International Journal on Robotics, Automation and Sciences

Genetic Algorithm-Based Multitier Ensemble Classifier for Diagnosis of Heart Disease

Thirumalaimuthu Thirumalaiappan Ramanathan, Md. Jakir Hossen*, and Md. Shohel Sayeed

***Abstract* - Designing a hybrid or ensemble data mining system appropriate to the application is a research challenge. Heart disease is a life threatening disease that need to be recognized correctly in the starting stage before it becomes more complex. Using artificial intelligence techniques in a hybrid and ensemble architecture can support the prediction of heart disease more effectively based on the given sample cases. This paper proposes a classification system called genetic algorithm-based ensemble classification system (GA-ECS) for the identification of heart disease. As feature selection is the crucial step before applying the data mining techniques, the genetic algorithm is used in GA-ECS to identify the best features in a given dataset. The Cleveland heart disease dataset is used for testing GA-ECS. The performance of GA-ECS is compared with different machine learning classifiers for the prediction of heart disease. GA-ECS showed a promising outcome with an accuracy of 90% for the diagnosis of heart disease.**

Keywords—Ensemble Learning, Genetic Algorithm, AdaBoost, Data Mining, Medical Diagnosis.

I. INTRODUCTION

Artificial intelligence (AI) techniques are useful in healthcare applications in providing solutions for diagnosis of diseases through the development of software that can analyze like medical experts [1-4]. The AI and machine learning techniques are being investigated for the medical diagnosis problem by testing its performance on prediction of unknown cases [5-8]. Different machine learning algorithms produce different outcomes based on the application. The efficiencies of machine learning algorithms are being varied according to the training samples. Data preprocessing techniques such as feature extraction methods mainly affect the outcomes of AI classifiers applied for the data mining process [9-12]. An appropriate selection of feature extraction methods and AI classifiers based on the application can improve the performance of a medical prediction system. Heart disease is one of the critical challenges which remains to be the major cause of death and it is being a serious discussion in various healthcare organizations in Malaysia [13]. Many people of all types of ages living in Malaysia are being diagnosed with heart disease [14]. Many research works are being done in data mining to support the prediction of heart disease applying the accessible medical datasets [15].

The following described research works used the Cleveland heart disease dataset [16] for testing. The performance of support vector machine (SVM) [17] and artificial neural network (ANN) [18] were studied in the research work [19] for the classification of heart disease where SVM showed an accuracy of 83.1% and ANN showed an accuracy of 72.7%. The performances of different machine learning boosting techniques were studied individually in the research work [20] for the classification of heart disease the hybrid of AdaBoost [21] and SVM classification approach showed good performance on classification. The performances of decision tree [22], extreme learning machine [23], and

*Corresponding author. Email: jakir.hossen@mmu.edu.my
Thirumalaimuthu Thirumalaiappan Ramanathan is with the Faculty of Information Science and Technology, Multimedia University, Melaka 75450 Malaysia (e-mail: 1181402216@student.mmu.edu.my).
Md. Jakir Hossen is with the Faculty of Engineering and Technology, Multimedia University, Melaka 75450 Malaysia (phone: 06-2523425; fax: 303-555-5555; e-mail: jakir.hossen@mmu.edu.my).
Md. Shohel Sayeed is with the Faculty of Information Science and Technology, Multimedia University, Melaka 75450 Malaysia (e-mail: shohel.sayeed@mmu.edu.my).

International Journal on Robotics, Automation and Sciences (2024) 6,1:29-35
https://doi.org/10.33093/ijoras.2024.6.1.5
Manuscript received: 23 Aug 2023 | Accepted: 12 Dec 2023 | Published: 30 Apr 2024
Published by MMU PRESS. URL: http://journals.mmupress.com/ijoras

gradient boosting [24] were studied in the research work [25] where the decision tree showed the accuracy of 84.2%, extreme learning machine showed the accuracy of 81%, and gradient boosting showed the accuracy of 81.6% for the classification of heart disease. The performance of hybrid approach using particle swarm optimization (PSO) [26] and SVM were studied in the research work [27] where the SVM-PSO classifier showed the accuracy of 84.81% for classification of heart disease. The performance of PSO based multi-layer perceptron (MLP) [28] was studied in the research work [29] where the PSO based MLP showed the accuracy of 84.6% for classification of heart disease.

An effective hybrid kind of approach is required for the classification of heart disease data. This paper presents a machine learning classification system called genetic algorithm-based ensemble classification system (GA-ECS). This paper is structured as follows: GA-ECS is described in the second section, the results and discussion of GA-ECS is given in the third section, and finally the conclusion about GA-ECS is given in the fourth section.

