ISSN 2089-8673 (Print) | ISSN 2548-4265 (Online)
Volume 13, Issue 1, March 2024

PARTICLE SWARM OPTIMIZATION FOR OPTIMIZING PUBLIC SERVICE SATISFACTION LEVEL CLASSIFICATION

Tyastuti Sri Lestari[1], Ismaniah[2], Wowon Priatna[3]

[1,3]Informatika, Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya
[2]Teknik Industri, Fakultas Teknik, Universitas Bhayangkara Jakarta Raya

email: tyas@ubharajaya.ac.id[1], ismaniah@ubharajaya.ac.id[2], wowon.priatna@dsn.ubharajaya.ac.id[3]

Abstract

This research aims to categorize survey data to determine the level of satisfaction with the services provided by the village government as a public service provider. Villages or sub-districts currently offer services in response to community demand, although only partially or as efficiently as possible. The data collection technique used was distributing questionnaires to the village community. The method used for classification is the machine learning method. Before the classification process, feature selection is carried out at the data pre-processing stage using Particle Swarm Optimization (PSO), which has been proven to increase the accuracy of the classification values. The classification methods employed include Decision Tree (DT), Naive Bayes, Support Vector Machine (SVM), and K-Nearest Neighbor (KNN) algorithms for classification purposes. This study achieves the maximum level of accuracy in decision tree classification, attaining an accuracy rate of 97.74%. Subsequently, the KNN algorithm achieved an accuracy of 77.90%, the Nave Bayes algorithm achieved 64.4%, and the SVM algorithm, which yielded the lowest accuracy value, achieved 59.90%. Following the application of Particle Swarm Optimization (PSO) for optimization, the accuracy of the SVM and KNN algorithms improved to 98.3%. The Decision Tree algorithm achieved a value of 97.77%, while the Naive Bayes technique yielded a value of 69.30%.

Keywords: public employees' satisfaction, Classification, Feature Selection, Particle Swarm Optimization, Machine Learning.

Received: 26-10-2023 | **Revised:** 10-01-2024 | **Accepted:** 06-03-2024
DOI: https://doi.org/10.23887/janapati.v13i1.69612

INTRODUCTION

Public services are provided to members of the broad public who are either citizens or permanent residents of the country in question. Its execution helps achieve civic objectives. Law No. 32 of 2004 deals with regional government. In the framework of regional autonomy, government administration must prioritize community interests, particularly when it comes to the provision of public infrastructure and public administration. The village head is in responsible of managing the local administration, and village representatives work under him or her. Today's society is changing, and as a result, there is a definite need for better, faster, and more precise services as well as more sophisticated services.

The manner that villages or subdistricts are now meeting the requirements of their community may not be optimal; as a result, the community may suffer misery and material harm. To enhance the caliber of the services they receive, raise user satisfaction, and boost the efficacy of their complaint handling, it is essential to swiftly identify the users and provide a mechanism for classifying them[1]. The sole method service providers use to classify issues with the provision of public services is machine learning [2][3].

Several studies on categorizing public service satisfaction have been conducted in the past, including A decision tree approach used in research [4] on the classification of village service satisfaction, which yields an accuracy of 90.66%. Using LDA-SVM to classify complaint reports produces an accuracy of 79.85% [5], and Nave Bayes to classify complaints about public services. The Nave Bayes, KNN, SVM, and boosting algorithms are used to classify public service complaints, and they produce the best SVM accuracy when compared to other approaches[6].

There is a feature selection technique for classification that can increase accuracy by selecting a subset of features from the total dataset[7]. The classification findings are considered when choosing a subset of the feature space for the feature selection process

ISSN 2089-8673 (Print) | ISSN 2548-4265 (Online)
Volume 13, Issue 1, March 2024

[8]. When employing the feature optimization method Particle Swarm Optimization (PSO), the accuracy values of the Naive Bayes accuracy optimization method[7], the Support Vector Machine accuracy value optimization method[9] , and the KNN algorithm [10] can all be improved. Random Forest, Decision Tree, Nave Bayes, and KNN have all been enhanced with PSO for categorizing diabetic datasets [11]. Particle swarm optimization can increase the accuracy of Support Vector Machine and Decision Tree classification by selecting the right features[12].

The goal of this research is to use one o f the feature selection techniques, the Particle S arm Optimization Algorithm, to add optimization to improve the classification accuracy value of cl assification algorithms such as Decision Tree (DT), Nave Bayes, SVM, and KNN. Classify survey data to ensure satisfaction with the services offered by the village government as a public service provider[4].

METHOD

The flow of this research is shown in the figure 1. The dataset for this study came from a survey that was done and given out to residents of the villages in Bekasi's Tambun Selatan subdistrict. Table 1 lists the questionnaire statements. A total of 9999 individuals and fifteen people replied to the survey, which measures respondents' satisfaction with public services through a series of questions, in particular, that service providers provide. Table 2 lists the variables, measurements, descriptions, grades, and timescales for the research data used The values in Table I correspond to the following categories: represents 1 = Do Not Concur (DNC), 2 = Disagree (DA), 3 = Agree (A), 4 = Strongly Concur (SC).

