JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9435
Hal 60-70
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Optimized Fault Prediction in Power Distribution Transformers Using Grey Wolf Optimizer-Based SVM and MLP Models

Rosena Shintabella[1*], Meyer Mega Eklesia Silaban[2], Muhammad Ichsan Basir[3]

[1] Department of Informatics Engineering, D4 Software Engineering, Polytechnic State of Indramayu, Indramayu, Indonesia
[2] Department of Informatics Engineering, D3 Informatics Engineering, Polytechnic State of Indramayu, Indramayu, Indonesia
[3] Department of Health, D4 Medical Laboratory Technology, Polytechnic State of Indramayu, Indramayu, Indonesia
Email: [1*]rosenasbella@polindra.ac.id, [2]meyermega@polindra.ac.id, [3]ichsanbasir@polindra.ac.id
Submitted **22-12-2025;** Accepted **27-01-2026;** Published **24-02-2026**

Abstract

Distribution transformers are critical components of power distribution systems, and their reliability directly affects the continuity and quality of electrical energy supply. However, early-stage transformer faults are difficult to detect because their operational characteristics often closely resemble normal operating conditions, which can lead to undetected degradation and unexpected failures. This study aims to improve the accuracy and robustness of fault prediction in distribution transformers by proposing a hybrid approach that integrates the Grey Wolf Optimizer (GWO) with Support Vector Machine (SVM) and Multilayer Perceptron (MLP) models. The main contribution of this research is a direct and systematic performance comparison between baseline machine learning models and their GWO-optimized counterparts, highlighting the effectiveness of metaheuristic optimization in enhancing classification performance. GWO is employed to optimize key model parameters, enabling improved convergence behavior, higher classification accuracy, and better generalization capability. The proposed models are evaluated under four transformer operating conditions, namely Light Load Imbalance, Light Overload, Normal, and Normal High Temperature, which represent practical scenarios in power distribution networks. Model performance is assessed using standard classification metrics, including Accuracy, Precision, Recall, and F1-Score. Experimental results show that the baseline SVM achieved an accuracy of 68%, while the baseline MLP reached 87% accuracy. After GWO-based optimization, the SVM–GWO model demonstrated a significant improvement, achieving 92% accuracy, whereas the MLP–GWO model produced the best overall performance, achieving 93% accuracy, precision, recall, and F1-score. These findings confirm that GWO-based optimization substantially enhances transformer fault prediction performance and demonstrates the strong potential of the proposed hybrid models for real-time monitoring and preventive maintenance of power distribution transformers.

Keywords: Distribution Transformer; Fault Prediction; Grey Wolf Optimizer; Multilayer Perceptron; Support Vector Machine

1. INTRODUCTION

Distribution transformers are one of the most crucial components in power systems because they serve as a direct link between the medium-voltage distribution network and the end consumers. The reliability of distribution transformers determines the continuity of the electricity supply, system stability, and the quality of service to customers [1]. Failure or disturbance in distribution transformers not only causes power outages but also results in significant economic losses, decreased customer trust, and increased operational and maintenance costs of the distribution network. As the complexity and load of modern electricity distribution systems increase, the need for more accurate condition monitoring methods and transformer damage prediction becomes increasingly urgent. Various recent studies show that conventional approaches based on periodic inspections and reactive maintenance are no longer adequate to ensure the reliability of power systems in the modern era [2][3][4].

Distribution transformer damage is generally triggered by a combination of various operational and environmental factors, such as prolonged overloading, current imbalance between phases, excessive operating temperature rise, insulation system degradation, and the influence of extreme environmental conditions. Conventional maintenance methods that are periodic and based on manual inspection tend to be reactive because they are only able to identify damage after a disturbance occurs [5][6]. Consequently, the potential for early failure is often not detected in a timely manner, thereby increasing the risk of system disturbance and permanent damage to the transformer [7], [8]. To overcome these problems, predictive maintenance approaches based on operational data have begun to be widely developed. Several studies show that the utilization of transformer historical data and operational parameters through machine learning-based approaches can significantly improve the effectiveness of early detection and the accuracy of damage prediction compared to conventional methods [9][10].

Among the various machine learning methods applied in distribution transformer fault prediction and diagnosis, the Support Vector Machine (SVM) and Multilayer Perceptron (MLP) algorithms are two of the most studied approaches. SVM is known for its good generalization ability, effectiveness in handling high-dimensional data, and ability to produce optimal decision boundaries [11][12]. Meanwhile, MLP, as part of artificial neural networks, has the advantage of modeling complex non-linear relationships between transformer operational parameters and damage conditions. Various studies show that both algorithms are capable of producing promising prediction performance in power system applications [13][14][15]. However, most existing studies evaluate SVM and MLP independently and rely on conventional or manually tuned parameter settings, which may not be optimal for dynamic distribution transformer operating conditions. Since the performance of SVM and MLP is highly dependent on the selection of appropriate parameters and hyperparameters, suboptimal configurations can lead to reduced accuracy, slow convergence, and model instability [16]. Therefore, a clear gap remains in the literature regarding the systematic optimization and comparative

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9435
Hal 60-70
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

evaluation of SVM and MLP models using advanced optimization techniques under identical experimental settings for distribution transformer fault prediction.

