IMPROVED ARTIFICIAL NEURAL NETWORK THROUGH METAHEURISTIC METHODS AND ROUGH SET THEORY FOR MODERN MEDICAL DIAGNOSIS

Ambily Merlin Kuruvilla

Research Scholar, Department of Computer Science,
Karpagam Academy of Higher Education, Coimbatore.
E mail: ambilykuruvilla@gmail.com

Dr. N.V. Balaji

Dean, Faculty of Arts, Science and Humanities,
Karpagam Academy of Higher Education, Coimbatore.
E mail: balajinv@karpagam.com

Abstract- A novel meta-heuristic soft computing model with feature selection implemented using Rough set (RS) theory for the diagnosis of coronary artery disease in diabetes patients is proposed in this study. The binary classification method in multiclass classification problems is applied by the One Versus Rest approach (OVR) is incorporated. To avoid the redundancy problem, a mathematical approach known as rough-set theory (RS) is applied to identify the most significant features from the dataset. The Artificial Neural Network with improved hidden layers is used as the classifier which is optimized through a metaheuristic population-based method, known as the Grasshopper Optimization Algorithm (GOA) with a single objective optimization approach integrated for improving the accuracy of the model. Mean Square Error (MSE) is taken as the objective function and the result shows that the accuracy of the model has been improved significantly from 89.1% to 95.25% after optimization.

Keywords: **One Versus Rest (OVR); Artificial Neural Network (ANN); Rough Set (RS); Grasshopper Optimization Algorithm (GHO); Mean Square Error (MSE).**

1.Introduction

Diabetes is a chronic condition that has the potential to lead to many side effects and heart disease is one among them. According to the report of the World Health Organization, diabetics patients have a higher probability of heart diseases. The statistics show that the death rate due to heart diseases are on a rise. Diabetes can cause damage to the blood vessels as well as the nerves which control the functions of the heart. This will lead to cardiac disorders and coronary artery problems. Heart disease being a major health hazard is an issue that immediately requires more academic research. The risk of heart disease can be avoided or reduced by controlling the diet, blood glucose level, and regular physical activity. Therefore, early prediction of the possibility of heart disease in diabetes patients will help to improve their health conditions. The main objective of the study is to find out the possibility of heart disease especially among diabetes patients. Artificial Neural network is used as the classifier for the proposed study. It is widely accepted that backpropagation in the Neural Network is a very strong tool for forecasting and estimation. The main power of the Neural Network is that we can do many complex problems with large datasets that can be estimated without any complex mathematical expressions. In Neural Networks, weight optimization plays an important role in the accuracy obtained. By using the activation function and error value obtained in each iteration, the weights of the Neural Network are adjusted until it reaches the optimum value. But during these steps, the local optimization problem can arise. To overcome this issue, metaheuristic population-based algorithms such as grasshopper optimization algorithms are used to optimize the weight of the Artificial Neural Network and the optimal weight value is used for the prediction.

Grasshopper Optimization is one among the stochastic methods of search optimization [Janmenjoy Nayak, (2020)] which will help to overcome the issue of local optima problem in Neural Network. It is developed based on the life cycle and food search pattern of insects known as Grasshoppers. During their life cycle, grasshoppers follow

two stages: The nymph and adult stage. During the Nymph stage, it follows slow movement, and during the adult stage, it follows the abrupt movement. These two movements will be adjusted in the algorithm with the help of mathematical methods. This feature of GOA helps to find out the optimum results by overcoming the stochastic gradient descent problems in the Neural Network.

Another main concept used in the study is dimensionality reduction. A mathematical foundation tool called Rough Set is used for feature selection by finding the most relevant subsets. A reduced subset retains all the features of the parent feature set. In the study, the Framingham dataset is used for classification purposes.

2.   Literature Review

Ashir Javeed and Sanam Shahla Rizvi proposed a floating window with adaptive size [Ashir Javeed, (2020)] method for dimensionality reduction purposes. They have proposed this method to create a subset feature and is given as the input to the Neural Network architecture. ANN and Deep Neural Networks are used as the classifiers. The cleveland dataset was used for the analysis. The result shows that the proposed model achieves an accuracy rate of 91.11% for ANN and 93.33% by the Deep Neural Network.

A hybrid Model of Emotional Neural Networks (EmNNs) and Particle Swarm Optimization (PSO) [Afzal Hussain Shahid, (2020)] was introduced for the detection of Coronary Artery Disease (CAD) by Afzal Hussain Shahid. Feature selection was carried out by four different methods. To improve the learning capability of the Neural Network, emotional parameters are also taken into consideration along with the conventional weights and bias.

Nithyavishnupriya in her article used DNN for heart disease prediction. They used the statistical analysis Chi-Square analysis [Nithyavishnupriya and Ramprakash (2020)] along with DNN architecture. They utilized the Dataset collected from the UCI Repository which consists of 303 instances with 14 columns.