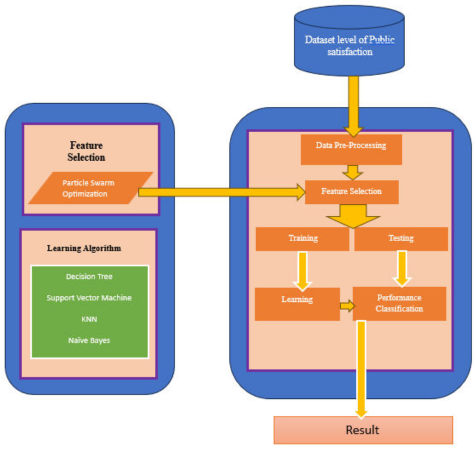

Figure 1. Research Design

The next stage is data pre-processing, where the dataset must be correctly processed before applying the classification model. This allows the dataset to be modeled with a classification algorithm. Microsoft Excel was used to process the survey's result data. The dataset was then split into two clusters and trained Google colaboratory using the k-Mean clustering algorithm. The dataset may be used to create a classification model displayed in Table 2, using the clustering result generated by two clusters to determine the class or target.

Table 1. Questionnaire

No.	Statement
1	The Village Chief's Office has a clean, comfortable space.
2	The completeness of the Village Head's Office's supporting facilities (such as parking lots and waiting rooms) is very good.
3	The employees appear very professional in providing services.
4	Have clear service time standards.
5	Readiness of employees on-site or in the space
6	The suitability of the services provided by Village Office employees with existing procedures is very appropriate.
7	Village Office service employees are responsive in dealing with problems that arise.
8	The readiness of Village Office employees to provide information is very clear and easy to understand
9	Village Office employees respond quickly, precisely, and efficiently to customer requests and complaints.
10	Village Office employees in providing services in accordance with agreed promises
11	The time and cost of service are very clear and definite.
12	The applicant's ability to contact authorized employees is very good.
13	The attentive attitude of Village Office employees in handling applicants' complaints is very good.
14	Village Office employees provide services and be friendly and polite to the community.

ISSN 2089-8673 (Print) | ISSN 2548-4265 (Online)
Volume 13, Issue 1, March 2024

15	Village Office employees serve fairly and non-discriminatively

Table 2. Characteristics of Research Data

Variable	Description	Possible Value
X1	Facilities and infrastructure	1= DNC; 2 = DA; 3 = A; 4= SC
X2	Cleanliness	1= DNC; 2 = DA; 3 = A; 4= SC
X3	Employee appearance	1= DNC; 2 = DA; 3 = A; 4= SC
X4	On-Time	1= DNC; 2 = DA; 3 = A; 4= SC
X5	Employee appearance	1= DNC; 2 = DA; 3 = A; 4= SC
X6	Conformity	1= DNC; 2 = DA; 3 = A; 4= SC
X7	Responsive	1= DNC; 2 = DA; 3 = A; 4= SC
X8	Accuracy of information	1= DNC; 2 = DA; 3 = A; 4= SC
X9	Attitude	1= DNC; 2 = DA; 3 = A; 4= SC
X10	Service Suitability	1= DNC; 2 = DA; 3 = A; 4= SC
X11	Cost Suitability	1= DNC; 2 = DA; 3 = A; 4= SC
X12	Easy to contact	1= DNC; 2 = DA; 3 = A; 4= SC
X13	Attention	1= DNC; 2 = DA; 3 = A; 4= SC
X14	Fair	1= DNC; 2 = DA; 3 = A; 4= SC
X15	Friendly	1= DNC; 2 = DA; 3 = A; 4= SC

Table 3. Characteristics of Research Data

Cluster	Sum	Information
Cluster 1	7043	Satisfied
Cluster 2	2956	Dissatisfied

Table 1's clustering results show that 7043 people are satisfied with cluster 1 and 2956 people are dissatisfied with cluster 2. Targets, labels, and classes for classification can be developed based on the outcomes of data clustering[13].

Three scenarios are used to divide training and testing data at the beginning of the classification process[14]. In scenario 1, the training data is partitioned into 70% and the testing data into 30%. In another case, the training data is split into 80% and the testing data into 20%. Lastly, the training data is separated into 90%. The proportion of testing data in the scenario is 10%.

In this study, four algorithms—decision trees, Naïve Bayes, support vector machines, and KNN—are used for the algorithm learning stage of classification. There are two stages of classification: the first uses PSO-free classification, and the second uses PSO-based classification.

