

## Leveraging the Firefly Algorithm for Dimensionality Reduction and Accurate Diabetes Prediction

Paul Olujide Adebayo<sup>1</sup>, Micheal Abayomi Olatunji<sup>2</sup>, Oluwakemi Oduwole<sup>3</sup>, Olayiwola Stephen Ojo<sup>4</sup>

<sup>1</sup>Department of Computer Science, University of Ilorin, Nigeria

<sup>2</sup>Department of Cybersecurity, Redeemer's University, Ede, Nigeria

<sup>3</sup>Department of Computer Science, University of Ilesa, Nigeria

<sup>4</sup>Computer Science Programme, National Mathematical Centre, kwali-Sheda, Abuja, Nigeria

<sup>1</sup>[adebayo.po@unilorin.edu.ng](mailto:adebayo.po@unilorin.edu.ng); <sup>2</sup>[olatunjim@run.edu.ng](mailto:olatunjim@run.edu.ng); <sup>3</sup>[ooduwole@nmc.edu.ng](mailto:ooduwole@nmc.edu.ng)

<sup>4</sup>[layi\\_ojo@unilesa.edu.ng](mailto:layi_ojo@unilesa.edu.ng)

**ABSTRACT:** Accurate and early diagnosis of diabetes remains a significant challenge in healthcare, compounded by the high dimensionality of medical datasets. These datasets often contain numerous irrelevant or redundant features, which can negatively impact model performance and increase computational complexity. Feature selection, the process of identifying the most relevant features, is critical in addressing these challenges. This study explores the application of the Firefly Algorithm (FA), a metaheuristic optimization technique inspired by the natural behavior of fireflies, for efficient feature selection in diabetes prediction models. The FA simulates fireflies' attraction behavior to navigate complex search spaces and identifies key features that improve model accuracy. Using a diabetes dataset, the study evaluates the performance of the Firefly Algorithm compared with traditional feature selection methods. Seven classification models were tested, and performance metrics, including accuracy, precision, recall, and F1-score, were used to evaluate the results. The findings demonstrate that the Firefly Algorithm effectively reduces dimensionality while maintaining or improving prediction accuracy. This method shows great potential to enhance diabetes diagnosis by developing more efficient and accurate machine learning models, ultimately leading to improved patient outcomes. The study also highlights the advantages of the Firefly Algorithm over conventional techniques in healthcare applications.

**KEYWORDS:** Firefly Algorithm, Feature Selection, Diabetes Prediction, Machine Learning, Dimensionality Reduction, Metaheuristic Optimization.

### I. INTRODUCTION

Early and accurate diagnosis of diabetes is a critical challenge in healthcare as the global increase of this chronic condition continues to strain healthcare systems and resources. Detecting the disease early and intervening quickly can significantly improve patient outcomes and reduce long-term complications associated with diabetes (Smoljo, 2023). However, the complexity of medical datasets, which often contain numerous features, poses a significant obstacle to developing effective predictive models. Specifically, irrelevant or redundant features in these datasets can reduce model performance, increase computational costs, and complicate the interpretation of results. Therefore, feature selection—choosing and retaining only the most relevant features—

becomes essential for enhancing model accuracy and efficiency (Cheng, 2024; I.V. Dwaraka Srihith et al., 2023).

Feature selection is a crucial step in machine learning, especially when dealing with large and complex datasets standard in medical diagnostics. The goal is to improve model performance by concentrating on the most significant features that affect the outcome, while eliminating noise that can reduce predictive accuracy. Traditional methods such as filter, wrapper, and embedded techniques are commonly used across various fields, including healthcare (Bashir et al., 2022; Khalid et al., 2023). However, these methods often face challenges related to computational efficiency and managing complex, high-dimensional spaces. In recent years, metaheuristic optimization methods have gained attention for their ability to enhance feature selection in machine learning tasks (Kamal et al., 2024). One such approach is the Firefly Algorithm (FA), inspired by the natural behavior of fireflies (Yang, 2009). The FA simulates the attraction between fireflies based on their brightness, which indicates the quality of solutions in the search space. This algorithm has been successfully applied to various optimization problems, yielding promising results in fields such as engineering, finance, and biology. Its ability to explore complex solution spaces and identify optimal or near-optimal solutions makes it a strong choice for feature selection in predictive models. This study examines the use of the Firefly Algorithm for feature selection in diabetes prediction models. Its primary goal is to evaluate the effectiveness of FA in reducing dimensionality while improving model accuracy. Specifically, the study compares FA's performance to traditional feature selection methods, assessing its impact on diabetes prediction accuracy across various machine learning models. Through multiple experiments, this research demonstrates how FA can optimize feature selection, improve prediction accuracy, and provide a more efficient approach to handling high-dimensional medical datasets. The findings in this paper contribute to the expanding body of research on the use of metaheuristic algorithms in healthcare. By exploring the Firefly Algorithm's potential for feature selection in diabetes prediction, this study offers valuable insights into its effectiveness and highlights its advantages over traditional methods. Additionally, the research underscores the FA's broader capability to enhance the accuracy and efficiency of machine learning models across various healthcare applications. Ultimately, the insights from this study could help develop more accurate, efficient, and cost-effective diagnostic tools, leading to better patient outcomes in the fight against diabetes. The following are the contribution to the study:

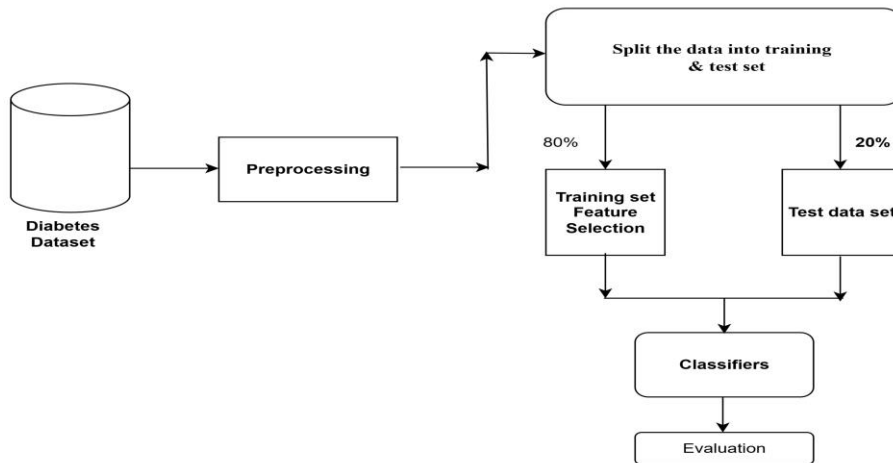
- (a). **Enhanced Model Performance:** This study demonstrates how using the Firefly Algorithm for feature selection can optimize machine learning models for diabetes prediction, enhancing key performance metrics like accuracy, precision, recall, and F1-score.
- (b). **Application of FA in Healthcare:** The study demonstrates the practical use of the Firefly Algorithm in healthcare, specifically for diagnosing diabetes, providing an innovative way to enhance diagnostic accuracy and patient results.
- (c). **Potential for Broader Healthcare Applications:** The findings indicate that the Firefly Algorithm has considerable potential for application in various healthcare fields, particularly for tasks involving complex datasets with irrelevant or redundant features, thereby expanding its role in medical prediction tasks.

## II. RELATED WORK

Accurately predicting diabetes is a significant challenge in healthcare because medical datasets are complex. It's essential to select the right features from these datasets to improve predictive model performance (Datta et al., 2023; Sreehari & Babu, 2024). Although commonly utilized, conventional feature selection techniques such as filter, wrapper, and embedded methods are often complemented by metaheuristic algorithms, which are increasingly recognized for their ability to manage complex datasets more effectively (Jain et al., 2023; Sadeghian et al., 2025). In recent times, Swarm intelligence-based algorithms, such as Particle Swarm Optimization (PSO) (Jeyafzam et al., 2021), ensemble PSO-Grey Wolf Optimizer (GWO) (Ulutas et al., 2024), Hybrid Gradient Descent Gray Wolf Optimizer (HGD-GWO) (Baharie et al., 2025). The Firefly Algorithm (FA) has demonstrated significant results in feature selection. Also, studies, which include that by Rahman et al. (2025), indicate that metaheuristic approaches often outperform conventional feature selection methods for diabetes prediction. The Firefly Algorithm is important because it can identify key features in high-dimensional datasets. This helps improve accuracy and efficiency in classification tasks. Many studies have also examined hybrid approaches that combine FA with other algorithms and machine learning classifiers to further improve performance. For example, (Deepalakshmi et al., 2025; Geo Jenefer & Deepa, 2023) used FA on the Pima Indian Diabetes dataset, finding it enhanced classification accuracy compared to more traditional methods. Researchers have combined a technique called FA with cuckoo search to predict diabetes. This shows that the algorithm is flexible and could further improve prediction accuracy. Although FA shows promise, there is still a need for direct comparisons between FA and other well-established metaheuristic algorithms across multiple diabetes prediction datasets. This gap in literature remains. This study aims to address this by offering a comprehensive evaluation of FA alongside traditional feature selection methods, with a specific focus on enhancing prediction accuracy for diabetes. In summary, while metaheuristic algorithms, especially FA, show great potential, there is still much to explore in refining these techniques for practical healthcare applications. This research builds on existing work and aims to provide new insights into how FA can be used to develop more efficient and accurate diabetes prediction models.