In the model, proposed by Archana Singh, they utilized Python programming in Anaconda (Jupiter) notebook [Archana Singh (2020)] as a simulation environment. They have tested various supervised machine learning algorithms and stated that the KNN algorithm is the most suited method compared with the other classifiers they have tested.

A Machine learning-based medical intelligence system was made by Amin Ul Haq [Amin Ul Haq, (2018)]. In his study, Relief, minimal-redundancy-maximal-relevance (mRMR), Shrinkage and Selection Operator (LASSO) were used to find out the most relevant and highly correlated features. Seven supervised classification algorithms are used for the prediction. Reduced feature sets obtained from each feature selection process were given as input to the classifiers. Time for prediction is also calculated.

Awais Mehmood proposed a new method of study known as cardio to help their study. CNN algorithm [Awais Mehmood, (2021)] was used for the construction of the model. The proposed model is used to study the probability of the presence of cardiovascular disease in a patient and the model achieved an accuracy of 97%.

Zubar in his studies proposes a hybridized method using KNN and the Spiral optimization method [A. H. Zubar and R. Balamurugan, (2020)] for the classification of cardiovascular diseases. The method enhances the clustering quality and also PCA is used for dimensionality reduction.

The author of the article [Irfan Javid, (2020)] used deep learning methods along with conventional machine learning techniques to predict heart disease. A voting-based method is adopted to strengthen the weak classifiers. LSTM is added along with RNN and is used as a deep neural network. 85.71% accuracy is achieved by the proposed model.

Kalaselvi and G N Nasira analysed the complex relationship between cancer and diabetes in their study. They propose a novel method known as ANIFS [Kalaiselvi and G.N. Nasira, (2014)] in their study. They have achieved 80% accuracy in the PIMA dataset using ANIFIS with adaptive KNN approach.
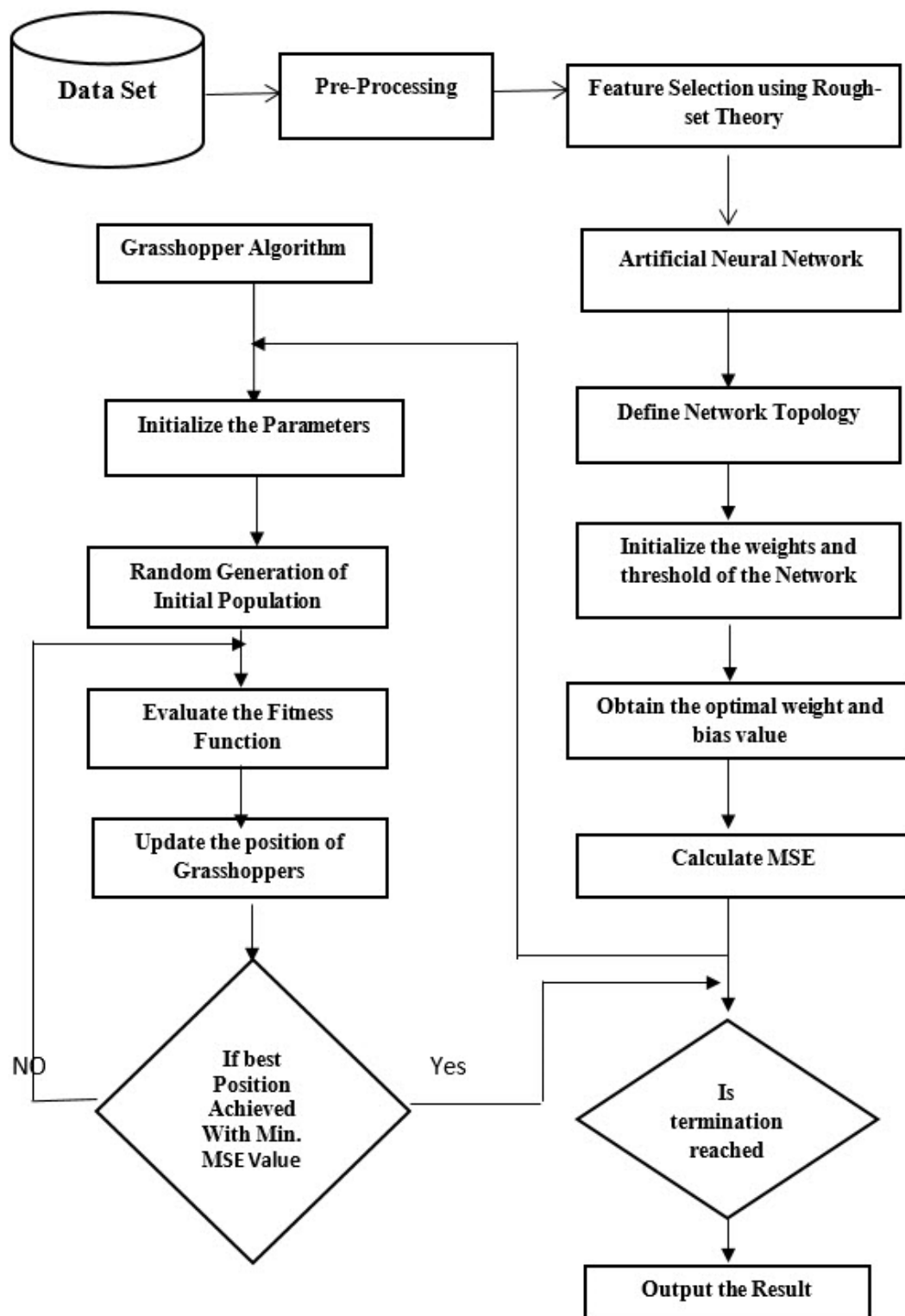
3.Proposed Methodology

