will use the KNN method. In the KNN method, the results which used for comparison are the best accuracy values in 10 trials to find the K value. The test results in finding the best K value could be seen in Table 2.

Table 2. Accuracy of K Score Testing on Training Data

K	Accuracy
5	96.84%
6	96.74%
7	96.84%
8	96.84%
9	96.74%
10	96.84%
11	97.12%
12	96.47%
13	96.93%
14	97.02%
15	97.02%

From Table 2, it could be seen that from 10 trials, the highest accuracy value was when the value of K = 11 with an accuracy value of 97.12%. These results were obtained in cross validation of testing. The training and testing data were divided by itself and the level of accuracy which produced was influenced by the dataset [22]. From the value of K=11, the precision value was 95.96% and the recall value was 96.40%.

In the KNN-PSO method, the PSO method was used for attribute weighting with the best K search process based on the KNN algorithm as discussed in the research method [22]. In the KNN-PSO process, several experiments would be carried out based on the number of populations. The population size was determined to initialize the number of particles to be run in the PSO process. The bigger population will produce the better result, but the processing time takes a long time. Therefore, the number of populations which used were 5, 10, 15, 20, 25. The results could be seen in Table 3.

Table 3. Results Experiment of KNN-PSO for Some Populations

Number of Population	Accuracy	Process Duration
5	98.42%	10s
10	98.51%	14s
15	98.98%	22s
20	98.88%	27s
25	98.98%	48s

In Table 3, the highest accuracy value was showed when the population is 15-25, but the selected population was 15 because it had a shorter processing time. In the process, to find the global best on PSO, it was necessary to evaluate fitness. The fitness value was obtained from the accuracy value at the time of classification using the KNN method, then continued by determining the local best and global best. Then PSO would give weight to the attributes that affect the classification results in air quality predictions based on ISPU. The attribute weights on the ISPU can be seen in Table 4.

Table 4. Attribute Weight of ISPU Dataset Using KNN-PSO Method

Attribute	Weight
PM_{10} ((*Particulate*)	1
SO_2 (*Sulfur Dioxide*)	0,306
CO (Carbon Monoxide)	0
O_3 *(ozone)*	1
NO_2 (*Nitrogen Dioxide*)	0

From Table 4, it could be seen that the most influential attribute is O_3 *(ozon)*, because the parameter that most often had the highest value at ISPU was O_3 *(ozon)*. The *CO* and NO_2 attributes had a weight of 0 which means that the two attributes had no effect on the air quality category because in the dataset the two parameters never had a high value and affect air quality. The precision value for KNN-PSO was 98.86% and the recall value was 98.45%. From the two methods, KNN and KNN-PSO, the accuracy values have increased as well as the precision and recall values which have also increased. The comparison of the results of the KNN method and the KNN-PSO method could be seen in Figure 2.

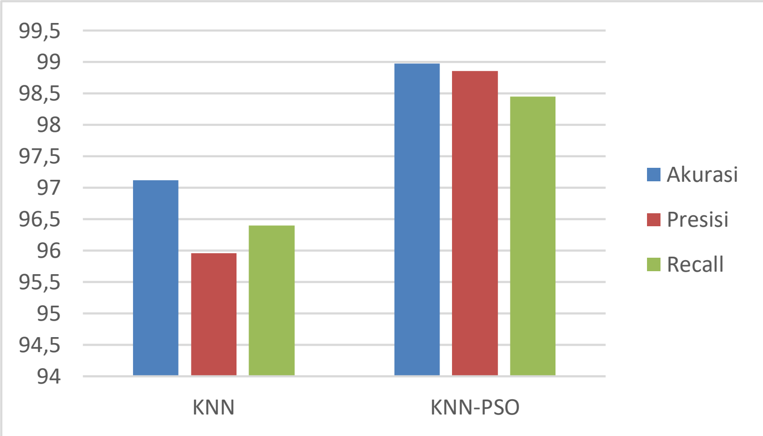

Figure 2. Results Comparison of KNN and KNN-PSO on The Study

To analyze the results further, the results of this study were compared with several previous studies. The comparison of research results can be seen in Table 5.