## III. METHODOLOGY

This study uses both analytical and quantitative methods to create a diabetes prediction model. The process starts with gathering diabetes datasets, including patient data such as medical records. Data preprocessing then involves handling missing values, normalizing data, and removing outliers to improve data quality. The core of the methodology is feature selection using the Firefly Algorithm (FA), inspired by the flashing behavior of fireflies. FA repeatedly selects the most relevant features by optimizing classification metrics like accuracy and precision. Its ability to avoid local optima and reduce dimensionality by discarding irrelevant features improves model accuracy and efficiency. This methodology aims to develop a reliable diabetes prediction model, enhancing both prediction accuracy and computational efficiency.



**Figure 1:** Framework of the Feature Selection Method, FA, and Classifiers

Figure 1 illustrates the complete process for classifying diabetes. First, the diabetes dataset is preprocessed and then divided into two parts: 80% for training and 20% for testing. To improve the model's accuracy, the Firefly Algorithm (FA) is applied to the training data to select the most important features. These features are then used to train various classification models. Finally, the models are evaluated on the test data to assess their performance and generalization to new data.

### A. Firefly Algorithm for Feature Selection

Objective Function:

The objective function  $f(x)$  evaluates fireflies, where  $x = (x_1, x_2, \dots, x_d)^T$  represents a firefly's position in the feature space. The light intensity  $I_i$  at position  $x_i$  is determined by:

$$I_i = f(x_i) \tag{Equation 1}$$

Step 1: Initialize Population

Generate an initial population of  $n$  fireflies  $x_i$ , where  $i = 1, 2, \dots, n$ , and evaluate their brightness:

$$I_i = f(x_i) \tag{Equation 2}$$

Step 2: Define Light Absorption Coefficient

Set the light absorption coefficient  $\gamma$ .

Step 3: Iterative Process

Repeat the following until the termination criterion is met (e.g., maximum generations or convergence):

$$\text{While } t < \text{MaxGeneration} \tag{Equation 3}$$

Step 4: Loop Through Each Firefly  $i$

For each firefly  $i = 1, 2, \dots, n$ :

Step 5: Loop Through Other Fireflies  $j$

For each other, firefly  $j = 1, 2, \dots, n$ :

Step 6: Compare Brightness

If firefly  $i$  is less bright than a firefly  $j$  ( $I_i < I_j$ ), move  $i$  towards  $j$ :

$$\text{If } I_i < I_j \text{ (Equation 4)}$$

Step 7: Update the Position of Firefly  $i$

The new position is calculated as:

$$x_i(t + 1) = x_i(t) + \beta_0 e^{-\gamma r_{ij}^2} (x_j(t) - x_i(t)) + \alpha \epsilon_i(t) \tag{Equation 5}$$

Where:

- $x_i(t)$  and  $x_j(t)$  are positions of fireflies  $i$  and  $j$  at time  $t$
- $r_{ij} = \|x_i(t) - x_j(t)\|$  is the distance between fireflies  $i$  and  $j$
- $\beta_0 e^{-\gamma r_{ij}^2}$  is the attractiveness based on brightness?
- $\alpha \epsilon_i(t)$  introduces randomness

Step 8: Vary Attractiveness with Distance

$$\text{Attractiveness} = e^{-\gamma r_{ij}^2} \tag{Equation 6}$$

Step 9: Evaluate New Solutions

$$I_i = f(x_i(t + 1)) \tag{Equation 7}$$

Step 10–11: Continue Loops

Repeat for all fireflies  $j$  (inner loop) and  $i$  (outer loop).

