Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871

Vol. 6, No. 3, Juni 2025, Page. 1503-1516
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4690

Optimization Of Extreme Learning Machine Models Using Metaheuristic Approaches For Diabetes Classification

Gilang Sulaeman[1], Yohani Setiya Rafika Nur*[2], Adanti Wido Paramadini[3], Dasril Aldo[4], M. Yoka Fathoni[5]

[1,2,4]Department of Informatics, Telkom University, Indonesia
[3]Biomedical Engineering, Telkom University, Indonesia
[5]Information Technology, University Kuala Lumpur, Malaysia

Email: yohanin@telkomuniversity.ac.id

Received : May 6, 2025; Revised : Jun 2, 2025; Accepted : Jun 4, 2025; Published : Jun 23, 2025

Abstract

Proper classification of diabetes is a significant challenge in contemporary healthcare, especially related to early detection and clinical decision support systems. This study aims to optimize the Extreme Learning Machine (ELM) model with a metaheuristic approach to improve performance in diabetes classification. The data used was an open dataset containing the patient's medical attributes, such as age, gender, smoking status, body mass index, blood glucose level, and HbA1c. The initial process includes data cleansing, one-hot coding for categorical features, MinMax normalization, and unbalanced data handling with SMOTE. The ELM model was tested with four activation functions (Sigmoid, ReLU, Tanh, and RBF) each combined with three metaheuristic optimization strategies, namely Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and Bat Algorithm. The results of the evaluation showed that the combination of the Tanh activation function with GA optimization obtained the highest accuracy of 87.98% and an F1-score of 0.5489. Overall, GA optimization appears to be superior to all other measurement configurations in consistent classification performance. The main contribution of this study is to offer a systematic approach to select the best combination of activation functions and optimization algorithms in ELM, as well as to provide empirical evidence to support the application of metaheuristic strategies to improve the accuracy of disease classification based on health data. This research has direct implications for the development of a more precise and data-based medical diagnostic classification system for diabetes.

Keywords : *Activation Function, Diabetes, Extreme Learning Machine, Metaheuristic, SMOTE.*

1. INTRODUCTION

Diabetes mellitus is one example of a chronic disease that is growing exponentially around the world [1]-[3]. The International Diabetes Federation estimates that by 2045, 783 million people will suffer from Diabetes, a significant increase from 537 million in 2021 [4]-[6]. Early intervention is essential in an effort to reduce the risk of long-term complications [7], [8]. This highlights the need for a constitutional automatic classification system based on artificial intelligence in the medical field. In the medical context, where accuracy and speed of diagnosis are crucial, the integration of algorithms such as the Extreme Learning Machine (ELM) can provide an efficient solution for automatically classifying medical data, thus supporting faster and more precise clinical decision-making.

The Extreme Learning Machine (ELM) is one of the highly efficient fast classifier algorithms [9] -[12] The main advantages of ELM are its fast training speed and its simple architecture [13], [14]. ELM performance is often unstable and highly sensitive to initial configurations due to randomly assigned initial weights and biases [15] This drives the need to incorporate optimization techniques aimed at improving the predictive accuracy of ELMs.

Some studies have integrated Extreme Learning Machines (ELMs) with metaheuristic algorithms such as Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and Bat Algorithm to improve classification accuracy. Ding et al. used an ELM optimized with PSO for the classification of syndromes in primary liver cancer, but this study applied only one type of activation function without further exploration of the other functions [16] Liu et al. utilized Genetic Algorithm in combination with ELM for the classification of Android apps, but the study did not address a comparison of various activation functions [17] Fan in 2021 developed a hybrid ELM model optimized with a combination of the Bat Algorithm and Cuckoo Search Algorithm for regression and multiclass classification tasks. This study optimized the input weights and hidden thresholds of the ELM, but did not consider data balancing in the classification process [18] Meanwhile, a study by Liu et al. and Tan (2023–2024) evaluated the performance of activation functions in ELM and proposed improvements to the model structure using the PSO strategy, but neither combined the exploration of activation functions with other optimization algorithms [19], [20].

Of all these studies, there has been no clear attempt to systematically combine multiple activation functions with multiple optimization algorithms in a single experimental framework, especially in unbalanced medical datasets such as diabetes. This study aims to assess the performance of the activation functions of Sigmoid, ReLU, Tanh, and RBF in combination with PSO, GA, and Bat Algorithm metaheuristic algorithms in the ELM model for diabetes classification. The study also applied the Synthetic Minority Oversampling Technique (SMOTE) to address class imbalances in the dataset.

The main contribution of this study is the design and comparison of 12 combinations of ELM-based models with integrated imbalance data characterized by metaheuristic activation and optimization to achieve a more accurate, stable, and reliable diabetes classification system for real-world clinical settings.

2. METHOD

To illustrate the stages of the research, Figure 1 presents the process flow from data processing to model evaluation. Each stage is described systematically to ensure replication and validation of the experimental results.