Table 5. Results Comparison of Proposed Method and Previous Researches

Author	Dataset	Method	Accuracy	Precision	Recall
Lin et al [24]	Air Quality Index in Wuhan, China	Cloud Model Granulation	71,43%	-	-
Zhao and Song [25]	Meteorological Data on Heating and $PM_{2.5}$ emission in Shenyang	Logistic Regrresion	76,43%	-	-
Ruiyun et.al [20]	Air Quality Index in Shenyang, China	Random Forest	-	81,6%	-
Yunus [23]	Chronic Kidney Disease (UCI)	KNN-PSO	97,25%	-	-
Mahardika et al [26]	Dataset Citrus Pest Symptoms	KNN-PSO	96,25%	-	-
Proposed Method	ISPU Dataset DKI Jakarta	KNN-PSO	98,98%	98,86%	98,45%

From Table 5, it can be seen some comparisons from several studies. Based on the dataset used, the research with the proposed method had a value of accuracy that greater than the research of Lin et al [24] and research by Zhao and Song [25], The two studies did not mention other assessments besides the accuracy value. Based on the method used, almost all studies with the same method have a fairly high accuracy value, although the highest value was in this study. In Yunus' research [23] and Mahardika et al [26], there were no other assessments, although it has a high accuracy value,. From this comparison, it can be concluded that the KNN-PSO method is very good for aiq quality predicting.

4. CONCLUSIONS

Based on the study results of air quality prediction, it showed that KNN-PSO was a good method for air quality prediction. The performances of KNN-PSO gave great results for accuracy, recall and precision. The performance results of KNN-PSO were above 90%. For evaluating performance of the proposed study, the KNN-PSO results were compared with conventional KNN results. The comparison showed that KNN-PSO results get the higher accuracy, precision, and recall score compared with the results of conventional

KNN and compared with the other studies. It can be concluded the combination of KNN and PSO namely KNN-PSO has been given alternative robust method for air quality prediction problem.

REFERENCES

[1] J. Ma, J. C. P. Cheng, C. Lin, Y. Tan, and J. Zhang, "Improving air quality prediction accuracy at larger temporal resolutions using deep learning and transfer learning techniques," *Atmos. Environ.*, vol. 214, 2019, doi: 10.1016/j.atmosenv.2019.116885.

[2] H. Maleki, A. Sorooshian, G. Goudarzi, Z. Baboli, Y. Tahmasebi Birgani, and M. Rahmati, "Air pollution prediction by using an artificial neural network model," *Clean Technol. Environ. Policy*, vol. 21, 2019, doi: 10.1007/s10098-019-01709-w.

[3] A. Kurniawan, "Pengukuran Parameter Kualitas Udara (Co, No2, So2, O3 dan Pm10) di Bukit Kototabang Berbasis Ispu," *J. Teknosains*, vol. 7, no. 1, pp. 1–82, 2018, doi: 10.22146/teknosains.34658.

[4] A. Pandey and A. Jain, "Comparative Analysis of KNN Algorithm using Various Normalization Techniques," *Int. J. Comput. Netw. Inf. Secur.*, vol. 9, no. 11, pp. 36–42, 2017, doi: 10.5815/ijcnis.2017.11.04.

[5] H. Liu, X. Li, and S. Zhang, "Learning Instance Correlation Functions for Multilabel Classification," *IEEE Trans. Cybern.*, vol. 47, no. 2, pp. 499–510, 2017, doi: 10.1109/TCYB.2016.2519683.

[6] X. Zhu, X. Li, and S. Zhang, "Block-Row Sparse Multiview Multilabel Learning for Image Classification," *IEEE Trans. Cybern.*, vol. 46, no. 2, pp. 450–461, 2016, doi: 10.1109/TCYB.2015.2403356.

[7] A. Munter, R. Razif, M. AbuAlhaj, M. Anbar, and S. Nizam, "A Preliminary Performance Evaluation of K-means, KNN and EM Unsupervised Machine Learning Methods for Network Flow Classification," *Int. J. Electr. Comput. Eng.*, vol. 6, no. 2, p. 778, 2016, doi: 10.11591/ijece.v6i2.8909.

[8] W. Xing and Y. Bei, "Medical Health Big Data Classification Based on KNN Classification Algorithm," *IEEE Access*, vol. 8, pp. 28808–28819, 2020, doi: 10.1109/ACCESS.2019.2955754.

[9] L. Xiong and Y. Yao, "Study on an adaptive thermal comfort model with K-nearest-neighbors (KNN) algorithm," *Build. Environ.*, vol. 202, p. 108026, 2021, doi: https://doi.org/10.1016/j.buildenv.2021.108026.

[10] S. B. Imandoust and M. Bolandraftar, "Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events : Theoretical Background," *Int. J. Eng. Res. Appl.*, vol. 3, no. 5, pp. 605–610, 2013.

[11] J. Y. Lee and M. P. Styczynski, "NS-kNN: a modified k-nearest neighbors approach for imputing metabolomics data," *Metabolomics*, vol. 14, no. 12, p. 153, 2018, doi: 10.1007/s11306-018-1451-8.

