GWO-SVM: AN APPROACH TO IMPROVING SVM PERFORMANCE USING GREY WOLF OPTIMIZER IN INTELLECTUAL DISABILITY CLASSIFICATION

Muhammad Afifudin[1*], Ahmad Junaedi[2], Andreas Nugroho[3], Izzatul Fithriyah[4]

[1,2,3]Informatika, Universitas Pembangunan Nasional Veteran Jawa Timur; Jalan Rungkut Madya No.1 Gunung Anyar, Surabaya, Indonesia
[4]Instalasi Kesehatan Jiwa RSUD Dr. Soetomo

Received: 17 September 2024
Accepted: 5 Oktober 2024
Published: 12 Oktober 2024

Keywords:
Intellectual Disability; Grey Wolf Optimizer; Support Vector Machine

Corespondent Email:
20081010187@student.upnjatim.ac.id

Abstrak. Retardasi mental (RM) merupakan gangguan neurodevelopmental yang membutuhkan diagnosis dini dan akurat. Penelitian ini bertujuan untuk meningkatkan efisiensi diagnosa RM menggunakan pendekatan machine learning. Model Support Vector Machine (SVM) yang dioptimasi dengan Grey Wolf Optimizer (GWO) dikembangkan dan dilatih menggunakan data dari kuesioner yang diisi oleh 101 keluarga/wali dari pasien RM di RSUD Dr. Soetomo Surabaya. Fitur-fitur yang digunakan meliputi riwayat keluarga, kemampuan kognitif, dan perilaku adaptif. Hasil penelitian menunjukkan bahwa model GWO-SVM mencapai akurasi 95% dalam mengklasifikasikan pasien RM, meningkat sekitar 5% dibandingkan dengan SVM konvensional. Algoritma GWO berhasil mengoptimalkan parameter pada SVM, sehingga diperoleh model dengan kinerja terbaik. Temuan ini mengindikasikan potensi GWO-SVM sebagai alat bantu diagnosa RM yang efektif dan efisien.

Abstract. Intellectual disability (ID) is a neurodevelopmental disorder that requires early and accurate diagnosis. This study aims to improve the efficiency of ID diagnosis using a machine learning approach. A Support Vector Machine (SVM) model optimized with Grey Wolf Optimizer (GWO) was developed and trained using data from questionnaires completed by 101 families/guardians of ID patients at RSUD Dr. Soetomo Surabaya. The features used include family history, cognitive abilities, and adaptive behaviors. The results showed that the GWO-SVM model achieved an accuracy of 95% in classifying ID patients, an improvement of 5% compared to the conventional SVM. The GWO algorithm successfully optimized the parameters in SVM, resulting in a model with the best performance. These findings indicate the potential of GWO-SVM as an effective and efficient tool for assisting in the diagnosis of ID.

1. INTRODUCTION

WHO reports that around 7-10% of children in the world experience disabilities, including Intellectual Disability. In Indonesia, the number of children with Intellectual Disability is quite significant. The American Association on Intellectual Disability or AAMR explains that Intellectual Disability is an individual's limitations in functioning intellectually which are known through intelligence and behavioral tests that refer to social and practical conceptual abilities [1]. Intellectual Disability can occur in various forms, including Intellectual Disability which is accompanied by mental or physical disorders, and Intellectual Disability which is not accompanied by mental or physical disorders [2]. During development, if a child has a developmental disorder characterized by Intelligence Quotiet (IQ) low levels, this is an early symptom of children with Intellectual Disability [3]. Intellectual Disability as a disorder of the central nervous system results in an individual's IQ being measured below 70, thus having an impact on the ability to fulfill basic needs for socialization skills, academic abilities, and communication skills [4].

Intellectual Disability (ID) lasts a lifetime and cannot be cured. However, special treatment and therapy are provided to develop the patient's ability to be able to carry out daily activities independently. Providing treatment or medication for people with Intellectual Disability is carried out based on each type the patient suffers. This means that the treatment for each type of ID sufferer is different. Therefore, accuracy in providing treatment is greatly influenced by the accuracy of the classification that has been carried out. To help parents find out ID disorders in children early.

Therefore, a system is needed to help health workers transform the conventional diagnostic process into a more efficient and accurate process. Current technological developments have given rise to a variety of classification methods. Classification itself is the process of grouping data or objects into separate classes that have previously been determined based on certain characteristics or variables. The classification process can be carried out by applying a classification algorithm. The classification algorithm will be used to learn patterns from the data that has been provided and produce a classification model. The classification model will later be used to classify new data that is not yet included in any class into the most appropriate class. This classification method has been widely applied in various fields, the simplest application of classification is E-mail spam classification. E-mail spam classification has the aim of grouping email messages into two classes, namely spam (unsolicited messages) and non-spam (actual messages). In the economic and financial fields, classification is applied in determining the feasibility of providing credit to individuals or Credit Scores. In the world of health, classification can be applied to patient medical record data such as symptoms, diagnosis, and treatment. This can help health workers understand the patient's condition better and more accurately, and make more appropriate treatment decisions.

Optimization machine learning aims to improve the performance of machine learning models. Optimization can improve model performance when evaluated using various matrices such as accuracy, precision, recall, and f1-score. Optimization can also increase the efficiency of the model being developed, such as shortening training time and reducing the memory requirements used. Various machine learning optimization methods have been developed, including the Gradient Descent, Grid Search, and Metaheuristic methods. One optimization method that applies metaheuristic principles is the Gray Wolf Optimizer method.

In previous research, Hilda Apriyani and Kurniati revealed that the SVM algorithm with a polynomial kernel had an accuracy of 96.2704%. This value is higher when compared to other methods used in this research, namely Naïve Bayes with an accuracy value of 92.0746% [5]. Thus, in their research, they revealed that SVM was superior to Naïve Bayes in classifying patients with diabetes mellitus. This level of accuracy is also supported by the data pre-processing stage which uses technical assistance K-Fold Cross Validation, SVM kernel selection, and evaluation methods Confusion Matrix.

Helena Nurramdhani Irmanda and Ria Astriratma in their research found that SVM can produce satisfactory scores in the classification of pantun types. SVM can properly classify rhyme types with an accuracy of 81,915% [6], This research uses a dataset of 470 rhymes which are divided into 3 classes, namely children's rhymes, youth rhymes, and elderly rhymes. Of the 3 classes available, children's rhymes are the class with the highest precision, recall, and specificity scores with scores of 90.63%, 87.88%, and 95.08%. This is influenced by the greater number of children's rhymes in the dataset compared to other classes. This research proves that SVM can be applied well in dealing with multiclass data.