Step 12: Rank Fireflies

Rank fireflies based on brightness and determine the current global best solution.  $g^*$ :

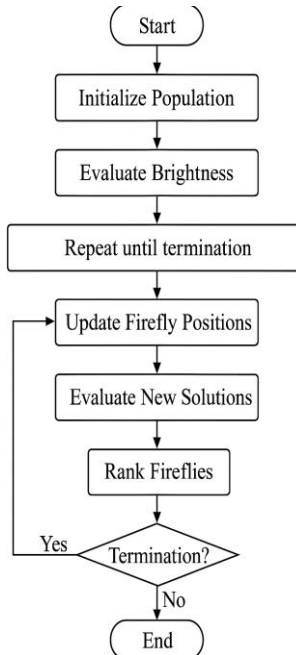
$$g^* = \arg \max I_i \tag{Equation 8}$$

Step 13: Termination

Stop when the stopping condition is satisfied (e.g., after a specified number of generations or upon convergence).

Step 14: Postprocessing and Visualization

After termination, process the results and visualize the selected feature subset for diabetes prediction.



**Figure 2:** A Typical Flowchart for Firefly Algorithm

### B. Classifier

After feature selection, classification models analyze the chosen features to identify whether instances are diabetic or non-diabetic. The models used include Support Vector Machines (SVM), Random Forests, k-Nearest Neighbors, Light Gradient Boosting Machines, Neural Networks,

Logistic Regression, and Multilayer Perceptron. These models are trained on the optimized feature set derived from the Firefly Algorithm (FA) and evaluated with metrics such as accuracy, precision, recall, F1 score (see Table 2), and the area under the Receiver Operating Characteristic (ROC) curve (Figure 4). The performance of the FA-optimized model is compared to models that use all features and other feature selection methods. Statistical tests and graphical representations, such as ROC curves, highlight improvements in classification accuracy, processing time, and computational efficiency, providing a comprehensive evaluation of the approach.

### C. Performance Metrics for Classification Algorithms

The performance of classification models is evaluated using several key metrics that provide insights into their ability to distinguish classes, particularly on imbalanced datasets such as diabetes prediction. These metrics assess different aspects of model performance, including overall accuracy, the ability to correctly identify positive cases, and the trade-off between precision and recall. Table 1.1 summarizes the key performance metrics used to evaluate the classification algorithms for diabetes diagnosis.

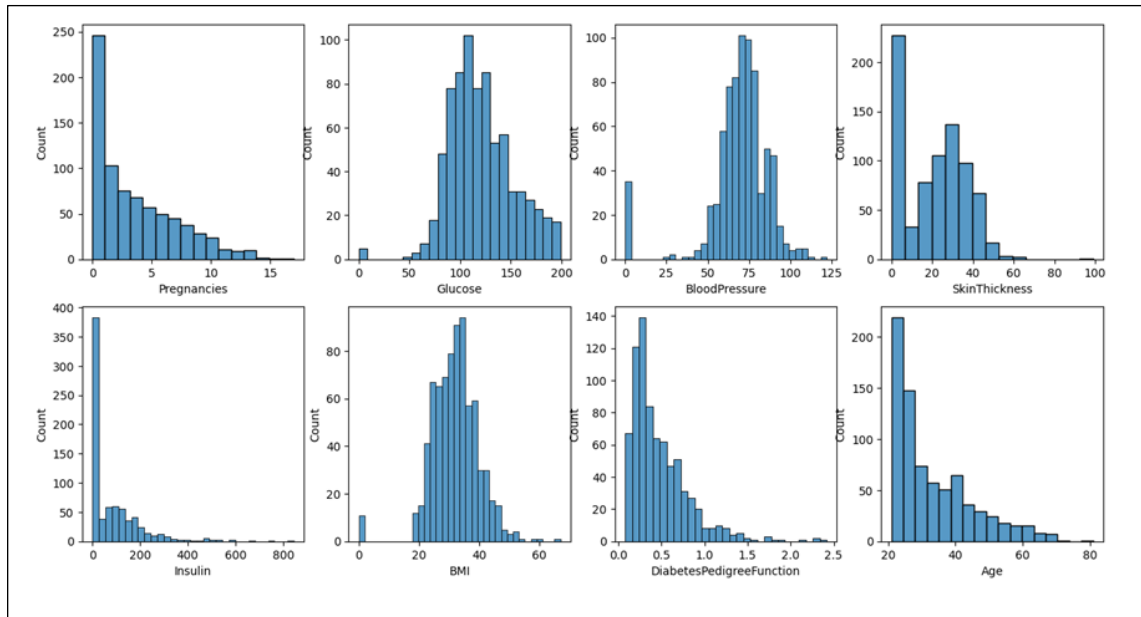
**Table 1:** Metrics for classification algorithms