Eliminating redundant and redundant characteristics from a dataset is the goal of the feature selection step [14], which is essential for processing high-dimensional sample data[15]. Particle Swarm Optimization (PSO) feature selection is used in this work at the learning step to optimize each piece of data generated[16]. PSO is used to select optimal attributes to use in the classification process. PSO requires fitness values when selecting features, to find the best solution candidates. In this process, the fitness value is the accuracy of the classification algorithm, which in the classification process uses the attributes represented by each particle. The representation of the attributes in each part of the dimension in each particle can be made from particles, where the representation is in the form of a binary string 0 and position 1, a value of 0 means that the value of the particle will not be used which is called an inactive value and vice versa for a value of 1 which will be used which is called an active value. Algorithm 1 outlines the steps of the PSO approach.

The DT stage uses[17], which is utilized in e-commerce fraud detection in both the learning stages for the DT classification technique and the DT stage. Algorithm 2 specifically states the steps of the DT method. The second classification uses the Naive Bayes method, which operates under the premise of probabilistically evaluating possibilities[18].

Algorithm 3 specifies the naive Bayes algorithm's phases. The SVM algorithm is the third classification algorithm. Finding a model with the best performance for training data is the main objective of SVM classification[19]. SVM classification can solve linear and non-linear problems using the kernel concept in high-dimensional space, and the best hyperplane will be found in this dimensional space to maximize the distance between classes[20].

The K-nearest neighbors (KNN) algorithm is classified as the fourth algorithm for classification. The concept underlying KNN classification is to assign a classification to a new query from an agency by considering the proximity to the majority category. The classification is defined by the category that occurs most frequently[21]. Algorithm 4 specifies the steps of the KNN.

The evaluation of the performance of the classification is the last step. A confusion matrix is used to describe how often the model is produced correctly, and it is based on the accuracy value [22]. Recall and precision are two additional performance indicators used in classification evaluation. Equation (1), equations for precision and recall, and equation (3) are used to calculate the accuracy value. A classification model can additionally include Receiver Operating Characteristics (ROC) [22][23] and Area Under the Curve (AUC) [24] in addition to the confusion matrix, depending on how well the predictions performed.

Algorithm 1 Particle Swarm Optimization[25]

1: Assume that there are N particles in the group or swarm.
2: To obtain X1, create a random initial population X with a range of X (B) and X (A). Then, particle j and its velocity in iteration I are designated as X(i) j and V(i) j, respectively, making the initial particles X1(0), X2(0),... XN (0) Assume that there are N particles in the group or swarm.
3: Find out each particle's speed. Set iteration i to be 1.
4: Find two crucial parameters for each particle in the ith iteration.
5: Calculate the velocity of particle j in iteration 1 of vi, where m=w. vi,m+c 1R(pbesti,m) +c 2R(gbestm).
6: Calculate the particle j's speed in iteration I of the formula vi,m=w. vi,m+c 1R(pbesti,mxi,m) +c 2R(gbestmxi,m).

Algorithm 2 Decision Tree [25]

1: Create attributes
2: First, calculate the entropy value, and then choose the properties

$$entropi = \sum_{i=1} -pi * log_2 pi$$

3: Gaining information by mending
4: Determine the knowledge obtained from the output.
5: Repetition of Step 2

Algorithm 3 Naïve Bayes [26]

1: To practice and compare your vocabulary, create a document.
2: The equation formula can be used to find probability values.

$$p(A|B) = \frac{P(B|A*P(A)}{P(B)}$$

3: Documents are classified so that the largest value of the class is visible.

Algorithm 4 KNN[27]

1: Determine the K value.
2: Utilize the Euclidean distance formula to determine the separation between the data.

$$Ed = \sqrt{(a1 - b1)^2 + ..(an - bn)^2}$$

3: The distance computation results are used to determine the K nearest neighbors.

RESULT AND DISCUSSION

Using the Python programming language and its built-in machine learning library, classification models using decision trees, Naive Bayes, SVM, and KNN are created [28].

Making a Decision Tree Model

Python was used to implement the DT classification model without PSO optimization, and the results were an accuracy score of 0.9774, a precision score of 1, and a recall score of 0.943. The confusion matrix DT without PSO is shown in Table 4, and the AUC and Roc values are shown in Figure 2.

Classification optimization using PSO in Python programming with a composition of 70% training and 30% testing data. The results of several trials carried out the best results for classification optimization with PSO using particles of 50, dimensions of 14, and a fitness value of 0.5 at 1300 iterations. This resulted in a reduction in features/variables from 15 to 4 features. Next, the PSO results are made into a classification model using decision trees, naïve Bayes, SVM, and KNN.The accuracy, precision, and recall values for DT classification with additional PSO optimization were 0.9777, 1, and 0,944, respectively. AUC and Roc values are shown in Figure 3, and Table 5 shows a confusion matrix with PSO.