Pallavi Sharma and Gurmanuk Kaur in their research prove that integrating GWO into SVM is an effective step in creating a machine learning model. Proven by the case of chronic kidney disease classification, the model can produce an accuracy value of 97.58% and an IDSE of 0.1581 after applying GWO in SVM parameter optimization [7]. The data used in this research is data from the UCI machine learning repository known as the Chronic_kidney_disease Dataset. This dataset has 24 variables with 1 target feature. This research proves that SVM optimized with GWO can produce satisfactory results for high-dimensional datasets.

M. Bahrul Subkhi et al in research on feature selection using the Hybrid Binary Gery Wolf Optimizer for Arabic Hadith Classification proved that, selecting with Gray Wolf Optimizers, classification using SVM got a score of 84% and was superior to KNN classification with a score of 76% [8]. This research uses 844 hadith data which is divided into 5 classes. This research uses GWO in the feature selection stage with a fitness function. This research shows that GWO can not only be used for parameter optimization but can also be used at the data pre-processing stage, namely the feature selection stage.

In malaria data classification research, Nur Ghaniaviyanto and Azka proved that the SVM method was superior with a value of 92.3% compared to Naïve Bayes as a comparison method. This result was obtained with data that had previously gone through the normalization stage in pre-processing and cross-validation in the pre-processing stage. [9] This research proves that normalization and cross-validation can improve the performance of the SVM model that has been developed. The research was conducted with a dataset originating from Nigeria, namely the malaria dataset.

Based on the research that has been mentioned, this research was conducted to develop a classification model to assist doctors in making diagnostic decisions for Intellectual Disability patients. The algorithm used in this research is Algorithm Support Vector Machine which is reinforced with Grey Wolf Optimizer. Later the model that has been developed will be evaluated using a confusion matrix It is hoped that the model developed can be considered helpful and integrated into the process of diagnosing patients with Intellectual Disability.

2. LITERATURE REVIEW

2.1. Machine Learning

Machine Learning or machine learning is a branch of computer science that aims to learn patterns from a set of data to improve performance in various processes [10]. The learning process to make machines intelligent is achieved in two stages, namely the training and testing stages [11]. The field of machine learning is closely related to the problem of how to develop a computer program so that it can learn and develop automatically based on previous experience. This makes machines not only able to behave in making decisions but also adapt to changes that occur.

Machine learning is generally divided into two parts, namely supervised learning and unsupervised learning. Supervised learning is a machine learning method used to make predictions based on data that has been labeled [12]. This method works by learning patterns and relationships in data and then using these patterns to predict new data [13]. In simple terms, unsupervised learning is a machine learning process that runs without the need for supervision [14]. The two main forms of supervised learning are classification and regression. Unsupervised learning is a machine learning method used to learn patterns and relationships in data without the help of labels. The data is not given categories or explicit instructions, so the algorithm must learn and find correlations between the available data on its own. One of the main applications of unsupervised learning is clustering. Clustering is the process of grouping several data into small groups (clusters) based on the degree of similarity and vice versa, respectively cluster is differentiated based on the dissimilarities between the data [15].

2.2. Classification

The process of finding a function that can distinguish or describe class differences between data with the aim that it can be used to predict the class of new data that has not been labelled [16]. Classification generally has various purposes, including fraud detection, image recognition, and document classification. The classification training process itself is carried out by giving the model labeled data. These labels will be used to indicate the correct group or class for the data. In the final stages of development, the model will be used to predict categories or classes for new data.

2.3. Support Vector Machine

Support Vector Machines or SVM is a powerful and robust regression and classification algorithm that has been widely applied in various technical and scientific fields [17]. Support Vector Machine (SVM) uses predictive methods to learn more optimal multivariate patterns in classifying data in groups [18]. The working principle of SVM is simply the best hyperplane to separate classes in data. This hyperplane will act as a supporting vector [19].

2.4. Grey Wolf Optimizer

The Grey Wolf Optimizer algorithm is an algorithm inspired by the hunting behavior of wolves in nature. The gray wolf is considered an apex predator, meaning that it is at the top of the food chain. Gray wolves also have a high social dominance hierarchy. The leaders at the first level will be referred to as alpha, the second level is beta, the third level is delta, and the last level is omega. In addition to the social hierarchy of wolves, hunting in packs is another interesting behavior of gray wolves.

2.5. Confussion Matrix

Evaluation is the process of measuring and assessing the quality of a machine learning model that has been developed to determine the extent to which the model is reliable, relevant, accurate and complete [20]. Confusion matrix is a table that states the classification of the number of correct test data and the number of incorrect test data [21]. It is a good idea to carry out an in-depth evaluation until the model developed is following the initial development objectives [22].

Confusion Matrix is a table with 4 (four) different combinations of predicted values and actual values. On performance measurement using Confusion Matrix, There are 4 (four) terms as representations of the results of the classification process. These four terms are True Positive (TP), True Negative (TN), False Positive (FP) dan False Negative (FN). Where Value True Negative (TN) is the number of negative data detected correctly, whereas False Positive (FP) is negative data but is detected as positive data. Meanwhile, True Positive (TP) is positive data that is detected correctly. False Negative (FN) is the opposite of True Positive, so that the data is positive, but is detected as negative data.

3. RESEARCH METHOD

This research will be carried out according to previously determined stages. In general, this research is divided into 5 stages, namely the initial preparation and analysis stage, the data collection stage, the implementation stage, and the testing scenario stage.

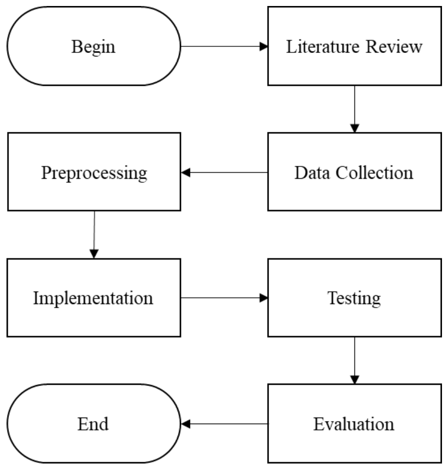


Figure 1 . Flowchart of research stages**.**

The initial preparation and analysis stage is divided into literature study and research design analysis. The data collection stage is where the researcher determines the population and sample of the research, prepares the data collection instruments, and collects the data. Data pre-processing is a stage where data that has been collected in the previous stage is processed, cleaned, and prepared so that it can be processed for the needs of developing machine learning models. After GWO and SVM are successfully implemented in program form, various test scenarios will be carried out to find the model with the best performance.