II. The Proposed Solution

The architecture of GA-ECS is shown in Figure 1. The ensemble classification system of GA-ECS is an ensemble of ensemble classifier and AdaBoost classifier. In GA-ECS, the genetic algorithm is used to identify best features from the dataset. The identified features given by the genetic algorithm is used to train the ensemble classifier. The sub ensemble classifier that consists of decision tree, k-nearest neighbor (KNN) [30], naïve bayes [31], random forest [32], and SVM is hybridized with the AdaBoost classifier in GA-ECS. Here, the random forest classifier is used as the base classifier for the AdaBoost classification. The sub ensemble classifier in GA-ECS uses a voting based approach to compute the output from different classifiers. The GA-ECS gives the final output through getting voting from the sub ensemble classifier and AdaBoost classifier. The integration of AdaBoost and sub ensemble classifier in GA-ECS will contribute towards an enhanced classification system as it combines the outputs of several machine learning classifiers. The AI methods employed in GA-ECS are described below.

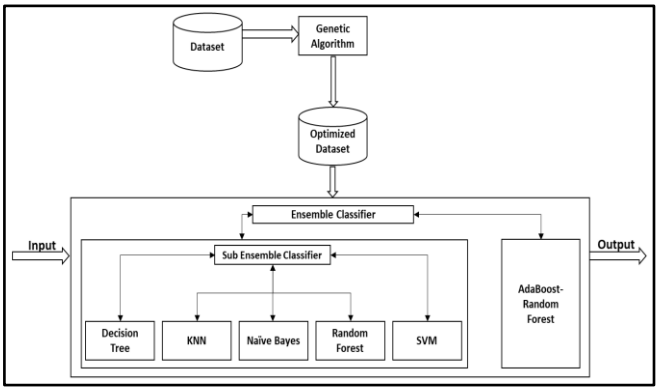

FIGURE 1. Architecture of GA-ECS.

A. Genetic Algorithm

The genetic algorithm is based on the idea of natural selection given by Charles Darwin [33]. Only the fittest individuals from different generations are strictly chosen by the natural selection process. The machine learning techniques adopt genetic algorithm for choosing the finest variables in applications related to prediction. In a series of generation, the genetic algorithm picks the successful genetics (variables) from each generation. The genetic algorithm is mostly used in data mining applications like feature selection from dataset, parameter tuning in machine learning models, etc. Figure 2 shows the process of genetic algorithm. Below are the steps of genetic algorithm in feature selection problem.

- Present a population (generation of chromosomes) from the available set of features
- Estimate the value of fitness for each individual based on some measures.
- The crossover and mutation process take place in the population.
- A new set of population is produced.
- Considering the fitness values, the finest individuals are selected from the population.
- The steps 3 to 5 are iterated till reaching the desired count of generation.

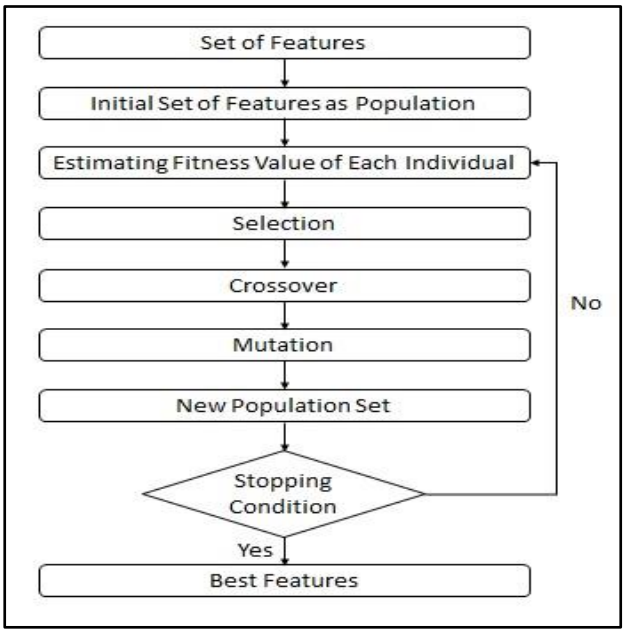

FIGURE 2. Flowchart of genetic algorithm,

B. Decision Tree

Decision tree is a classification process which uses a tree like design to make decision on the input features. The features of the training dataset are represented as nodes. The internal node and root node of decision tree represent different input features of the training dataset. The leaf or final nodes of the decision tree are considered as the output categories of training dataset. The decision tree follows a hierarchical way of making decisions on the input features by using a test condition on each node. The test condition leads to splitting of the training dataset into binary or multiple parts. The selection of node or attribute for splitting is the important step in generating decision tree. The information gain method is used widely for the selection of attributes for splitting procedure. Information gain can be computed by equation (1).

