ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ARTIFICIAL NEURAL NETWORK WITH SPIDER MONKEY OPTIMIZATION ALGORITHM FOR CARDIOVASCULAR DISEASE PREDICTION

K. L. ANUSHA[1], JANARDHAN G[1], MAHANKALI SARITHA[2], MANNE ARCHANA[2]

[1, 2] Assistant Professor, Department of CSE, Vignan Institute of Technology and Science, Deshmukhi, Telangana-508284, India

E-mail: [1]anushakothamasu21@gmail.com

ABSTRACT

Globally, among the leading causes of death is coronary heart disease. Heart disease cases are rising quickly every day, thus early disease prediction is both important and dangerous. It is important to complete this diagnosis accurately and quickly because it is a challenging task. Specifically to improve diagnosis efficiency and accuracy, Machine Learning (ML) is increasingly popularity in the healthcare sector. By analyzing vast amounts of healthcare data, ML can predict diseases. Providing people and health professionals with the information they need to integrated content choices regarding preventing disease. However, the most difficult task in the area of clinical data analysis, is predicting a cardiac disease. Hence in order to solve these issues, Artificial Neural Network with Spider Monkey Optimization Algorithm for Cardio Vascular Disease Prediction is presented in this work. In a recent addition to the group of swarm intelligence-based optimization methods is the Spider Monkey optimization (SMO) algorithm. The combination of SMO and BSA (Bird Swarm optimization Algorithm) namely SMBS (Spider Monkey- Bird Swarm) will be used along with ANN (Artificial Neural Network) to predict the CVD (Cardio Vascular Disease. This System predicts the possibilities (i.e. presence, absence) of CVD. Parameters of prediction accuracy, precision, Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) are used to assess the performance of the presented approach. This approach will achieve better results than earlier prediction models.

Keywords: *Cardio-Vascular Disease (CVD), Prediction, Artificial Neural Network (ANN), Spider Monkey Optimization (SMO), Machine Learning*

1. INTRODUCTION

People today rarely have time to take care of themselves because they are so busy with work and daily activities. Due to their hectic lives, most people experience stress, anxiety, depression, and many other mental health issues. However, there are a lot of data about patients, diseases, and diagnoses produced by the healthcare industry. But, this data does not provide the value it should because it is not properly analyzed.

In both underdeveloped and developing nations, cardiovascular diseases have increasing rapidly as the leading cause of death over the past few decades. Clinical staff monitoring and early diagnosis of heart disorders are primary components that can decrease mortality rates. Cardio Vascular Diseases (CVDs) include coronary artery disease, vascular disease, and a number of issues with the heart and blood vessels. Strokes or heart attacks are too responsible for four out of every five CVD fatalities [1].

The heart is an essential component of the human body. Every part of the body receives oxygen and other nutrients from the blood, which is carried throughout the body by the heart. If the heart does not function properly, the body's other organs may stop working properly. As a result, taking care of the heart and other organs becomes difficult. People are more likely to develop heart-related conditions as a result of their hectic lifestyles and poor eating habits [11]. The precise cause of CVDs is still unknown; diabetes, Body Mass Index (BMI), high blood pressure, smoking, age, family history, and cholesterol, are all risk factors for these diseases. These factors differ from person to person. A few other important risk factors for CVDs, include age, gender, stress, and an unhealthy lifestyle [7].

Heart disease, which includes stroke, hypertension, and heart attacks, is a serious cause of death. In a situation like this, a cardiac disease can be diagnosed early, allowing for prompt treatment and the prevention of death. In addition, cardiovascular disease weaker patients, particularly elderly peoples, by having an impact on several human organs and blood vessels The main risk factors for heart disease are sex, smoking, family history, age, poor diet, cholesterol, physical inactivity, getting overweight, high blood pressure, and alcohol consumption. High blood pressure and diabetes are risk factors for heart disease that are hereditary. A few of the secondary factors that raise the risk are inactivity at the physical level, obesity, and a diet that is unhealthy. Overall weakness, chest pain, palpitations, shoulder pain, sweating, back pain, fatigue, and arm pain are the most common symptoms. Chest pain is still the most common sign that the heart doesn't get enough blood.