To improve the performance of SVM and MLP algorithms, various metaheuristic optimization methods have been developed, one of which is the Grey Wolf Optimizer (GWO). GWO is an optimization algorithm inspired by the social behavior and hunting strategies of grey wolves, which has a good balance between exploration and exploitation in the search for optimal solutions [17][18][19]. Compared to other commonly used metaheuristic algorithms, such as Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), GWO has a simpler structure, requires fewer control parameters, and exhibits more stable convergence behavior, making it suitable for complex optimization problems. Several recent studies have reported that integrating GWO with machine learning algorithms, particularly SVM and MLP, can enhance prediction accuracy, stability, and reliability across various application domains, including power systems. However, most existing works focus on applying GWO to a single learning model or evaluate its performance independently, without providing a systematic and fair comparison between GWO-optimized SVM and MLP models under identical experimental conditions [20][21]. Furthermore, limited attention has been given to comprehensive performance evaluation using multiple classification metrics in the context of distribution transformer fault prediction. As a result, a clear research gap remains regarding the comparative effectiveness of GWO-based optimization over different machine learning models and its practical impact on transformer fault prediction accuracy and robustness [22][23]

Therefore, this study aims to develop and evaluate distribution transformer fault prediction models based on SVM and MLP optimized using the Grey Wolf Optimizer. The proposed approach emphasizes systematic parameter optimization and a fair comparative analysis between baseline and GWO-optimized models under identical experimental conditions, using multiple performance evaluation metrics. By explicitly assessing the impact of GWO-based optimization on different learning models, this study addresses the existing research gap in comparative transformer fault prediction studies. The results are expected to enhance predictive accuracy, robustness, and model reliability, thereby contributing to more effective predictive maintenance strategies and supporting improved decision-making in modern power distribution networks [24].

2. RESEARCH METHODOLOGY

2.1 Research Stages

This study proposes a hybrid machine learning approach for predicting faults in distribution transformers by integrating the Grey Wolf Optimizer (GWO) with Support Vector Machine (SVM) and Multi-Layer Perceptron (MLP) models. The research methodology is structured into three main stages: data exploration and preprocessing, parameter optimization using GWO, and development and performance evaluation of hybrid SVM–GWO and MLP–GWO models, as illustrated in the research flowchart.

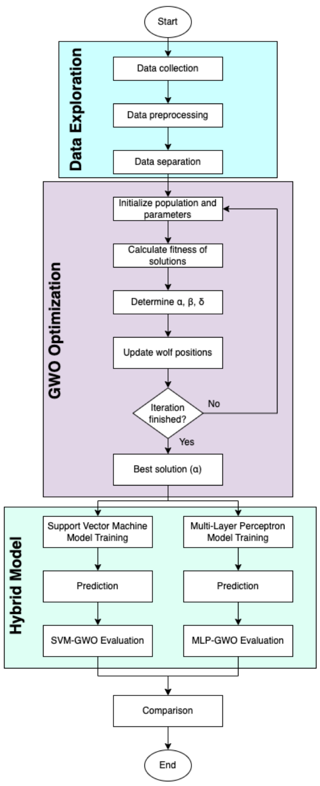


Figure 1. Research Flowchart

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9435
Hal 60-70
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

2.2 Data Exploration

a. Data Collection

The data used in this study comes from operational data and conditions of distribution transformers. The dataset includes important parameters representing the transformer's working conditions and indications of disturbances, which are used as input for the prediction model [7].

b. Data Pre-processing

The pre-processing stage is carried out to improve data quality and model reliability. This process includes handling missing values, removing outliers, and normalizing numerical data to be within a uniform range. Additionally, transformer condition labels are encoded into numerical form to support the machine learning process [7].

c. Data Splitting

The pre-processed data is then divided into training data and test data. The training data is used to build and optimize the model, while the test data is used to evaluate the model's generalization ability in predicting transformer faults [7].

2.3 Grey Wolf Optimizer Optimization

Compared to other metaheuristic optimization algorithms such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Differential Evolution (DE), the Grey Wolf Optimizer (GWO) offers a simpler structure with fewer control parameters while maintaining an effective balance between exploration and exploitation. GA requires complex genetic operators and extensive parameter tuning, whereas PSO may suffer from premature convergence in complex search spaces, and DE often involves higher computational cost. In contrast, GWO utilizes a hierarchical leadership mechanism and cooperative hunting strategy to guide candidate solutions toward the optimal position [25][26]. This efficient search behavior enables stable and fast convergence when optimizing SVM and MLP hyperparameters. The fundamental working principle of GWO, including the roles of alpha (α), beta (β), and delta (δ) wolves in updating candidate solution positions toward the estimated prey, is illustrated in Figure 2.

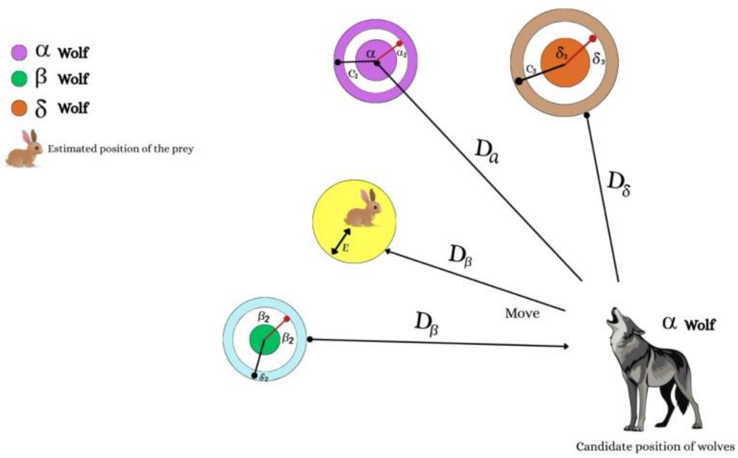


Figure 2. Research Flowchart

a. Population and Parameter Initialization

At this stage, the Grey Wolf Optimizer algorithm is initialized by determining the population size, the maximum number of iterations, and the algorithm's control parameters. Each individual (wolf) represents a combination of parameters for the SVM and MLP models [22].

b. Fitness Value Calculation

The fitness value of each individual is calculated based on the performance of the resulting prediction model. Performance metrics such as classification accuracy are used as the fitness function to guide the optimization process [22].

c. Determination of Alpha (α), Beta (β), and Delta (δ) Wolves

Based on the fitness value, the three best individuals are designated as alpha (α), beta (β), and delta (δ) wolves. These three individuals act as leaders in the search for the optimal solution [22].

d. Wolf Position Update

The position of each wolf is updated using the Grey Wolf Optimizer mathematical equations that mimic the hunting behavior and social hierarchy of grey wolves. This process enables effective exploration and exploitation of the solution space [22].