Metric	Description	Formula
<b>Accuracy</b>	Percentage of accurately predicted instances out of the total cases.	$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$
<b>Precision</b>	Proportion of true positives among all optimistic predictions made by the model.	$\text{Precision} = \frac{TP}{TP + FP}$
<b>Recall (Sensitivity)</b>	Proportion of actual positive cases correctly predicted by the model.	$\text{Recall} = \frac{TP}{TP + FN}$
<b>F1 Score</b>	Harmonic mean of precision and recall, useful when there is a trade-off between them.	$\text{F1 Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}}$
<b>ROC Curve and AUC</b>	Measures the model's ability to differentiate between classes by plotting the actual positive rate against the false positive rate. AUC indicates the area under the ROC curve.	

## IV. RESULTS AND DISCUSSION

In this section, the study presents and discusses the results obtained from applying the Firefly feature selection algorithm and various classifiers to the Pima Indians Diabetes Database. The effectiveness of each model was evaluated using standard performance metrics such as accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (AUC-ROC). Additionally, the study provides an analysis of the dataset, visualizing each feature to assess its contribution to diabetes diagnosis and better understand the distribution of key variables. The results are further enhanced with visualizations, specifically histograms of eight different features from the Pima Diabetes dataset. Figure 3 presents histograms of key variables, including

pregnancies, glucose levels, blood pressure, skin thickness, insulin, BMI, diabetes pedigree function, and age. The distribution patterns offer insights into the data, including the positively skewed distribution of pregnancies, glucose, and age, as well as the skewed distributions of insulin and skin thickness, which may indicate missing or erroneous data. For example, many insulin and skin thickness values are close to zero, suggesting missing or unrecorded data for some individuals. Conversely, features such as BMI and diabetes pedigree function tend to follow more normal distributions, providing valuable insights into the demographic characteristics of the study population. These visualizations deepen understanding of the dataset's structure and the importance of each feature in predicting diabetes.



**Figure 2** Histogram to Visualize the Diabetes Features

**A. Classifiers Performances**

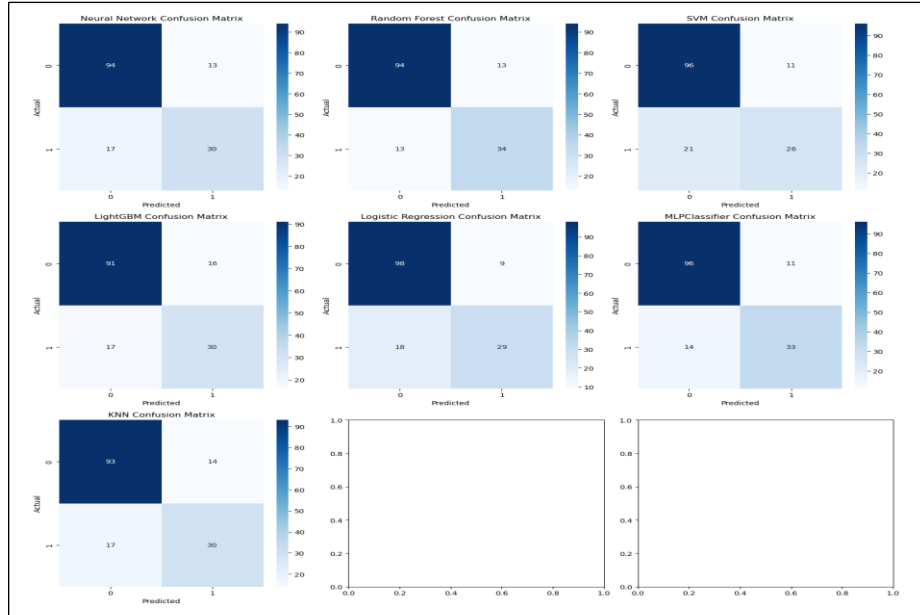
The table summarizes the performance of several machine learning classifiers on a diabetes dataset, evaluating them based on metrics like accuracy, precision, recall, specificity, F1 score, AUC, and misclassification error. The performance of the classifiers demonstrates notable variation across evaluation metrics. Artificial Neural Networks (ANN) achieved an accuracy of 80.52%, with a precision of 69.77% and a recall of 63.83%, indicating moderate effectiveness in identifying positive cases while maintaining strong specificity (87.85%). Its F1 score (66.67%) reflects a reasonable balance between precision and recall, and its AUC (84.47%) confirms good discriminative ability, though the misclassification error (19.48%) suggests room for improvement. Random Forest (RF) emerged as one of the strongest models, attaining an accuracy of 83.12%, balanced precision and recall (72.34%), and high specificity (87.85%). Its F1 score (72.34%), robust AUC (86.41%), and the lowest misclassification error (16.88%) collectively indicate highly reliable performance.