This can be difficult to remotely monitor on a patient's health and to recommend people who have less chronic health issues to live separately. This is due to an increase in cardiovascular/heart disease patients due to abnormal health parameters. As a result, one more difficult challenge in the healthcare sector is the early detection of heart disease prior to a stroke or heart attack using the remote monitoring dataset [5]. Tools for the early detection of cardiac abnormalities and for the identification of heart problems have the potential to significantly reduce cardiovascular disease mortality and save a great number of lives. Numerous patient data are now accessible due to the development of advanced healthcare systems (i.e. the electronic health record system makes use of big data.) could be utilized to create cardiovascular disease predictive models [16]. In order to effectively treat a patient prior to a heart attack, it is essential and challenging to accurately predict cardiovascular disease.

The use of artificial intelligence and machine learning in the medical field is now generally regarded [6]. Implicit information can be manipulated and extracted through machine learning, previously unknown or known information about the data that may be useful. The application and scope of machine learning are expanding at an ever-increasing rate. In order to predict and determine the accuracy of the provided dataset, machine learning uses several types of classifiers from supervised, unsupervised, and ensemble learning [4]. The self-organized, intelligent, and collaborative type of behavior of a wide range of insects (ants, antlions, termites, bees, and wasps), animals (monkeys, lions, camels), Swarm Intelligence (SI), a type of artificial intelligence that is more advanced, takes into account both birds (eagles, cuckoos, and crows) and birds (eagles, cuckoos, and crows) that have been created to solve a wide range of problems in the real world.

Hence in this analysis, Artificial Neural Network with Spider Monkey Optimization Algorithm for Cardio Vascular Disease Prediction is presented. The following is the structure of the remaining work: The literature survey is described in section 2. The section 3 demonstrates the Artificial Neural Network with Spider Monkey Optimization Algorithm for Cardio Vascular Disease Prediction. The results analysis is presented in Section 4. Finally, section 5 introduces the work to a conclusion.

2. LITERATURE SURVEY

Manuscripts Ali Al Bataineh and Sarah Manacek et. al. [2] describes MLP-PSO (Multi-Layer Perceptron- Particle Swarm Optimization) Hybrid Algorithm for Heart Disease Prediction. With the support of the publicly accessible Cleveland Heart Disease datasets, in this research will develop and evaluate a number of intelligent systems for predicting a person's likelihood of developing heart disease. The PSO methodology is used in this research to propose an alternative MLP training technique for the recognition of heart problems. This analysis uses ten different ML algorithms to predict cardiac disease in addition to the presented MLP-PSO hybrid approach. The accuracy of this MLP-PSO is 84.61%, outperforming all other algorithms. The results indicate that the current MLP-PSO classifier supports practitioners in making earlier, more precise, and more accurate diagnoses of cardiac disease.

Yar Muhammad, Maqsood Hayat Muhammad Tahir, and Kil To Chong et. al. [12] describes an intelligent computational model can detect heart disease early and accurately. This analysis examines various classification algorithms for machine learning. The extracted feature space is eliminated using four different feature selection algorithms. An analysis is performed on the results of classifiers as well as the data of each feature selection technique. A number of performance metrics, including: the developed model's

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

effectiveness and strength can be observed by examining its sensitivity, accuracy, specificity, F1-score, AUC (Area under the RoC Curve), RoC (Receiver Operating Characteristics curve) and MCC (Matthews correlation coefficient) curve. The top two classification algorithms, ET (Extra Tree classifier) and GB (Gradient Boosting), achieved 92.09% and 91.34% classification accuracies on all features, respectively.

