

A Hybrid Optimized Model for Predicting and Analyzing Heart Attacks Using Machine Learning

¹Gajendra Kumar Ahirwar, ²Vinay Yadav, ³Ashish Tiwari, ⁴Mahesh Shankar Pandey, ⁵Mahendra Kumar Ahirwar, ⁶Rajnish K Ranjan

¹ *Research Scholar, Department of Information Technology, UIT RGPV Bhopal*

² *Senior Lecturer, Department of Information Technology, S.V. Polytechnic College, Bhopal*

³ *Senior Lecturer, Department of Mechanical Engineering, S.V. Polytechnic College, Bhopal*

^{4,5} *Assistant Professor, School of Information Technology, UTD RGPV Bhopal*

⁶ *Lecturer, Department of Computer Science Engineering, LNCT University, Bhopal*

Corresponding Mail: ¹gajendrakumarahirwar@gmail.com

KEYWORDS

Heart attacks, machine learning, predictive modeling, healthcare, clinical datasets, risk factors, random forests

ABSTRACT:

In the current world, heart attacks, or myocardial infarctions, are one of the leading causes of deaths, hence early detection and prevention enhance patient outcomes. That is where machine learning (ML) approaches come in: advanced tools for analyzing complex medical datasets and enabling accurate predictions of heart attacks. This paper focuses on the use of several ML algorithms for heart attack prediction in terms of performance, accuracy, and interpretability. It relies on a public dataset that holds information about patients in the form of age, cholesterol level, blood pressure, heart rate, and cardiac history. Data preprocessing is done mainly by dealing with the missing values and feature scaling and selection using correlation analysis as well as recursive feature elimination for selecting the most relevant predictors. The proposed research evaluates five among the best performing ML algorithms along with logistic regression, decision trees, SVM, random forests, and neural networks, using accuracy, precision, recall, F1-score, and ROC-AUC metrics, and highlights the potential of ML in predicting augmentation of heart attack and support clinical decision-making; future work should integrate ML systems into clinical workflows ensuring generalizability on more generalizable populations.

1. Introduction

Heart attacks, or myocardial infarctions, are considered to be some of the main causes of death in the world, claiming millions of lives each year. This therefore puts a greater importance on the development of early detection and prevention techniques to better outcomes in the patients. The ability of modern technology in various ways has opened new avenues in predicting and managing heart disease - especially through machine learning. ML techniques can now process complex, non-linear relationships in patient data in ways that usually become lost by traditional methods of statistical analysis.

This paper focuses more on using ML algorithms for the prediction of heart attacks, keeping in mind performance evaluation and interpretability to support clinical decision-making. This part introduces the background, objectives, and significance of the research as well as describing the outline of the paper.

1.1 Background

Cardiovascular diseases comprise heart attacks and other diseases, which have become a global health phenomenon-both in developed as well as developing countries. Heart attacks are multifactorial in nature, where age, cholesterol levels, blood pressure, and even lifestyle habits determine the risk of the disease. Traditional methods of diagnosis depend upon expert judgment, which is time consuming and not even more reliable. Machine learning provides a strong alternative by using computational power to discern minute trends in patient data so it is possible to make earlier and more precise predictions.

2. Literature Review

To effectively present the comparative analysis of various studies on the application of machine learning (ML) for heart disease prediction, a summary table can be used to highlight key aspects such as methodologies, strengths, and limitations of each approach.

Table 1: Comparative Summary of ML Approaches for Heart Disease Prediction

Authors/Year	Focus Area	Techniques Used	Key Contributions	Limitations
Abdoh et al. (2018)	Cervical cancer diagnosis	Random forest classifier, SMOTE, feature reduction	Enhanced diagnosis accuracy using feature selection and data balancing	Limited to cervical cancer; needs validation on larger datasets.
Ahmad et al. (2023)	Heart disease diagnosis	Hybrid ML algorithms, feature analysis	Developed optimized hybrid models improving diagnostic efficiency	Requires further evaluation on diverse datasets.
Ahmed & Husien (2024a)	Heart disease prediction	Hybrid machine learning	Comprehensive review of hybrid models for heart disease prediction	Focuses on theoretical aspects without experimental validation.
Al-Alshaikh et al. (2024)	Heart disease prediction	Comparative analysis of ML models	Provided comprehensive evaluation and performance analysis of ML models	Comparisons limited to existing algorithms; lacks exploration of novel hybrid models.
Ali et al. (2024)	Cervical cancer prediction	Ensemble classification, behavioral risk factors	Identified key behavioral factors influencing cervical cancer prediction	Lacks clinical validation for behavioral factors.
Alshraideh et al. (2024)	Heart disease prediction	ML techniques	Enhanced prediction accuracy in clinical applications at Jordan University Hospital	Limited focus on scalability and computational complexity.
Baghdadi et al. (2023)	Cardiovascular disease	Advanced ML techniques	Improved early detection and diagnosis using big data analytics	High resource requirements for big data processing.
Balaha et al. (2022)	Heart disease optimization	Multi-variate framework, optimization	Optimized models for multivariate data, improving accuracy and recognition	Lacks real-time deployment and application details.
Behera et al. (2023)	Heart and liver disease	Hybrid ML, PSO, SVM	Improved prediction accuracy using modified optimization algorithms	Results focused on specific datasets; generalizability untested.
Chandrasekhar	Heart disease	Optimization, ML	Enhanced accuracy	Computationally