Figure 1: Work Flow of GOA based Artificial Neural Network for Coronary artery disease.

The proposed work consists of three phases. In the first phase, the Rough Set-based feature selection approach finds the most significant subset attributes from the dataset. In the second phase, Artificial Neural Network with improved hidden layers is used for the classification. LSTM is added along with the hidden layers of the Neural Network. In the third phase where a metaheuristic algorithm known as the Grasshopper Optimization algorithm is used for optimizing the model.

By using Rough-set theory the most relevant features were extracted [Youness, (2018)] from the dataset for the classification. The reduced dataset is given as input to the Neural Network. Weights and bias of the hidden layers are initialized by the random initialization methods. Mean Square Error (MSE) is calculated. A single-objective optimization method is followed in the study. Grasshopper Optimization algorithm, optimizes the model, by setting the objective function as MSE.

In the beginning, the initial population is constructed using the random initialization method, and the fitness value for the population is calculated. Positions of grasshoppers are adjusted and the new fitness value is calculated. The process repeated until it received the best fit value. The best fit value is returned to the Neural Network. GOA computes the optimized weights by adjusting the initial parameters and the fitness value for each solution is evaluated. The best fit solution is returned to the Neural Network.

3.1. Rough set-based feature selection

Rough Set Theory (RS) is a mathematical tool [Javad Rahimipour Anaraki, (2013)] that successfully determines the data dependencies among attributes of a dataset and reduces the dataset through structural approach. The features selected by the rough set produce very useful and informative results while other attributes can be removed with less loss of information, as they have a high degree of dependency with the other attributes in the dataset. This will help to reduce redundancy. In this work, the dataset with discretized values of attributes are used to discover the subset of all attributes using the concept of RS, which selects the attributes that provide more information helping to predict the most relevant class attributes accurately.

3.1.1. Fundamental notations for attribute reduction in rough set theory.

In this study, all information is stored in the table format. The dataset contains all the information corresponding to the Information System [Richard Jensen, (2007)]. Assume that the Information System IS = (DS, A) where DS is the set of instances in the dataset A and Att is the non-empty set of finite attributes and att: $DS{\rightarrow}SV_{att}$ for each att $\in$ Att, where SV_{att} is the set of values that an attribute att may have. With any IP→Att, it has its associated equivalence relation, ie (Independency) which is denoted as IND(IP) [Kanchan Shailendra Tiwari (2013)] as shown in the equation below [Javad Rahimipour Anaraki, (2013)]

$$IND(IP) = \{(u, v){\in}DS^2| \ \forall att \in IP, att(u) = att(\ v)\} \qquad (1)$$

IND(IP) partitions the instances corresponding to the dataset and it is represented as DS / IND (IP) (or *DS / IP*). If (u, v)) $\in$ IND(IP) then u and v are indiscernible by attribute of IP. In rough set-based feature selection dependency between the attributes is calculated, and if there is interdependency between the attributes exists, these features are removed. So that it can avoid redundancy in the dataset.

The attribute in an information system is classified into different categories using roughest. Decision attributes [Omnia S. Elazab and Hany M. Hasanien, (2020)] decides in which class the attributes belong at the same time. Conditional attributes, do not decide the class of an object, but it will help to decide the class which belongs.

Let $Y \subseteq DS$. Then y can be approximated by calculating the information contained in IP by constructing IP Upper and IP lower approximations [Richard Jensen, (2007)]. IP-lower. and IP-upper approximations [Richard Jensen, (2007)] of Y as formulated as

$$\underline{IP}Y = \{y|[Y]_{IP} \subseteq Y\} \qquad (2)$$

$$\overline{IP}Y = \{y|[Y]_{IP} \cap Y \neq \emptyset\} \qquad (3)$$

$\underline{IP}Y$ represents the lower approximation and $\overline{IP}Y$ represents the upper approximations. The attribute values belonging to these regions can also be considered in the subset.

