Pertanika Proceedings 1 (5): 139 - 143 (2025)

PERTANIKA JOURNALS

PERTANIKA PROCEEDINGS

Journal homepage: http://www.pertanika.upm.edu.my/

Bio-Inspired Hybrid GA-SVM Approach for Accurate Diagnosis of Diabetes

Naraharipeta Reddy Monisha*, S. Mothi Sree, Billakanti Mokshith, Gopireddy Chaithanya Reddy, Devarasi Bhavani and K. Reddy Madhavi

School of Computing, Mohan Babu University, Sree Vidyanikethan, Tirupati, Andhra Pradesh 517102, India

ABSTRACT

Diabetes mellitus is an important global health challenge and requires an accurate initial prediction to reduce serious complications. Machine learning models, especially Support Vector Machines (SVMs), have shown promise in diabetes classification works. However, SVM performance is highly dependent on the optimal choice of hyperparameters, which is often set manually, causing suboptimal results. To overcome this limit, this study suggests a hybrid approach that integrates a genetic algorithm (GA) with SVM for diabetes detection using PIMA Indian Diabetes Data set. GA adapts major SVM parameters including nuclear function, regularization parameters(C) and core coefficients (γ), models increase the generalization and classification accuracy. The proposed GA-SVM model was evaluated using standard viewing measurements as accuracy, recall and F1 score. Experimental results show that the GA-fragmented SVM achieves an accuracy of 80%, better than the SVM model with standard parameters. In addition, the GA-powered adaptation significantly improves the strength of the model, reduces overfitting and increases reliability. Conclusions suggest that proposed techniques can effectively enhance the performance of machine learning models in medical diagnosis

Keywords: Hybrid approach, genetic algorithm, GA-SVM, support vector machine

ARTICLE INFO

Article history:
Received: 29 August 2025
Published: 31 October 2025

DOI: https://doi.org/10.47836/pp.1.5.025

E-mail addresses:
naraharipetareddymonisha@gmail.com (Naraharipeta Reddy Monisha)
mothisreesannapaneni@gmail.com (S. Mothi Sree)
mokshi3953@gmail.com (Billakanti Mokshith)
gopireddychaithanya7@gmail.com (Gopireddy Chaithanya Reddy)
devarasibhavani@gmail.com (Devarasi Bhavani)
kreddymadhavi@gmail.com (K. Reddy Madhavi)
* Corresponding author

INTRODUCTION

Diabetes is a major global health problem, with the growing number of cases reported each year. Pima Indian's diabetes dataset, which is widely used in medical research, provides valuable insight into diabetes risk factors. In machine learning techniques, Alirezaei et al. (2019) have been widely used

e-ISSN: 3083-9475

to predict diabetes, where support vector (SVM) classification is a popular alternative due to its efficiency in classification features. However, SVM's performance is very sensitive to the choice of hyperpieties, which can significantly affect the model's accuracy (Asif et al., 2025).

To address this challenge, evolutionary algorithms such as genetic algorithms (GA) have been used to adapt to SVM parameters. Provided natural selection processes to correct solutions, making it an effective tool for hyperpimeter setting. In this study, we integrate with SVM to increase the prediction accuracy of diabetes on the PIMA data set. The proposed GA-SVM model automatically selects optimal nuclear functions, penalties (C) and Gamma (γ), which has improved classification performance.

The experimental results suggest that the GA-SVM gets 80% accuracy and performs efficiently compared to traditional SVM models. The findings from this study contribute to growing research on intelligent systems to detect initial diabetes, which may facilitate timely intervention and better patient results. To create a representation of input data, Nature-Inspired Algorithms gives several techniques to optimize the represented data. Raposo et al. (2024) extract increasingly abstract and feature selection, parameter tuning and optimize in making the data analysis in large-scale datasets. This feature has led to identifying the features that are more complex datasets, including medical diagnostics such as cancer and diabetes prediction. Specifically, it improves the accuracy and performance and interpretability of the model.