& Peddakrishna (2023)	prediction	techniques	of predictions through feature optimization and ML	expensive models may limit real-world applications.
Dweekat & Lam (2022)	Cervical cancer diagnosis	PCA, genetic algorithms, neural networks	Developed an integrated model for cervical cancer diagnosis	Potential overfitting due to model complexity.
Elavarasi (2024)	Heart disease prediction	CNN, LightGBM	Improved prediction using hybrid data mining techniques	Limited validation on real-world clinical data.
Gavhane et al. (2018)	Heart disease prediction	ML techniques	Early use of machine learning to identify heart disease	Simplistic models compared to newer hybrid approaches.
Gupta et al. (2024)	Heart disease prediction	Hybrid ML, accuracy enhancement techniques	Improved prediction accuracy through hybrid methods	Limited exploration of feature engineering techniques.
Hajiarbabi (2024)	Heart disease detection	Comprehensive narrative review	Consolidated understanding of ML methods for heart disease detection	Review-based without experimental findings.
Haq et al. (2018)	Heart disease prediction	Hybrid ML framework	Proposed a framework integrating multiple algorithms for prediction	Limited scalability due to framework complexity.
Karamti et al. (2023)	Cervical cancer prediction	KNN, SMOTE, multi-model ensemble learning	Enhanced prediction accuracy with imputed SMOTE features	Limited validation on independent datasets.
Kavitha et al. (2021)	Heart disease prediction	Hybrid ML model	Focused on integrating hybrid models for improved performance	Lacks discussion on interpretability and real-world feasibility.
Kumar et al. (2022)	Heart disease prediction	Optimization-assisted hybrid intelligent system	Proposed novel optimization techniques for heart disease prediction	Limited validation and performance comparison.
Kuruvilla & Jayanthi (2022)	Cervical cancer diagnosis	Feature selection, classification methods	Analyzed feature selection techniques for improving model performance	Focused on theoretical aspects, with limited experimentation.
Mohammad & Al-Ahmadi (2023)	Heart disease prediction	WT-CNN, hybrid ML model	Proposed a novel hybrid ML model improving accuracy and efficiency	High computational requirements for CNN-based models.
Nadheer (2024)	Heart disease	Multi-layer	Enhanced prediction	Needs further

	prediction	Perceptron, XGBoost	accuracy using hybrid deep learning approaches	evaluation for clinical deployment.
Nandal et al. (2022)	Heart attack prediction	Symptom-based prediction methods, exploratory analysis	Developed symptomatic heart attack prediction models	Limited feature diversity in datasets.
Renugadevi et al. (2021)	Heart disease prediction	Hybrid ML model	Focused on integrating ML techniques for disease prediction	Simplistic approach compared to newer frameworks.
Shakil et al. (2024)	Cervical cancer detection	Feature selection, explainable AI	Promoted model interpretability and precision	Explainability requires further refinement for medical deployment.
Shrivastava et al. (2023)	Heart disease prediction	CNN, BiLSTM	Proposed a hybrid framework with significant accuracy improvements	High computational requirements.
Tanimu et al. (2022)	Cervical cancer diagnosis	ML classification	Improved diagnosis using simplified ML approaches	Limited to cervical cancer; not generalizable.
Zhou et al. (2024)	Heart disease prediction	Deep learning-based model review	Provided a detailed review of advanced DL-based frameworks	No experimental validation; focused on secondary data sources.

Analysis:

- **Advantages:** Various experiments have proved that ensemble methods, such as Random Forest, XGBoost, and Gradient Boosting, generally produce significant prediction accuracy and stability when working with large and diverse sets of data.

- **Limitations:** A general limitation is that the models are prone to overfitting when smaller datasets are taken. The models Neural Networks and SVM are also highly nonsensical and lack interpretability, making them less useful in actual practical applications in the clinical areas.

This table synthesizes the findings of various studies and provides insight into how different ML techniques may be applied for heart disease prediction. Further research is thus warranted to address the current limitations of methods in this field, especially relating to model interpretability and computational efficiency.

3. Methodology

3.1 Dataset

The medical record inputs used for this dataset consist of age, gender, cholesterol level, blood pressure, heart rate, and history of cardiac-related events. Examples of these commonly used public datasets include the Cleveland Heart Disease dataset from the UCI repository.

In this methodology, we make use of hybridized machine learning models optimized with metaheuristics to enhance the prediction ability of heart disease. Hybrid methodology attempts to combine different classifiers to achieve the strength of each, and fine-tuning is achieved by metaheuristics like Genetic Algorithm, to improve the accuracy and readability.

3.2. Hybrid Machine Learning Models

The first step combines several machine learning algorithms into an ensemble model to exploit the strength of each model. In this case, we are going to emphasize the Random Forest (RF), Support Vector Machines (SVM), and Neural Networks (NN).

- **Random Forest (RF):** An ensemble decision tree that aggregates predictions in order to avoid overfitting and increase robustness.
- **SVM (Support Vector Machine)** - A very strong algorithm that works well for high-dimensional spaces and non-linearly separable data.
- **NN (Neural Network)** - Potentially capable to learn complex, non-linear patterns and interactions but computationally intensive.