$$IG(X, A) = E(X) - \sum_a \frac{|X_a|}{|X|} E(X_a) \quad (1)$$

Here, 'X' is a set of data tuples used for training, 'A' is an attribute and 'a' is its value, 'X_a' is a subset of X consisting of the instances with A = a, and E(X) is the entropy of X. E(X) can be represented by equation (2).

$$E(X) = - \sum_{i=1}^{T} P_X(c_i) \log P_X(c_i) \quad (2)$$

Here, '$P_S(c_i)$' is calculated by the percentage of instances belonging to c_i in X, and 'T' is the number of target classes in the dataset. Some of the popularly used decision tree are classification and regression tree (CART), ID3, C4.5. CART is used as decision tree classifier in GA-ECS.

C. KNN

KNN algorithm is a classification algorithm that depends on learning by analogy. In KNN algorithm, a comparison is made between the given test data and trained data for finding the similarity. Assuming that 'n' number of attributes describe the training data tuples where each data tuple is visualized as a data point in the m-dimensional space. The KNN classifier finds the closeness of the test data with the trained data that holds position in n-dimensional space when a test data is given, the algorithm. The 'K' in KNN classifier indicates the number of closest neighbors that can be involved in voting process to determine the class of the unknown test data. The closeness between the data tuples is measured by distance function such as Euclidean distance. The Euclidean distance among the two data tuples, say, $X_1 = (x_{11}, x_{12}, x_{13}, \ldots\ldots, x_{1n})$ and $X_2 = (x_{21}, x_{22}, x_{23}, \ldots\ldots, x_{2n})$, is computed using equation (3).

$$dist(X_1, X_2) = \sqrt{\sum_{i=1}^{n} (x_{1i} - x_{2i})^2} \quad (3)$$

The KNN classifier with 'K' value equal to five is used in GA-ECS.

D. Naïve Bayes

Naive bayes algorithm is relied on bayes theorem which estimates a set of probabilities for a training dataset by examining the frequency and combination of values. The naïve bayes algorithm considers that all the features of training dataset are not dependent on each other given the value of class variable. The algorithm is named 'naïve' because of the conditional independence assumption does not suit best to all real-world applications although it achieves good accuracy and learns quickly in different supervised classification problems. Consider there are m target classes, T_1, T_2, … T_m for a given dataset, then the naïve bayes classifier identifies the data tuple 'X' belonging to class 'T_i' only if it satisfies the condition, $P(T_i|X) > P(T_j|X)$ for $1 \leq j \geq m$, $j \neq I$.

Here, $X = (x_1, x_2, x_3, \ldots\ldots, x_n)$ from n attributes $A_1, A_2, \ldots\ldots, A_n$. $P(T_i|X)$ is the posterior probability of the target class T_i, given a tuple X. $P(T_i|X)$ is computed using equation (4).

$$P(T_i|X) = \frac{P(X|T_i)\, P(T_i)}{P(X)} = (P(T_i) \prod_{k=1}^{n} P(x_k|T_i)) / P(X) \quad (4)$$

Here, $P(T_i)$ is the prior probability of the target class. $P(X|T_i)$ is probability of tuple in class T_i. P(X) is the prior probability of X. As x_k refers to the value of attribute A_k when computing $P(X|T_i)$, the attribute should be checked whether it is categorical or continuous-valued. The attribute value given to naïve bayes classifier could be numerical or categorical. If the attribute contains continuous value, then it is considered to possess gaussian distribution with a mean μ and standard deviation σ. When using continuous valued attributes, the $P(x_k|T_i)$ is computed using equation (5).

$$P(x_k|T_i) = g(x_k, \mu_{T_i}, \sigma_{T_i})) \quad (5)$$

$$\text{Here, } g(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (6)$$

The gaussian naïve bayes classifier is used in GA-ECS.