Assume that the IP and IQ are equivalence relations over DS, then their positive area can be represented as

$$PSV_{IP}(IQ) = \bigcup_{Y \epsilon DS/IQ} \underline{IP}Y \qquad (4)$$

The positive region is comprised of all instances of DS which can be classified to classes of DS/IQ with the information in attributes IP. With this RST defines the degree of dependency of a set of attributes IQ on a set of IP attributes is formulated as For IP, IQ $\subset$ Att, it is signified as IQ depends on IP in a degree l $(0 \le l \le 1)$ represented as IP→IQ if

$$l = \gamma_{IP}(IQ) = \frac{|PSV_{IP}(IQ)|}{|DS|} \qquad (5)$$

The dimensionality of attributes is accomplished by associating equivalence relations produced by sets of attributes. Once the redundant attributes are removed as the reduced set of attributes provides the same predictive ability of decision feature as the original dataset.

3.1.2 Algorithm.

For attribute reduction, an RS-based Quick Reduct [Javad Rahimipour Anaraki, (2013)] Algorithm is implemented. Att is a set of all features, CF is the conditional variable and DF is the decision features [Richard Jensen (2007)] RST denotes the reduced dataset corresponding to the conditional attributes CF.

RST ← {}

Do

T ← RST

$\forall y \in (CF - RST)$

If $(\gamma_{RST \cup \{y\}}(DF) > \gamma_T(DF))$ then

T← RST ∪ {y}

RST ← T

until $\gamma_{RST}(DF) = \gamma_{CF}(DF)$

Return RST

The reduced set of features were computed by the algorithm without exhaustively producing all probable subsets. It begins with an empty set and starts adding attribute one in the incremental order [Javad Rahimipour Anaraki, (2013)], attributes are chosen, which consequence in the greatest increase in rough set dependency, until it generates maximum possible value for the dataset.

3.2 Artificial Neural Network

A kind of computing approach and artificial intelligence which simulates the functioning and analyzing the behavior of the human brain is known as an artificial neural network (ANN). It has the ability of self-learning which enables it to generate better results when the dataset is voluminous. ANN is composed of three different layers. Each layer in turn comprised of processing unit sets. ANN has an input layer, output layer, and hidden layer. Figure2 shows the overall structure of a standard Artificial Neural Network. Each layer nodes are connected to the forwarded layer nodes through the links. The link among neural nodes or perceptrons indicates the stream of information passing from the previous layer node to the forwarding layer node. The neural network function receives the predictor values at the left side and the computations are done on the hidden layer and the output layer shows the predicted outcome.
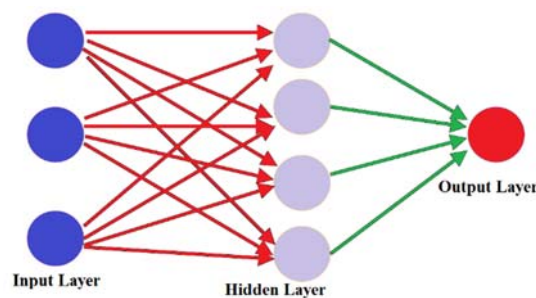

Figure2: General Structure of an Artificial Neural Network

The Figure 3 depicts the weight and bias assignment for each Neuron in an Artificial Neural Network. The hidden nodes receive the value from the previous layer and which is multiplied by weights, whose values are predetermined. The weighted input values received from all the nodes of the previous layer are summed together to generate a single value as shown in Figure 3. Next, an activation function known as ReLU is applied to the summed values which control the output of the nodes. Same as the hidden layer, the output layer's active node also does the relevant computations and produces the output. If the output value is binary then the sign of positive represented the presence and negative represented the absence which is highly dependent on the input data. The input nodes are passive, its function is to receive the input from the outside world and pass it to the next layer,

the hidden layer performs the multiplication of the received value with its concern weight assigned on each.
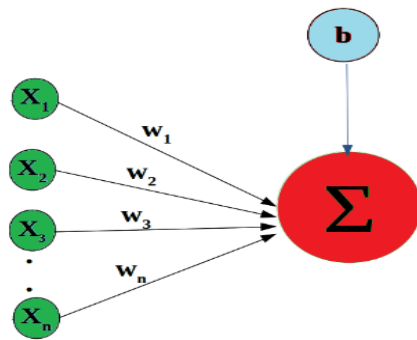

Figure 3: Assignment of weights and bias [20] of a Neuron in an Artificial Neural Network.