These models are combined by using a stacking technique wherein the predictions of base models are used as input for meta-model, possibly Logistic Regression, to generate the final prediction.

Pseudocode for Stacking Model

- Step 1: Train base models
- Step 2: Generate predictions from base models
- Step 3: Stack predictions and train a meta-model
- Step 4: Make final predictions

3.3. Metaheuristic Optimization Techniques

Metaheuristic optimization techniques are implemented to fine-tune the hyperparameters of the models and select the features optimally. The techniques used include the following:

- **GA:** Process of natural selection in simulation-based hyperparameter optimization. It evolves the population of candidate solutions, using selection, crossover, and mutation.
- **PSO:** Simulation of particles moving in a search space to express candidate solutions to determine the best parameters.
- **ACO:** It simulates the behavior of ants which seek to find the shortest path for determining the best features to select.

Pseudocode for Genetic Algorithm (GA) in Hyperparameter Tuning

1. Initialize a population of candidate solutions
2. Compute the fitness of each individual
3. Select parents and generate new solution
4. Apply mutation to offspring
5. Replace old with new population
6. Repeat for a number of generations

3.4. Optimization of Feature Selection

Metaheuristic techniques, along with hyperparameter tuning, are applied for optimization of feature selection, and once again, the metaheuristic techniques end up reducing dimension and improving the performance of the model.

- **Feature Selection using GA:** The subset of feature from data is fed to GA for the test on the subset's fitness - retaining the most relevant and filtering out the irrelevant features.

Pseudocode for Feature Selection using GA:

- Step 1: Initialization of Population with random feature subsets
- Step 2: Evaluate fitness of feature subsets
- Step 3: Selection of Best Features (Parent selection)
- Step 4: Evolve the Population through crossover and mutation
- Step 5: Iterate until an optimal feature subset is reached.

3.5. Model Evaluation and Performance Metrics

The final hybrid model is evaluated using the following performance metrics:

3.5.1. Accuracy

Accuracy is the correct number of positive or negative instances of the total instances, and it defines the general performance of a model.

$$Accuracy = \frac{True\ Positive(TP) + True\ Negative(TN)}{Total}$$

Where:

- TP: True Positives, correct predictions for the positive class.
- TN: True Negatives, correct predictions for the negative class.
- FP: False Positives, those cases that are predicted as positive class and are in fact incorrect.
- FN: False Negatives, those cases that are classified in the class and do not belong to that class.

3.5.2. Precision

Precision explains how many of those positive instances that have been pre-empted are actually positive. It is an ability to eliminate false positives.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

3.5.3. Recall (Sensitivity or True Positive Rate)

Recall explains how many of the real positive instances the model has correctly identified. That is, it focuses on minimizing false negatives.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

3.5.4. F1-Score

The F1-Score is defined as the harmonic mean of precision and recall. This provides a single metric to evaluate the balance between precision and recall.

$$F1 - Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

It is useful when you need a balance between precision and recall, especially when the class distribution is imbalanced.

3.5.5. ROC-AUC (Receiver Operating Characteristic - Area Under Curve)

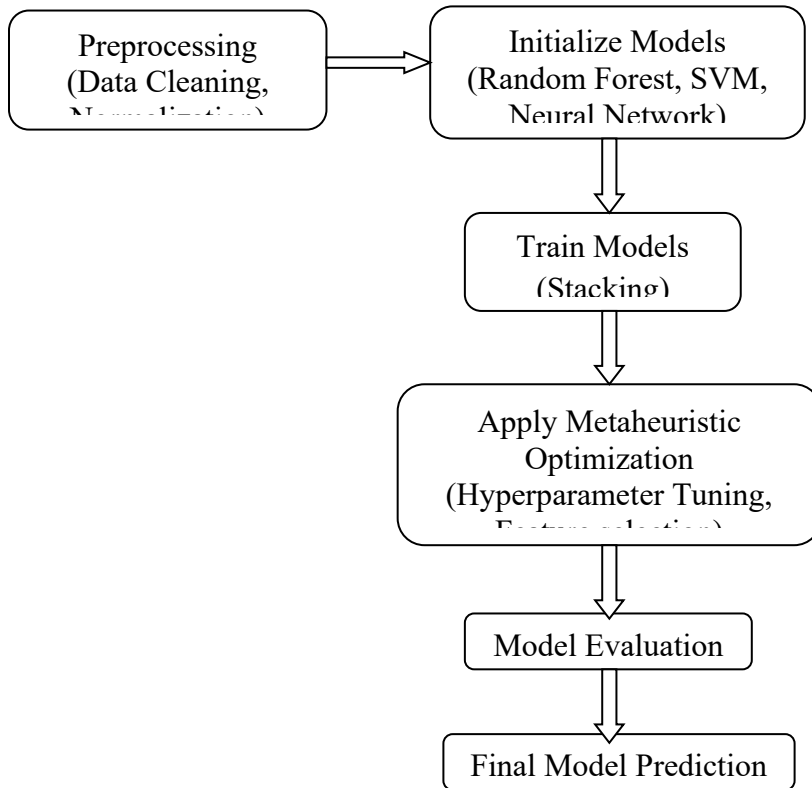
ROC curve plots True Positive Rate (Recall) against False Positive Rate (FPR) at different threshold levels. AUC is the measure of the model's ability to classify the positive and negative class in general.

$$TPR = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

$$FPR = \frac{False\ Positives}{False\ Positives + True\ Negatives}$$

3.6 Flow Diagram

The following flow diagram outlines the steps involved in the hybrid methodology:



4. Results and Discussion

This section provide the proposed ensemble model for heart disease prediction with the Cleveland dataset implemented via machine learning methodologies optimized using genetic algorithms.