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9435
Hal 60-70
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

e. Iteration Termination Criteria

The optimization process is carried out iteratively until the termination criteria are met, namely reaching the maximum number of iterations or obtaining solution convergence [22].

2.3 Hybrid model

a. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised learning classification method that aims to find the optimal hyperplane to separate data between classes with a maximum margin. Mathematically, SVM is formulated as an optimization problem:

$$\min_{w,b,\xi} \frac{1}{2} ||w||^2 + C \sum_{i=1}^{n} \xi_i \tag{1}$$

with constraints:

$$y_i(w \cdot \phi(x_i) + b) \geq 1 - \xi_i \tag{2}$$

The CC parameter regulates the balance between the margin and classification error, while the function $\phi(\cdot)$ is represented through a kernel function. In this study, the RBF kernel is used:

b. Multi-Layer Perceptron (MLP)

Multi-Layer Perceptron (MLP) is a feedforward artificial neural network consisting of an input layer, hidden layers, and an output layer. Each neuron calculates the weighted sum of inputs:

$$K(x_i, x_j) = \exp(-\gamma ||x_i - x_j||^2) \tag{3}$$

and produces output through an activation function:

$$z_j = \sum_{i=1}^{n} w_{ij}\, x_i + b_j \tag{4}$$

The MLP training process is performed using the backpropagation algorithm with the aim of minimizing the error function:

$$L = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{5}$$

c. SVM–GWO

In the SVM–GWO model, the Grey Wolf Optimizer algorithm is used to optimize important SVM parameters, namely C and γ. Each wolf in GWO represents one parameter combination:

$$X = [C, \gamma] \tag{6}$$

The performance of each solution is evaluated using a fitness function in the form of classification accuracy. The position of the wolves is updated following the GWO mechanism, and the best solution (alpha wolf) is selected as the optimal parameter for SVM.

d. MLP–GWO

MLP–GWO integrates GWO to optimize MLP parameters such as the number of neurons in the hidden layer and the learning rate. Each solution is represented as:

$$X = [n_{hidden}, \eta] \tag{7}$$

The fitness value is calculated based on the MLP prediction performance. The best parameters resulting from GWO optimization are used to train the final MLP model, thereby improving accuracy and reducing prediction errors.

3. RESULT AND DISCUSSION

3.1 Data Exploration

a. Data Collection

The data used in this study comes from operational data and conditions of distribution transformers. The dataset includes important parameters representing the transformer's working conditions and indications of disturbances, which are used as input for the prediction model.

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9435
Hal 60-70
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

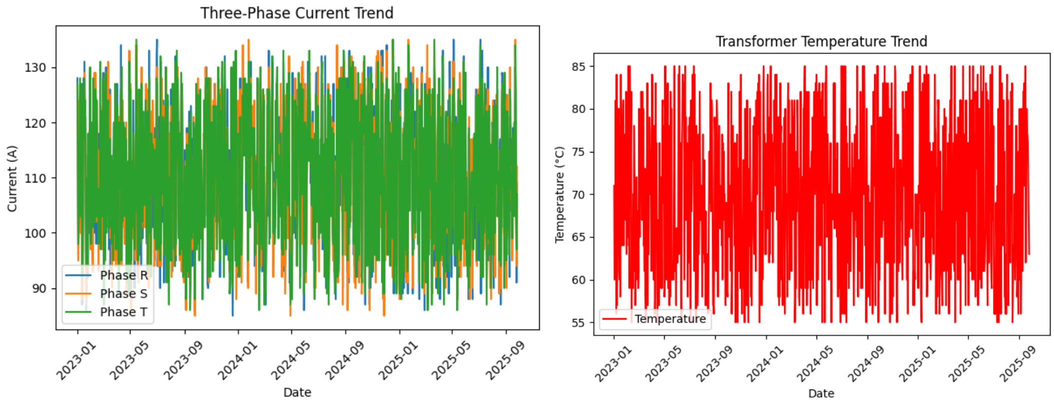


Figure 3. Distribution of Phase R, S, T Current Data and Temperature

Figure 3 represents the data collection stage in the distribution transformer condition prediction study. The graph on the left shows the trend of three-phase currents (R, S, and T) recorded periodically throughout the observation period from 2023 to 2025. This current data was obtained from the measurement of transformer loads on each phase and reflects daily and seasonal load variations. Fluctuations in current values between phases illustrate the transformer's operating conditions, including indications of load imbalances that could potentially affect the performance and reliability of the distribution transformer.

Meanwhile, the graph on the right displays the transformer temperature trend over the same period. Temperature data was obtained from the transformer's thermal sensors and used as an indicator of thermal conditions due to loading. The temperature variations that occurred show the transformer's thermal response to changes in current and environmental conditions. At this data collection stage, all parameters of phase current and temperature were recorded continuously and integrated into one dataset, thus forming a representative historical database. This dataset is then used in the pre-processing and machine learning modeling stages to predict operational conditions and potential faults in the distribution transformer.

b. Data Pre-processing

The pre-processing stage is carried out to improve data quality and model reliability. This process includes handling missing values, removing outliers, and normalizing numerical data to be within a uniform range. Additionally, transformer condition labels are encoded into numerical form to support the machine learning process.