Muktevi Srivenkatesh et. al. [13] presents Predicting Cardiovascular Disease Using Algorithms from Machine Learning. In order to diagnose or increase awareness of cardiovascular disease, this research presents a prediction model to determine whether a person has the condition or not. To present a reliable model for predicting cardiovascular diseases, the accuracy of applying rules to each result is compared to the individual outcomes of Random forest, support vector machine, naive bayes classifier, and logistic regression. The study's machine learning algorithms had an accuracy range of 58.71 % to 77.06% when predicting cardiovascular disease in patients.

Youness Khourdifi, Mohamed Bahaj et. al., [14] describes a Hybrid Machine Learning Model for Predicting Heart Disease that is based on Random Forest and is optimized by PSO and ACO (Ant Colony Optimization). The primary goal of this analysis is to utilize the weka data-mining tool to predict heart disease and to classify data in medical bioinformatics. The best algorithm for heart disease diagnosis and prediction is chosen after the data set is classified. Prediction begins with the identification of patient symptoms, followed by the separation of sick patients from a large number of healthy patients. SVM (Support Vector Machine) and NB (Naïve Bayes) have a precision average value of 83.6%, while RF (Random Forest) has an average value of 81.35%, providing performance improvement without any optimization.

Jae Kwon Kim and Sanggil Kang et. al. [17] describes Feature Correlation Analysis for Neural Network-Based Coronary Heart Disease (CHD) Risk Prediction. A CHD risk prediction made in two stages using NN (Neural Network)-based Feature Correlation Analysis (NN-FCA). The feature correlation analysis phase determines the each NN predictor output's data and feature relations are correlated, while the stage of feature selection evaluates features based upon the way they predict the risk of CHD. In that order, these two stages are performed. In terms of CHD risk prediction, the feature correlation analysis-based on presented NN-FCA, which outperforms state of art approaches, is found to be more effective. In addition it forecasts the CHD risk in the Korean population and is more precise and had a larger ROC curve.

Chun-yan Zhu, Yu Tian, Run-ze Li, Dan-yang Tong, Sheng-qiang Chi, Jing-song Li et. al. [19] describes the process of developing a cardiovascular disease readmission risk assessment system. The risk assessment model is divided into three parts: clustering analysis, using risk prediction, and regression analysis, in the risk level and risk factors for patients who are discharged within thirty days can be automatically predicted. The accuracy of the model was 90.62%. With the assistance of follow-up doctors and nurses, this strategy has the potential to enhance the self-management skills of discharged patients.

Nowadays, methods for predicting and diagnosing heart disease are mostly dependent on a patient's medical history, their symptoms, and the results of their physical examination. Considering the diagnosis of any disease is now predicated on the similar symptoms displayed in patients who have already been diagnosed, it is frequently challenging for healthcare experts to predict a patient's heart condition with accuracy. However, sometimes they can do it with a 92% accuracy rate. As a result, the medical industry requires a system for accurately predicting CVD.

3. ARTIFICIAL NEURAL NETWORK WITH SPIDER MONKEY OPTIMIZATION ALGORITHM

This section presents an Artificial Neural Network with Spider Monkey Optimization Algorithm for Prediction of Cardio Vascular Disease. The work flow diagram of presented approach is shown in Figure 1. The Cardiovascular Disease dataset was gathered from Kaggle this analysis. 70,000 patient data records, 11 features, and a target are included in the CVD dataset. There are three types of input features in this dataset: Objective (Factual Data); Exam (Results of the Medical Exam); Subjective (patient-provided information).