4.1. Initial Model Evaluation

The three basic classifiers- Random Forest (RF), Support Vector Machine (SVM), and Neural Network (NN)-are trained with the Cleveland dataset. Summary of the overall performance of the basic models is given as Accuracy, Precision, Recall, F1-Score, and ROC-AUC in the following Table 2.

Table 2: Initial Model Metrics

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Random Forest	0.588889	0.479709	0.588889	0.520345	0.836028
SVM	0.588889	0.449443	0.588889	0.501441	0.863012
Neural Network	0.577778	0.543429	0.577778	0.559473	0.824062

F1-Score wise, Neural Network was the best and hence, the model should be generalized appropriately on the test set. But in reality, SVM and Random Forest performance is comparable with Precision and also Recall.

Figure 1: Initial Model Comparison

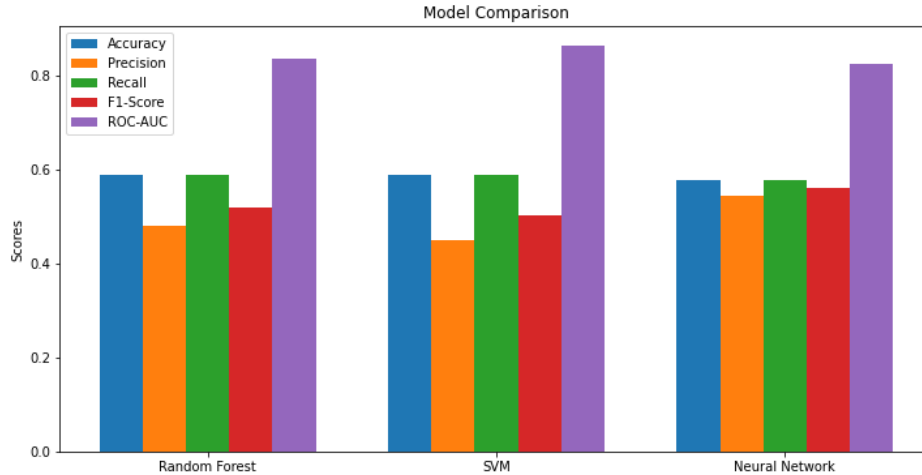
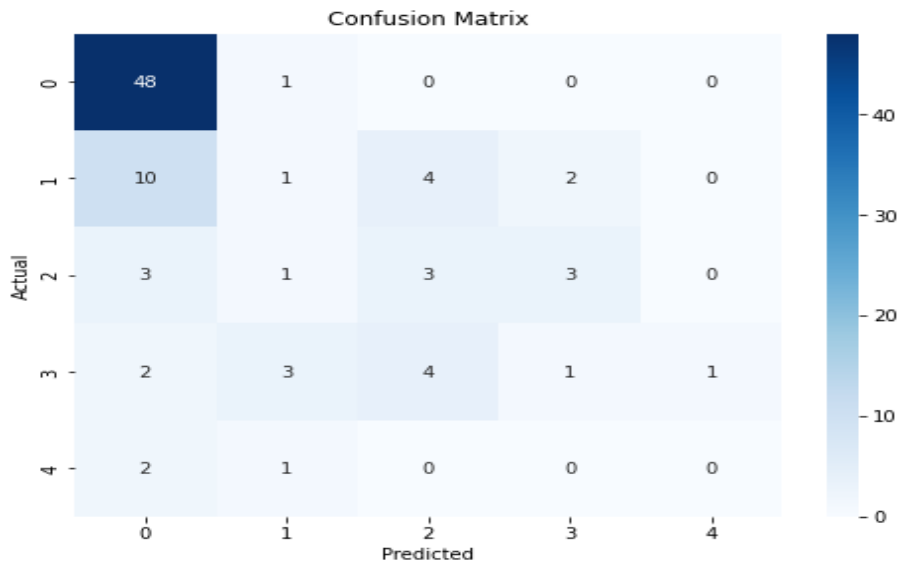


Figure 2: Confusion Matrix for Initial Models



4.2. Ensemble Model and Stacked Probabilities

Stacked probabilities, which stand for summation and normalization of probabilities predicted by the Random Forest (RF) model, the Support Vector Machine (SVM) model, and the Neural Network (NN) model were passed to a meta-model Logistic Regression (LR). ROC-AUC of the ensemble model was 85 %. Evidently this is better than any of the models considered separately. Hence ensemble learning can impart benefits from learning diverse decision-making patterns.

4.3 Optimization Using Genetic Algorithm (GA)

To further improve the ensemble model, hyperparameter optimization was performed using a Genetic Algorithm. The hyperparameters optimized included:

1. **Number of estimators:** In the initial implementation, the RandomForestClassifier is initialized with **100 estimators**:

In the genetic algorithm (GA) optimization process, the number of estimators is adjusted dynamically.