**Table 2:** Performance Metrics of Classification Models on Diabetes Dataset

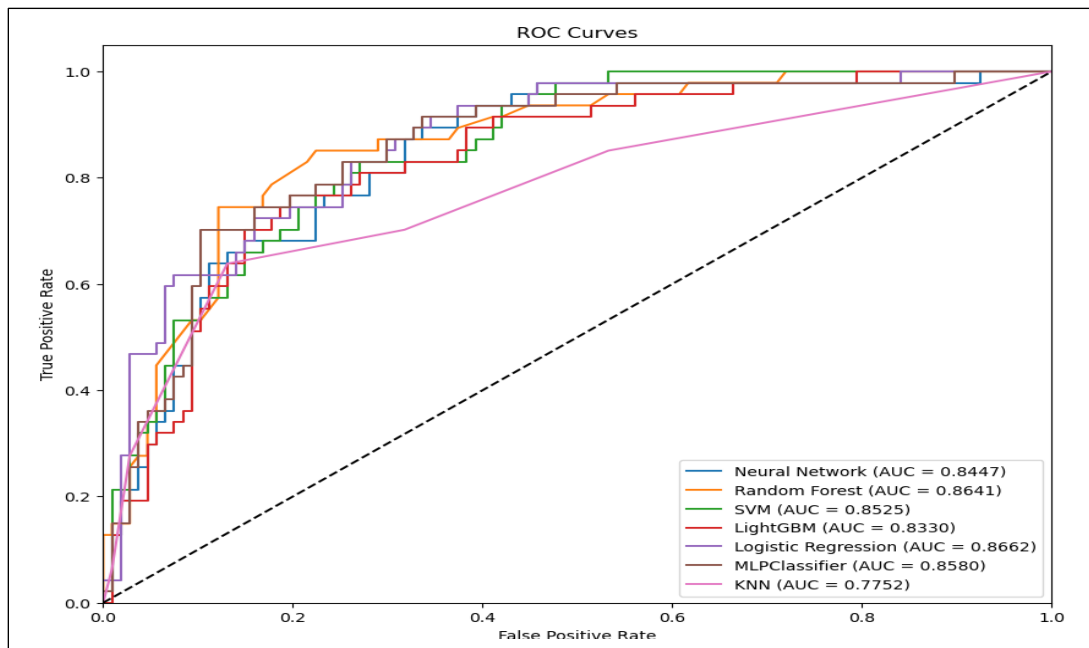
Classifier	Accuracy	Precision	Recall	Specificity	F1 Score	AUC	M. Error
ANN	0.8052	0.6977	0.6383	0.8785	0.6667	0.8447	0.1948
RF	0.8312	0.7234	0.7234	0.8785	0.7234	0.8641	0.1688
SVM	0.7922	0.7027	0.5532	0.8972	0.6190	0.8525	0.2078
LightGBM	0.7857	0.6522	0.6383	0.8505	0.6452	0.8330	0.2143
LR	0.8247	0.7632	0.6170	0.9159	0.6824	0.8662	0.1753
<b>MLP</b>	<b>0.8377</b>	<b>0.7500</b>	<b>0.7021</b>	<b>0.8972</b>	<b>0.7253</b>	<b>0.8580</b>	<b>0.1623</b>
KNN	0.7987	0.6818	0.6383	0.8692	0.6593	0.7752	0.2013