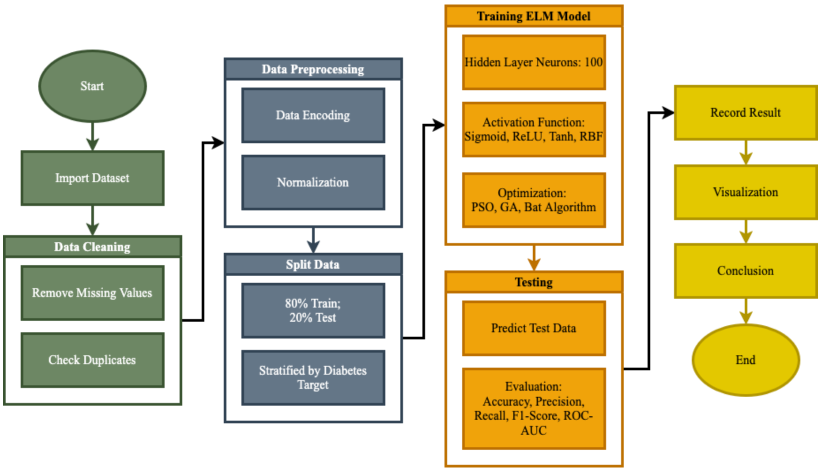

Figure 1. Flow of Research Process

This study aims to evaluate the performance of the Extreme Learning Machine (ELM) model in diabetes risk classification by combining various activation functions and metaheuristic optimization algorithms. The research method can be seen in Figure 1, which illustrates the overall stages of the methodology employed in this study. Each step includes data preprocessing, model training with different activation functions, optimization using metaheuristic algorithms, and performance evaluation.

2.1. Data Import and Cleanup

The dataset used is diabetes_prediction_dataset.csv, which contains medical information such as age, gender, body mass index (BMI), glucose levels, HbA1c, and diabetes status.

Table 1. Diabetes Data

No	X1	X2	X3	X4	X5	X6	X7	X8	Y
1	Female	80	0	1	never	25.19	6.6	140	0
2	Female	54	0	0	No Info	27.32	6.6	80	0
3	Male	28	0	0	never	27.32	5.7	158	0
4	Female	36	0	0	current	23.45	5	155	0
5	Male	76	1	1	current	20.14	4.8	155	0
6	Female	20	0	0	never	27.32	6.6	85	0
7	Female	44	0	0	never	19.31	6.5	200	1
8	Female	79	0	0	No Info	23.86	5.7	85	0
…	…	…	…	…	…	…	…	…	…
99998	Male	66	0	0	former	27.83	5.7	155	0
99999	Female	24	0	0	never	35.42	4	100	0
100000	Female	57	0	0	current	22.43	6.6	90	0

Table 1 shows the first five rows of the dataset used in this study. Each row represents a patient's data, with nine key attributes. Column X1 represents the patient's gender (*gender*), which has been coded in numerical form. X2 is the patient's age in units of years, while X3 expresses smoking *history*, also in the form of numerical codes resulting from one-hot encoding. Furthermore, X4 is a blood pressure value or other clinical parameter that reflects the patient's metabolic status (depending on the original data). X5 to X7 represent body mass index *(BMI),* HbA1c levels, and blood glucose levels, respectively. X8 shows other numerical additional features that have been normalized. Finally, column Y is a binary target label that indicates diabetes status: 1 if the patient is detected to have diabetes, and 0 if not. After the data is successfully imported, the initial process of data cleansing is carried out to ensure optimal input quality. Checking for empty values is performed using the isnull() function on each column, and the results show that the dataset does not contain missing values. Next, the duplicated data is identified using df.duplicated(), and all rows detected as duplicates are immediately deleted so as not to affect the model training process.

2.2. Data Pre-processing

Data pre-processing is an important stage in machine learning to improve data quality, reduce noise, and ensure models can learn optimally from available data [21], [22]. The cleaned data is further processed to be compatible with the ELM model:

a. One-hot encoding is applied to categorical fields such as gender and smoking_history.
b. Normalization uses MinMaxScaler on numerical features such as age, bmi, HbA1c_level, and blood_glucose_level.

2.3. Data Splitting

The data was then divided into 80% for training and 20% for testing using Stratified Split based on diabetes labels. Because the dataset has an unbalanced distribution of classes, an oversampling process is carried out with SMOTE (Synthetic Minority Over-sampling Technique) to balance the classes on the training data [23].

2.4. ELM Model Training

The model used in this study is Extreme Learning Machine (ELM), which is a feedforward neural network-based supervised learning algorithm with one hidden layer [24], [25]. The main advantage of ELM is its high training speed because the learning process is only carried out once (non-iterative) through linear solutions [26], [27]. In this study, the number of neurons in the hidden layer was set to 100 neurons for all experiments.

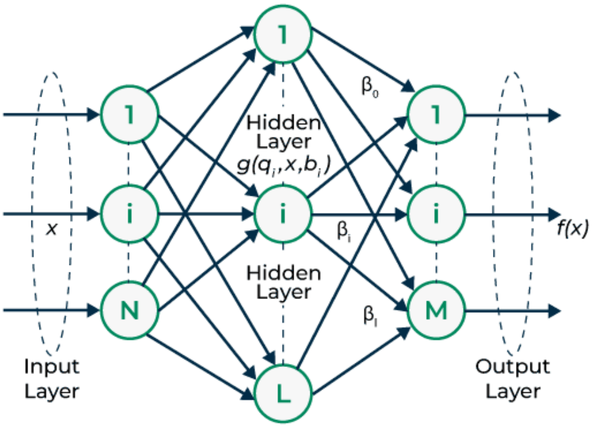

Figure 2. ELM Architecture

Figure 2 illustrates the general architecture of an ELM network. The model consists of three main components:

a. An input layer with N units representing the input features.
b. A hidden layer with L neurons and a specific activation function.
c. An output layer that provides the final classification result.