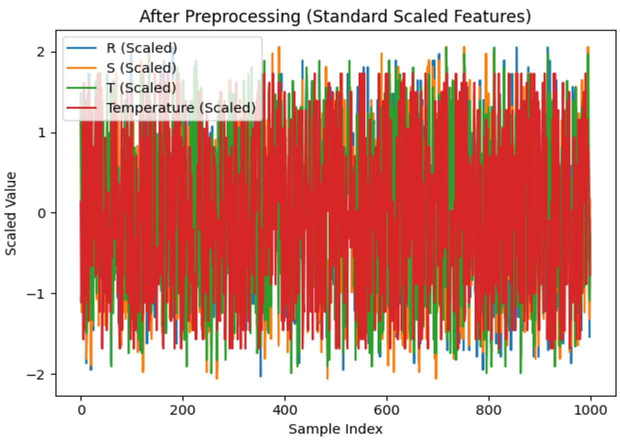


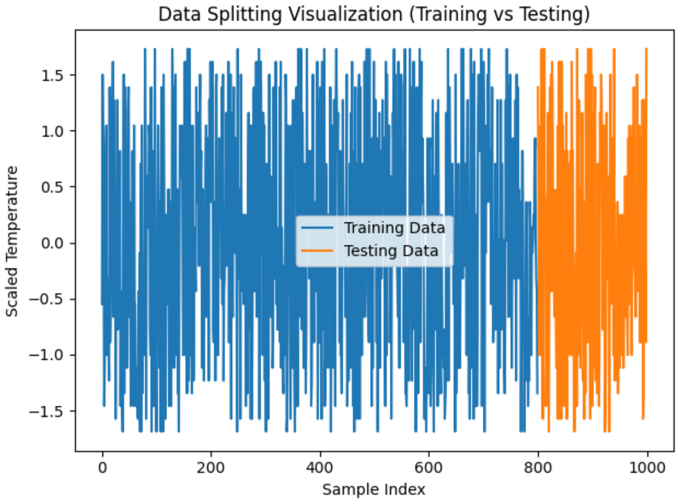
Figure 4. Distribution of Data After Pre-processing

Figure 4 shows the condition of the distribution transformer features after the pre-processing stage using the Standard Scaler. All input variables, namely phase R, S, T currents and Temperature, have been normalized to have a mean near zero and a standard deviation near one. This is visible from the data distribution which is in the range of approximately -2 to +2 on the standardized value axis. After the scaling process, no features dominate due to scale differences, so each parameter has a balanced contribution in the model learning process. This condition is crucial, especially for kernel-based SVM and MLP algorithms, as it helps improve optimization stability, speeds up convergence, and produces more optimal classification performance in predicting distribution transformer conditions.

c. Data Splitting

The pre-processed data is then divided into training data and test data. The training data is used to build and optimize the model, while the test data is used to evaluate the model's generalization ability in predicting transformer faults.

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9435
Hal 60-70
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom



Gambar 5. Data Splitting Visualization

The figure displays a visualization of data splitting between training data and testing data based on the normalized (scaled) Temperature feature. The blue part represents the training data, which includes about 80% of the total samples, while the orange part shows the test data, which includes the remaining 20%. This separation is carried out in a controlled manner to ensure that the model is trained using most of the historical data but is still tested on data it has never seen before. The distribution of standardized temperature values in both data subsets is in a relatively similar range, which is about -1.5 to +1.7, indicating that the data splitting process was carried out representatively and did not cause significant distribution differences between training and test data. This condition is important to ensure that the model's performance evaluation, such as in SVM-GWO, reflects good generalization capability and is not biased towards certain data subsets.

3.2 Grey Wolf Optimizer Optimization

Several previous studies have demonstrated the effectiveness of metaheuristic algorithms, such as Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), in optimizing machine learning models for fault prediction and classification tasks in power systems [27], [28]. However, these methods often require complex parameter tuning or exhibit premature convergence in high-dimensional search spaces. In comparison, Grey Wolf Optimizer (GWO) has been reported to achieve competitive or superior performance with fewer control parameters and more stable convergence behavior when applied to SVM and MLP optimization problems[29]. These characteristics make GWO particularly suitable for distribution transformer fault prediction, where robustness and generalization across multiple operating conditions are essential.

a. Population and Parameter Initialization

The Grey Wolf Optimizer process begins with random initialization of the wolf population in the specified parameter solution space. Each wolf represents one candidate solution in the form of a combination of SVM or MLP model parameters. Initialization is performed by generating random parameter values between the established lower and upper bounds to ensure the diversity of initial solutions.

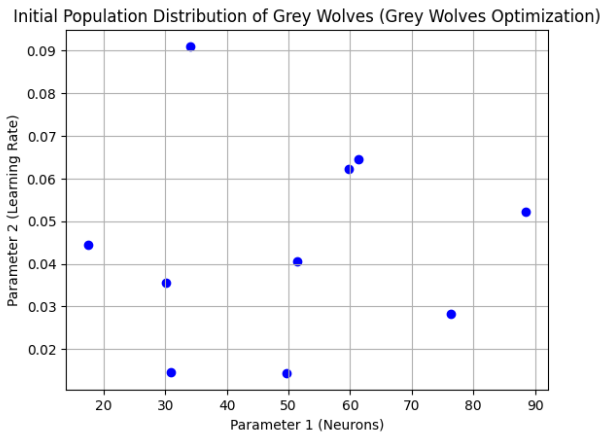


Figure 6. Initial Population Distribution of Grey Wolf Optimizer

Figure 6 shows the initial population distribution of the Grey Wolf Optimizer in a two-dimensional search space. From the image, it can be seen that the wolf positions are spread evenly across the solution space, signifying the

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9435
Hal 60-70
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

algorithm's early exploration capability in exploring various possible parameter combinations. This broad initial distribution is important to avoid premature convergence to local solutions.

b. Fitness Value Calculation

After the initial population is formed, each solution is evaluated using an objective function that represents the performance of the machine learning model. The fitness value is calculated based on the prediction error, which is 1 - Accuracy, so a solution with a smaller fitness value indicates better model performance