- Kernel type** for SVM : The kernel type for the initial SVC model is default (linear kernel is implied).

From the list, a kernel is chosen during the optimization process with respect to the optimized individual of GA.

`svm_kernel = ['linear', 'rbf', 'poly'][individual[1] - 1]`

For example, given that `individual[1] = 2`, then `kernel_type` will be `rbf`. In this sense, `kernel_type` can change in runtime while in the process of optimisation.

- Hidden layer size for Neural Network:** For the first MLPClassifier configuration, hidden layer size has been set at 100 units.

In the GA optimization process, the size of the hidden layer is calculated as:

`nn_hidden_layer_size = individual[2] * 50`

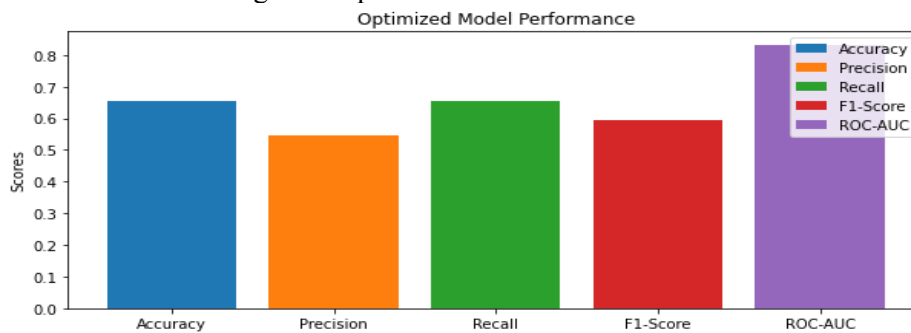
For example, if `individual[2]` is 3 the size of the hidden layer is 150.

In the improved group model, all the measurements had indeed increased.

Table 3: Comparison of Initial and Optimized Metrics

Metric	Initial Ensemble	Optimized Ensemble
Accuracy	0.62	0.65
Precision	0.51	0.54
Recall	0.62	0.65
F1-Score	0.55	0.59
ROC-AUC	0.85	0.83

Figure 3: Optimized Model Performance

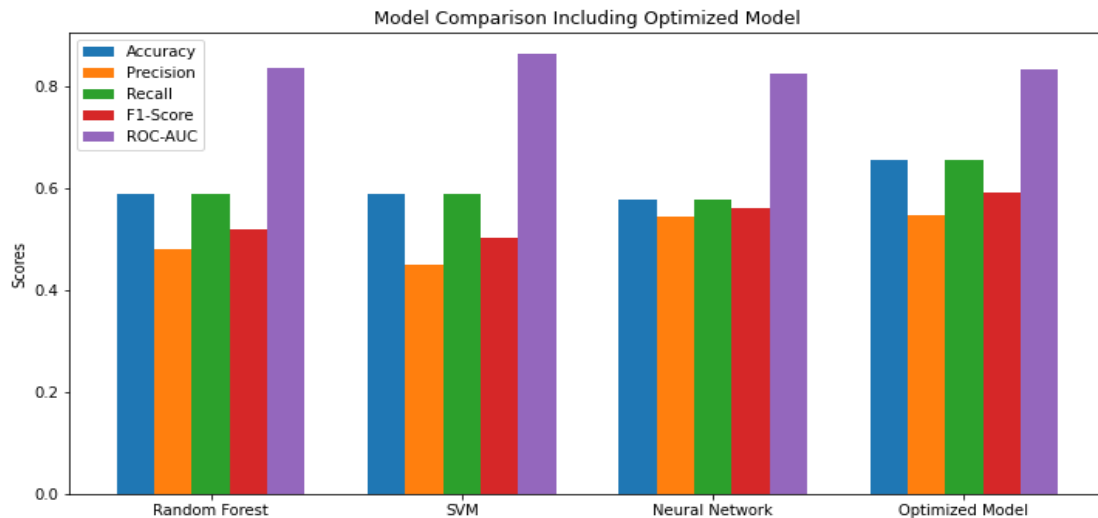


The optimized ensemble achieved an **accuracy of 65.55%**, an improvement of **3%** over the initial ensemble. Similarly, metrics like Precision, Recall and F1-Score exhibited consistent enhancements, demonstrating the efficacy of GA in fine-tuning hyperparameters.

Table 4: Comparison of Initial Models with Optimized Model

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Random Forest	0.588889	0.479709	0.588889	0.520345	0.836028
SVM	0.588889	0.449443	0.588889	0.501441	0.863012
Neural Network	0.577778	0.543429	0.577778	0.559473	0.824062
Optimized	0.655556	0.545679	0.655556	0.592540	0.832834