The ELM model is implemented with a single hidden layer comprising 100 neurons. The activation functions used include those described in [28], [29], [30], [31]:

1. Sigmoid:

$$g(x) = \frac{1}{1+e^{-x}} \qquad (1)$$

2. ReLu (Rectified LinearUnit):

$$g(x) = max(0, x) \qquad (2)$$

3. Tanh (Hyperbolic Tangent):

$$g(x) = \tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \qquad (3)$$

4. RBF (Radial Basis Function)

$$g(x) = \exp\left(-\gamma\left|\left|x - c\right|\right|^{2}\right) \qquad (4)$$

Where c is the center and γ is the shape parameter. After the activity is computed, the output weight β is obtained using the formula:

$$\beta = H'T \qquad (5)$$

where:

a. H: hidden layer output matrix,

Vol. 6, No. 3, Juni 2025, Page. 1503-1516
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.3.4690

b. T: target,
c. H': moore-penrose pseudoinvers of H.

To enhance prediction performance and accuracy, the input weights and hidden layer biases are not randomly initialized but instead optimized using three metaheuristic algorithms [32], [33], [34]: Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Bat Algorithm.

To support the implementation of this study, several computational tools were employed. The entire modeling and evaluation process was conducted using Python programming language. Key libraries included NumPy and Pandas for data manipulation, scikit-learn (sklearn) for preprocessing and evaluation metrics, and custom-built modules for implementing the ELM architecture and the metaheuristic algorithms. Visualization of results, such as performance comparisons and convergence graphs, was done using Matplotlib and Seaborn.

2.5. Testing and Evaluation

After the model training is completed, testing is performed on the test dataset. The evaluation is conducted using five key metrics [35], [36]:

1. Accuracy:

$$Accuracy = \frac{TP + TN}{TP+TN+FP+FN} \quad (6)$$

2. Precision:

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

3. Recall (Sensitivity):

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

4. F1-Score:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision+Recall} \quad (9)$$

5. ROC-AUC

$$\text{ROC} - \text{AUC} = \int_0^1 TPR(FPR)dFPR \quad (10)$$

2.6. Visualization and Result Storage

The results of all experiment combinations (4 activation functions × 3 optimization algorithms = 12 configurations) are stored and visualized. The confusion matrix and ROC curve were used to show the model's ability to classify diabetes cases.

3. RESULT

Experiments were conducted to evaluate the performance of 12 ELM model combinations based on variations in activation functions and metaheuristic optimization algorithms. The evaluation results of each model combination are shown in Table 2, which includes the accuracy, precision, recall, F1-score, and ROC-AUC metrics.

Table 2. ELM Model Evaluation Results

No	Activation	Optimization	Accuracy	Precision	Recall	F1-Score	ROC-AUC
1	Sigmoid	PSO	0.8764	0.3961	0.8641	0.5432	0.9517
2	Sigmoid	GA	0.8780	0.3994	0.8629	0.5461	0.9514
3	Sigmoid	Bat	0.8717	0.3846	0.8488	0.5293	0.9441
4	ReLU	PSO	0.8740	0.3922	0.8759	0.5418	0.9521

No	Activation	Optimization	Accuracy	Precision	Recall	F1-Score	ROC-AUC
5	ReLU	GA	0.8783	0.3999	0.8612	0.5462	0.9507
6	ReLU	Bat	0.8770	0.3975	0.8659	0.5449	0.9505
7	Tanh	PSO	0.8788	0.4009	0.8612	0.5471	0.9505
8	Tanh	GA	0.8798	0.4029	0.8606	0.5489	0.9503
9	Tanh	Bat	0.8531	0.3485	0.8376	0.4922	0.9339
10	RBF	PSO	0.8769	0.3975	0.8694	0.5456	0.9508
11	RBF	GA	0.8663	0.3763	0.8718	0.5257	0.9501
12	RBF	Bat	0.7649	0.1912	0.5465	0.2832	0.7391

In general, the model with the Tanh activation function showed quite good performance on all optimization methods, where the combination of Tanh + GA obtained the highest accuracy of 0.8798 and an F1-score of 0.54889. Other combinations, such as ReLU + GA and Sigmoid + GA, also showed good results and consistency was shown with a ROC-AUC value of over 0.95. In contrast to this, the model with RBF activation was not good at predicting its energy positives, especially in combination with the Bat Algorithm, where the accuracy obtained was only 0.7649 and the lowest F1-score was 0.2832. These findings show that BPF RBF is controlled by ELM settings and the data faced is slightly detrimental to the activation of RBF in this method.

To complement the tabular evaluation results, a heatmap visualization is provided in Figure 3, displaying the F1-Score performance for all combinations of activation functions and optimization algorithms. This visual aid allows for rapid identification of top-performing configurations and highlights the consistency of certain methods. As shown in the heatmap, the Tanh + GA, Tanh + PSO, and ReLU + GA combinations achieved the highest F1-Scores, confirming their superior classification capability in the context of diabetes prediction using ELM.