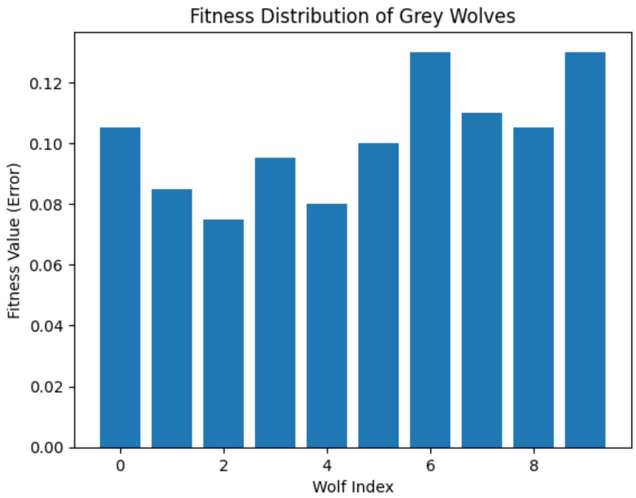


Figure 7. Fitness Value Distribution of the GWO Population

Figure 7 shows the fitness value distribution of all wolves at the evaluation stage. The large variation in fitness values indicates that the quality of initial solutions is still diverse, so selection and position update mechanisms are needed to improve solution quality. This stage becomes the basis for determining the hierarchy of alpha, beta, and delta wolves in the next iteration.

c. Determination of Alpha (α), Beta (β), and Delta (δ) Wolves

Based on the results of the fitness evaluation, the population is sorted from best to worst solutions. The three solutions with the lowest fitness values are designated as alpha (α), beta (β), and delta (δ) wolves. The alpha wolf represents the best parameter combination in that iteration, while the beta and delta wolves serve as supporting solutions that help guide the optimization process.

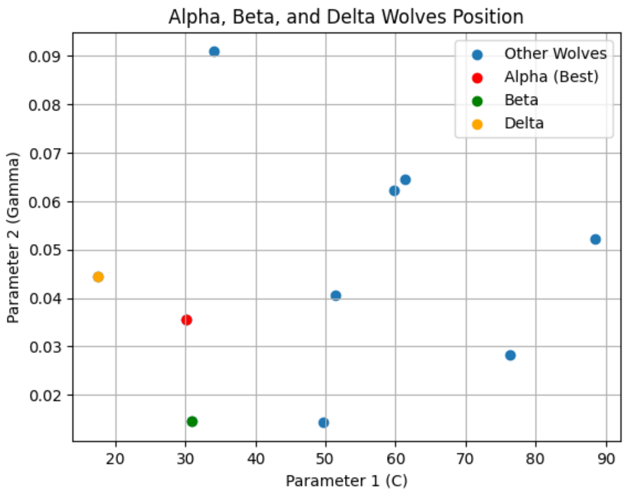


Figure 8. Positions of Alpha, Beta, and Delta Wolves

Figure 8 shows the positions of alpha, beta, and delta wolves among the rest of the population. It can be seen that these three wolves are in the solution area with the lowest fitness values. This hierarchy mimics the social behavior of wolves in nature and is the primary mechanism for directing the population movement toward the optimal solution.

d. Wolf Position Update

The position update stage is the core of the Grey Wolf Optimizer process. The position of each wolf is updated based on the combined influence of the alpha, beta, and delta wolves. This process is done by balancing exploration and exploitation through a control parameter 'a', whose value decreases linearly as the number of iterations increases.

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9435
Hal 60-70
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

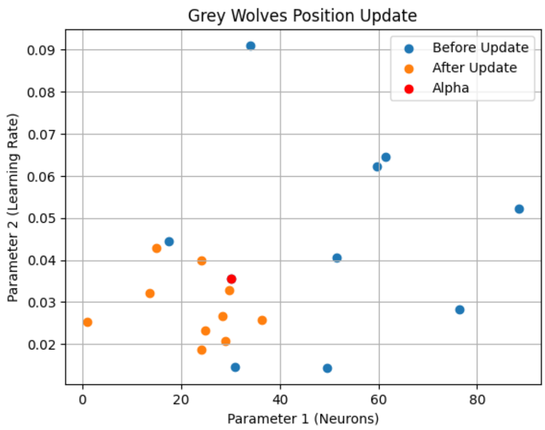


Figure 9. Wolf Population Movement

Figure 9 shows the changes in wolf positions before and after the update process. It is visible that the population moves closer to the alpha wolf position, indicating the exploitation process of the best solution. This position update causes the quality of the solutions to improve gradually at each iteration.

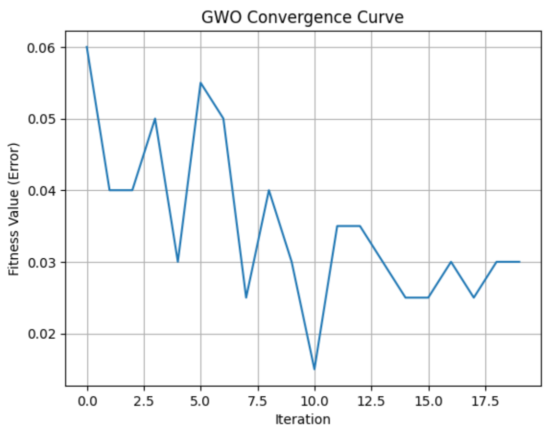


Figure 10. Grey Wolf Optimizer Convergence Curve

Furthermore, the effectiveness of the position update process is visualized through the GWO convergence curve in Figure 10. The curve shows a significant decrease in error values from the initial iteration to the final iteration. In the initial iterations, the error decreases relatively quickly due to global exploration, while in the final iterations, the curve tends to be stable, indicating that the algorithm has reached a convergent condition.

e. Iteration Termination Criteria

The optimization process is carried out iteratively until the termination criteria are met, namely reaching the maximum number of iterations or obtaining solution convergence. The best solution produced by the alpha (α) wolf is selected as the optimal model parameter. Based on the visualization results at each stage of optimization, the Grey Wolf Optimizer is proven capable of performing parameter searches systematically and efficiently. Wide-spreading population initialization supports global exploration, fitness evaluation enables selection of the best solution, alpha–beta–delta hierarchy determination guides the optimization process, and position updates produce stable convergence. This shows that GWO is effectively used as a parameter optimization method for hybrid GWO–SVM and GWO–MLP models to improve prediction performance.