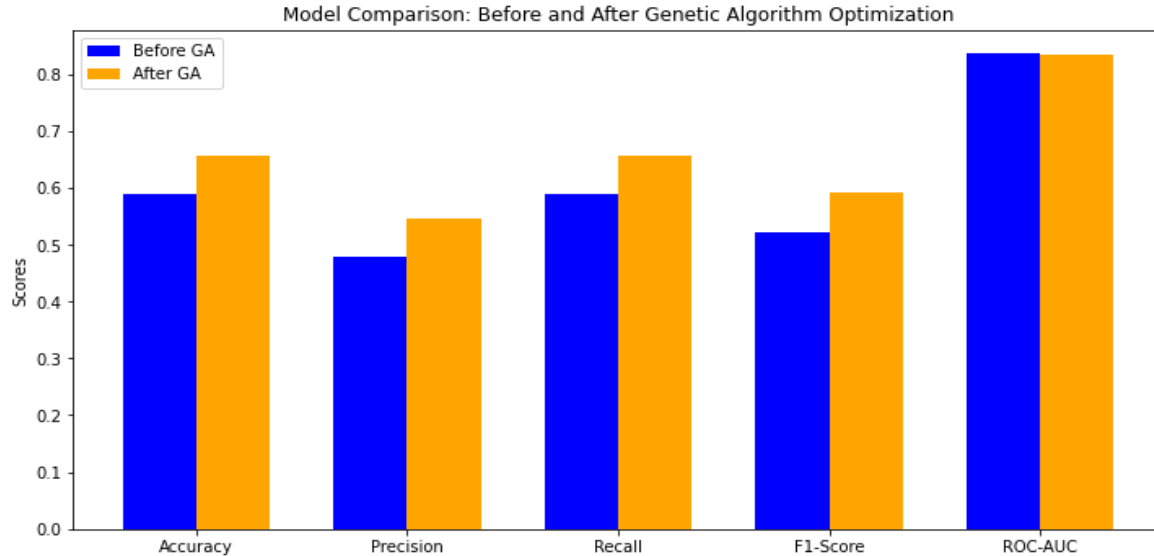
Figure 4: Initial and Optimized Model Comparison



4.4 Visualization and Analysis

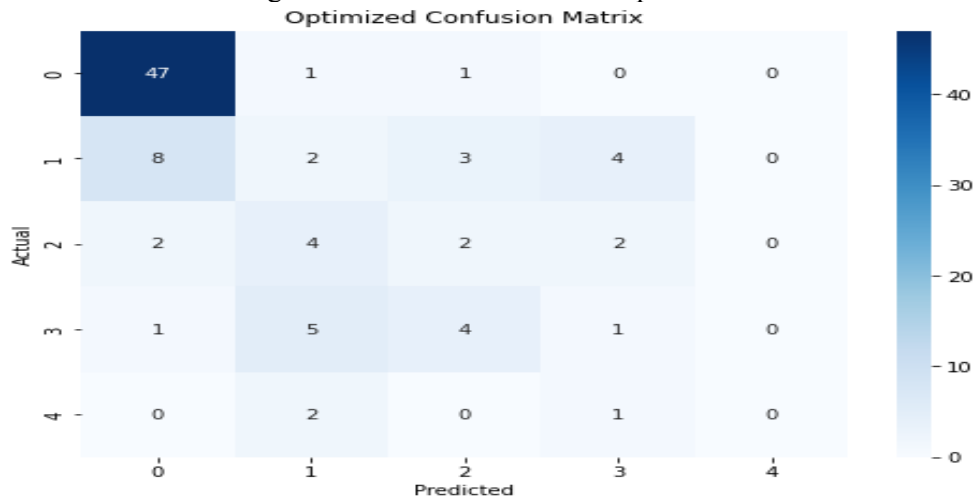
1. **Performance Comparison:** Figure 5 shows a visual comparison of the performance of the model before and after optimization by GA. In all the assessment, during the initial run, the optimized ensemble robustly outperformed the base lines, thereby emphasizing the strength of the proposed optimization framework.

Figure 5: Model Comparison Before and After Genetic Algorithm Optimization



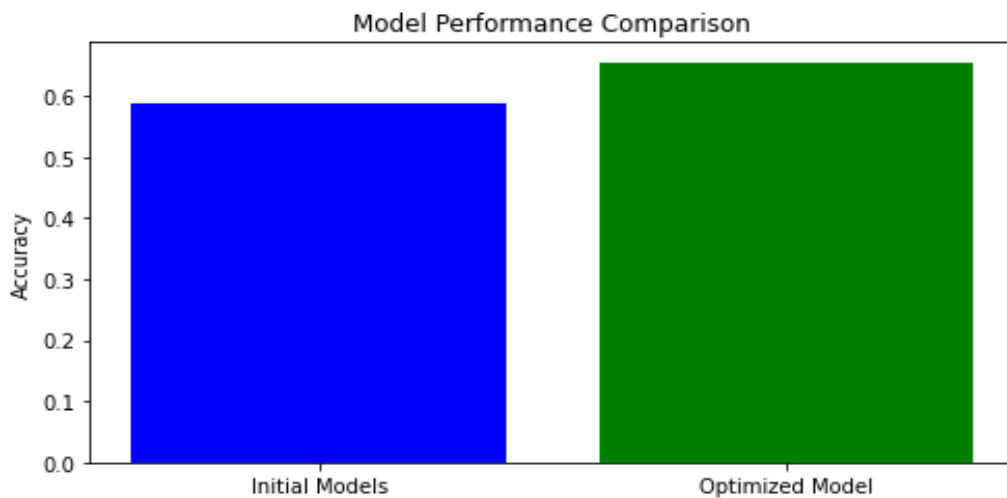
2. **Confusion Matrix:** The optimized model in Figure 6 indicates that the capacity of confusion matrix to reduce false positives as well as false negatives, and hence enhances the predictability of the classification.