3.1. Literature Review

Literature studies are used to provide an understanding of the theoretical basis for research and identify current research which is considered to support the topic raised in this research, namely the application of the Gray Wolf Optimizer (GWO) and Support Vector Machine for the Classification of Mentally Retarded Patients. The literature used in study materials comes from various sources such as journals, books, papers, and websites.

3.2. Data Collection

The data used in this research is primary data collected from family/guardian questionnaire interviews from outpatients with a diagnosis of Intellectual Disability at the Poli Daycare, Instalasi Kesehatan Jiwa, RSUD Dr Soetomo, Surabaya, East Java. The data amounts to 101 data which covers family history, cognitive abilities, and adaptive behavior. Details of each item in the questionnaire can be seen in the following table.

Table 1. Research questionnaire question list

Item	Description
Kode	Unique Code for each respondent
Profile And Family History	
Age	The patient's age when the data was taken
Gender	Patient's gender
IQ	The patient's IQ value is based on an intelligence quotient test
Profile 1	Patient's birth condition (Premature/Normal)
Profile 2	History of relatives with similar conditions (Yes/No)
Profile 3	The method of delivery experienced by the patient (Normal/Normal with assistance/Caesarean)
Patient Behavior,	
All statements below will be answered on a Likert scale (Very Unsuitable/Not Appropriate/Agreeable/Very Appropriate)	
Q1	Children do not experience delays in physical growth
Q2	Children can communicate/speak well
Q3	Children can help with light household work
Q4	Children can dress without help
Q5	Children can take care of themselves in the bathroom (defecating, urinating, bathing)
Q6	Children can serve themselves at the dinner table
Q7	Children are able to travel independently and back to and from familiar places
Q8	Children can write simple words
Q9	Children are aware and able to avoid anything dangerous
Q10	Children can be trusted/understand the value of money
Q11	Children can understand the concept of time simply
ID	Intellectual Disability Severity Category (Mild, Moderate, Severe)

3.3. Preprocessing

3.3.1. Missing Value Analysis

Missing values are undesirable in a machine learning and data mining model because missing values will cause many problems [23]. Missing values can occur for various reasons, such as data collection errors, data entry errors, or incomplete observations. Handling missing values is an important aspect of data preprocessing and can significantly impact the reliability and accuracy of data analysis. missing value can be categorized into several types, namely: MCAR (missing at random), MNAR (missing not at random), and MAR (missing depending on other variables).

Several techniques are generally used in handling missing values, including deleting missing data, replacing missing data with certain values (imputation), or using several replacement values (multiple imputation). If the variable or case contains a missing value of less than 30% then the data can be eliminated and not included in the analysis process. The choice of appropriate technique depends largely on the type of missing data, the nature of the other data, and the objectives of the study.

3.3.2. Label Encoding

Label encoding is a data pre-processing step to convert categorical data or non-numerical data (usually represented as strings/text) into a numeric format. This method will involve assigning a unique numerical value to each category/text that appears in the categorical variable [24]. Machine learning algorithms generally perform well on numerical data. For this reason, this process is a stage that cannot be missed in developing a machine-learning model. However, it is important to remember that the Label Encoding process will create an ordinal relationship between data, for this reason, sufficient insight and understanding of the data to be processed is needed before Label Encoding is finally carried out.

3.3.3. Normalization

Normalization is a method for making several variables have the same range of values, not too large and not too small so that statistical analysis becomes easier to apply [25]. Normalization is a feature scaling method that is applied to a range of data to avoid overfitting in machine learning models developed with that data. Based on the findings from Aziz et al.'s research, differences in the value range for each attribute or variable in the transformation process will result in a reduction in the significance level of attributes that have a small value range compared to other attributes, so that models with data that have gone through the feature scaling process will have better results than models trained without feature scaling [26].

3.4. Grey Wolf Optimizer

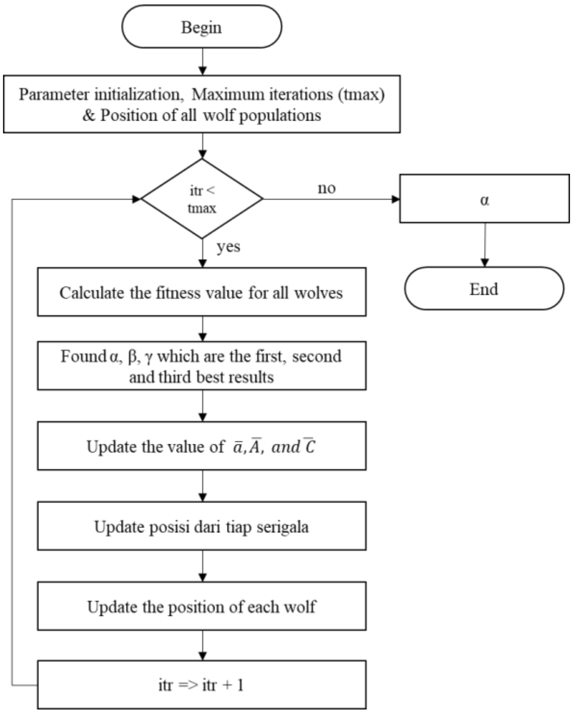


Figure 2. Flowchart of GWO

The process of finding the optimal solution is carried out by updating the position of each wolf based on the positions α, β, and γ. This position update is influenced by random vectors A and C that control the balance between exploration and exploitation. Wolf position updates are done using the following formula:

$$\overrightarrow{D_\alpha} = \left|\overrightarrow{C_1} \cdot \overrightarrow{X_\alpha} - \vec{X}\right| \quad (1)$$

$$\overrightarrow{D_\beta} = \left|\overrightarrow{C_2} \cdot \overrightarrow{X_\beta} - \vec{X}\right| \quad (2)$$

$$\overrightarrow{D_\gamma} = \left|\overrightarrow{C_3} \cdot \overrightarrow{X_\gamma} - \vec{X}\right| \quad (3)$$

D is the distance between the wolf and the prey. Once the distance is found, the wolf's position will be updated with the following formula:

$$\overrightarrow{X_1} = \overrightarrow{X_\alpha} - \overrightarrow{A_1} \cdot \overrightarrow{D_\alpha} \quad (4)$$

$$\overrightarrow{X_2} = \overrightarrow{X_\beta} - \overrightarrow{A_2} \cdot \overrightarrow{D_\beta} \quad (5)$$

$$\overrightarrow{X_3} = \overrightarrow{X_\gamma} - \overrightarrow{A_3} \cdot \overrightarrow{D_\gamma} \quad (6)$$

Wolf positions are updated as the average of the α, β, and γ contributions.

$$\vec{X}(t+1) = \frac{\overrightarrow{X_1}+\overrightarrow{X_2}+\overrightarrow{X_3}}{3} \quad (7)$$

At each iteration, the best fitness value will be updated if a wolf finds a better solution. At the end of the iteration, the wolf position α is considered the optimal solution found by the GWO algorithm. GWO works by simulating the behavior of a wolf pack in searching for prey. The wolves will learn from each other and work together to find the best solution.