3.3 Hybrid Model

The hybrid SVM–GWO and MLP–GWO models are built by integrating the Grey Wolf Optimizer (GWO) as a parameter optimization method for SVM and MLP models. In SVM–GWO, GWO is used to find optimal parameter combinations, such as kernel and regularization parameters, to improve class separation capability. Meanwhile, in MLP–GWO, GWO optimizes neural network parameters, including weights and learning parameters, thereby enhancing the model's ability to capture non-linear patterns in transformer operational data. The integration of GWO aims to improve the accuracy, sensitivity, and stability of the model in predicting transformer disturbance conditions.

a. Analysis of Confusion Matrix

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9435
Hal 60-70
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

A confusion matrix is used to compare the performance of four classification models, namely SVM and MLP as baseline models and SVM–GWO and MLP–GWO as hybrid models, in classifying four distribution transformer operational conditions: Light Load Imbalance, Light Overload, Normal, and Normal High Temperature. This evaluation focuses on the model's ability to reduce classification errors in early disturbance conditions.

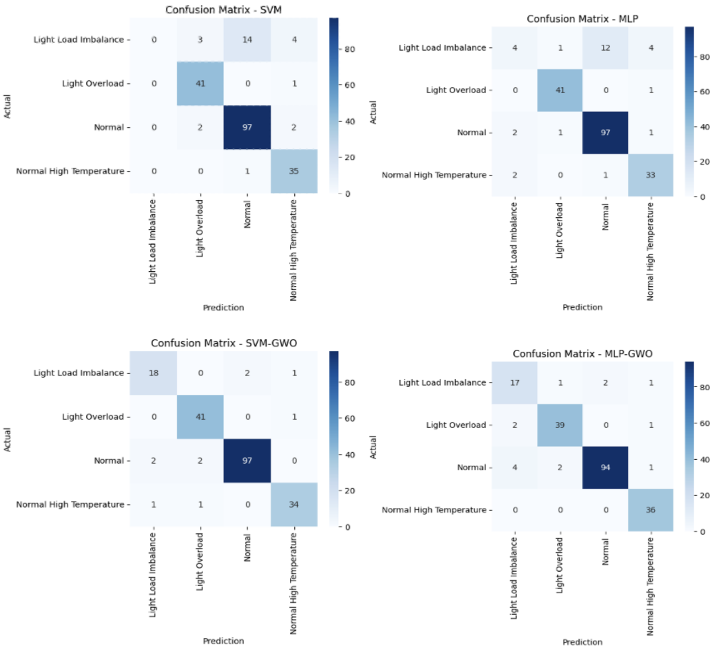


Figure 11. Confusion Matrix

The baseline SVM model shows high accuracy in the Normal class but still experiences significant errors in the Light Load Imbalance class, which is mostly misclassified as a normal condition. This indicates the limitation of SVM without optimization in distinguishing light disturbance patterns. Meanwhile, the baseline MLP shows fairly good performance in the Normal and Light Overload classes but is still less sensitive to Light Load Imbalance and Normal High Temperature, due to suboptimal network parameters. After applying the Grey Wolf Optimizer, both models show significant performance improvements. SVM–GWO is able to drastically increase classification accuracy in the Light Load Imbalance class without decreasing performance in the Normal class. This shows that optimizing SVM parameters using GWO is effective in increasing the model's sensitivity to light disturbances. On the other hand, MLP–GWO provides the most balanced performance, especially in the Normal High Temperature class, with very low error rates, indicating an improvement in the model's generalization capability.

b. Model Comparison

Model comparison is carried out to evaluate the effectiveness of the hybrid approach based on Grey Wolf Optimizer (GWO) in improving classification performance. The compared models include Support Vector Machine (SVM) and Multilayer Perceptron (MLP) as base models, as well as the hybrid SVM–GWO and MLP–GWO models. Performance evaluation is performed using Accuracy, Precision, Recall, and F1-Score metrics to assess the ability of each model to predict fault conditions in distribution transformers.

Tabel 1. Model Comparison

Model	Accuracy	Precision	Recall	F1-Score
SVM	68	73	68	60
SVM-GWO	92	93	92	92
MLP	87	88	87	87
MLP-GWO	93	93	93	93

The performance table presents a comparison of the performance of four classification models: SVM, SVM–GWO, MLP, and MLP–GWO, evaluated using Accuracy, Precision, Recall, and F1-Score metrics. These four metrics are used to provide a comprehensive overview of the model's ability to detect fault conditions in distribution transformers. The SVM model without optimization produced an accuracy of 68%, with a precision of 73%, recall of 68%, and F1-score of 60%. These results show that although SVM is capable of basic classification fairly well, the model is still inconsistent in detecting all fault conditions, especially light disturbances. The relatively low F1-score indicates an imbalance between precision and recall, making the model less than optimal for early detection systems.

After optimization using Grey Wolf Optimizer, the SVM performance improved significantly. The SVM–GWO model reached 92% accuracy, 93% precision, 92% recall, and 92% F1-score. This improvement shows that GWO

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9435
Hal 60-70
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

successfully found the optimal combination of SVM parameters, thereby improving the model's sensitivity to faults while reducing classification errors. The MLP model without optimization showed better performance than the baseline SVM, with 87% accuracy and relatively balanced precision, recall, and F1-score values. This indicates that MLP is more adaptive in capturing non-linear patterns in transformer condition data. However, there is still room for performance improvement through parameter optimization. The MLP–GWO model showed the best performance among all models, with accuracy, precision, recall, and F1-score values of 93% each. Consistency across all metrics shows a very good balance between the accuracy and completeness of predictions. This confirms that GWO is effective in optimizing MLP weights and parameters, thus improving the model's generalization capability.