Support Vector Machine (SVM) achieved an accuracy of 79.22% and a precision of 70.27%, but its recall (55.32%) was the lowest among the models, indicating reduced sensitivity to positive cases despite excellent specificity (89.72%). LightGBM showed moderate performance, with 78.57% accuracy, lower precision (65.22%), and average recall (63.83%), along with a comparatively higher misclassification error (21.43%). Logistic Regression demonstrated strong performance, reaching 82.47% accuracy and high precision (76.32%), though its recall (61.70%) was modest. It achieved the highest specificity (91.59%) and an excellent AUC (86.62%), with a low misclassification error (17.53%). The MultiLayer Perceptron (MLP) had the highest accuracy (83.77%) and strong precision (75.00%) and recall (70.21%), supported by high specificity (89.72%) and an F1 score of 72.53%. Its low misclassification error (16.23%) makes it the best overall performer. Meanwhile, KNearest Neighbors (KNN) showed moderate performance across metrics, with 79.87% accuracy and balanced precision (68.18%) and recall (63.83%), but with the lowest AUC (77.52%) and a higher error rate (20.13%).

Overall, Random Forest and MLP emerge as the most effective classifiers for this dataset, combining high accuracy, strong F1 scores, and low error rates. Logistic Regression also performs well, especially in terms of specificity and AUC. Conversely, SVM and LightGBM show limitations in recall, while KNN has the weakest discriminative ability. These comparison results clearly support choosing the most suitable classification model based on the desired performance criteria.



**Figure 3:** Confusion Matrices for the Classifiers explored in Diabetes Classification



**Figure 4:** ROC Curve Graphs for the Classification Models

In Figure 4, the ROC curve provides a comparative analysis of the performance of several classifiers on the dataset. Logistic Regression (AUC = 0.8662) and Random Forest (AUC =

0.8641) exhibited the highest ability to differentiate between classes, indicating their superior performance. The MLP Classifier (AUC = 0.8580) and SVM (AUC = 0.8525) also demonstrated strong performance, though slightly behind the top two models. Neural Network (AUC = 0.8447) and LightGBM (AUC = 0.8330) performed reasonably well but were outperformed by other classifiers. KNN (AUC = 0.7752) had the lowest AUC, suggesting its relatively weaker classification ability. Overall, the most effective models, based on AUC, were Logistic Regression, Random Forest, and MLP Classifier.

## V. CONCLUSION

This study demonstrates that the Firefly Algorithm (FA) offers an effective and efficient approach to feature selection in diabetes classification. By reducing the dimensionality of the dataset and isolating the most informative features, FA enhances the predictive performance of several machine learning classifiers, including Neural Networks, Random Forests, Support Vector Machines, and Logistic Regression. This reduction in feature space not only improves model accuracy and computational efficiency but also contributes to more interpretable models—an essential requirement in clinical settings where understanding key determinants of disease is critical. The findings further show that FA-based feature selection consistently outperforms models trained on the full feature set, providing measurable gains in classification performance. The approach presented here highlights a scalable framework for optimizing machine learning workflows in medical diagnosis, where both precision and rapid decision-making are vital. The successful integration of FA with various classifiers underscores its potential for broader application across other medical and high-dimensional datasets, reinforcing its value as a practical tool for advancing data-driven decision-making in healthcare.