3.5. GWO-SVM

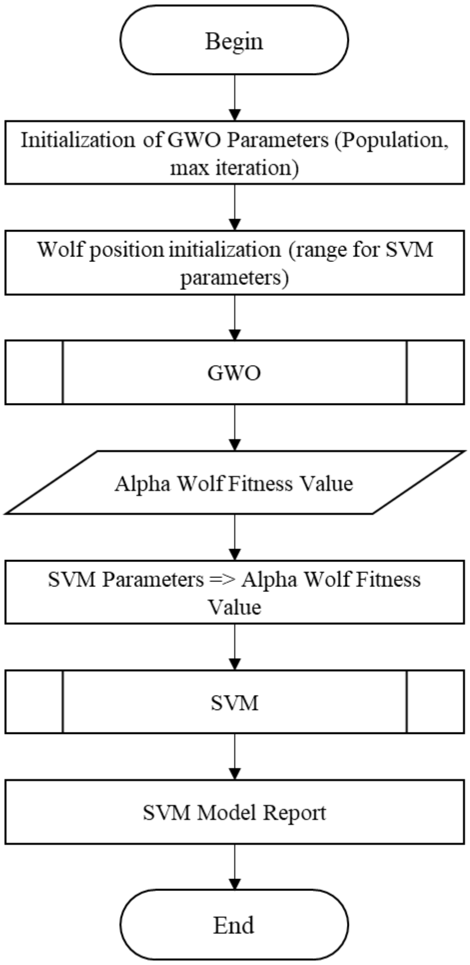


Figure 3. Flowchart of GWO-SVM method

The Gray Wolf Optimizer (GWO) algorithm can be used to optimize parameters in the Support Vector Machine (SVM) model. GWO simulates the behavior of a wolf pack in searching for prey. In the context of SVM parameter optimization, these wolves represent various combinations of parameter values. GWO will look for the best combination of parameters that produces the most optimal SVM performance. In other words, GWO helps SVM achieve higher accuracy by finding the most appropriate parameter values for the given data. This process involves randomly initializing the wolf population, evaluating the performance of each wolf, and iteratively updating the wolf positions until an optimal solution is found. GWO acts as an intelligent search tool to find the optimal combination of SVM parameters. This allows SVM to achieve the best performance in classification or regression.

4. RESULT AND DISCUSSION

4.1. Data Analysis

Before entering data pre-processing, descriptive analysis will be carried out from the dataset that has been obtained. Descriptive analysis is used to summarize and describe the dataset in a form that is easier to understand. Descriptive analysis can help understand data in a clearer and easier-to-digest way and become more meaningful information.

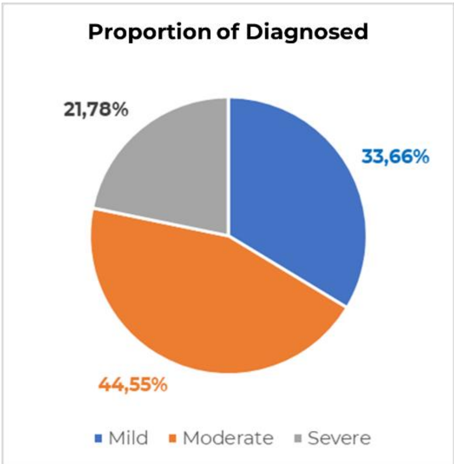


Figure 4. Pie chart of Patient Proportion

The diagnosis of moderate Intellectual Disability was the one that was most often found when collecting data. However, this will not be a problem at the model development stage, because the differences in the number of labels are not too extreme. It can be seen that the other two labels that will be used in model training each have a percentage of 33.66% for Mild and 21.78% for Severe. With this information, oversampling methods to balance labels are not necessary.

Patients aged 7 years are the largest age demographic. Followed by ages 8 and 5 years. This is thought to occur because this age is the age at which children enter school age, both at the Early Childhood Education and Elementary School levels. With the start of the education period, parents/guardians can immediately identify children who they feel have different behavior from their peers, making this age the age at which Intellectual Disability or other psychological conditions occur.

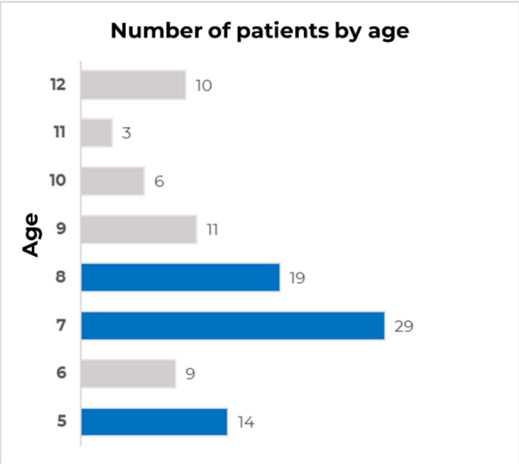


Figure 5. Count of patient based on age

It was found that the IQ of the patients had a mean and median value of 45 and 46 respectively. These values are quite close and it can be interpreted that the data is normally distributed and perhaps close to symmetrical. A range of 50 indicates that the data is spread over a fairly wide range, the lowest IQ is 20 and the highest is 70. A standard deviation value of 13.4 indicates that the data has quite large variations around the average value.

Table 2. Descriptive Statistic of the Dataset

N	101
Mean	45
Median	46
Std. Deviation	13,4
Range	50
Min value	20
Max value	70

4.2. SVM Classification

As a basis for measurement. A conventional SVM classification model will be developed. The performance of the model produced at this stage will be compared with the model that applies GWO-SVM. The model will be developed with different data-sharing schemes, namely 60:40, 70:30, and 80:30.