The weight for an input corresponding to an individual node is calculated as in the Equation (6)

$$netwrk_j = \sum_{i=1}^{n} w\,gt_{ij}.\,O_i \qquad (6)$$

where wgt_{ij} signifies weights between two nodes i and j, and output from the unit 'i' is denoted as O_i. To generate the final output 'O_t' the sum is passed to the nonlinear filter φ called activation function which generates the output given in Equation (7).

$$O_t = \varphi(netwrk_j) \qquad (7)$$

where ϕ refers to activation function and the weighted sum of input values prior to the signal processes to the subsequent layer. Finally, the generated output is compared with the expected output and their difference is assigned as an error. The standard neural network uses the backpropagation method to update the weights on trial and error basis repeatedly until the accepted error rate is obtained. This result determining the appropriate weights to produce more accurate results. The error rate is defined in equation (8)

$$\text{Error} = \frac{1}{2}\sum_k (desireout_k - resultout_k) \qquad (8)$$

3.3 Grass Hopper Optimization Algorithm.

The Grasshopper Algorithm is a bio inspired metaheuristic method which mimics the life cycle of grasshopper insects. The life cycle of grasshopper consists of two different stages: Nymph and Adult. Even though the grasshoppers are seen individually, they will foam large swarms during their Nymph and Adult stage. The movement of this swam during the Nymph stage is very slow compared to that of the adult stage. At the adult stage, they will show long-range abrupt movements [Ali Asghar Heldari, (2018)].

Based on these movements of grasshoppers, the search process in the GOA algorithm usually follows two methods: exploration and exploitation. During exploration, an abrupt movement is shown by the search agents in exploration [Peng Chen and Xiaoqiang Zhao, (2020)] but a local movement is followed during exploitation. The swarming [Ali Asghar Heldari, (2018)] behavior of Grasshoppers can be mathematically expressed in the formulae given below:

$$Y_j = S_j + G_j + W_j \qquad (9)$$

Where Y_j defines the position [Ali Asghar Heldari, (2018)] of the individual grasshopper" j". S_j defines the social interaction of the jth individual, G_j defines the gravitational force of the jth individual and W_j defines the advection. The social interaction between the grasshopper is one of the main parameters for the search mechanism. It can be computed by the mathematical formulae,

$$S_i = \sum_{j=1\ j\#i}^{N} S(A_{ij})\hat{A}_{ij} \qquad (10)$$

A_{ij} is the distance between [Ahmed A. Ewees and Zhang Jianhua (2020)] the grasshopper i and j which can be mathematically expressed as $A_{ij} = |\ Y_i - Y_j\ |$. $\hat{A}_{ij}$ is a singular vector between grasshopper i and j and is mathematically denoted as

$$\hat{A}_{ij} = (Y_j - Y_i)/\ A_{ij} \qquad (11)$$

The social force [Omnia S. Elazab and Hany M. Hasanien, (2020)] between grasshoppers are mathematically calculated as

$$S(r) = f\,e^{(-r/l)} - e^{-r} \qquad (12)$$

The intensity of gravitation is represented by f and l denotes the length of the social interaction. The impact of the interaction between the grasshoppers is denoted as S. A parameter C represents the mathematical formulation that can be used for balancing the movements [Omnia S. Elazab and Hany M. Hasanien, (2020)] of search agents.

$$C = (\text{Maximum Iteration}) - (\text{Current Iteration})\frac{\text{Maximum Iteration} - \text{Mininmum Iteration}}{\text{Total No of Iteration}} \qquad (13)$$

3.3.1Fitness Evaluation.

The model follows a single objective optimization method. For finding the best fit solution, the fitness value of each solution is calculated. MSE is taken as the parameter for fitness evaluation. The experiment is repeated for 1000 iterations and the best fit value shows minimum MSE. MSE can be calculated using the Equation (14)

$$\text{MSE} = \frac{1}{n}\,(\,r_j - \hat{r})^2 \qquad (14)$$

Where '**r**' represents the actual result and '**r̂**' represents the predicted result, the total number of instances is denoted as 'n'. The experiment is repeated for 1000 iterations and the best fit value is taken as the one with minimum MSE value. The fitness value corresponding to different iterations is plotted in Figure 4.
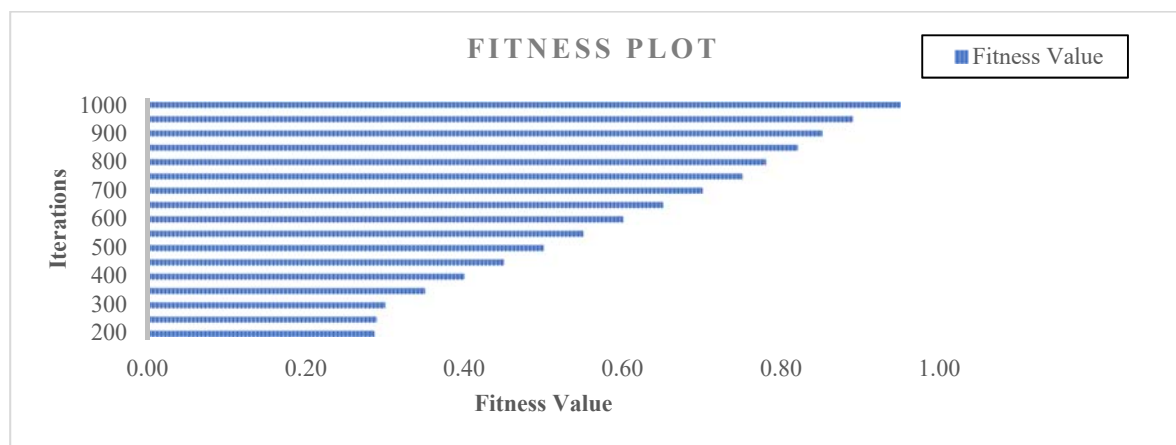