Table 4. Confusion Matrix DT without PSO

Class	Predictive Positive	Predictive Negative
Actual Positive	1136	68
Actual Negative	0	1805

Table 5. Confusion Matrix DT with PSO

Class	Predictive Positive	Predictive Negative
Actual Positive	1135	67
Actual Negative	0	1807

ISSN 2089-8673 (Print) | ISSN 2548-4265 (Online)
Volume 13, Issue 1, March 2024

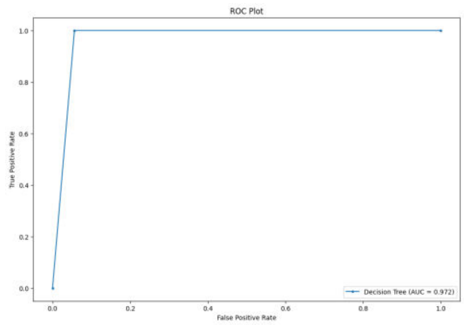

Figure 2. AUC Dan ROC DT Without PSO

An AUC of 0.972 (good classification) is generated by the data in Figure 2 of the ROC graph.

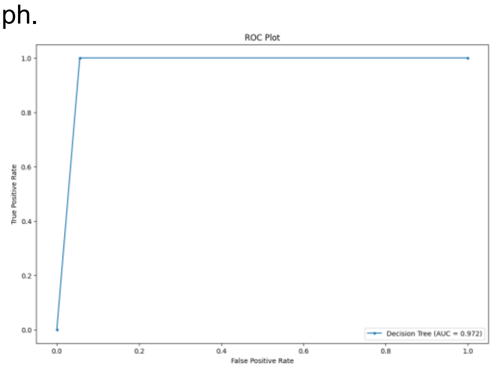

Figure 3. AUC Dan ROC DT With PSO

An AUC of 0.972 (excellent classification) is generated by the data in Figure 3 of the ROC graph.

Making a Naïve Bayes Classification Model

In this study, naive Bayes classification is performed using Python programming with 70% training data and 30% test data. Without PSO optimization, the Naive Bayes classification model yielded accuracy, precision, and recall values of 0.644, 0.548, and 0.623, respectively. The confusion matrix DT without PSO is shown in Table 6, and the AUC and Roc values are shown in Figure 4.

After feature selection, there were 9 features left out of the initial 15 features before PSO optimization. Accuracy was 0.693, precision was 0.960, and recall was 0.245 for naive Bayes classification with additional PSO optimization. AUC and Roc values are shown in Figure 5 and are part of the confusion matrix in Table 7 with PSO.

Table 6. Confusion Matrix Naïve Bayes without PSO

Class	Predictive Positive	Predictive Negative
Actual Positive	751	453
Actual Negative	618	1187

Table 7. Confusion Matrix Naïve Bayes with PSO

Class	Predictive Positive	Predictive Negative
Actual Positive	295	909
Actual Negative	12	1793

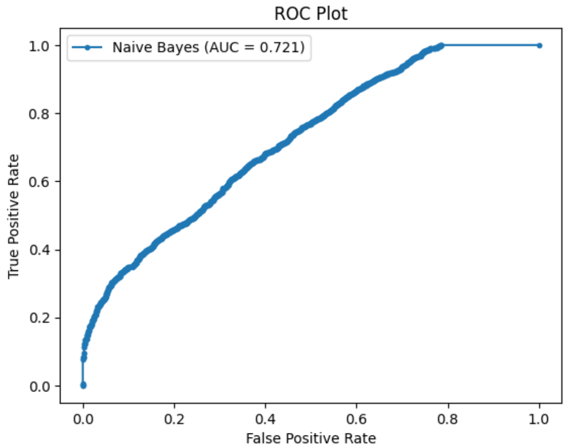

Figure 4. AUC Dan ROC Naïve Bayes Without PSO

An AUC of 0.721 (good classification) is generated from the data in Figure 4 of the ROC graph.

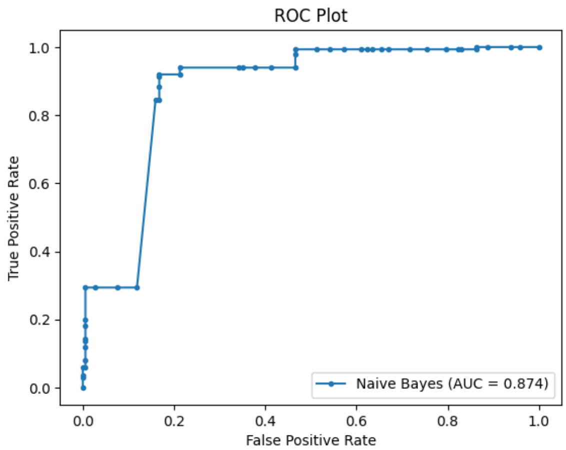

Figure 5. AUC Dan ROC Naïve Bayes With PSO

The ROC graph results in Figure 5 with an AUC of 0.874 (excellent classification).