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

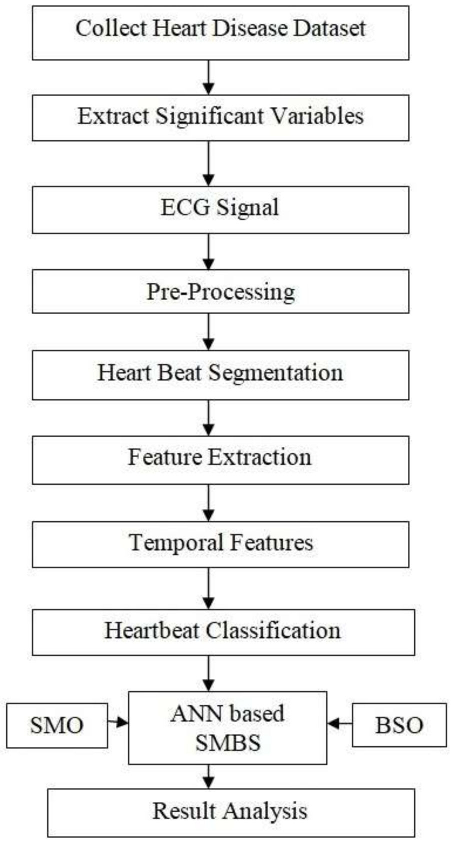

Figure 1: Work Flow Diagram Of ANN With SMO Algorithm For CVD Prediction

The objective features of dataset are Age, Height, Weight and Gender. Cholestrol, diastolic blood pressure, and systolic blood pressure are among the dataset's examination features(1: normal; 2: above average; 3: well above average), and glucose (1: normal; 2: above average; 3: well above average). Smoking is among the subjective characteristics(1: smoker; 0: nonsmoker), alcohol consumption, and regular exercise. The target indicates whether CVD is present or not. During the medical examination, all of the dataset values are collected. This dataset allows for the extraction of the pattern that identifies those who are at heart disease risk.

To check for various cardiac diseases, an ElectroCardioGram (ECG) measures the electrical signal from the heart. Using electrodes that are inserted into the chest, the heart's electrical signals are recorded and displayed as waves on a computer screen or printer that is connected to the device. Millions of individual cell action potentials and the recorded signal is determined by their activation order. The P, Q, R, S, and T waves make up the ECG. The Atrio-Ventricular node (AV) generates the QRS (Ventricular depolarization) wave, which is one of the five waves in the ElectrocardioGram (ECG). An ECG complex's P wave indicates atrial depolarization. The T wave provided ventricular repolarization, whereas the QRS generated ventricular depolarization. Atrial depolarization is absent in the absence of a P wave. During the stage of signal preprocessing, the baseline is initially removed from the input ElectrocardioGram (ECG) signal. Pre-processing the signals is important because it allows removal of the parts of the signal that aren't needed to ensure that the heart rate can be determined using the extracted features in an effective and accurate manner.

The removal of artifacts like ECGs, noise filtering, and all of these steps are included in the pre-processing stage, as is resampling the signal to meet detector input requirements. This is accomplished by implementing an adaptive signal processing, low pass filter and an artifact removal algorithm. The QRS complex will be used to divide the beats in the pre-processed signal during the segmentation step. The segmented result will continue on to the feature extraction stage, where temporal and wave features, will be extracted, among other methods. The global average interval, post-RR (Inter Beat) interval, pre-RR interval, local average and RR interval, are some of the temporal features. The RR interval, PP (the time between successive P waves) interval, PR (the time between atrial depolarization and ventricular depolarization) interval, QT (ventricular electrical systole) interval, and R intervals are among the wave features that will be properly considered.

Heartbeats are categorized using the received features in the classification step using the presented artificial neural network based on the Spider Monkey-Bird Swarm (SMBS). The spider monkey-bird swarm algorithm, which is a combination of the Bird Swarm Algorithm (BSA) and the Spider Monkey Optimization (SMO), will be utilized in this approach.

The development of optimization algorithms has traditionally involved research into the foraging behaviour of social animals. A global optimization technique called Spider Monkey Optimization (SMO) was inspired by the fission-fusion social structure of spider monkeys during foraging behaviour. Self-organization and labour division are two essential swarm intelligence concepts that SMO beautifully illustrates. SMO, a swarm intelligence-based algorithm, has grown in prominence in recent years and is now being used to solve numerous engineering optimization issues. The Spider Monkey Optimization algorithm is thoroughly explained in this chapter.