Figure 4: Fitness Plot generated by the Grasshopper Algorithm

4. Experimental Results and Discussions

4.1 Dataset description

Framingham Dataset from Kaggle Machine Learning Repository is used for the study. The dataset is comprised of 4240 instances with 16 attributes. They are gender information, age of the person, education of the person, current Smoker or not, cigarettes Per Day, Blood Pressure, prevalent Stroke, Prevalent Hyper Tension, diabetes, total Cholesterol, Systolic blood pressure, Diastolic blood pressure, Body Mass Index, heartbeat Rate, Blood glucose level with the class attribute Chronic Heart Disease   present or not.

To predict diabetic heart disease, a new class variable is generated to label the patients in four different risk levels such as without diabetic without heart disease as 0, presence of diabetic and absence of heart disease as 1, absence of diabetic, and presence of heart disease as 2 and presence of both diabetes and heart disease as 3. Inter Quartile Range is used to preprocess the dataset in this study. Ambiguous and erroneous data is replaced by the mathematical and statistical approach used in IQR.

4.2 Results & discussions

This work is implemented with the help of two different class variables in the dataset, they are responsible for generating the new class variable which determining the presence or absence of diabetes and the consequent heart failure possibilities among the patients. A new class variable is generated based on the values of these two class variables.

The proposed model RS-GANN is designed using Python and its performance is examined and compared with other prediction models like Naive Bayes and Logistic Regression. The parameters used for evaluating the performance of the model are Precision, Recall, F1-Score, Specificity and, Accuracy. The results are discussed in Table 1. The comparison of results is represented in a graphical format using Figure5. The proposed model is compared with the model before optimization (RS-ANN). Comparison results are shown in Table2 and Figure 6. The mathematical formulae for calculating the performance measurement parameters are given below.

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}. \tag{15}$$

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{16}$$

$$\text{Specificity} = \frac{True\ Negative}{True\ Negative + False\ Positive} \tag{17}$$

$$\text{F1 Score} = 2*(\ (\text{Precision}* \text{Recall}) / (\text{Precision} + \text{Recall})\ ). \tag{18}$$

$$\text{Accuracy} = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + FAlse\ Positive + False\ Negative} \tag{19}$$

Algorithm	Precision	Recall	Specificity	F1-Score	Accuracy
Navie Bayes	97.8	79.9	84.5	88.0	80.41
Logistic Regression	98.3	86.1	87.4	91.8	86.28
RS-GANN	99.1	95.6	92.2	97.3	95.25

Table1: showing the Performance of Comparison of Classifiers

The rate of positive predicted value can be explained by the parameter Precision. The recall is the result of the analysis which explains the percentage of relevant values predicted by the model. Specificity measures the ability of a classifier to identify non-disease individuals. The weighted average of Precision and Recall can be represented as the F1-Score of the Classifier. The accuracy of the model is computed by the mathematical formulae given in equation (19).

Algorithm	Precision	Recall	Specificity	F1 -Score	Accuracy
RS-ANN (Before Optimization)	98.8	88.9	90.9	93.6	89.1
Proposed Model	99.1	95.6	92.2	97.3	95.25