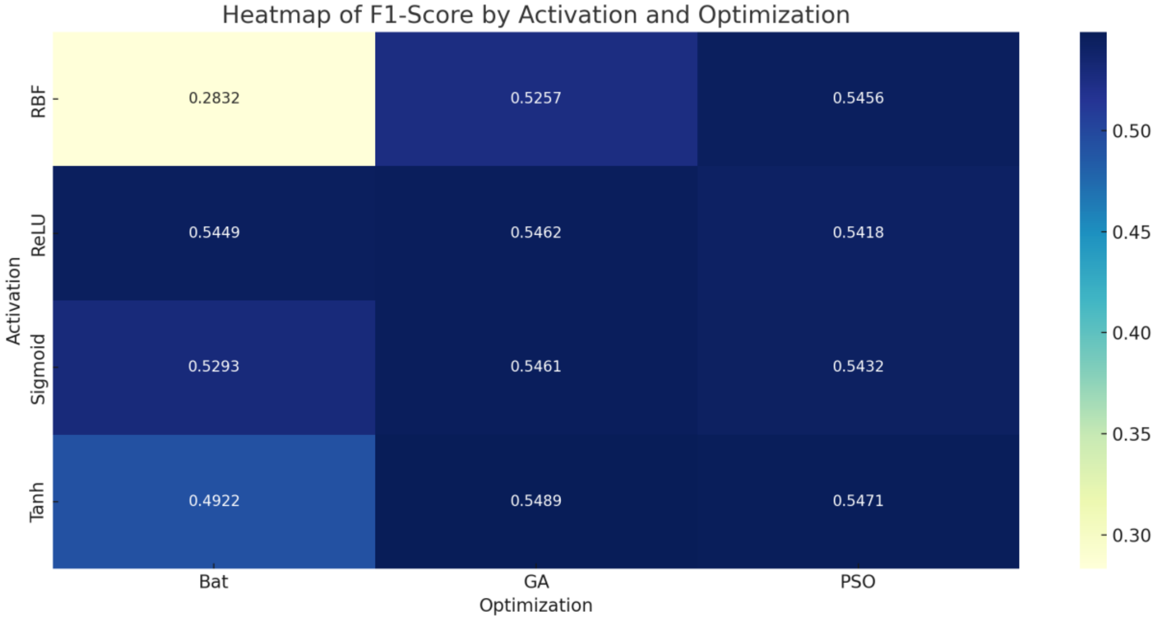

Figure 3. Heatmap Visualization Of F1-Score Across Activation Functions And Optimization Algorithm Combinations In ELM.

From the recall, it seems that currently all models can classify diabetes cases into positive classes. However, the decrease in precision values indicates that there are so many false positives. Many of these errors may be caused by class imbalance even though SMOTE has been carried out. In this case, the activation of Tanh and Sigmoid is better in controlling ni Sensitivity and Accuracy compared to RBF. To further analyze the performance of all optimization combinations on a given activation function, an evaluation visualization of each activation function is performed. In Figure 4, the performance of the ELM model with Sigmoid activation is shown in three optimization methods, namely PSO, GA, and Bat Algorithm, and using five main evaluation metrics.

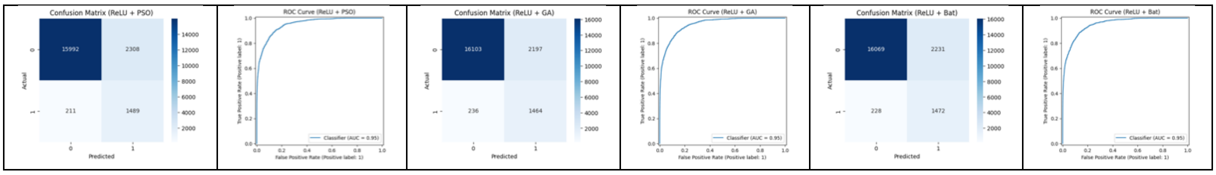

Figure 4. Evaluation Results of Sigmoid Activation Function

From the evaluation results, the performance of Sigmoid activation function was found to be good and stable across the three tested optimization algorithms. Accuracy value ranged from 0.8717 to 0.8780, with ROC-AUC score all above 0.94, indicating strong model discrimination. Best combination within these configurations was Sigmoid + GA achieving 0.8780 of accuracy and 0.5461 of F1-score which indicates fair balance of sensitivity and specificity. Subsequently, to assess the contribution of ReLU activation function, similar analysis was conducted on three optimization combinations (PSO, GA, and Bat Algorithm) as shown in Figure 5.

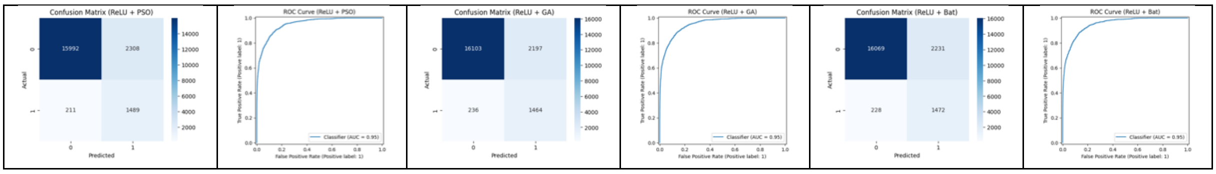

Figure 5. Evaluation Results of ReLU Activation Function

The application of the ReLU activation function results is quite impressive, especially in the combination of ReLU + GA where accuracy achieved was 0.8780 with F1-score 0.5456 and consistently ROC-AUC over 0.95 for all optimization methods. This indicates that ReLU has good pattern recognition capability for classification, although performance slightly deteriorates with Bat Algorithm combinations. For comparison purposes, Tanh activation function was studied further as its smooth and bounded nature often contributes to model stability and generalization. We present visualizations of the results below.