E. Random Forest

Random forest is a machine learning algorithm which is widely used for feature selection. Random forest can be viewed as enhancement of decision tree algorithm. Producing numerous small decision tree from random subsets of the data is the main concept in random forest algorithm. A random selection of attribute is used to decide the split at each node to produce the individual decision trees. Due to random selection of attribute, a biased classification can be expected from each of the decision tree. Various trends in the data are captured by each of the individual decision tree. An ensemble of trees is used by random forest for producing complicated decision boundaries. This ensemble of trees can be thought of a group of specialists where each of them are expertise in their area but don't have complete knowledge about the entire subject. The advantage is that the individual decision tree is not correlated as it does not capture all the features and hence the random forest is less inclined to over fitting. When using random forest for classification problems, each individual tree involves in voting, the class which has the majority voting is returned. The reduction of node impurity influenced by the probability of attaining that node is used to calculate feature value in random forest. The probability of node is computed from the

number of instances reaching the node divided by total number of instances. The features with high value are given higher importance.

F. SVM

SVM is a support vectors-based classification algorithm that works well for classification of both linear and nonlinear data. According to SVM algorithm, initially, the given dataset is transformed into higher dimension through nonlinear mapping. Then, the tuples of one class are separated from another class using a linear optimal separating hyperplane which is identified in the new dimension. SVM identifies the hyperplane using the support vectors and margins. The data points nearer to the hyperplane are called support vectors which defines the margins. Finally, two classes of data are separated by a hyperplane. SVM classifier is based on a kernel which is a set of computational functions that transforms the input data into the required form. SVM classifiers applies various kernel functions based on the applications. Some of the popular kernel function used by SVM are polynomial, sigmoid, nonlinear, radial basis function, and linear. A separating hyperplane can be represented using equation (7).

$$W \cdot X + c = 0 \tag{7}$$

Here, 'W' = $\{w_1, w_2, \ldots, w_n\}$ represents the weight assigned to the data point, 'n' is the number of attributes, 'X' is a set of data tuples, and 'c' is a scalar which is referred to as bias. Linear kernel based SVM classifier is used in GA-ECS.

G. AdaBoost

AdaBoost is kind of ensemble classification system that uses a weak classifier to get trained multiple times with the different datasets. The outputs of all the classification models produced from different trained weak classifiers are ensembled to give the final output. According to the AdaBoost classification method, initially, a random weight is assigned to all the instances of training dataset. Then, the training dataset is used to train the weak classifier. This process is followed by testing the trained weak classifier. Then, the training dataset is updated by weights of instances are modified according to the training error rate. The weights are incremented for the instances that are classified wrongly. The weights are decremented for the instances that are classified correctly. Then, the updated training dataset is used to train the weak classifier. This process continues with multiple repetitions. Finally, a voting method is used with all the trained weak classifiers which decides the final output for unknown cases. The weights of instances are updated in each repetition using equation (8).

$$w_k^{(r+1)} = \frac{w_k^r}{z_l} \times \begin{cases} e^{-\beta_r} \; if \; b_r(x_k) = c_i \\ e^{\beta_r} \; if \; b_r(x_k) \neq c_i \end{cases} \tag{8}$$

$$\text{Here, } \beta_r = \frac{1}{2} \ln\left(\frac{1-e_r}{e_r}\right) \tag{9}$$

In equation (8), 'r' is the repetition level, 'z_l' is the normalization factor, 'w_k' is the weight of instance 'x_k', 'b_r' is the weak classifier, 'c_i' is the output class label of the training dataset, and 'e_r' is the training error of 'b_r'. In equation (9), e_r is the training error. The final output of AdaBoost classifier for an unknown case x' is computed using equation (10).

$$\text{AdaBoost} = \underset{c}{\operatorname{argmax}} \sum_{r=1}^{q} \beta_r \, I(b_r(x') = c) \tag{10}$$

The AdaBoost classifier with random forest as base classifier is used in GA-ECS.

The Cleveland heart disease dataset is used for testing the GA-ECS. The Cleveland heart disease dataset contains 303 samples where the input features are serum cholesterol, angina, number of major vessels, gender, electrocardiographic results, blood pressure, heart rate, slope of the peak exercise, types of chest pain, thalassemia, fasting blood sugar, age, and oldpeak. The output categories of the heart disease dataset are 'below 50% narrowing' and 'above 50% narrowing'.