Table 2: Comparison of Classifier Performance
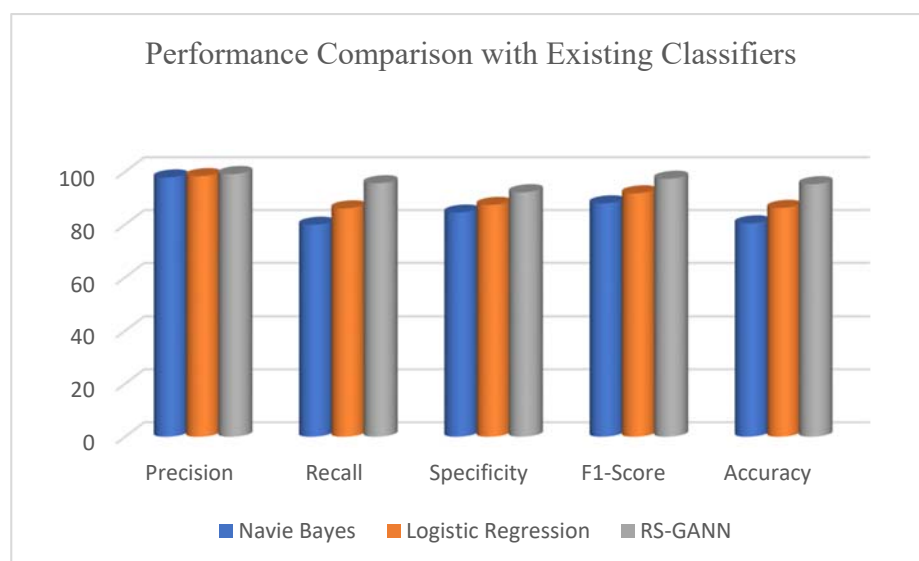
Figure 5:   Performance Comparison with Existing Classifiers

Ambily Merlin Kuruvilla et al. / Indian Journal of Computer Science and Engineering (IJCSE)
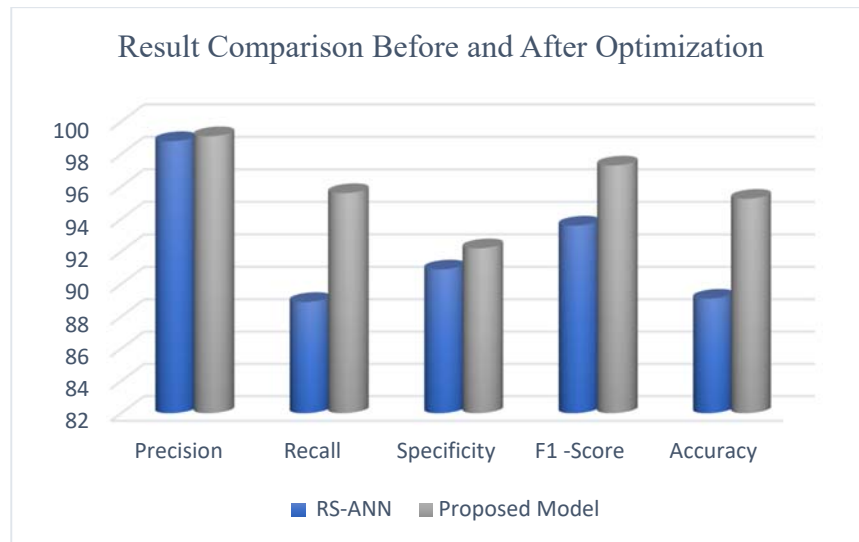

Figure 6: Comparison of Result Before and After Grasshopper Optimization

5. Conclusion.

In this study, a novel soft computing model is proposed for heart disease prediction among diabetes patients based on Rough set-based feature selection with an Artificial Neural Network optimized using a metaheuristic stochastic algorithm known as grasshopper optimization. Evaluation of results is plotted in tabular as well as in the graphical format. The results revels that the proposed system shows better performance than the existing algorithms. After executing the proposed methodology, it is found that the classifier shows an improvement of 6.15% accuracy, 0.3% in precision, 6.7% recall, 3.7% in F1-Score, and 1.3% in Specificity while compared with the performance before optimization and that is a significant improvement.