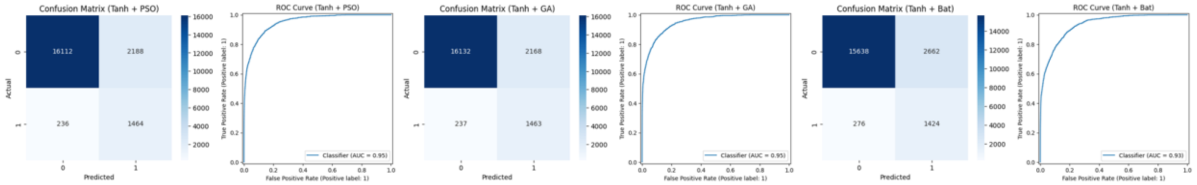

Figure 6. Evaluation Results of Tanh Activation Function

The Tanh activation function performed consistently across all optimization combinations. The best combination was found on Tanh + GA where accuracy was 0.8798 with F1-score of 0.5489 making it one of the most balanced configurations between accuracy, sensitivity, and precision. High values of ROC-AUC rank across all variants demonstrate that Tanh activation has the capacity to support ELM in class discrimination even with data that was previously imbalanced. As a concluding part of the analysis the RBF activation function was also tested to assess how well it could handle the non-linear complexity of the ELM model.

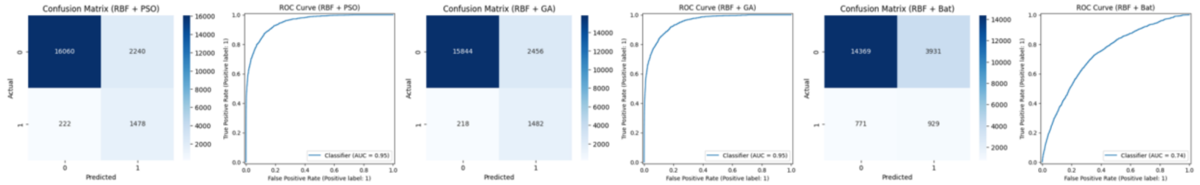

Figure 7. Evaluation Results of RBF Activation Function

RBF activation functions tend to be outperformed relative to other activation functions. Its best combination, RBF + GA, achieved an accuracy of 0.8733 and F1 score of 0.5366, still under the optimal Tanh and ReLU results. Particularly with RBF + Bat Algorithm, performance suffered sharply with accuracy dropping to 0.7649, F1 score of 0.2832, and ROC-AUC of 0.74, the worst of all the experiments. These results indicate that RBF is limited in adapting to the architectural complexity of the ELM model for diabetes classification.

Evaluating the performance of the twelve combinations of activation functions and optimization algorithms on the Extreme Learning Machine (ELM) model revealed very different results in accuracy, precision, recall, F1-score, and ROC-AUC. Overall, the best results came from using Tanh activation and Genetic Algorithm (GA) as optimization algorithm. The model achieved an accuracy of 0.8798, F1 score of 0.5489, and ROC-AUC 0.9503, indicating robust and precise classification performance.

Combining sigmoid and GA optimizations also produced powerful performers along with ReLU + GA, demonstrating accuracy values surpassing 0.878 and F1-scores approaching 0.55. The three also had ROC-AUCs above 0.95, indicating that the model could reliably distinguish between diabetes and non-diabetes patients across various classification thresholds. This suggests that the application of Genetic Algorithm as an optimization strategy consistently improves ELM performance regardless of the activation function employed. In general, Tanh and ReLU activation functions volatilely outperformed Sigmoid, which often exhibited erratic dynamics in precision and recall, dominating the stability of evaluation metrics.

On the opposite end, Radial Basis Function (RBF) activations seem to underperform. The RBF + Bat Algorithm combination recorded the worst result, achieving only an accuracy of 0.7649, F1-score of 0.2832, and ROC-AUC of 0.7391, illustrating the model's weak ability in accurately identifying the minor class. This is likely due to the local nature of RBF activation that is poorly suited for ELM's global mapping design, in addition to the Bat Algorithm‘s limited performance when attempting to explore optimal solutions within complex search spaces.

These findings confirm that the effectiveness of ELM does not depend only on the number of neurons or the underlying architecture of the ELM, but on the specific optimization system and activation functions that are applied. In the classification of diabetes based on health data, the most optimal combination is ELM with Tanh activation function and GA optimization, The best consideration in the optimization of sensitivity (recall), precision, and generalization ability (ROC-AUC) is obtained here. These findings contribute to the development of machine learning-based models for efficient and accurate early detection of diabetes.

4. DISCUSSIONS

This research demonstrates that the combination of activation function and optimization algorithm significantly contributes to the performance of the Extreme Learning Machine (ELM) model in diabetes classification. The Tanh + GA combination provided the best result consistently across multiple metrics including accuracy, F1, and ROC-AUC. This is in agreement with the findings in [37] which showed that the incorporation of a Genetic Algorithm into ELM increased the stability and the overall generalization capability of the model for data classification tasks, including those relevant to medical applications. Also, ReLU + GA and Sigmoid + GA combinations demonstrated high-performed accuracy above 0.878 and ROC-AUC above 0.95. This supports the study conducted by [38] which reported that the ReLU activation function tends to produce faster and more stable convergence during non-linear modelos Non-linear ELMs. The reason for this dip in performance could be attributed to the local nature of the RBF activations in mapping global data distributions. These findings support [39] who pointed out that RBF functions are better suited for local classification tasks but are ineffective for global classifications, such as a data-table-driven diagnostic medicine. This research extends the application of metaheuristic-integrated neural networks in medical informatics, particularly for imbalanced health datasets, by demonstrating how activation function and optimization selection in ELM can directly improve predictive accuracy in clinical classification tasks.