Overall, the evaluation results show that the integration of Grey Wolf Optimizer consistently improves the performance of SVM and MLP models. MLP–GWO provides the most optimal and stable performance, while SVM–GWO shows a significant improvement over the baseline model. These findings prove that the metaheuristic-based hybrid approach is very effective for supporting reliable and high-precision distribution transformer fault prediction systems.

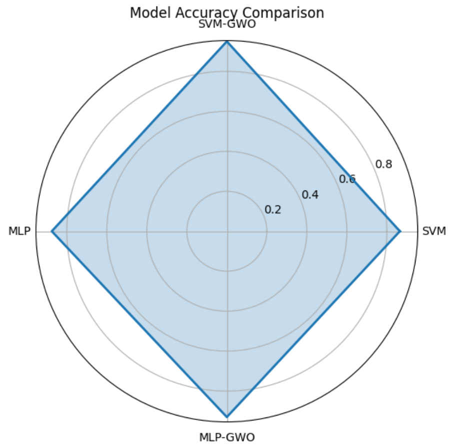


Figure 12. Model Accuracy Comparison

Figure 12 shows a radar chart comparing the accuracy values of four distribution transformer fault prediction models: SVM, MLP, SVM–GWO, and MLP–GWO. The distance of a point from the center of the chart represents the level of accuracy, where values further from the center indicate better performance. The visualization results show that models without optimization, especially SVM, have the lowest accuracy. The MLP model shows improved performance compared to SVM, indicating its ability to capture non-linear data patterns. Hybrid models based on GWO, namely SVM–GWO and MLP–GWO, show higher accuracy, with MLP–GWO occupying the outermost position as the best-performing model. Overall, this graph confirms that the application of Grey Wolf Optimizer consistently improves model accuracy and demonstrates the superiority of MLP–GWO as the most optimal approach for distribution transformer fault prediction.

4. CONCLUSION

This study concludes that the application of a hybrid model based on Grey Wolf Optimizer (GWO) can significantly improve the fault prediction performance of distribution transformers. Based on the evaluation results, the SVM model without optimization only achieved 68% accuracy with a 60% F1-score, indicating limitations in detecting light disturbances. The baseline MLP model provided better performance with 87% accuracy and 87% F1-score, but still has room for improvement in prediction consistency. The integration of GWO into SVM drastically improved performance with 92% accuracy, 93% precision, 92% recall, and 92% F1-score, signifying an improvement in model sensitivity and stability. Meanwhile, MLP–GWO produced the best and most balanced performance with accuracy, precision, recall, and F1-score of 93% each. Analysis of the confusion matrix and radar chart shows that the GWO-based model is able to reduce classification errors, especially in light disturbance and high-temperature conditions. Overall, MLP–GWO is recommended as the most optimal model for distribution transformer fault prediction systems, while SVM–GWO remains competitive with significant performance improvements. This hybrid approach has the potential to support early fault detection as well as improve the reliability and operational life of distribution transformers.

REFERENCES

[1] R. Shintabella, C. E. Widodo, and A. Wibowo, “A Hybrid Machine Learning-Genetic Algorithm (ML-GA) to Predict Power Tranformer Life,” International Journal of Health, Education &Social (IJHES), vol. 6, no. 12, p. 12, 2023, [Online]. Available: www.ijhes.com

[2] V. Rokani, S. D. Kaminaris, P. Karaisas, and D. Kaminaris, “Power Transformer Fault Diagnosis Using Neural Network Optimization Techniques,” Mathematics, vol. 11, no. 22, Nov. 2023, doi: 10.3390/math11224693.

e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9435
Hal 60-70
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

[3] G. Santamaria-Bonfil, G. Arroyo-Figueroa, M. A. Zuniga-Garcia, C. G. Azcarraga Ramos, and A. Bassam, “Power Transformer Fault Detection: A Comparison of Standard Machine Learning and autoML Approaches,” Energies (Basel)., vol. 17, no. 1, Jan. 2024, doi: 10.3390/en17010077.

[4] A. Esmaeili Nezhad and M. H. Samimi, “A review of the applications of machine learning in the condition monitoring of transformers,” Feb. 01, 2024, Springer Science and Business Media Deutschland GmbH. doi: 10.1007/s12667-022-00532-5.

[5] A. Esmaeili Nezhad and M. H. Samimi, “A review of the applications of machine learning in the condition monitoring of transformers,” Feb. 01, 2024, Springer Science and Business Media Deutschland GmbH. doi: 10.1007/s12667-022-00532-5.

[6] H. Çuhadaroğlu and Y. Uyaroğlu, “Detection of Transformer Faults: AI-Supported Machine Learning Application in Sweep Frequency Response Analysis,” Energies (Basel)., vol. 18, no. 10, May 2025, doi: 10.3390/en18102481.

[7] P. Vaishali Niranjane et al., “Machine Learning for Fault Identification in Power Transformers,” 2025.

[8] M. Alenezi, F. Anayi, M. Packianather, and M. Shouran, “Enhancing Transformer Protection: A Machine Learning Framework for Early Fault Detection,” Sustainability (Switzerland), vol. 16, no. 23, Dec. 2024, doi: 10.3390/su162310759.

[9] G. A. Hamoud, “Reliability Assessment of Distribution Power Transformers Considering Load Transfer Capability,” IEEE Transactions on Power Systems, vol. 38, no. 2, pp. 1655–1662, Mar. 2023, doi: 10.1109/TPWRS.2022.3171693.