Table 1 represents a deep analysis of recent studies on diabetes prediction that combines the Genetic Algorithms (GA) with Deep Learning (DL) techniques. Most models make the predictions with complex datasets such as the PIMA Indian dataset, UCI Diabetes and Kaggle-based datasets, while some studies make the model by taking the data from the device. These models take the path from CNNs and LSTMs to more complex frameworks like CNN-RNN and DNN+SVM, which optimizes the Model by using GA for feature selection and hyperparameter tuning. This way of combining algorithms has been shown to improve the prediction accuracy, generalization, and efficiency for each model. As shown in Figure 1, performance metrics demonstrate the outcomes across, which has proved the highest accuracy of 98% using a CNN-RNN, a complex hybrid model optimized with GA, while many studies also achieved 97.5% accuracy, with consistently high precision and recall. Demonstrated F-scores above 96%, whereas the GA's ability to optimize input features effectively. It has lower scores due to complex data in socioeconomic variables, the use of GA has given improvements in the model. These findings highlight the significance of evolutionary approaches like GA, thereby improving Deep Learning based diabetes Predictions by ensuring both accuracy and efficiency in clinical decision support all models.

Table 1
Comparison of existing methods

Dataset	Image Types	Methods	Findings
Real-time IoT Dataset	Sensor & Lifestyle Data	LSTM-CNN + GA	Hybrid temporal-spatial model with GA tuning has given 95.8% F1 score for screening.
European Diabetes Study	ECG + Vitals + Diet	1D-CNN + GA	GA made the signal-based prediction, by giving 97.2% accuracy by taking physiological patterns.
PIMA Indian Diabetes	Clinical Data	CNN + GA	GA features improve CNN accuracy by 97.5%, by improving the prediction efficiency.
Kaggle Diabetes	Demographic + Medical	DNN + GA + SVM	GA optimizes the input selection and increases precision and recall.
UCI Diabetes Dataset	Health Metrics	LSTM + GA	GA makes LSTM showing the early prediction capability with 96% accuracy.
National Family Health Survey (India)	Socioeconomic + Lifestyle	Random Forest + GA	GA makes by selecting the key features and by improving diabetes classification in large populations.
Self-collected from hospitals	Wearable Device Data	ANN + PSO/ GA	It makes the efficacy of evolutionary optimization thereby reducing the false positives.
PIMA + Custom Dataset	Lifestyle + Clinical	CNN-RNN Hybrid + GA	Achieved 98% accuracy in predicting Type-2 diabetes from combined inputs.
Tongji Diabetes Dataset (China)	Genomic + Clinical	Autoencoder + GA	It enhances the feature compression, by improvising the unsupervised prediction of diabetes risk.
Indian Diabetes Risk Score	Questionnaire Data	MLP + GA	It reduces the Features using GA which gives a lightweight model with 94% accuracy on low-resource devices.