It aligns with the findings of [40], which stated that Genetic Algorithm is efficient in solving non-convex optimization problems with high stability and accuracy on classification, especially in model structures which are sensitive to weight initialization like ELM. This is further corroborated by the study of [41] which advocates for the ensemble or hybrid approach for better classification accuracy on the imbalanced medical datasets as well as overcoming the SMOTE limitation of generating high positive false rate. Thus, these experiment results not only reinforce the findings from previous research but also contribute practically by aiding the selection of the optimal ELM configuration for disease classification, particularly in the early detection of diabetes using patients' past health records.

Despite yielding promising results, this study is not without limitations. First, all experiments were conducted using a single open-access dataset, which does not guarantee the model's generalizability to data from other populations. Second, the application of SMOTE for data balancing presents the risk of producing non-representative synthetic samples, particularly for the severely restricted minority class. Moreover, this study has yet to test the model's sensitivity to variations in the number of neurons in the hidden layer or the parameters of the optimizing algorithm. Further evaluation using statistical testing or cross-validation also remains incomplete, presenting opportunities for improvement in future studies.

5. CONCLUSION

This research evaluates theperformance of Extreme Learning Machine (ELM) model with different activation functions combinations (Sigmoid, ReLU, Tanh, and RBF) and metaheuristic optimization algorithms (Genetic Algorithm, Particle Swarm Optimization, and Bat Algorithm) for diabetes classification. The experiments conducted proved that the combination of Tanh + GA performed best achieving an accuracy of 0.8798, F1 score of 0.5489, and ROC-AUC of 0.9503, outperforming other configurations in polarization stratification balance. Genetic Algorithm (GA) proved to be the most consistent optimization method, significantly increasing accuracy in all but two activation functions. The Tanh and ReLU activation functions also demonstrated robust stability in supporting classificaion performance; RBF, on the other hand, showed the poorest results, most notably when paired with Bat Algorithm. These findings reinforce that the choice of combination of the activation function and optimization algorithm is essential for improving the accuracy and generalization of ELM model, especially in disease classification. This study could be used as the basis for developing more accurate systems for early detection of diabetes, while also being expanded upon for other applications in medical classification. For further work, this research could be extended by testing the model's sensitivity to its parameters, cross-validation, and testing with other medical datasets to improve generalizability.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper.

ACKNOWLEDGEMENT

The author expressed his gratitude to the Institute for Research and Community Service (LPPM) Telkom University for the support and facilities provided so that this research could be carried out properly.

REFERENCES

[1] J. S and U. H. R., ‘A Comparative Study Of Diabetes Trends: India And Selected Countries’, *Global Journal For Research Analysis*, pp. 50–54, Feb. 2025, doi: 10.36106/gjra/5404887.

[2] K. Manglani *et al.*, ‘Correlation of Leptin in Patients With Type 2 Diabetes Mellitus’, *Cureus*, Apr. 2024, doi: 10.7759/cureus.57667.

[3] L. S. Satin, S. A. Soleimanpour, and E. M. Walker, ‘New Aspects of Diabetes Research and Therapeutic Development’, *Pharmacological Reviews*, vol. 73, no. 3, pp. 1001–1015, Jul. 2021, doi: 10.1124/pharmrev.120.000160.

[4] S. Akhtar, ‘Diabetes-induced peripheral neuropathy: Is prescribing physical exercise the answer?’, *Biomol Biomed*, vol. 24, no. 3, pp. 436–439, May 2024, doi: 10.17305/bb.2023.10188.

[5] J. Patel, N. Null, N. Null, N. Null, N. Null, and N. Null, ‘Diabetes, A Global Epidemic’, *SACAD*, 2023, doi: 10.58809/HBJV7965.

[6] A. Kumar, R. Gangwar, A. Ahmad Zargar, R. Kumar, and A. Sharma, ‘Prevalence of Diabetes in India: A Review of IDF Diabetes Atlas 10thEdition’, *CDR*, vol. 20, no. 1, p. e130423215752, Jan. 2024, doi: 10.2174/1573399819666230413094200.

[7] X. An *et al.*, ‘Early effective intervention can significantly reduce all-cause mortality in prediabetic patients: a systematic review and meta-analysis based on high-quality clinical studies’, *Front. Endocrinol.*, vol. 15, p. 1294819, Mar. 2024, doi: 10.3389/fendo.2024.1294819.

[8] K. C. Alaga and J. Haddad, ‘Lifestyle intervention is superior to metformin in prevention of microvascular complications of diabetes in women, but socioeconomic barriers may lead to use of combination’, *Clin Res Pract*, vol. 7, no. 2, p. crp.eP2626, Oct. 2021, doi: 10.22237/crp/1625098080.