References

[1] Javad Rahimipour Anaraki, Mahdi Eftekhari(2013): Rough set-based feature selection: A Review IEEE, The 5th Conference on Information and Knowledge Technology.
[2] Ashir Javeed, Sanam Shahla Rizvi, Shijie Zhou, Rabia Riaz, Shafqat Ullah Khan, and Se Jin Kwon (2020): Heart Risk Failure Prediction Using a Novel Feature Selection Method for Feature Refinement and Neural Network for Classification. Personal Communication Technologies for Smart Spaces, Hindawi.
[3] Afzal Hussain Shahid, M.P. Singh (2020): A Novel Approach for Coronary Artery Disease Diagnosis using Hybrid Particle Swarm Optimization based Emotional Neural Network. IEEE.
[4] Janmenjoy Nayak, Bighnaraj Naik, Paidi Dinesh, Kanithi Vakula, Pandit Byomakesha Dash2(2020): FireflyX Algorithm in Biomedical and Health Care: Advances, Issues and Challenges. springer, SN Computer Science.
[5] Youness Khourdifi, Mohamed Bahaj (2018): Heart Disease Prediction and Classification Using Machine Learning Algorithms Optimized by Particle Swarm Optimization and Ant Colony Optimization. International Journal of Intelligent Engineering & Systems.
[6] S. Nithyavishnupriya, P. Ramprakash,, Sarumathi (2020): "Disease Prediction Using Deep Neural Network ", IEEE.
[7] Archana Singh, Rakesh Kumar (2020). Heart Disease Prediction Using Machine Learning Algorithms. International Conference on Electrical and Electronics Engineering (ICE3-2020), ieeexplore.ieee.org / document/9122958.
[8] Rahul Katarya, Sunit Kumar Meena (2020): Machine Learning Techniques for Heart Disease Prediction A Comparative Study and Analysis. Health and Technology, Springer Nature.
[9] Amin Ul Haq, Jian Ping Li, Muhammad Hammad Memon Shah Nazir and Ruinan Sun (2018): A Hybrid Intelligent System Framework for the Prediction of Heart Disease Using Machine Learning Algorithms. Wearable Technology and Mobile Applications for Healthcare, Hindawi.
[10] Richard Jensen, Qiang Shen (2007): Rough Set-Based Feature Selection: A Review. ResearchGate.
[11] Kanchan Shailendra Tiwari, Ashwin Kothari (2013). Design and Implementation of Rough Set Algorithms on FPGA: A Survey. International Journal of Advanced Research in Artificial Intelligence, 2013.
[12] Awais Mehmood, Munwar Iqbal, Zahid Mehmood, Aun Irtaza, Marriam Nawaz, Tahira Nazir & Momina Masood (2021): Prediction of Heart Disease Using Deep Convolutional Neural Networks. Springer.
[13] A. H. Zubar, R. Balamurugan(2020): Green Computing Process and its Optimization Using Machine Learning Algorithm in Healthcare Sector. Mobile Networks and Applications, Springer.
[14] Irfan Javid, Ahmed Khalaf Zager Alsaedi, Rozaida Ghazali (2020): Enhanced Accuracy of Heart Disease Prediction using Machine Learning and Recurrent Neural Networks Ensemble Majority Voting Method. (IJACSA) International Journal of Advanced Computer Science and Applications.
[15] C. Kalaiselvi, Dr. G.M. Nasira(2014): A New Approach for Diagnosis of Diabetes and Prediction of Cancer using ANFIS. 2014 World Congress on Computing and Communication Technologies.
[16] Peng Chen, Xiaoqiang Zhao, Qixian Zhu (2020): A novel classification method based on ICGOAKELM for fault diagnosis of rolling bearing. Applied Intelligence.
[17] Omnia S. Elazab, Hany M. Hasanien, Ibrahim Alsaidan, Almoataz Y. Abdelaziz, S. M. Muyeen (2020): Parameter Estimation of Three Diode Photovoltaic Model Using Grasshopper Optimization Algorithm. Energies.
[18] Ali Asghar Heldari, Hossam Faris (2018): An Efficient Hybrid Multilayer Perceptron Neural Network with Grasshopper Optimization. Soft Computing, Springer.

[19] Ahmed A. Ewees, Zhang Jianhua (2020): Improving multilayer perceptron neural network using chaotic grasshopper optimization algorithm to forecast iron ore price volatility, Resources Policy, Elsevier.

[20] https://towardsdatascience.com/whats-the-role-of-weights-and-bias-in-a-neural-network-4cf7e9888a0f

Authors Profile

Ambily Merlin Kuruvilla is pursuing her research in Computer Science in the area of "Data Mining and Machine Learning" from, Karpagam Academy of Higher Education, Coimbatore. She holds a Master Degree in Computer Applications from Sri Ramakrishna College of Arts and Science for Women affiliated to Bharathiar University, Coimbatore, and M.Phil from Vinayaka Missions University, Salem. She has more than 19 years of graduate teaching experience and has published research papers in Journals and Proceedings of International repute.

Dr. N.V. Balaji is a highly acclaimed academician with dedication and commitment towards education. As a veteran educationalist for more than a decade he renders his service as an academic contriver with zeal for 21 years. He enriched training and placements with immense enthusiasm scaling the academy to great heights. Dr. Balaji a pertinacious personality brought laurels upon the institutions by representing it at Cambridge University for Business English Certifications. He is a honorable recipient of the award of Ambassador for Computer based Learning and Assessment category in the year 2015.

Adding to his reputation he has accumulated 15 years of research experience in the field of computer science. He exhibits deep interest towards various genres such as Neural Networks, Fuzzy Logic, Image Processing, Classification and Data Mining. Embellishing the annals of education at Karpagam institutions, he has been an influential face between prime IT firms namely Infosys, Wipro, Cognizant and Zoho. Through his constant coordination with competent companies, he incorporated industry collaborated electives in the curriculum of computer science.

He is meticulously marching towards success through significant strategies of prior planning and precise amalgamation of academics with practical experience. Emphasizing industry-based syllabus, he evolved a new theory of learning in order to elevate education as a wholesome experience. As a part of his educational expedition, he has visited United Kingdom and Israel for academic assignments. He is committed to the continuous growth of the institution by presently serving as a Dean of Arts, Science and Humanities at KAHE.