4.2.1. Data split 60:40

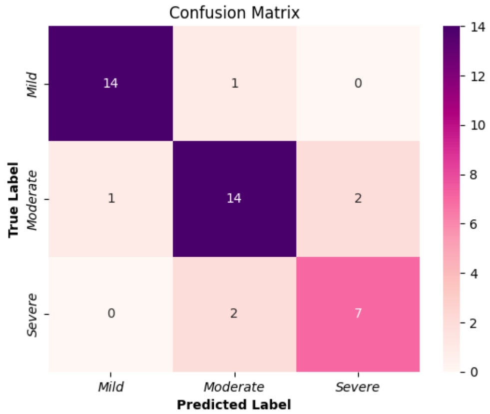


Figure 6. SVM 60:40 Confussion Matrix

Evaluation of the confusion matrix from this test scenario shows that the model developed predicted 35 data correctly out of 41 existing test data. There are 14 predictions for the Mild Category (Mild), 14 data for the Moderate Intellectual Disability category, and 7 data for the Severe Intellectual Disability category. The classification report of this model can be seen in the following table

Table 3. 60:40 SVM model report

Category	Precision	Recall	F1-Score
Mild	0.93	0.93	0.93
Moderate	0.82	0.82	0.82
Severe	0.78	0.78	0.78

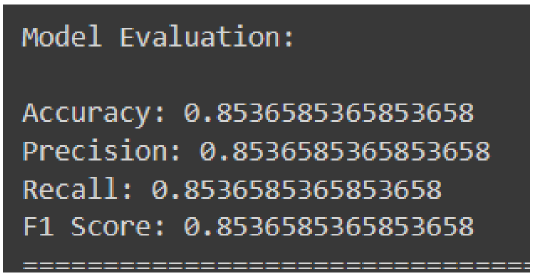


Figure 7. SVM 60:40 Model Evaluation

4.2.2. Data split 70:30

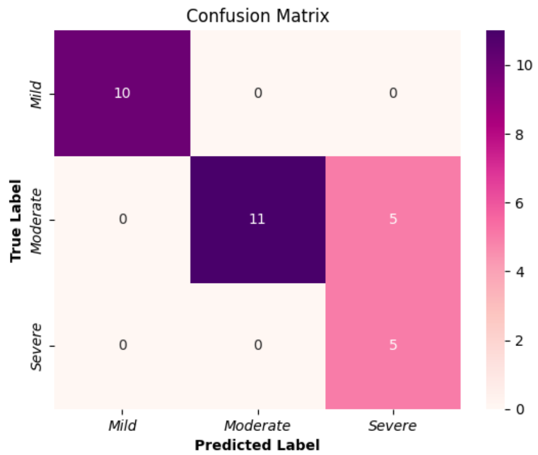
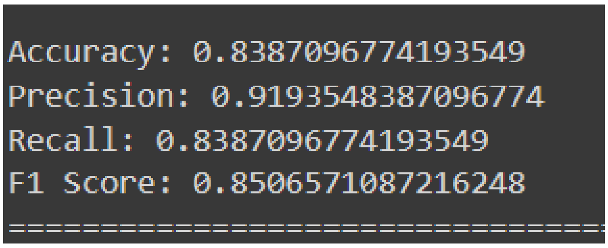


Figure 8. 70:30 SVM Confussion Matrix

Evaluation of the confusion matrix from this test scenario shows that the model developed predicted 16 data correctly out of 21 existing test data. There are 10 predictions for the Mild Category (Mild), 11 data for the Moderate Intellectual Disability category, and 5 data for the Severe Intellectual Disability category. The Classification Report of this model can be seen in the following table.

Table 4. 70:30 SVM Model Report

Category	Precision	Recall	F1-Score
Mild	0.93	0.93	0.93
Moderate	0.82	0.82	0.82
Severe	0.78	0.78	0.78



```
Accuracy: 0.8387096774193549
Precision: 0.9193548387096774
Recall: 0.8387096774193549
F1 Score: 0.8506571087216248
==================================
```

Figure 9. 70:30 SVM Model Evaluation

4.2.3. Data Split 80:20

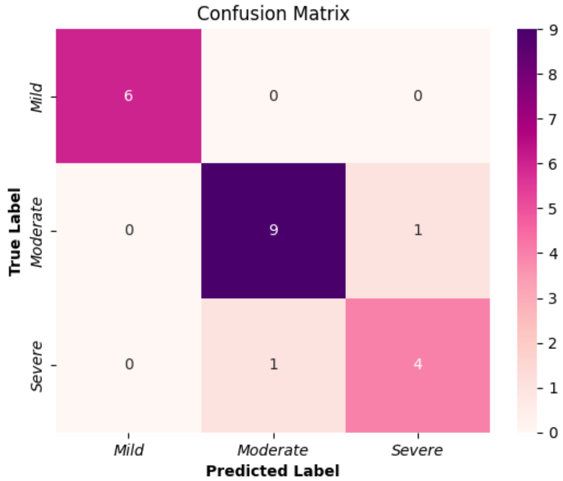
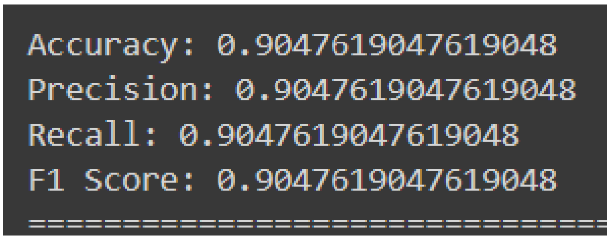


Figure 10. 80:30 SVM Confussion Matrix

Evaluation of the confusion matrix from this test scenario shows that the model developed predicted 19 data correctly out of 21 existing test data. There are 6 predictions for the Mild Category (Mild), 9 data for the Moderate Intellectual Disability category, and 4 data for the Severe Intellectual Disability category.

Table 5. 80:20 SVM Model Report

Category	Precision	Recall	F1-Score
Mild	1	1	1
Moderate	0.90	0.90	0.90
Severe	0.80	0.80	0.80



```
Accuracy: 0.9047619047619048
Precision: 0.9047619047619048
Recall: 0.9047619047619048
F1 Score: 0.9047619047619048
==================================
```

Figure 11. 80:30 Model Evaluation

4.3. GWO-SVM Classifier

The SVM parameters that will be optimized in this research are a combination of parameter C, learning rate, and maximum iteration. The parameters used by GWO in conducting exploration are a wolf population of 10 and a maximum search iteration of 40 times.