[9] B. P. Ooi, N. A. Rahim, M. J. Masnan, and A. Zakaria, ‘A study of extreme learning machine on small sample-sized classification problems’, *J. Phys.: Conf. Ser.*, vol. 2107, no. 1, p. 012013, Nov. 2021, doi: 10.1088/1742-6596/2107/1/012013.

[10] Y. Zhang *et al.*, ‘An Improved Extreme Learning Machine (ELM) Algorithm for Intent Recognition of Transfemoral Amputees With Powered Knee Prosthesis’, *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 32, pp. 1757–1766, 2024, doi: 10.1109/TNSRE.2024.3394618.

[11] T. Chouikh, H. Hellara, R. Barioul, F. Sandid, and O. Kanoun, ‘Comparative Study of Optimization Methods Proposed for the Extreme Learning Machine (ELM)’, in *2021 International Workshop on Impedance Spectroscopy (IWIS)*, Chemnitz, Germany: IEEE, Sep. 2021, pp. 137–140. doi: 10.1109/IWIS54661.2021.9711902.

[12] P. Sun and L. Yang, ‘Generalized eigenvalue extreme learning machine for classification’, *Appl Intell*, vol. 52, no. 6, pp. 6662–6691, Apr. 2022, doi: 10.1007/s10489-021-02654-2.

[13] L. Tobar-Valenzuela, M. Mora, F. Silva-Pavez, I. Torres-Gonzalez, and P. Barria-Valdebenito, ‘Fast Tuning of Extreme Learning Machine Neural Networks based with Simple Optimization Algorithms’, in *2022 IEEE International Conference on Automation/XXV Congress of the Chilean Association of Automatic Control (ICA-ACCA)*, Curicó, Chile: IEEE, Oct. 2022, pp. 1–5. doi: 10.1109/ICA-ACCA56767.2022.10005961.

[14] J. Wang, S. Lu, S. Wang, and Y. Zhang, ‘A review on extreme learning machine’, *Multimedia Tools and Applications*, vol. 81, pp. 41611–41660, 2021, doi: 10.1007/s11042-021-11007-7.

[15] A. Al-Btoush, M. Fernandez-Delgado, E. Cernadas, and S. Barro, ‘Extreme learning machine with confidence interval based bias initialization’, in *2021 Second International Conference on Intelligent Data Science Technologies and Applications (IDSTA)*, Tartu, Estonia: IEEE, Nov. 2021, pp. 23–30. doi: 10.1109/IDSTA53674.2021.9660822.

[16] L. Ding, X. Zhang, D. Wu, and M. Liu, ‘Application of an extreme learning machine network with particle swarm optimization in syndrome classification of primary liver cancer’, *Journal of Integrative Medicine*, vol. 19, no. 5, pp. 395–407, Sep. 2021, doi: 10.1016/j.joim.2021.08.001.

[17] R. Gupta, A. Agarwal, D. Dua, and A. Yadav, ‘Android Malware Detection Using Extreme Learning Machine Optimized with Swarm Intelligence’, in *Cyber Security and Digital Forensics*, vol. 73, K. Khanna, V. V. Estrela, and J. J. P. C. Rodrigues, Eds., in Lecture Notes on Data Engineering and Communications Technologies, vol. 73. , Singapore: Springer Singapore, 2022, pp. 31–43. doi: 10.1007/978-981-16-3961-6_4.

[18] Q. Fan and T. Fan, ‘A Hybrid Model of Extreme Learning Machine Based on Bat and Cuckoo Search Algorithm for Regression and Multiclass Classification’, *Journal of Mathematics*, vol. 2021, pp. 1–11, Nov. 2021, doi: 10.1155/2021/4404088.

[19] R. Liu, L. Wei, and P. Zhang, ‘Extreme Learning Machine Classifier Based on Novel Particle Swarm Optimization Algorithm’, in *2023 38th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, Hefei, China: IEEE, Aug. 2023, pp. 1005–1009. doi: 10.1109/YAC59482.2023.10401835.

[20] K. Tan, ‘Extreme Learning Machine based on Improved Multi-Objective Particle Swarm Optimization’, in *2024 5th International Conference on Computer Engineering and Application (ICCEA)*, Hangzhou, China: IEEE, Apr. 2024, pp. 333–337. doi: 10.1109/ICCEA62105.2024.10603496.

[21] J. Deng, ‘The Data Preprocessing Technique in Machine Learning’, *TE*, vol. 1, no. 10, Dec. 2024, doi: 10.61173/0hfqqz07.

[22] I. Chahid, A. K. Elmiad, and M. Badaoui, ‘Data Preprocessing For Machine Learning Applications in Healthcare: A Review’, in *2023 14th International Conference on Intelligent Systems: Theories and Applications (SITA)*, Casablanca, Morocco: IEEE, Nov. 2023, pp. 1–6. doi: 10.1109/SITA60746.2023.10373591.

[23] N. Sigeef, ‘An Oversampling Algorithm combining SMOTE and RF for Imbalanced Medical Data’, *IJRASET*, vol. 11, no. 6, pp. 2429–2434, Jun. 2023, doi: 10.22214/ijraset.2023.54074.

[24] Y. Lv, B. Li, J. Yu, and Y. Ding, ‘Reduce Training Error of Extreme Learning Machine by Selecting Appropriate Hidden Layer Output Matrix’, *J. Syst. Sci. Syst. Eng.*, vol. 30, no. 5, pp. 552–571, Oct. 2021, doi: 10.1007/s11518-021-5502-8.