To solve the optimization problem, the global optimization algorithm known as Spider Monkey Optimization (SMO), is based on the Fission-Fusion Social (FFS) structure that spider monkeys use when foraging. The behavior of social iterations of birds in nature served as inspiration for the development of the new Bird Swarm optimization Algorithm (BSA), the global intelligent swarm optimization method. A new metaheuristic algorithm for solving optimization issues is called BSA. Address issues that requiring global optimization, it mimics the foraging, vigilance, and flight behaviours of birds. The bird swarms foraging and intelligent behaviour can be effectively extracted using SMO and BSA. The optimization's goal is ensure that the ANN is effectively tuned for accurate CVD prediction and classification.

A computer can learn new information and make decisions in a manner that is comparable to that of a human by simulating the neurons of an artificial neural network. ANNs are made by programming regular computers to function like a network of interconnected brain cells. In terms of their structure and function, artificial neural networks focus to closely resemble the biology of neurons: A mathematical object known as an artificial neuron is made up of inputs, the output and the total of which are processed by an activation function (or transfer function).

A group's social organization in which a female leader decides whether to combine or split, impacts the SMO algorithm's features. Here, the global leader of the group is identified, while the smaller groups local leaders are identified. The absence of an SMO algorithm improvement when referring to the phenomenon of food scarcity is a defining feature. Each small group should have at least one monkey because SMO is an algorithm based on swarm intelligence. As a result, they define the time for fusion as the time at which a subsequent fission produces at least one group of monkeys with smaller than the minimum number. The SMO algorithm offers a resolution in the form of a Spider Monkey (SM). There are six stages to the SMO process: the Learning phase of the Local Leader, the Local Leader phase, the Decision phase of the Global Leader, the Decision phase of the Local Leader, the Learning phase of the Global Leader, and the Global Leader phase.

Spider Monkey Optimization (SMO) is a nature-inspired metaheuristic algorithm that mimics the behavior of spider monkeys in search and optimization processes. This algorithm is based on the social interaction and foraging behavior of spider monkeys in the wild. SMO has been developed to solve complex optimization problems by simulating the social behavior of spider monkeys.

Here's a general overview of how Spider Monkey Optimization works:

i. Initialization: The algorithm starts by initializing a population of spider monkeys. Each spider monkey represents a potential solution to the optimization problem.

ii. Spider Monkey Movement: Spider monkeys move in search of food and interact with each other. In SMO, the movement of spider monkeys is guided by several factors, including exploration, exploitation, and communication.

iii. Exploration: Spider monkeys explore the search space by randomly searching for potential solutions. This helps in discovering new regions of the search space that might contain better solutions.

iv. Exploitation: Spider monkeys exploit the discovered regions of the search space by focusing their search on promising solutions. This is done by applying local search operators to refine and improve the solutions.

v. Communication: Spider monkeys communicate with each other by exchanging information about their current positions and solutions. This social interaction helps in sharing knowledge and guiding the search towards better solutions.

vi. Fitness Evaluation: After each movement, the fitness of each spider monkey's solution is evaluated using an objective function specific to the optimization problem. The objective function determines how good a solution is based on the problem's requirements.

vii. Update: Based on the fitness evaluations, spider monkeys update their positions and solutions. The update is performed by considering the best solutions found so far and incorporating the knowledge gained from the exploration and exploitation steps.

viii. Termination: The algorithm continues the movement, exploration, exploitation, and communication steps until a termination condition is met. This condition could be a maximum number of iterations, reaching a desired solution quality, or a predefined computational time limit.

SMO aims to find the best solution to the optimization problem by effectively balancing exploration and exploitation. By simulating the social behavior of spider monkeys, the algorithm combines both individual and collective intelligence to guide the search process.