Making an SVM classification mode

The Python programming language is used to implement the SVM classification in this study, which uses 70% training data and 30% testing data. The accuracy value was 0.599, the precision was 0, and the recall was 0, after using the SVM classification model without PSO optimization. Figure 6 shows the AUC and Roc values, while Table 8 contains the DT confusion matrix without PSO.

ISSN 2089-8673 (Print) | ISSN 2548-4265 (Online)
Volume 13, Issue 1, March 2024

After feature selection, there were only 8 features left out of the initial 15 features before PSO optimization. A score of 0.983 was obtained for accuracy, 0.960 for precision, and 0.959 for recall for SVM classification with PSO optimization. Figure 7 displays the AUC and Roc values, while Table 9 is a confusion matrix with PSO.

Table 8. Confusion Matrix SVM without PSO

Class	Predictive Positive	Predictive Negative
Actual Positive	0	1204
Actual Negative	0	1805

Table 9. Confusion Matrix SVM with PSO

Class	Predictive Positive	Predictive Negative
Actual Positive	1155	49
Actual Negative	0	1805

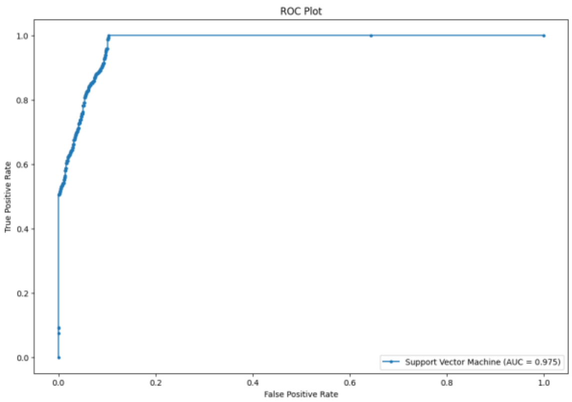

Figure 6. AUC Dan ROC SVM Without PSO

An AUC of 0.975 (excellent classification) is generated by the data in Figure 6 of the ROC graph. The results of the ROC graph in Figure 7 result in an AUC of 0.984 (excellent classification).

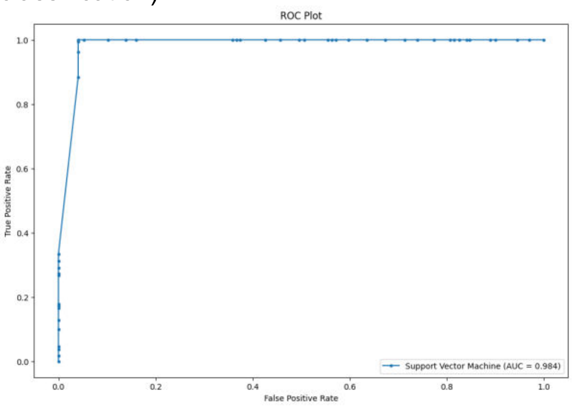

Figure 7. AUC Dan ROC SVM With PSO

Modeling KNN Classification

Using Python programming, the KNN classification in this study uses 70% training data and 30% testing data. Without PSO optimization, the KNN classification model produced accuracy values of 0.779, precision values of 1, and recall values of 0.522. The confusion matrix DT without PSO is shown in Table 10, and the AUC and Roc values are shown in Figure 8.

Prior to PSO optimization, there were 15 initial features; after feature selection, there were only 6. The accuracy, precision, and recall metrics for KNN classification with additional PSO optimization were 0.983, 0.960, and 0.959, respectively. Figure 9 shows the AUC and Roc values from Table 11's confusion matrix with PSO.

According to figure 8, the ROC curve from the outcomes of the KNN computation without PSO optimization yields an AUC of 0.795 (good classification).

Table 10. Confusion Matrix KNN without PSO

Class	Predictive Positive	Predictive Negative
Actual Positive	629	575
Actual Negative	89	1716

Table 11. Confusion Matrix KNN with PSO

Class	Predictive Positive	Predictive Negative
Actual Positive	1155	49
Actual Negative	0	1805