[25] R. Kaloni, T. R. Nayak, M. Sankhe, and Mr. G. Wakure, ‘Optimized Extreme Learning Machine’, *IJRASET*, vol. 10, no. 4, pp. 1334–1345, Apr. 2022, doi: 10.22214/ijraset.2022.41514.

[26] X. Liu, G. Zhou, Y. Zhou, and Q. Luo, ‘Functional extreme learning machine’, *Front. Comput. Neurosci.*, vol. 17, p. 1209372, Jul. 2023, doi: 10.3389/fncom.2023.1209372.

[27] A. Akusok, L. Espinosa-Leal, and K.-M. BjöRk, ‘Sharing non-reversible data statistics for fast and secure Federated learning native to Extreme Learning Machine’, in *Proceedings of the 17th International Conference on PErvasive Technologies Related to Assistive Environments*, Crete Greece: ACM, Jun. 2024, pp. 674–675. doi: 10.1145/3652037.3663891.

[28] Monisha. P, A. ArunRaja, and J. Judeson Antony Kovilpillai, ‘Structural Health Monitoring Using Machine Learning Techniques’, in *2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN)*, Salem, India: IEEE, Jun. 2023, pp. 275–282. doi: 10.1109/ICPCSN58827.2023.00052.

[29] A. D. Roshan, P. Guha, and G. Trivedi, ‘Hardware-Optimized Regression Tree-Based Sigmoid and Tanh Functions for Machine Learning Applications’, *IEEE Trans. Circuits Syst. II*, vol. 72, no. 1, pp. 283–287, Jan. 2025, doi: 10.1109/TCSII.2024.3485493.

[30] O. Pantalé, ‘Comparing Activation Functions in Machine Learning for Finite Element Simulations in Thermomechanical Forming’, *Algorithms*, vol. 16, no. 12, p. 537, Nov. 2023, doi: 10.3390/a16120537.

[31] R. Çekik, ‘Categorization of Turkish Text Documents Using Extreme Learning Machine’, in *2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP)*, Malatya, Turkiye: IEEE, Sep. 2024, pp. 1–7. doi: 10.1109/IDAP64064.2024.10710721.

[32] S. Manjula, N. Hema Rajini, and K. Chokkanathan, ‘Optimizing Chronic Kidney Disease Diagnosis: A Comprehensive Review of Metaheuristic Algorithms and Their Applications in Detection and Classification’, in *2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, India: IEEE, Nov. 2024, pp. 1654–1660. doi: 10.1109/ICECA63461.2024.10801087.

[33] C. Kaya, Z. H. Kilimci, M. Uysal, and M. Kaya, ‘A Review of Metaheuristic Optimization Techniques in Text Classification’, *IJCESEN*, vol. 10, no. 2, Apr. 2024, doi: 10.22399/ijcesen.295.

[34] G. Phatai and T. Luangrungruang, ‘A Comparative Study of Hybrid Neural Network with Metaheuristics for Student Performance Classification’, in *2023 11th International Conference on Information and Education Technology (ICIET)*, Fujisawa, Japan: IEEE, Mar. 2023, pp. 448–452. doi: 10.1109/ICIET56899.2023.10111495.

[35] K. Takahashi, K. Yamamoto, A. Kuchiba, and T. Koyama, ‘Confidence interval for micro-averaged F1 and macro-averaged F1 scores’, *Appl Intell*, vol. 52, no. 5, pp. 4961–4972, Mar. 2022, doi: 10.1007/s10489-021-02635-5.

[36] B. Shannaq, O. Ali, S. A. Maqbali, and A. Al-Zeidi, ‘Advancing user classification models: A comparative analysis of machine learning approaches to enhance faculty password policies at the University of Buraimi’, *J. Infras. Policy. Dev.*, vol. 8, no. 13, p. 9311, Nov. 2024, doi: 10.24294/jipd9311.

[37] L. Bai, H. Li, W. Gao, and J. Xie, ‘A cooperative genetic algorithm based on extreme learning machine for data classification’, *Soft Comput*, vol. 26, no. 17, pp. 8585–8601, Sep. 2022, doi: 10.1007/s00500-022-07202-9.

[38] Nur Azizah Janad, Junaidi, and Iman Setiawan, ‘Application Of The Extreme Learning Machine (Elm) Method In Predicting The Combined Stock Price Index (Ihsg) In Indonesia’, *Tadulako Soc. Sci. Hum. J.*, vol. 2, no. 2, pp. 22–34, Jul. 2022, doi: 10.22487/sochum.v2i2.17382.

[39] R. Siouda, M. Nemissi, and H. Seridi, ‘Diverse activation functions based-hybrid RBF-ELM neural network for medical classification’, *Evol. Intel.*, vol. 17, no. 2, pp. 829–845, Apr. 2024, doi: 10.1007/s12065-022-00758-3.

[40] A. Nozaripour and H. Soltanizadeh, ‘Sparse Representation Based on Modified Genetic Algorithm for Classification’, *Iran J Sci Technol Trans Electr Eng*, vol. 46, no. 3, pp. 819–827, Sep. 2022, doi: 10.1007/s40998-022-00499-0.

[41] K. Kim, ‘Noise Avoidance SMOTE in Ensemble Learning for Imbalanced Data’, *IEEE Access*, vol. 9, pp. 143250–143265, 2021, doi: 10.1109/ACCESS.2021.3120738.