It's worth noting that Spider Monkey Optimization is just one of many metaheuristic algorithms available for optimization problems. Its effectiveness and performance depend on the nature of the problem being solved and the specific implementation details.

SMO achieves a better balance between exploitation and exploration in the development of optimal conditions. Throughout the local leader phase, which has significant dimensional perturbation, each member of the group changes their locations, this stage involves searching the search area. Whereas the phase of the global leader encourages exploitation, more opportunities are given to better candidates to update their positions. This characteristic makes SMO a better selection when compared to search-based optimization techniques. Additionally, SMO incorporates a check for stagnation. To determine whether the search process has stagnated, it uses both the local leader learning phase and the global leader learning phase. In the event of either local or global stagnation, the decision phases of the local leader and the global leader function. A decision on Cardiovascular Disease (CVD) or non-CVD (healthy) is taken at the global leader decision phases. The decision-making process for the local leader leads to a further analysis.

The primary new control parameters of SMO are: the Maximum number of Groups (MG), the global leader limit, the perturbation rate pr, and the local leader limit. The following are the suggested parameter settings:

$$MG = \frac{N}{10} \quad (1)$$

i.e., It's also collected that a group of SMs must contain at least 10 members.

$$Global\ leader\ limit\ \epsilon \left[\frac{N}{2}, 2 \times N\right] \quad (2)$$

$$Local\ leader\ limit\ is\ set\ to\ D \times N \quad (3)$$

where D represents dimension and N is the overall number of SM

$$pr \in [0.1.0.8] \quad (4)$$

BSO (Bird Swarm Optimization) uses the foraging, vigilance, and to address problems with global optimization, BSO mimics the foraging, vigilance, and flight behaviors of birds. Each bird forages for food using both its own experience and the experience of the population collectively. The following mathematical description can be given for this behavior:

$$x_{i,j}^{t} = x_{i,j}^{t-1} + \left(p_{i,j} - x_{i,j}^{t-1}\right) * C * rand(0,1) + \left(g_{best,j} - x_{i,j}^{t-1}\right) * S * rand(0,1) \quad (5)$$

where rand(0, 1) is a uniform distribution function, $g_{best,j}$ stands for the j-th element of the global optimal solution, $p_{i,j}$ stands for the best previous location for the j-th element of the i-th bird, and $x^{t}_{i,j}$ stands for the value of the j-th solution element at generation t-th. The terms "cognitive accelerated coefficient" and "social accelerated coefficient" are two positive numbers.

Every bird would compete with other birds while remaining vigilant as it attempted to move toward the center of the swarm. The following describes the vigilant behavior:

$$x_{i,j}^{t} = x_{i,j}^{t-1} + \left(mean_j - x_{i,j}^{t-1}\right) * A1 * rand\ (0,1) + \left(p_{i,j} - x_{i,j}^{t-1}\right) * A2 * rand\ (0,1) \quad (6)$$

$$A1 = a1 * \exp\left(-\frac{pFit_i}{SumFit + \varepsilon} * N\right) \quad (7)$$

$$A2 = a2 * \exp\left((\frac{pFit_i - pF_k}{|pFit_i - pFit_k| + \varepsilon}) - \frac{pF_{i*N}}{SumFit + \varepsilon} \right) \quad (8)$$

Where the number k ($k \neq i$) is a random positive integer.

The sum of the swarm's highest fitness values is represented by sumFit, while the term $pFit_i$ stands for " i-th bird's best fitness value", Two positive constants in the range [0, 2] are N. a_1 and a2, and they are both positive. The smallest constant in a computer, which ε is used to prevent zero-division errors. The j-th component of the overall swarm's average position is identified by the term $Mean_j$ mathematically; the following can be said about the producers and scroungers flight patterns:

$$x_{i,j}^{t} = x_{i,j}^{t-1} * 1(1 + randn(0,1)) \quad (9)$$

$$x_{i,j}^{t} = x_{i,j}^{t-1} + FL * (x_{k,j}^{t-1} - x_{i,j}^{t-1}) * rand\ (0,1) \quad (10)$$