III. RESULTS AND DISCUSSIONS

The performance measures: accuracy, precision, sensitivity, specificity, and F-measure are used to test the performance of machine learning classifiers for the optimized medical datasets obtained from RF-EMLC method. The performance measures: accuracy, sensitivity, and specificity are estimated using equations (11), (13), and (14), respectively [34]. The precision is estimated using equation (12) [35]. Here, TNE, TPO, FNE, and FPO represent the total observations that are recognized correctly for the negative class, total observations that are recognized correctly for the positive class, total observations that are recognized wrongly for the negative class, and total observations that are recognized wrongly for the positive class, respectively.

$$\text{Accuracy} = \frac{\text{TPO+TNE}}{\text{TPO+TNE+FPO+ FNE}} \tag{11}$$

$$\text{Precision} = \frac{\text{TPO}}{\text{TPO+FPO}} \tag{12}$$

$$\text{Sensitivity} = \frac{\text{TPO}}{\text{TPO+FNE}} \tag{13}$$

$$\text{Specificity} = \frac{\text{TNE}}{\text{TNE+FPO}} \tag{14}$$

Figures 3, 4, 5, 6, 7, and 8 show the receiver operating characteristics (ROC) graphs [36] of decision tree, KNN, naïve bayes, random forest, SVM, and AdaBoost classifiers, respectively. Table I shows the performances of machine learning algorithms without the feature selection process. Table II shows the performances of single machine learning classifiers, sub ensemble classifier, AdaBoost classifier, and GA-ECS after the feature selection process. As shown in Table I and Table II, the machine learning classifiers when testing with whole dataset without any feature selection method showed lower performance than the classifiers that uses genetic algorithm-based feature selection method. The GA-ECS shows the classification accuracy of 90% for classification of heart disease.

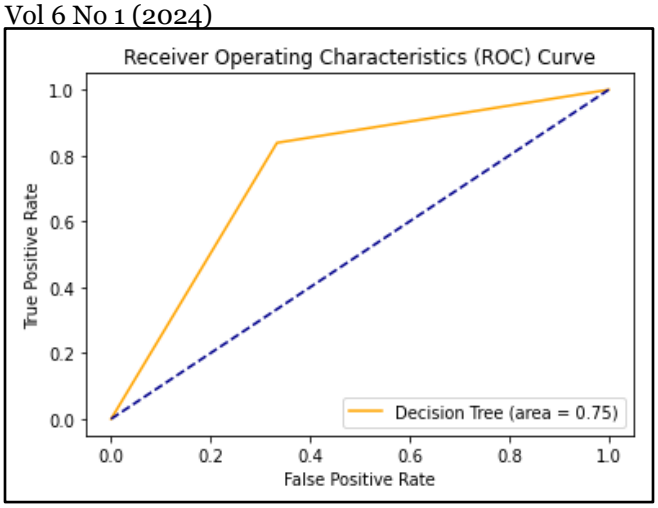

FIGURE 3. Performance graph of decision tree.

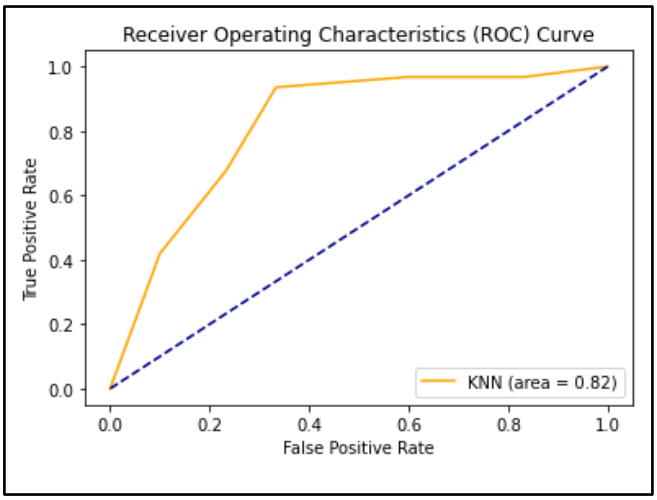

FIGURE 4. Performance graph of KNN.

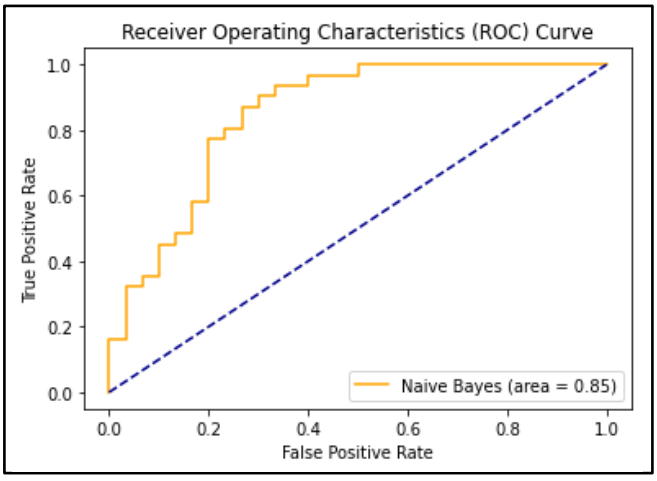

FIGURE 5. Performance graph of naïve bayes.