[12] J. Huang *et al.*, "Cross-Validation based K Nearest Neighbor Imputation for Software Quality Datasets: An Empirical Study," *J. Syst. Softw.*, vol. 132, pp. 226–252, 2017, doi: 10.1016/j.jss.2017.07.012.

[13] Okfalisa, I. Gazalba, Mustakim, and N. G. I. Reza, "Comparative Analysis of K-Nearest Neighbor and Modified K-Nearest Neighbor Algorithm for Data Classification," *Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng.*, pp. 294–298, 2017, doi: 10.1109/ICITISEE.2017.8285514.

[14] S. Zhang, D. Cheng, Z. Deng, M. Zong, and X. Deng, "A novel kNN algorithm with data-driven k parameter computation," *Pattern Recognit. Lett.*, vol. 109, pp. 44–54, 2018, doi: https://doi.org/10.1016/j.patrec.2017.09.036.

[15] D. J. O'Neil, "Nearest Neighbor Problem," in *Encyclopedia of GIS*, S. Shekhar, H. Xiong, and X. Zhou, Eds. Cham: Springer International Publishing, 2017, pp. 1421–1426.

[16] K. Movagharnejad and N. Vafaei, "A Comparison Between GA and PSO Algorithms in Training ANN to Predict the Refractive Index of Binary Liquid Solutions," *J. Chem. Pet. Eng.*, no. October, pp. 125–135, 2018, doi: 10.22059/JCHPE.2018.238595.1208.

[17] N. Hasanova, "A Comparative Study of Particle Swarm Optimization and Genetic Algorithm," *Qubahan Acad. J.*, vol. 1, no. 1, pp. 33–45, 2020, doi: 10.48161/qaj.v1n1a7.

[18] I. Rahman, P. M. Vasant, B. S. M. Singh, and M. Abdullah-Al-Wadud, "On the performance of accelerated particle swarm optimization for charging plug-in hybrid electric vehicles," *Alexandria Eng. J.*, vol. 55, no. 1, pp. 419–426, 2016, doi: https://doi.org/10.1016/j.aej.2015.11.002.

[19] S. Gambhir, S. K. Malik, and Y. Kumar, "PSO-ANN based diagnostic model for the early detection of dengue disease," *New Horizons Transl. Med.*, pp. 1–8, 2017, doi: 10.1016/j.nhtm.2017.10.001.

[20] R. Yu, Y. Yang, L. Yang, G. Han, and O. A. Move, "RAQ–A Random Forest Approach for Predicting Air Quality in Urban Sensing System," *Sensors (Switzerland)*, vol. 16, no. 1, 2016, doi: 10.3390/s16010086.

[21] R. D. Liklikwatil, E. Noersasongko, and C. Supriyanto, "Optimasi K-Nearest Neighbor dengan Particle Swarm Optimization untuk Memprediksi Harga Komoditi Karet," *e-Jurnal JUSITI (Jurnal Sist. Inf. dan Teknol. Informasi)*, vol. 7, no. 2, pp. 172–182, 2018, doi: 10.36774/jusiti.v7i2.252.

[22] N. Wati, A. Syukur, and A. Z. Fanani, "Prediksi Kelulusan Mahasiswa Menggunakan K-Nearest Neighbour Berbasis Particle Swarm Optimization," 2015.

[23] W. Yunus, "Algoritma K-Nearest Neighbor Berbasis Particle Swarm Optimization untuk Prediksi Penyakit Ginjal Kronik," *J. Tek. Elektro CosPhi*, vol. 2, no. 2, pp. 51–55, 2018.

[24] Y. Lin, L. Zhao, H. Li, and Y. Sun, "Air Quality Forecasting Based on Cloud Model Granulation," *Eurasip J. Wirel. Commun. Netw.*, vol. 216, pp. 1–10, 2018, doi: 10.1186/s13638-018-1116-3.

[25] C. Zhao and G. Song, "Application of Data Mining to the Analysis of Meteorological Data for Air Quality Prediction: A Case Study in Shenyang," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 81, pp. 1–5, 2017, doi: 10.1088/1755-1315/81/1/012097.

[26] K. W. Mahardika, Y. A. Sari, and A. Arwan, "Optimasi K-Nearest Neighbour Menggunakan Particle Swarm Optimization Optimasi K-Nearest Neighbour Menggunakan Particle Swarm Optimization pada Sistem Pakar untuk Monitoring Pengendalian Hama pada Tanaman Jeruk," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 9, pp. 3333–3344, 2018.