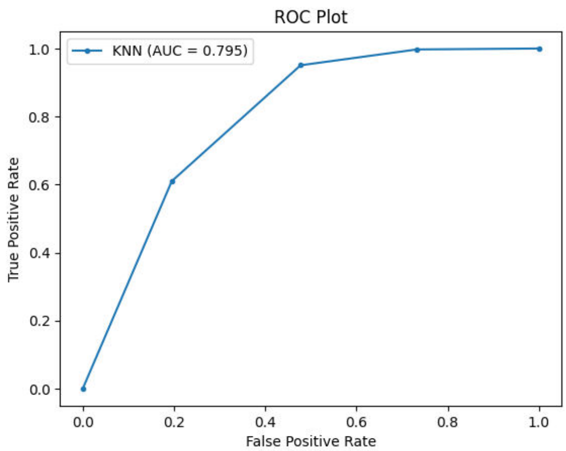

Figure 8. AUC Dan ROC KNN Without PSO

ISSN 2089-8673 (Print) | ISSN 2548-4265 (Online)
Volume 13, Issue 1, March 2024

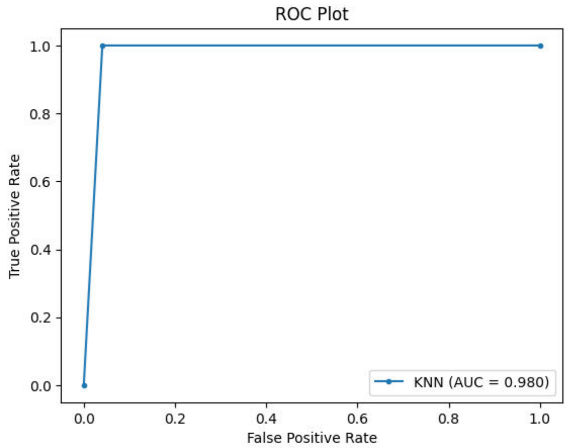

Figure 9. AUC Dan ROC KNN With PSO

The AUC of 0.980 (excellent classification) is produced by the ROC curve using the results of the KNN calculation with PSO optimization based on figure 9.

Models of Classification Algorithm Comparison

The graph in Figure 10 can be viewed as a comparison before feature optimization using PSO based on the output of the classification models built by DT, Naive Bayes, SVM, and KNN.

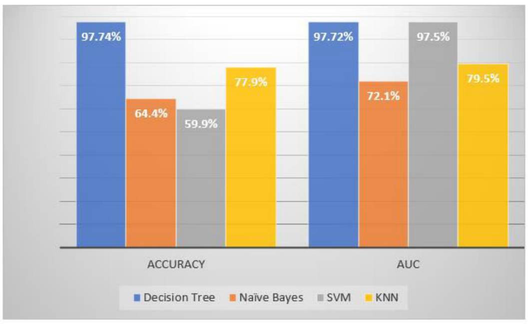

Figure 10. Comparison Chart for Classification Models

In comparison to the classification models created by Naive Bayes, SVM, and KNN, the Decision Tree algorithm in Figure 10's graph had the highest accuracy, with an accuracy value of 0.9774 and an AUC value of 0.972. The classification model is then optimized using PSO for each algorithm. Following optimization with PSO, Figure 11 shows a comparison graph of 4 (algorithms).

The SVM method in the classification model had an increase in accuracy following PSO of up to 38.4%, whereas DT saw the smallest increase in accuracy at 0.03%. The classification model for public service satisfaction with services offered by the village administration in Tambun Selatan sub-district is compared in the following, as shown in Table 11.

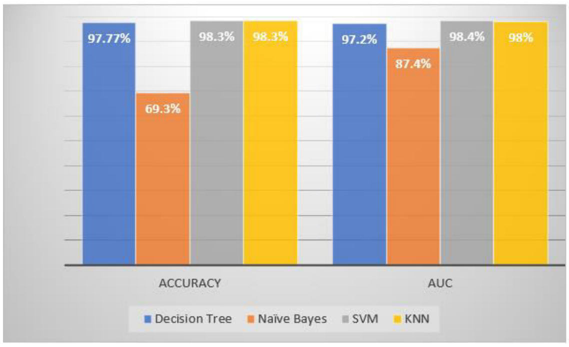

Figure 11. Comparative Models of Classification + PSO Chart

Table 11 demonstrates the customer service satisfaction categorization model's results, which reveal that the Decision Tree method had the best accuracy (97.74%) and the SVM approach had the lowest (59.90%). The SVM algorithm's performance in the classification model after optimization with PSO increased significantly, rising by 38.4% from the results of the prior classification model to 59.90% after optimization with PSO to 98.3%. Based on Table 9, the DT method was selected as the best algorithm for classifying public service happiness without PSO optimization, while the SVM algorithm was selected as the best algorithm for classifying public service satisfaction with PSO optimization.

Table 11. Models for Classifying Public Service Satisfaction Comparatively

Algorithm	No PSO	+PSO	Increase
Decision Tree	97.74%	97.77%	0.03%
Naïve Bayes	64.40%	69.30%	4.90%
Support Vector Machine	59.90%	98.3%	38.40%
K-Nearest Neighbor	77.90%	98.30%	20.40%

CONCLUSION

Before optimization, each algorithm in the classification model was evaluated. The decision tree algorithm achieved the highest accuracy of 97.74% and an AUC of 97.2% on the ROC graph. The KKN algorithm followed with an accuracy of 77.90% and an AUC of 79.5%. The Naïve Bayes algorithm had an accuracy of 64.4% and an AUC of 72.1%. The

SVM algorithm had the lowest accuracy of 59.90% and an AUC of 97.5%. By implementing Particle Swarm Optimization (PSO) for optimization, the SVM algorithm, which initially had the lowest accuracy value, experienced a significant increase of 38.4% in accuracy. As a result, the SVM algorithm achieved an accuracy value of 98.3%, and the ROC graph showed an Area Under the Curve (AUC) of 98.4%.