4.3.1. Data split 60:40

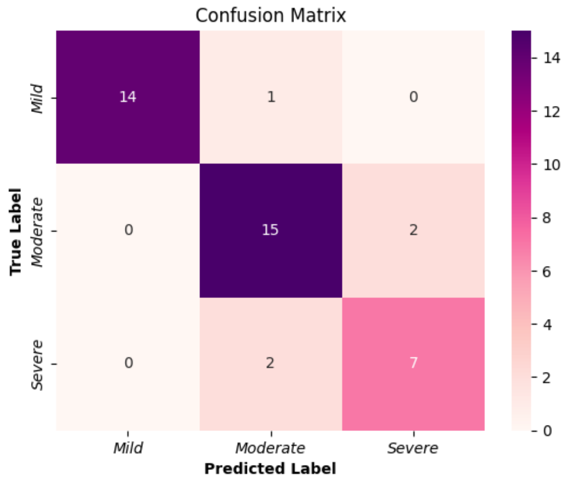
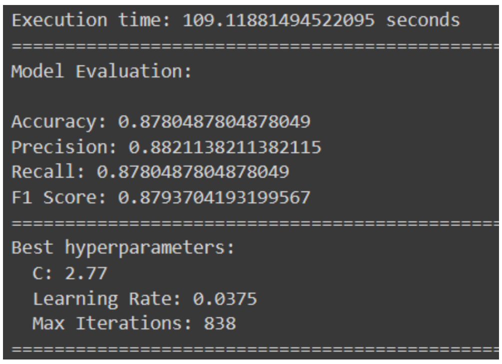


Figure 12. 60:40 GWO-SVM Confusion Matrix

Evaluation confusion matrix This test scenario shows that the model developed predicted 36 data correctly out of 41 existing test data. Predictions for the Light Category (Mild) a total of 14 data, the category Moderate Intellectual Disability (Moderate) as many as 15 data, and the Severe Intellectual Disability category (Severe) as many as 7 data. Here are the accuracy scores, precision scores, scores recall, and F1- scores from the model.



```
Execution time: 109.11881494522095 seconds
==========================================
Model Evaluation:

Accuracy: 0.8780487804878049
Precision: 0.8821138211382115
Recall: 0.8780487804878049
F1 Score: 0.8793704193199567
==========================================
Best hyperparameters:
  C: 2.77
  Learning Rate: 0.0375
  Max Iterations: 838
==========================================
```

Figure 13. 60:40 GWO-SVM Model Evaluation

Apart from the evaluation results of the trained model, the image above also presents the best parameter values produced by GWO for the classification of the Intellectual Disability category. GWO finds that the combination of C = 2.77, Learning Rate = 0.0375, and Max Iteration = 838 is the most optimal parameter in this scenario. Also presented is the execution time of this scenario, namely 109.12 seconds. Meanwhile, the accuracy score, precision score, score recall, and F1- score of each category can be seen in the figure and table classification report following.

Table 6. 60:40 GWO-SVM model report

Category	Precision	Recall	F1-Score
Mild	1	0.93	0.97
Moderate	0.83	0.88	0.86
Severe	0.78	0.78	0.78

4.3.2. Data split 70:30

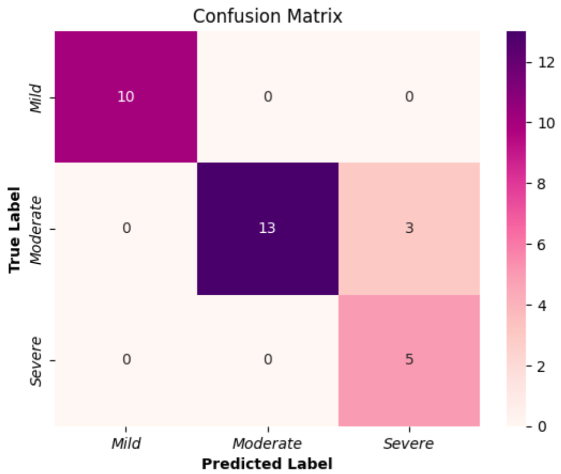


Figure 14. 70:30 GWO-SVM Confusion Matrix

Evaluation of the confusion matrix from this test scenario shows that the model developed predicted 28 data correctly out of 31 existing test data. There are 10 predictions for the Mild Category (Mild), 13 data for the Moderate Intellectual Disability category, and 5 data for the Severe Intellectual Disability category. The following are the accuracy scores, precision scores, recall scores, and F1 scores of the model.

```
Execution time: 122.48457646369934 seconds
==========================================
Model Evaluation:

Accuracy: 0.9032258064516129
Precision: 0.9395161290322581
Recall: 0.9032258064516129
F1 Score: 0.9093864978180884
==========================================
Best hyperparameters:
  C: 5.11
  Learning Rate: 0.0662
  Max Iterations: 379
==========================================
```

Figure 15. 70:30 GWO-SVM Model Evaluation

Apart from the evaluation results of the trained model, the image above also presents the best parameter values produced by GWO for the classification of the Intellectual Disability category. GWO finds that the combination of C = 5.11, Learning Rate = 0.0622, and Max Iteration = 379 is the most optimal parameter in this scenario. Also presented is the execution time of this scenario, namely 122.48 seconds. Meanwhile, the accuracy scores, precision scores, recall scores, and F1-scores for each category can be seen in the following classification report image and table.

Table 7. 70:30 GWO-SVM model report

Category	Precision	Recall	F1-Score
Mild	1	1	1
Moderate	1	0.81	0.90
Severe	0.62	1	0.77

4.3.3. Data split 80:20

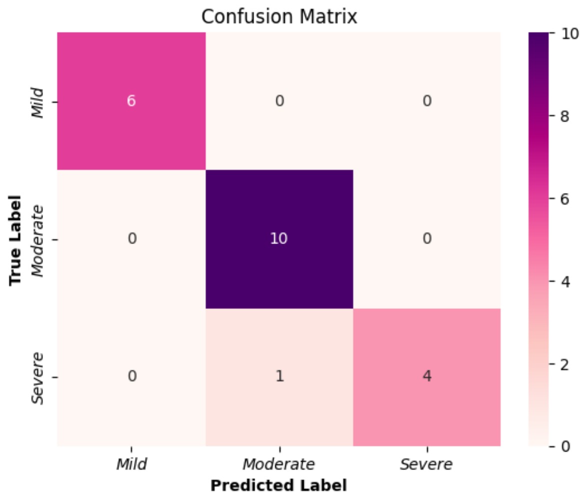


Figure 16. 80:20 GWO-SVM confussion matrix

Evaluation of the confusion matrix from this test scenario shows that the model developed predicted 20 data correctly from 21 existing test data. There are 6 predictions for the Mild Category (Mild), 10 data for the Moderate Intellectual Disability category, and 4 data for the Severe Intellectual Disability category. The following are the accuracy scores, precision scores, recall scores, and F1 scores of the model.

```
Execution time: 67.42314553260803 seconds
==========================================
Model Evaluation:

Accuracy: 0.9523809523809523
Precision: 0.9567099567099566
Recall: 0.9523809523809523
F1 Score: 0.9508692365835223
==========================================
Best hyperparameters:
  C: 2.40
  Learning Rate: 0.0560
  Max Iterations: 144
==========================================
```

Figure 17. 80:20 GWO-SVM Model Evaluation

Apart from the evaluation results of the trained model, the image above also presents the best parameter values produced by GWO for the classification of the Intellectual Disability category. GWO finds that the combination of C = 2.40, Learning Rate = 0.0560, and Max Iteration = 144 is the most optimal parameter in this scenario. Also presented is the execution time of this scenario, namely 67.42 seconds. Meanwhile, the accuracy scores, precision scores, recall scores, and F1-scores for each category can be seen in the following classification report image and table.