[10] S. Arumugam, “Failure diagnosis and root-cause analysis of in-service and defective distribution transformers,” Energies (Basel)., vol. 14, no. 16, Aug. 2021, doi: 10.3390/en14164997.

[11] N. Epriyanti, A. Meiriza, and D. Yunika Hardiyanti, “Perbandingan Kinerja SVM, Random Forest dan XGBoost pada Aplikasi Access by KAI Menggunakkan ADASYN,” Jurnal Riset Komputer), vol. 12, no. 5, pp. 2407–389, 2025, doi: 10.30865/jurikom.v12i5.9134.

[12] Z. B. Sahri and R. B. Yusof, “Support Vector Machine-Based Fault Diagnosis of Power Transformer Using k Nearest-Neighbor Imputed DGA Dataset,” Journal of Computer and Communications, vol. 02, no. 09, pp. 22–31, 2014, doi: 10.4236/jcc.2014.29004.

[13] J. L. Hernández, “Data-driven predictive maintenance of distribution transformers: class imbalance and Bayesian optimization,” Jun. 11, 2025. doi: 10.36227/techrxiv.174961906.69711574/v1.

[14] D. Bassi and H. Singh, “A Comparative Analysis of Metaheuristic Feature Selection Methods in Software Vulnerability Prediction,” E-Informatica Software Engineering Journal, vol. 19, no. 1, 2025, doi: 10.37190/e-Inf250103.

[15] B. C. Mateus, J. T. Farinha, and M. Mendes, “Fault Detection and Prediction for Power Transformers Using Fuzzy Logic and Neural Networks,” Energies (Basel)., vol. 17, no. 2, Jan. 2024, doi: 10.3390/en17020296.

[16] A. I. S Azis and B. Santoso, “Penerapan Grey Wolf Optimizer dalam Pelatihan Multi Layer Perceptron untuk Menangani Masalah Klasifikasi dan Regresi,” 2024.

[17] M. H. Nadimi-Shahraki, S. Taghian, and S. Mirjalili, “An improved grey wolf optimizer for solving engineering problems,” Expert Syst. Appl., vol. 166, Mar. 2021, doi: 10.1016/j.eswa.2020.113917.

[18] Y. Liu, A. As’arry, M. K. Hassan, A. A. Hairuddin, and H. Mohamad, “Review of the grey wolf optimization algorithm: variants and applications,” Neural Comput. Appl., vol. 36, no. 6, pp. 2713–2735, Feb. 2024, doi: 10.1007/s00521-023-09202-8.

[19] M. Husayn, O. R. Adegboye, and A. Alzubi, “GWO-Optimized Ensemble Learning for Interpretable and Accurate Prediction of Student Academic Performance in Smart Learning Environments,” Applied Sciences (Switzerland), vol. 15, no. 22, Nov. 2025, doi: 10.3390/app152212163.

[20] T. R. D. Saputri, E. Kurniawan, C. C. Lestari, and T. Antonio, “Nature-based Hyperparameter Tuning of a Multilayer Perceptron Algorithm in Task Classification: A Case Study on Fear of Failure in Entrepreneurship,” Journal of Applied Data Sciences, vol. 6, no. 2, pp. 969–980, May 2025, doi: 10.47738/jads.v6i2.539.

[21] M. Almseidin, A. Gawanmeh, M. Alzubi, J. Al-Sawwa, A. S. Mashaleh, and M. Alkasassbeh, “Hybrid Deep Neural Network Optimization with Particle Swarm and Grey Wolf Algorithms for Sunburst Attack Detection,” Computers, vol. 14, no. 3, Mar. 2025, doi: 10.3390/computers14030107.

[22] H. Rezaei, O. Bozorg-Haddad, and X. Chu, “Grey wolf optimization (GWO) algorithm,” in Studies in Computational Intelligence, vol. 720, Springer Verlag, 2018, pp. 81–91. doi: 10.1007/978-981-10-5221-7_9.

[23] N. Ji et al., “Transformer fault identification based on GWO-optimized Dual-channel M-A method,” PLoS One, vol. 19, no. 10 October, Oct. 2024, doi: 10.1371/journal.pone.0312474.

[24] M. Almseidin, A. Gawanmeh, M. Alzubi, J. Al-Sawwa, A. S. Mashaleh, and M. Alkasassbeh, “Hybrid Deep Neural Network Optimization with Particle Swarm and Grey Wolf Algorithms for Sunburst Attack Detection,” Computers, vol. 14, no. 3, Mar. 2025, doi: 10.3390/computers14030107.

[25] F. Pace, A. Raftogianni, and A. Godio, “A Comparative Analysis of Three Computational-Intelligence Metaheuristic Methods for the Optimization of TDEM Data”, doi: 10.1007/s00024.

[26] R. Mohakud and R. Dash, “Designing a grey wolf optimization based hyper-parameter optimized convolutional neural network classifier for skin cancer detection,” Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 8, pp. 6280–6291, Sep. 2022, doi: 10.1016/j.jksuci.2021.05.012.

[27] A. A. Bouzandani, B. Taheri, and S. A. Hosseini, “Optimized SVM and feature selection for fault detection and classification in GFM-IBR system,” Measurement: Energy, vol. 9, p. 100081, Mar. 2026, doi: 10.1016/j.meaene.2025.100081.

[28] M. Alenezi, F. Anayi, M. Packianather, and M. Shouran, “A Hybrid Machine Learning Framework for Early Fault Detection in Power Transformers Using PSO and DMO Algorithms,” Energies (Basel)., vol. 18, no. 8, Apr. 2025, doi: 10.3390/en18082024.

[29] Y. Wang, Y. Yin, H. Zhao, J. Liu, C. Xu, and W. Dong, “Grey wolf optimizer with self-repulsion strategy for feature selection,” Sci. Rep., vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-97224-8.