Similarly, the K-Nearest Neighbors (KNN) algorithm achieved an accuracy value of 98.3% with an AUC value of 98% after starting with an accuracy value of 20.40%. The DT algorithm achieved a 97.77% accuracy rate, with an AUC value of 97.2%. The lowest accuracy rate after optimization was observed with Naïve Bayes, which achieved a 69.30% accuracy rate with an AUC value of 87.4%. The utilization of Particle Swarm Optimization (PSO) for feature selection has effectively enhanced the accuracy of every model, hence enhancing the classification process for each model.

SVM, KNN, Naïve Bayes, and Decision Tree are optimized using PSO to identify the features or variables that have the greatest impact on public service satisfaction. The constructed model has been demonstrated to enhance the accuracy of each algorithm, enabling the classification model to assess the level of satisfaction with public services.

In order to get a greater level of precision, it is advisable to incorporate additional variables or characteristics into the data for future investigation. For advanced algorithms, it is optimal to integrate feature extraction, feature selection algorithms, and other classification approaches, along with feature selection algorithms.

REFERENCES

[1] D. Banga and K. Peddireddy, “Artificial Intelligence for Customer Complaint Management,” *Int. J. Comput. Trends Technol.*, vol. 71, no. 3, pp. 1–6, 2023, doi: 10.14445/22312803/ijctt-v71i3p101.

[2] K. Wabang, Oky Dwi Nurhayati, and Farikhin, “Application of The Naïve Bayes Classifier Algorithm to Classify Community Complaints,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 5, pp. 872–876, 2022, doi: 10.29207/resti.v6i5.4498.

[3] S. Khedkar and S. Shinde, “Deep Learning and Ensemble Approach for Praise or Complaint Classification,” *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 449–458, 2020, doi: 10.1016/j.procs.2020.03.254.

[4] M. Abdurohman, R. Husna, I. Ali, G. Dwilestari, and N. Rahaningsih, “Penerapan Model Klasifikasi Dalam Tingkat Kepuasan Layanan Publik Kelurahan Karyamulya Dengan Menggunakan Algoritma Decision Tree,” *Inf. Manag. Educ. Prof. J. Inf. Manag.*, vol. 6, no. 1, p. 81, 2022, doi: 10.51211/imbi.v6i1.1678.

[5] M. Alkaff, A. R. Baskara, and I. Maulani, “Klasifikasi Laporan Keluhan Pelayanan Publik Berdasarkan Instansi Menggunakan Metode LDA-SVM,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 6, pp. 1265–1276, 2021, doi: 10.25126/jtiik.2021863768.

[6] F. Caldeira, L. Nunes, and R. Ribeiro, “Classification of Public Administration Complaints,” *OpenAccess Series in Informatics*, vol. 104, no. 9. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany, pp. 9:1-9:0, 2022, doi: 10.4230/OASIcs.SLATE.2022.9.

[7] E. Purnamasari, D. Palupi Rini, and Sukemi, “Seleksi Fitur menggunakan Algoritma Particle Swarm Optimization pada Klasifikasi Kelulusan Mahasiswa dengan Metode Naive Bayes,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 3, pp. 469–475, 2020.

[8] S. A. Alsenan, I. M. Al-Turaiki, and A. M. Hafez, “Feature extraction methods in quantitative structure-activity relationship modeling: A comparative study,” *IEEE Access*, vol. 8, pp. 78737–78752, 2020, doi: 10.1109/ACCESS.2020.2990375.

[9] Y. R. Nugraha, A. P. Wibawa, and I. A. E. Zaeni, “Particle Swarm Optimization-Support Vector Machine (PSO-SVM) Algorithm for Journal Rank Classification,” *Proc. - 2019 2nd Int. Conf. Comput. Informatics Eng. Artif. Intell. Roles Ind. Revolut. 4.0, IC2IE 2019*, pp. 69–73, 2019, doi: 10.1109/IC2IE47452.2019.8940822.

[10] R. Hidayat, D. Kartini, M. I. Mazdadi, I. Budiman, and R. Ramadhani, “Implementasi Seleksi Fitur Binary Particle Swarm Optimization pada Algoritma K-NN untuk Klasifikasi Kanker Payudara,” *J. Sist. dan Teknol. Inf.*, vol. 11, no. 1, p. 135, 2023, doi: 10.26418/justin.v11i1.53608.