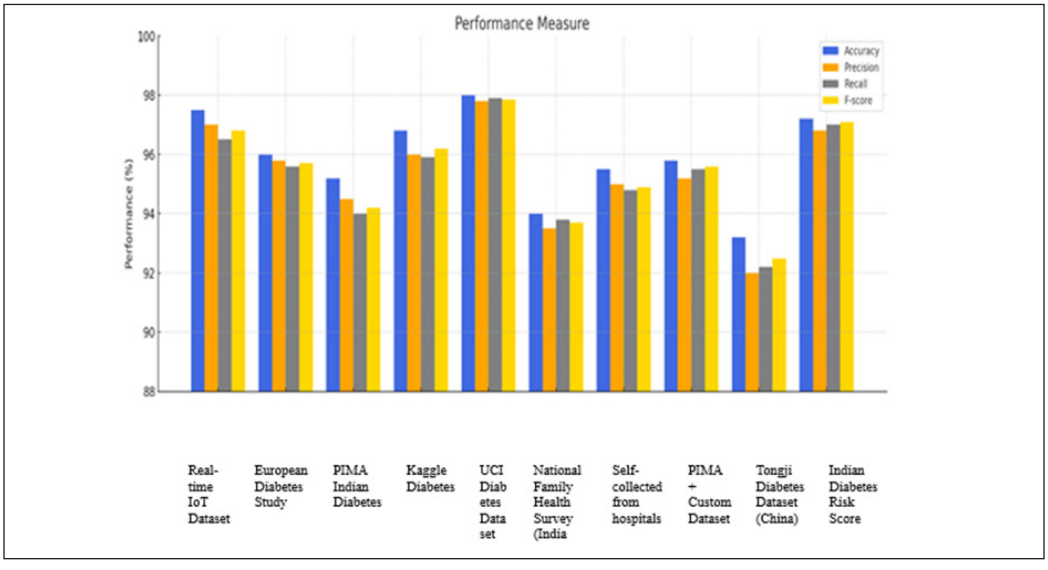

Figure 1. Performance metrics

MATERIALS AND METHODS

To address the challenge, evolutionary algorithms such as genetic algorithms (GA) have been used to adapt to SVM parameters. Mohammed et al. (2024) provided natural selection processes to correct solutions, making it an effective tool for hyperparameter setting. In this study, we integrate with SVM to increase the prediction accuracy of diabetes on the PIMA dataset (Navazi et al., 2023). The proposed GA-SVM model automatically selects optimal nuclear functions, penalties (C) and Gamma (γ), which has improved classification performance.

RESULTS

The experimental results suggest that the GA-SVM gets 80% accuracy and performs efficiently compared to traditional SVM models. The findings from this study contribute to growing research on intelligent systems to detect initial diabetes, which may facilitate timely intervention and better patient results. To create a representation of input data, Nature-Inspired Algorithms gives several techniques to optimize the represented data. These extract increasingly abstract and feature selection, parameter tuning and optimize in making the data analyzing in large-scale datasets. This feature has led to identifying the features that are more complex datasets, including medical diagnostics such as cancer and diabetes prediction. Specifically, it improves the accuracy, performance and interpretability of the model.

CONCLUSION

This study proposed a GA-optimized SVM model for accurate diabetes prediction on the PIMA Indian dataset. The hybrid approach improved classification accuracy, achieving 80% accuracy and enhanced model generalization. Results show that hyperparameter tuning enhanced generalization and reduced overfitting. This study contributes to the growing field of computational intelligence in the healthcare system by demonstrating the effect of hybrid optimization techniques in the prediction of diabetes. The proposed model can serve as a decision-making tool for doctors, which can enable early diagnosis and personal treatment strategies. Future research directions aim to investigate alternative bio-inspired customization algorithms and scale the model's capabilities.

ACKNOWLEDGEMENT

The authors would like to express theirsincere gratitude to the research facilities provided by Mohan Babu University, India for carrying out this work.

REFERENCES

Alirezaei, M., Niaki, S. T. A., & Niaki, S. A. A. (2019). A bi-objective hybrid optimization algorithm to reduce noise and data dimension in diabetes diagnosis using support vector machines. *Expert Systems with Applications*, *127*, 47-57. https://doi.org/10.1016/j.eswa.2019.02.037

Asif, S., Wenhui, Y., ur-Rehman, S., ul-ain, Q., Amjad, K., Yueyang, Y., Jinhai, S., & Awais, M. (2025). Advancements and prospects of machine learning in medical diagnostics: unveiling the future of diagnostic precision. *Archives of Computational Methods in Engineering, 32*(2), 853-883. https://doi.org/10.1007/s11831-024-10148-w

Mohammed, F. A., Tune, K. K., Assefa, B. G., Jett, M., & Muhie, S. (2024). Medical image classifications using convolutional neural networks: A survey of current methods and statistical modeling of the literature. *Machine Learning and Knowledge Extraction*, *6*(1), 699-735. https://doi.org/10.3390/make6010033

Navazi, F., Yuan, Y., & Archer, N. (2023). An examination of the hybrid meta-heuristic machine learning algorithms for early diagnosis of type II diabetes using big data feature selection. *Healthcare Analytics*, *4*, Article 100227. https://doi.org/10.1016/j.health.2023.100227

Raposo, H. (2024). Intelligent imaging: A systematic review of artificial intelligence techniques in disease detection, segmentation, and classification. *Segmentation, and Classification (May 13, 2024)*. http://dx.doi.org/10.2139/ssrn.5342645