Table 8. 80:20 GWO-SVM model report

Category	Precision	Recall	F1-Score
Mild	1	1	1
Moderate	0.91	1	0.95
Severe	1	0.80	0.89

4.4. Performance Comparison

Table 9. Model Comparison

Data Split	SVM	GWO-SVM
60:40	85.36%	87.80%
70:30	83.87%	90.32%
80:20	90.47%	95.23%

The table above is a summary of the evaluation results for each test scenario carried out. The table above was compiled to simplify the process of searching and comparing the best combination of applying the Gray Wolf Optimizer and Support Vector Machine for Classifying Mentally Retarded Patients.

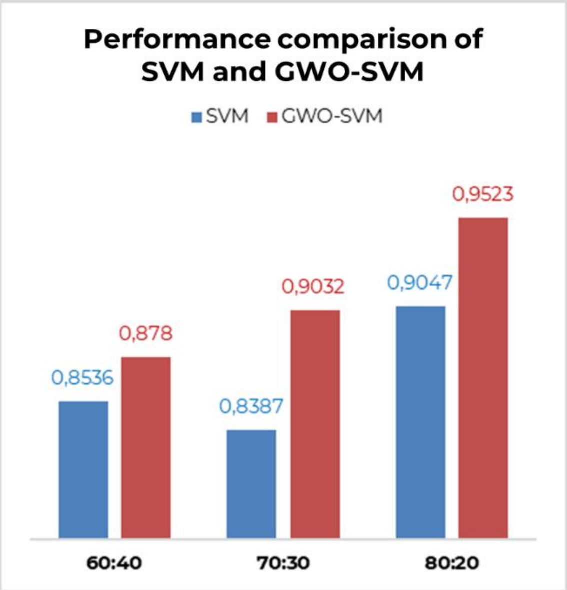


Figure 18. SVM and GWO-SVM performance Comparison

The image above is a bar diagram representation that shows a comparison of the accuracy of the SVM model and the average accuracy of the GWO-SVM model. It can be seen that GWO is proven to be able to increase the accuracy of the SVM model in all data-splitting schemes with varying increases in accuracy. The 60:40 data sharing scheme increased by 0.0244 or 2.4%, the 70:30 data sharing scheme increased by 0.0645 or 6.4% and the 80:20 data sharing scheme increased by 0.0476 or 4.7%.

5. CONCLUSION

Based on the results obtained from the completion of all research stages carried out by researchers, several conclusions can be drawn as follows:

a. The GWO method was proven to be able to improve the performance of the SVM model in terms of accuracy in the case of classifying patient Intellectual Disability categories. Determining the population value and iteration as well as the appropriate data-splitting scheme will produce good performance.
b. Based on various test scenarios used in this research, the GWO-SVM classification model was obtained with the best performance of 95.23% for accuracy values. This performance was produced using an 80:20 data sharing scheme, with a GWO population value of 10 and a Maximum iteration value of 40 times with a search time of 67.42 seconds. The best performance value obtained is an increase of 4.7% from the performance of the SVM model with an accuracy of 90.47%.
c. The best parameters produced by GWO in the GWO-SVM model with the best performance are a combination of the values C = 2.40, Learning Rate = 0.0560, and Max Iteration = 144.

ACKNOWLEDGMENT

The researcher would like to express deep gratitude to God Almighty for all His blessings and guidance that have been given so that the author succeeded in completing this research well. The author would also like to express his gratitude to his beloved parents for their very significant moral and material support, as well as to all parties who participated in the smooth running of this research.

REFERENCES

[1] W. M. Sholihah, "Strategi Mengajarkan Keterampilan Hidup Sehari-Hari Untuk Anak dan Remaja Dengan Retardasi Mental: Sebuah Review Literatur:," PINUS: Jurnal Penelitian Inovasi Pembelajaran, vol. 6, no. 2, pp. 50–61, Jul. 2021, doi: 10.29407/PN.V6I2.13840.

[2] M. Bagus Andrianto, A. Sartika, D. Septeah Ningrum, U. Muhammadiyah Bengkulu, and P. Pematang Tiga, "Pengalaman Orangtua dalam Merawat Anak Retardasi Mental," Jurnal Kesmas Asclepius, vol. 3, no. 1, pp. 9–16, Jun. 2021, doi: 10.31539/JKA.V3I1.2163.

[3] P. Panzilion, P. Padila, A. D. Setyawati, H. J, and A. Sartika, "Stimulation of Preschool Motor Development Through Brain Gym and Puzzle," JOSING: Journal of Nursing and Health, vol. 1, no. 1, pp. 10–17, Aug. 2020, doi: 10.31539/JOSING.V1I1.1166.

[4] R. Ruliati, "DUKUNGAN ORANG TUA TERHADAP FUNGSI SOSIAL PADA ANAK RETARDASI MENTAL," Jurnal Kesehatan Samodra Ilmu (JKSI), vol. 11, no. 2, pp. 200–208, Aug. 2020, doi: 10.55426/JKSI.V11I2.123.

[5] H. Apriyani and K. Kurniati, "Perbandingan Metode Naïve Bayes Dan Support Vector Machine Dalam Klasifikasi Penyakit Diabetes Melitus," Journal of Information Technology Ampera, vol. 1, no. 3, pp. 133–143, Dec. 2020, doi: 10.51519/JOURNALITA.VOLUME1.ISSSUE3.YEAR2020.PAGE133-143.

[6] H. N. Irmanda and R. Astriratma, "Klasifikasi Jenis Pantun Dengan Metode Support Vector Machines (SVM)," Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), vol. 4, no. 5, pp. 915–922, Oct. 2020, doi: 10.29207/RESTI.V4I5.2313.

[7] P. Sharma and G. Kaur, "Optimization of Support Vector Machine Classifier Using Grey Wolf Optimization Algorithm for Chronic Kidney Disease Prediction," Asian Pacific Journal of Health Sciences, vol. 9, no. 3, pp. 227–231, Apr. 2022, doi: 10.21276/APJHS.2022.9.3.46.