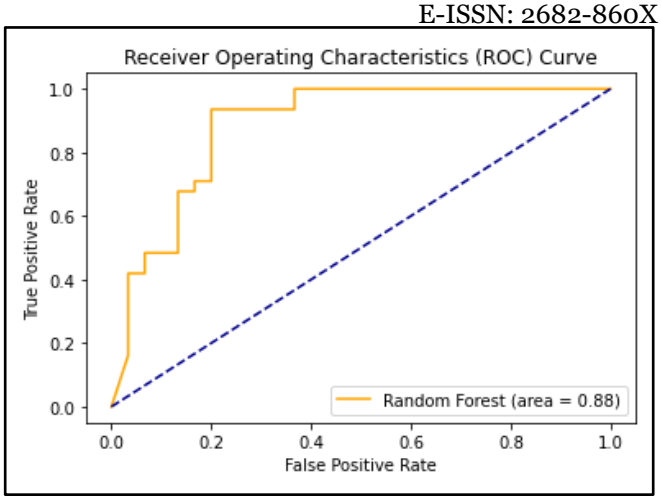

FIGURE 6. Performance graph of random forest.

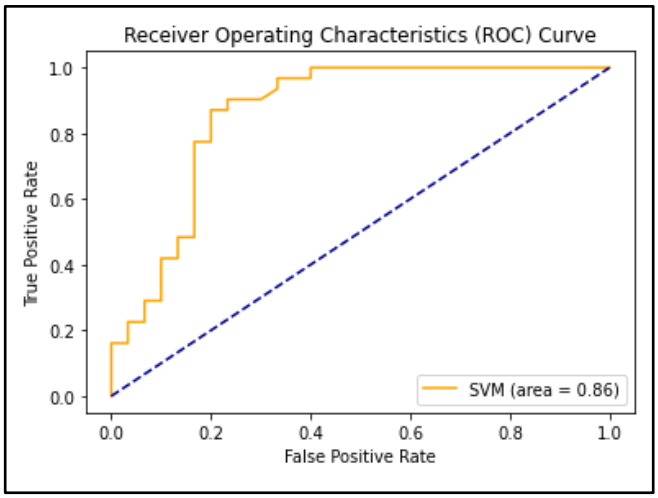

FIGURE 7. Performance graph of SVM.

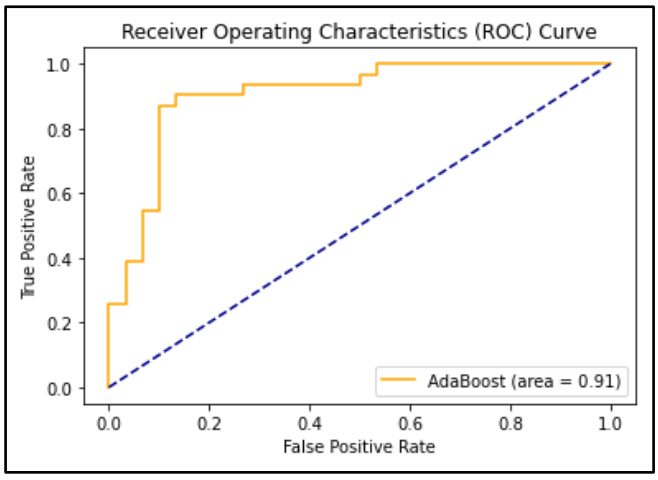

FIGURE 8. Performance graph of AdaBoost.

TABLE 1 Performance of Classifiers Before Feature Selection Process.

Classifier	Accuracy	Precision	Sensitivity	Specificity
Decision tree	0.72	0.73	0.71	0.73
KNN	0.57	0.57	0.68	0.47
Naïve bayes	0.77	0.76	0.81	0.73
Random forest	0.77	0.73	0.87	0.67
SVM	0.74	0.71	0.81	0.67

TABLE 2 Performance of Classifiers After Feature Selection Process.