[11] A. Fauzi and A. H. Yunial, “Optimasi Algoritma Klasifikasi Naive Bayes, Decision Tree, K – Nearest Neighbor, dan Random Forest menggunakan Algoritma Particle Swarm Optimization

pada Diabetes Dataset," *J. Edukasi dan Penelit. Inform.*, vol. 8, no. 3, p. 470, 2022, doi: 10.26418/jp.v8i3.56656.

[12] R. Poreddy and E. S. Gopi, "Improvement of accuracy of under-performing classifier in decision making using discrete memoryless channel model and Particle Swarm Optimization," *Expert Syst. Appl.*, vol. 213, no. September 2022, 2023, doi: 10.1016/j.eswa.2022.118929.

[13] R. G. Santosa, Y. Lukito, and A. R. Chrismanto, "Classification and Prediction of Students' GPA Using K-Means Clustering Algorithm to Assist Student Admission Process," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 7, no. 1, p. 1, 2021, doi: 10.20473/jisebi.7.1.1-10.

[14] Y. Wanli Sitorus, P. Sukarno, S. Mandala, F. Informatika, and U. Telkom, "Analisis Deteksi Malware Android menggunakan metode Support Vector Machine & Random Forest," *e-Proceeding Eng.*, vol. 8, no. 6, p. 12500, 2021.

[15] L. Zhang, "A Feature Selection Algorithm Integrating Maximum Classification Information and Minimum Interaction Feature Dependency Information," *Hindawi Comput. Intell. Neurosci.*, vol. 2021, 2021.

[16] Y. Zhang, S. Wang, P. Phillips, and G. Ji, "Binary PSO with mutation operator for feature selection using decision tree applied to spam detection," *Knowledge-Based Syst.*, vol. 64, pp. 22–31, 2014, doi: 10.1016/j.knosys.2014.03.015.

[17] H. H. Patel and P. Prajapati, "Study and Analysis of Decision Tree Based Classification Algorithms," *Int. J. Comput. Sci. Eng.*, vol. 6, no. 10, pp. 74–78, 2018, doi: 10.26438/ijcse/v6i10.7478.

[18] A. Saputra and Suharjito, "Fraud detection using machine learning in e-commerce," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 9, pp. 332–339, 2019, doi: 10.14569/ijacsa.2019.0100943.

[19] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, no. xxxx, pp. 189–215, 2020, doi: 10.1016/j.neucom.2019.10.118.

[20] S. H. Hasanah, "Classification Support Vector Machine in Breast Cancer Patients," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 16, no. 1, pp. 129–136, 2022, doi: 10.30598/barekengvol16iss1pp129-136.

[21] Y. Kustiyahningsih, "Feature Selection and K-nearest Neighbor for Diagnosis Cow Disease," *Int. J. Sci. Eng. Inf. Technol.*, vol. 5, no. 02, pp. 249–253, 2021, doi: 10.21107/ijseit.v5i02.10218.

[22] R. C. Chen, C. Dewi, S. W. Huang, and R. E. Caraka, "Selecting critical features for data classification based on machine learning methods," *J. Big Data*, vol. 7, no. 1, 2020, doi: 10.1186/s40537-020-00327-4.

[23] T. R. Shultz and S. E. Fahlman, *Encyclopedia of Machine Learning and Data Mining*. 2017.

[24] P. Sedgwick, "How to read a receiver operating characteristic curve," *BMJ*, vol. 350, no. May, 2015, doi: 10.1136/bmj.h2464.

[25] T. Arifin and A. Herliana, "Optimizing decision tree using particle swarm optimization to identify eye diseases based on texture analysis," *J. Teknol. dan Sist. Komput.*, vol. 8, no. 1, pp. 59–63, 2020, doi: 10.14710/jtsiskom.8.1.2020.59-63.

[26] F. Hasibuan, W. Priatna, and T. Sri Lestari, "Analisis Sentimen Terhadap Kementrian Perdagangan Pada Sosial Media Twitter Menggunakan Metode Naïve Bayes Sentiment Analysis Of The Ministry Of Trade On Twitter Social Media Using Naïve Bayes Method," *Techno.COM*, vol. 21, no. 4, pp. 741–752, 2022.

[27] H. Said, N. H. Matondang, and H. N. Irmanda, "Penerapan Algoritma K-Nearest Neighbor Untuk Memprediksi Kualitas Air Yang Dapat Dikonsumsi," *Techno.Com*, vol. 21, no. 2, pp. 256–267, 2022, doi: 10.33633/tc.v21i2.5901.

[28] M. R. S. Alfarizi, M. Z. Al-farish, M. Taufiqurrahman, G. Ardiansah, and M. Elgar, "Penggunaan Python Sebagai Bahasa Pemrograman untuk Machine Learning dan Deep Learning," *Karya Ilm. Mhs. Bertauhid (KARIMAH TAUHID)*, vol. 2, no. 1, pp. 1–6, 2023.