[8] M. B. Subkhi, C. Fatichah, and A. Z. Arifin, "Seleksi Fitur Menggunakan Hybrid Binary Grey Wolf Optimizer untuk Klasifikasi Hadist Teks Arab," Jurnal Teknologi Informasi dan Ilmu Komputer, vol. 10, no. 5, pp. 1115–1122, Oct. 2023, doi: 10.25126/JTIIK.2023106375.

[9] N. G. Ramadhan and A. Khoirunnisa, "Klasifikasi Data Malaria Menggunakan Metode Support Vector Machine," JURNAL MEDIA INFORMATIKA BUDIDARMA, vol. 5, no. 4, pp. 1580–1584, Oct. 2021, doi: 10.30865/MIB.V5I4.3347.

[10] T. Jiang, J. L. Gradus, and A. J. Rosellini, "Supervised Machine Learning: A Brief Primer," Behav Ther, vol. 51, no. 5, pp. 675–687, Sep. 2020, doi: 10.1016/J.BETH.2020.05.002.

[11] A. Roihan, P. Abas Sunarya, and A. S. Rafika, "Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper," IJCIT (Indonesian Journal on Computer and Information Technology), vol. 5, no. 1, pp. 75–82, May 2020, doi: 10.31294/IJCIT.V5I1.7951.

[12] D. Darwis, N. Siskawati, and Z. Abidin, "PENERAPAN ALGORITMA NAIVE BAYES UNTUK ANALISIS SENTIMEN REVIEW DATA TWITTER BMKG NASIONAL," Jurnal Tekno Kompak, vol. 15, no. 1, pp. 131–145, Feb. 2021, doi: 10.33365/JTK.V15I1.744.

[13] S. Suryani and M. Mustakim, "Estimasi Keberhasilan Siswa dalam Pemodelan Data Berbasis Learning Menggunakan Algoritma Support Vector Machine," Bulletin of Informatics and Data Science, vol. 1, no. 2, pp. 81–88, Nov. 2022, doi: 10.61944/BIDS.V1I2.36.

[14] A. Darmawan Sidik, A. Ansawarman, K. Kunci, J. Kendaraan Bermotor, M. Regresi, and F. Jalan, "Prediksi Jumlah Kendaraan Bermotor Menggunakan Machine Learning," Formosa Journal of Multidisciplinary Research, vol. 1, no. 3, pp. 559–568, Jul. 2022, doi: 10.55927/FJMR.V1I3.745.

[15] T. M. M. Tyas and A. I. Purnamasari, "Penerapan Algoritma K-means dalam Mengelompokkan Demam Berdarah Dengue Berdasarkan Kabupaten," Blend Sains Jurnal Teknik, vol. 1, no. 4, pp. 277–283, Mar. 2023, doi: 10.56211/BLENDSAINS.V1I4.231.

[16] S. Febriani and H. Sulistiani, "Analisis Data Hasil Diagnoga Untuk Klasifikasi Gangguan Kepribadian Menggunakan Algoritma C4.5," Jurnal Teknologi dan Sistem Informasi, vol. 2, no. 4, pp. 89–95, Dec. 2021, doi: 10.33365/JTSI.V2I4.1373.

[17] 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, pp. 189–215, Sep. 2020, doi: 10.1016/J.NEUCOM.2019.10.118.

[18] C. H. Miller, M. D. Sacchet, and I. H. Gotlib, “Support Vector Machines and Affective Science,” Emotion Review, vol. 12, no. 4, pp. 297–308, Jun. 2020, doi: 10.1177/1754073920930784.

[19] R. Tineges, A. Triayudi, and I. D. Sholihati, “Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM),” JURNAL MEDIA INFORMATIKA BUDIDARMA, vol. 4, no. 3, pp. 650–658, Jul. 2020, doi: 10.30865/MIB.V4I3.2181.

[20] L. Firdaus and T. Setiadi, “Perbandingan Algoritma Naive Bayes, Decision Tree, dan KNN untuk Klasifikasi Produk Populer Adidas US dengan Confusion Matrix,” Jurnal Sistem Komputer dan Informatika (JSON), vol. 5, no. 2, pp. 185–195, Dec. 2023, doi: 10.30865/JSON.V5I2.6124.

[21] D. Normawati and S. A. Prayogi, “Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter,” J-SAKTI (Jurnal Sains Komputer dan Informatika), vol. 5, no. 2, pp. 697–711, Sep. 2021, doi: 10.30645/J-SAKTI.V5I2.369.

[22] A. Khumaidi, “DATA MINING FOR PREDICTING THE AMOUNT OF COFFEE PRODUCTION USING CRISP-DM METHOD,” Jurnal Techno Nusa Mandiri, vol. 17, no. 1, pp. 1–8, Feb. 2020, doi: 10.33480/TECHNO.V17I1.1240.

[23] W. Sudrajat and I. Cholid, “K-NEAREST NEIGHBOR (K-NN) UNTUK PENANGANAN MISSING VALUE PADA DATA UMKM,” Jurnal Rekayasa Sistem Informasi dan Teknologi, vol. 1, no. 2, pp. 54–63, Nov. 2023, doi: 10.59407/JRSIT.V1I2.77.

[24] C. Herdian, A. Kamila, and I. G. A. M. Budidarma, “Studi Kasus Feature Engineering Untuk Data Teks: Perbandingan Label Encoding dan One-Hot Encoding Pada Metode Linear Regresi,” Technologia : Jurnal Ilmiah, vol. 15, no. 1, pp. 93–108, Jan. 2024, doi: 10.31602/tji.v15i1.13457.

[25] M. R. Kusnaidi, T. Gulo, and S. Aripin, “Penerapan Normalisasi Data Dalam Mengelompokkan Data Mahasiswa Dengan Menggunakan Metode K-Means Untuk Menentukan Prioritas Bantuan Uang Kuliah Tunggal,” Journal of Computer System and Informatics (JoSYC), vol. 3, no. 4, pp. 330–338, Sep. 2022, doi: 10.47065/JOSYC.V3I4.2112.

[26] A. R. Aziz, B. Warsito, and A. Prahutama, “PENGARUH TRANSFORMASI DATA PADA METODE LEARNING VECTOR QUANTIZATION TERHADAP AKURASI KLASIFIKASI DIAGNOSIS PENYAKIT JANTUNG,” Jurnal Gaussian, vol. 10, no. 1, pp. 21–30, Feb. 2021, doi: 10.14710/J.GAUSS.10.1.21-30.