Classifier	Accuracy	Precision	Sensitivity	Specificity
Decision tree	0.75	0.72	0.84	0.67
KNN	0.8	0.74	0.94	0.67
Naïve bayes	0.79	0.78	0.81	0.77
Random forest	0.82	0.76	0.94	0.7
SVM	0.8	0.74	0.94	0.67
Sub ensemble classifier	0.84	0.78	0.94	0.73
AdaBoost-random forest	0.87	0.85	0.9	0.83
GA-ECS	0.9	0.9	0.9	0.9

The decision tree classifier showed lower performance when compared to other machine learning classifiers for the classification of heart disease based on analyzing the ROC graphs and accuracy score given in Table II. The AdaBoost classifier showed better performance when compared to other machine learning classifiers for the classification of heart disease based on analyzing the ROC graphs and accuracy score given in Table II. The GA-ECS showed better accuracy score than AdaBoost classifier for the classification of heart disease because of the integration of AdaBoost with the sub ensemble classifier.

IV. CONCLUSION

This paper proposed an effective classification system, GA-ECS for the diagnosis of heart disease. It can be concluded that the feature selection process can enhance the performance of the classification system for the diagnosis of heart disease as shown through the experiments of GA-ECS with the Cleveland heart disease dataset. The genetic algorithm based on a random forest classifier was able to identify the best features for which the GA-ECS showed good performance on the classification of heart disease with an accuracy of 90% for the Cleveland heart disease dataset. It can also be concluded that the AdaBoost classification system with the random forest as a base classifier could enhance the overall efficiency of the classification system when the AdaBoost is hybridized with the ensemble machine learning classifier. The components of the ensemble classifier should be selected appropriately based on the given dataset. Future work will be extending the architecture of GA-ECS by integrating it with deep learning methods and analyzing its performance for the classification of various medical datasets.

ACKNOWLEDGMENT

This work has been supported by the Faculty of Engineering and Technology, Multimedia University, Melaka, Malaysia.

REFERENCES

[1] M. Y. Kamil, "A deep learning framework to detect covid-19 disease via chest X-ray and CT scan images," *International Journal of Electrical & Computer Engineering,* vol. 11, no. 1, pp. 2088-8708, 2021.

[2] M. Al-Smadi, M. Hammad, Q. B. Baker, and S. A. Al-Zboon, "A transfer learning with deep neural network approach for diabetic retinopathy classification," *International Journal of Electrical and Computer Engineering,* vol. 11, no. 4, p. 3492, 2021.

[3] T. T. Ramanathan, M. J. Hossen, and M. S. Sayeed, "Blockchain integrated multi-agent system for breast cancer diagnosis," *Indonesian Journal of Electrical Engineering and Computer Science,* vol. 26, no. 2, p. 998, 2022.

[4] Y. Dahdouh, A. Anouar Boudhir, and M. Ben Ahmed, "A new approach using deep learning and reinforcement learning in healthcare: skin cancer classification," *International Journal of Electrical and Computer Engineering Systems,* vol. 14, no. 5, pp.557-564, 2023.

[5] J. Wu and C. Hicks, "Breast cancer type classification using machine learning," *Journal of Personalized Medicine,* vol. 11, no. 2, p. 61, 2021.

[6] A. Ishaq, S. Sadiq, M. Umer, S. Ullah, S. Mirjalili, V. Rupapara, and M. Nappi, "Improving the prediction of heart failure patients' survival using SMOTE and effective data mining techniques," *IEEE Access,* vol. 9, pp. 39707-39716, 2021.

[7] H. Hamdani, H. R. Hatta, N. Puspitasari, A. Septiarini, and H. Henderi, "Dengue classification method using support vector machines and cross-validation techniques," *IAES International Journal of Artificial Intelligence,* vol. 11, no. 3, p. 1119, 2022.

[8] M. A. Elsadig, A. Altigani, and H. T. Elshoush, "Breast cancer detection using machine learning approaches: a comparative study," *International Journal of Electrical & Computer Engineering,* vol. 13, no. 1, pp. 2088-8708, 2023.

[9] S. Asadi, S. Roshan, and M. W. Kattan, "Random forest swarm optimization-based for heart diseases diagnosis," *Journal of Biomedical Informatics,* vol. 115, p. 103690, 2021.

[10] P. Ghosh, S. Azam, M. Jonkman, A. Karim, F. J. M. Shamrat, E. Ignatious, ... F. de Boer, "Efficient prediction of cardiovascular disease using machine learning algorithms with relief and LASSO feature selection techniques," *IEEE Access,* vol. 9, pp. 19304-19326, 2021.

[11] S. Suresh, D. T. Newton, T. H. Everett IV, G. Lin, and B. S. Duerstock, "Feature selection techniques for a machine learning model to detect autonomic dysreflexia," *Frontiers in Neuroinformatics,* vol. 16, p. 901428, 2022.