## REFERENCES

- Baharie, S. R. A. P., Surono, S., & Thobirin, A. (2025). Hybrid Gradient Descent Grey Wolf Optimizer for Machine Learning Performance Enhancement. *Jurnal RESTI*, 9(1), 146–152. <https://doi.org/10.29207/resti.v9i1.6203>
- Bashir, S., Khattak, I. U., Khan, A., Khan, F. H., Gani, A., & Shiraz, M. (2022). A Novel Feature Selection Method for Classification of Medical Data Using Filters, Wrappers, and Embedded Approaches. *Complexity*, 2022. <https://doi.org/10.1155/2022/8190814>
- Cheng, X. (2024). A Comprehensive Study of Feature Selection Techniques in Machine Learning Models. *Insights in Computer, Signals and Systems*, 1(1), 65–78. <https://doi.org/10.70088/xpf2b276>
- Datta, S., Mahedy Hasan, S. M., Faruk, M. F., Srizon, A. Y., Hossain, M. R., Efat, A. H., & Mitu, M. (2023). Improved Diabetes Prediction with Reduced Feature Sets: Evaluating Feature Selection Techniques in Machine Learning. *2023 International Conference on Information and Communication Technology for Sustainable Development, ICICT4SD 2023 - Proceedings*, 104–108. <https://doi.org/10.1109/ICICT4SD59951.2023.10303528>
- Deepalakshmi, M., Deepalakshmi, P., & Nagaraj, P. (2025). Efficient Diabetes Detection Using Ensemble Models Optimized by Firefly Algorithm. *Proceedings of 2025 2nd International Conference on Cognitive Robotics and Intelligent Systems, ICC - ROBINS 2025*, 573–578. <https://doi.org/10.1109/ICC-ROBINS64345.2025.11086333>
- Geo Jenefer, G., & Deepa, A. J. (2023). Diabetes disease prediction using firefly optimization-based catboost classifier in big data analytics. *Journal of Intelligent and Fuzzy Systems*, 44(6), 9943–9954. <https://doi.org/10.3233/JIFS-223105>
- I.V. Dwaraka Srihith, L. Rajjitha, K. Owdharya, A. David Donald, & G. Thippana. (2023). Trimming the

- Fat: An Insightful Exploration of Feature Selection and Dimensionality Reduction. *International Journal of Advanced Research in Science, Communication and Technology*, 11–24. <https://doi.org/10.48175/ijarsct-11403>
- Jain, S., Jain, A., & Jangid, M. (2023). Review of Metaheuristic Techniques for Feature Selection. *Lecture Notes in Networks and Systems*, 627 LNNS, 397–410. [https://doi.org/10.1007/978-981-19-9858-4\\_33](https://doi.org/10.1007/978-981-19-9858-4_33)
- Jeyafzam, F., Vaziri, B., Suraki, M. Y., Hosseinabadi, A. A. R., & Slowik, A. (2021). Improvement of grey wolf optimizer with adaptive middle filter to adjust support vector machine parameters to predict diabetes complications. *Neural Computing and Applications*, 33(22), 15205–15228. <https://doi.org/10.1007/s00521-021-06143-y>
- Kamal, R., Amin, E., Abdelminaam, D. S., & Ismail, R. (2024). A Comprehensive Survey on Meta-Heuristic Algorithms for Feature Selection in High-Dimensional Data: Challenges, Applications, and Future Directions. *4th International Mobile, Intelligent, and Ubiquitous Computing Conference, MIUCC 2024*, 84–90. <https://doi.org/10.1109/MIUCC62295.2024.10783538>
- Khalid, N. H. M., Ismail, A. R., Aziz, N. A., & Hussin, A. A. A. (2023). Performance Comparison of Feature Selection Methods for Prediction in Medical Data. *Communications in Computer and Information Science*, 1771 CCIS, 92–106. [https://doi.org/10.1007/978-981-99-0405-1\\_7](https://doi.org/10.1007/978-981-99-0405-1_7)
- Sadeghian, Z., Akbari, E., Nematzadeh, H., & Motameni, H. (2025). A review of feature selection methods based on meta-heuristic algorithms. *Journal of Experimental and Theoretical Artificial Intelligence*, 37(1), 1–51. <https://doi.org/10.1080/0952813X.2023.2183267>
- Smoljo, T. (2023). Improving quality of life for diabetic patients and delaying cardiovascular complications. *Cardiologia Croatica*, 19(3–4), 153–153. <https://doi.org/10.15836/ccar2024.153>
- Sreehari, E., & Babu, L. D. D. (2024). Critical Factor Analysis for prediction of Diabetes Mellitus using an Inclusive Feature Selection Strategy. *Applied Artificial Intelligence*, 38(1). <https://doi.org/10.1080/08839514.2024.2331919>
- Ulutas, H., Günay, R. B., & Sahin, M. E. (2024). Detecting diabetes in an ensemble model using a unique PSO-GWO hybrid approach to hyperparameter optimization. *Neural Computing and Applications*, 36(29), 18313–18341. <https://doi.org/10.1007/s00521-024-10160-y>
- Yang, X. S. (2009). Firefly algorithms for multimodal optimization. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5792 LNCS, 169–178. [https://doi.org/10.1007/978-3-642-04944-6\\_14](https://doi.org/10.1007/978-3-642-04944-6_14)