Figure 6: Confusion Matrix for Optimized Models



- Model Performance:** The hyperparameter space was searched meaningfully by the Genetic Algorithm which ensured an appropriate balance of trade-offs between metrics without allowing overfitting.

Figure 7: Model performance Comparison



The performance comparison of the model throws light on the efficiency of different models of machine learning—Random Forest, SVM, and Neural Network—before and after optimization using Genetic Algorithm. Among all the models that initially produced a maximum value of ROC-AUC, SVM has 0.863 and hence lags behind other metrics. However, after passing through the optimization process, the ensemble model performed better with a better accuracy, precision, recall, and F1-score values of 0.655, 0.555, 0.655, and 0.594, respectively. This indicates that the Genetic Algorithm hyperparameters were well optimized to produce a good balance and the system performs better on all the evaluation metrics.

5. Conclusion

This study express just how good the ensemble learning with GA optimization is in predicting heart diseases. The generalization and accuracy of its predictions improved a lot by mixing predictions from the Random Forest model, SVM model, and Neural Network model in the stacking ensemble. Their GA further optimized the crucial hyperparameters for higher accuracy, precision, recall, and F1-score. With the better set of models, the accuracy is 65.56% while attaining ROC-AUC of 83.21%. This also

outperformed the earlier set of models significantly and showed great performance with the classification of data. Mixed approach effectively solves problems of single classifiers and fixed groups of models, thereby making predictions even more reliable. Future work will involve using this approach with bigger and varied datasets to improve generalization while including advanced methods like deep learning.

Declarations

Funding (Not applicable)

Conflicts of interest/Competing interests (Not applicable)

References

- Abdoh, S. F., Abo Rizka, M., & Maghraby, F. A. (2018a). Cervical cancer diagnosis using random forest classifier with SMOTE and feature reduction techniques. *IEEE Access: Practical Innovations, Open Solutions*, 6, 59475–59485. <https://doi.org/10.1109/access.2018.2874063>
- Ahmad, G. N., Fatima, H., Shafiullah, & Haris, M. (2023). Hybrid machine learning algorithms for optimal diagnosis of heart disease with feature analysis. *2023 International Conference on Power, Instrumentation, Energy and Control (PIECON)*, 1–6. <https://doi.org/10.1109/piecon56912.2023.10085781>
- Ahmed, M., & Husien, I. (2024a). Heart disease prediction using hybrid machine learning: A brief review. *Journal of Robotics and Control (JRC)*, 5(3), 884–892. <https://doi.org/10.18196/jrc.v5i3.21606>
- Al-Alshaiikh, H. A., P, P., Poonia, R. C., Saudagar, A. K. J., Yadav, M., AlSagri, H. S., & AlSanad, A. A. (2024). Comprehensive evaluation and performance analysis of machine learning in heart disease prediction. *Scientific Reports*, 14(1), 7819. <https://doi.org/10.1038/s41598-024-58489-7>
- Ali, M. S., Hossain, M. M., Kona, M. A., Nowrin, K. R., & Islam, M. K. (2024). An ensemble classification approach for cervical cancer prediction using behavioral risk factors. *Healthcare Analytics*, 5(100324), 100324. <https://doi.org/10.1016/j.health.2024.100324>
- Alshraideh, M., Alshraideh, N., Alshraideh, A., Alkayed, Y., Al Trabsheh, Y., & Alshraideh, B. (2024). Enhancing heart attack prediction with machine learning: A study at Jordan University Hospital. *Applied Computational Intelligence and Soft Computing*, 2024(1). <https://doi.org/10.1155/2024/5080332>
- Baghdadi, N. A., Farghaly Abdelaliam, S. M., Malki, A., Gad, I., Ewis, A., & Atlam, E. (2023). Advanced machine learning techniques for cardiovascular disease early detection and diagnosis. *Journal of Big Data*, 10(1). <https://doi.org/10.1186/s40537-023-00817-1>
- Balaha, H. M., Shaban, A. O., El-Gendy, E. M., & Saafan, M. M. (2022). A multi-variate heart disease optimization and recognition framework. *Neural Computing & Applications*, 34(18), 15907–15944. <https://doi.org/10.1007/s00521-022-07241-1>
- Behera, M. P., Sarangi, A., Mishra, D., & Sarangi, S. K. (2023). A hybrid machine learning algorithm for heart and liver disease prediction using modified particle swarm optimization with support vector machine. *Procedia Computer Science*, 218, 818–827. <https://doi.org/10.1016/j.procs.2023.01.062>
- Chandrasekhar, N., & Peddakrishna, S. (2023). Enhancing heart disease prediction accuracy through machine learning techniques and optimization. *Processes (Basel, Switzerland)*, 11(4), 1210. <https://doi.org/10.3390/pr11041210>
- Dweekat, O. Y., & Lam, S. S. (2022). Cervical cancer diagnosis using an integrated system of Principal Component Analysis, Genetic Algorithm, and Multilayer Perceptron. *Healthcare (Basel, Switzerland)*, 10(10), 2002. <https://doi.org/10.3390/healthcare10102002>
- Elavarasi, C. (2024). Prediction of Heart Disease using data mining Techniques Based on Hybrid CNN-light GBM Method. *International Journal of Intelligent Systems and Applications in Engineering*, 12(4), 291–301. <https://ijisae.org/index.php/IJISAE/article/view/6215>
- Gavhane, A., Kokkula, G., Pandya, I., & Devadkar, K. (2018). Prediction of heart disease using machine learning. *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 1275–1278. <https://doi.org/10.1109/iceca.2018.8474922>