[12] J. H. Joloudari, H. Saadatfar, A. Dehzangi, and S. Shamshirband, "Computer-aided decision-making for predicting liver disease using PSO-based optimized SVM with feature selection," *Informatics in Medicine Unlocked,* vol. 17, p. 100255, 2019.

[13] S. B. Talip, "The knowledge, perception and attitude of year 1 undergraduate students in the faculty of medicine and health Sciences (FMHS), universiti Malaysia Sarawak (UNIMAS) on coronary artery disease (CAD)," *Trends in Undergraduate Research,* vol. 5, no. 1, b9-14, 2022.

[14] S. C. Ong and J. Z. Low, "Financial burden of heart failure in Malaysia: a perspective from the public healthcare system," *Plos One,* vol. 18, no. 7, e0288035, 2023.

[15] V. V. Ramalingam, A. Dandapath, and M. K. Raja, "Heart disease prediction using machine learning techniques: a survey," *International Journal of Engineering & Technology,* vol.

7, no. 2.8, pp. 684-687, 2018.
[16] D. Dua and C. Graff, "UCI machine learning repository," *Irvine, CA: University of California, School of Information and Computer Science,* 2019, accessed 4 Jul 2023, <http://archive.ics.uci.edu/ml>.
[17] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning,* vol. 20, no. 3, pp. 273-297, 1995.
[18] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain," *Psychological Review,* vol. 65, no. 6, p. 386, 1958.
[19] A. K. Faieq and M. M. Mijwil, "Prediction of heart diseases utilising support vector machine and artificial neural network," *Indonesian Journal of Electrical Engineering and Computer Science,* vol. 26, no. 1, pp. 374-380, 2022.
[20] C. Pan, A. Poddar, R. Mukherjee, and A. K. Ray, "Impact of categorical and numerical features in ensemble machine learning frameworks for heart disease prediction," *Biomedical Signal Processing and Control,* vol. 76, p. 103666, 2022.
[21] R. E. Schapire, "Explaining AdaBoost," Berlin/Heidelberg: Springer, 2013, pp. 37-52.
[22] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, "Classification and regression trees," 1st ed. Boca Raton, Florida, USA: CRC Press, 1984.
[23] G. -B. Huang, Q. -Y Zhu, and C. -K Siew, "Extreme learning machine: theory and applications," *Neurocomputing,* vol. 70, no. 1-3, pp. 489-501, 2006.
[24] J. H. Friedman, "Stochastic gradient boosting," *Computational Statistics and Data Analysis,* vol. 38, no. 4, pp. 367-378, 2002.
[25] P. S. Pandian, R. RubeshSelvaKumar, R., Muneeswaran, and S. V. Mayil, "Identification of heart disease in beginning stage using gradient boosting classification," *International Journal of Applied Engineering Research,* vol. 18, no. 2, pp. 136-139, 2023
[26] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE ICNN'95-International Conference on Neural Networks, Perth, WA, Australia,* vol. 4, pp. 1942-1948, 1995.
[27] D. Saputra, W. S. Dharmawan, and W. Irmayani, "Performance comparison of the SVM and SVM-PSO algorithms for heart disease prediction," *International Journal of Advances in Data and Information Systems,* vol. 3, no. 2, pp. 74-86, 2022.
[28] F. Murtagh, "Multilayer perceptrons for classification and regression," *Neurocomputing,* vol. 2, no, 5-6, pp.183-197, 1991.
[29] A. Al Bataineh and S. Manacek, "MLP-PSO hybrid algorithm for heart disease prediction," *Journal of Personalized Medicine,* vol. 12, no. 8, p. 1208, 2022.
[30] N. S. Altman, "An Introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician,* vol. 46, no. 3, pp. 175-185, 1992.
[31] G. I. Webb, "Naïve bayes," *Encyclopedia of Machine Learning,* vol. 15, pp. 713-714, 2010.
[32] L. Breiman, "Random forests," *Machine Learning,* vol. 45(1), pp. 5-32, 2001.
[33] D. L. Hull, "Darwin and His Critics: The Reception of Darwin's Theory of Evolution by the Scientific Community," *Chicago: University of Chicago Press,* 1973.
[34] W. Zhu, N. Zeng, and N. Wang, "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations," in *NESUG Proceedings: Health Care and Life Sciences, Baltimore, Maryland,* 2010.
[35] D. M. W. Powers, "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation," *Journal of Machine Learning Technologies,* vol. 2, no. 1, pp. 37–63, 2011.
[36] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters,*" vol. 27, no. 8, pp. 861-874, 2006.