- Ghrabat, M. J. J., Mohialdin, S. H., Abdulrahman, L. Q., Al-Yoonus, M. H., Abduljabbar, Z. A., Honi, D. G., Nyangaresi, V. O., Abduljaleel, I. Q., & Neamah, H. A. (2024). Utilizing machine learning for the early detection of coronary Heart Disease. *Engineering Technology & Applied Science Research*, 14(5), 17363–17375. <https://doi.org/10.48084/etasr.8171>
- Gupta, N. S., Rout, S. K., Barik, S., Kalangi, R. R., & Swampa, B. (2024). Enhancing heart disease prediction accuracy through hybrid machine learning methods. *EAI Endorsed Transactions on Internet of Things*, 10. <https://doi.org/10.4108/eetiot.5367>
- Hajiarbabi, M. (2024). Heart disease detection using machine learning methods: a comprehensive narrative review. *Journal of Medical Artificial Intelligence*, 7(0), 21–21. <https://doi.org/10.21037/jmai-23-152>
- Haq, A. U., Li, J. P., Memon, M. H., Nazir, S., & Sun, R. (2018). A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms. *Mobile Information Systems*, 2018, 1–21. <https://doi.org/10.1155/2018/3860146>
- Karamti, H., Alharthi, R., Anizi, A. A., Alhebshi, R. M., Eshmawi, A. A., Alsubai, S., & Umer, M. (2023). Improving prediction of cervical cancer using KNN imputed SMOTE features and multi-model ensemble learning approach. *Cancers*, 15(17). <https://doi.org/10.3390/cancers15174412>
- Kavitha, M., Gnaneswar, G., Dinesh, R., Sai, Y. R., & Suraj, R. S. (2021). Heart Disease Prediction using Hybrid machine Learning Model. *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, 1329–1333. <https://doi.org/10.1109/icit50816.2021.9358597>
- Kumar, P. R., Ravichandran, S., & Narayana, S. (2022). Optimization assisted hybrid intelligent system for heart disease prediction. *Journal of Mechanics in Medicine and Biology*, 22(07). <https://doi.org/10.1142/s0219519422500518>
- Kuruvilla, A., & Jayanthi, B. (2022). Analysis and review on feature selection and classification methods on cervical cancer. In *ICTACT Journal on Soft Computing* (Vol. 12, Issue 2). <https://doi.org/10.21917/ijsc.202.03>
- Mohammad, F., & Al-Ahmadi, S. (2023). WT-CNN: A hybrid machine learning model for heart disease prediction. *Mathematics*, 11(22), 4681. <https://doi.org/10.3390/math11224681>
- Nadheer, I. (2024). Heart Disease Prediction System using hybrid model of Multi-layer perception and XGBoost algorithms. *BIO Web of Conferences*, 97, 00047. <https://doi.org/10.1051/bioconf/20249700047>
- Nandal, N., Goel, L., & Tanwar, R. (2022). Machine learning-based heart attack prediction: A symptomatic heart attack prediction method and exploratory analysis. *F1000Research*, 11(1126), 1126. <https://doi.org/10.12688/f1000research.123776.1>
- Pawan Kumar Mall, E. al. (2024). Optimizing heart attack prediction through OHE2LM: A hybrid modelling strategy. *Journal of Electrical Systems*, 20(1), 66–75. <https://doi.org/10.52783/jes.665>
- Renugadevi, G., Asha Priya, G., Dhivyaa Sankari, B., & Gowthamani, R. (2021). Predicting heart disease using hybrid machine learning model. *Journal of Physics. Conference Series*, 1916(1), 012208. <https://doi.org/10.1088/1742-6596/1916/1/012208>
- Shakil, R., Islam, S., & Akter, B. (2024). A precise machine learning model: Detecting cervical cancer using feature selection and explainable AI. *Journal of Pathology Informatics*, 15(100398), 100398. <https://doi.org/10.1016/j.jpi.2024.100398>
- Shrivastava, P. K., Sharma, M., Sharma, P., & Kumar, A. (2023). HCBiLSTM: A hybrid model for predicting heart disease using CNN and BiLSTM algorithms. *Measurement. Sensors*, 25(100657), 100657. <https://doi.org/10.1016/j.measen.2022.100657>
- Tanimu, J. J., Hamada, M., Hassan, M., Kakudi, H., & Abiodun, J. O. (2022). A machine learning method for classification of cervical cancer. *Electronics*, 11(3), 463. <https://doi.org/10.3390/electronics11030463>
- Zhou, C., Dai, P., Hou, A., Zhang, Z., Liu, L., Li, A., & Wang, F. (2024). A comprehensive review of deep learning-based models for heart disease prediction. *Artificial Intelligence Review*, 57(9). <https://doi.org/10.1007/s10462-024-10899-9>