Where $k \in [0,N]$, $k \neq i$ $FL \in (0, 2)$ is the probability that scroungers would follow producers in their search for food and rand $n(0, 1)$ is a gaussian distribution with mean 0 and standard deviation '1' for each individuals. SMO utilizes intelligent behavior, whereas BSO mimics the foraging, flight, and vigilance behaviors. The problem can be effectively optimized using either of these algorithms. As a result, and the ANN can be used to predict CVD. Artificial Neural Network with Spider Monkey Optimization Algorithm for Cardio Vascular Disease Prediction approach predicted the presence and absence of CVD.

4. RESULT ANALYSIS

In this section, Artificial Neural Network with Spider Monkey Optimization Algorithm for Cardio Vascular Disease Prediction is implemented using python. In this analysis, CVD dataset from kaggle is used. This approach predicts the presence and absences of CVD using ANN with SMBS (Spider Monkey- Bird Swarm). This algorithm predicted the presence (CVD) and absence (non-CVD) of CVD.

This section discusses the presented approach's performance evaluation. The presented approach's performance is measured in terms of accuracy, Root Mean Square Error (RMSE), precision, and Mean Absolute Error (MAE).

Accuracy: According to this definition, accuracy is the proportion of accurately predicted instances to all instances.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+F} \times 100 \quad (11)$$

Precision: The accuracy rate of positive identifications is shown by precision. The following are some examples of precision:

$$Precision = \frac{TP}{TP+FP} \times 100 \quad (12)$$

Where the instances that are used to predict the CVD are TP (True Positive), TN (True Negative), False Positive (FP), and False Negative (FN).

RMSE: The standard way to measure a model's error in predicting quantitative data is in RMSE. The RMSE can be written as

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(y' - y)^2}{n}} \quad (13)$$

MAE (Mean Absolute Error): The arithmetic average of all absolute errors is the mean absolute error. It may be described as

$$|e| = |cy - x| \quad (14)$$

Where x is the actual value and y is the value that was predicted, and where e is the absolute error is a mod of the difference between the values of y and x. The table 1 represents the performance metrics evaluation of different ML algorithms and presented approach.

Table 1: Performance Metrics Evaluation

Performance Metrics	Naïve Bayes (NB)	Multilayer Perceptron (MLP)	Presented ANN with SMO Algorithm
Precision (%)	47	43	94.23
Accuracy (%)	90.10	75.42	96.32
RMSE	0.0989	0.2417	0.0424
MAE	0.090	0.2418	0.0423

Accuracy, precision, RMSE Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) are all measured in compared to the performance of the various ML algorithms.

Compared to LR and MLP algorithms, presented ANN with SMPO has better results. The Figure 2 shows the comparative graph for RMSE and MAE of different ML algorithms.

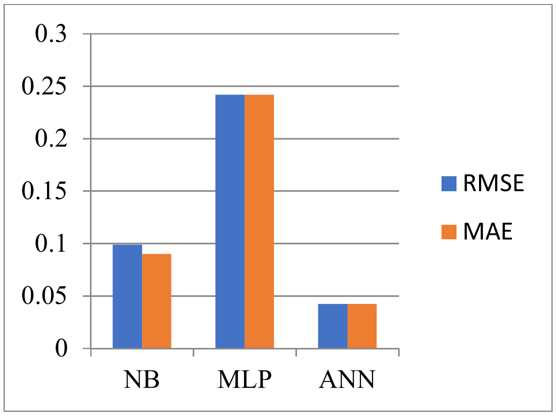

Figure 2: RMSE and MAE metrics Comparison

In Figure 2 the x-axis represents different ML algorithms whereas y-axis indicates performance values. Presented ANN with SMO algorithm has better RMSE and MAE than NB and MLP algorithms. The Figure 3 describes the precision and accuracy of the performance comparison.

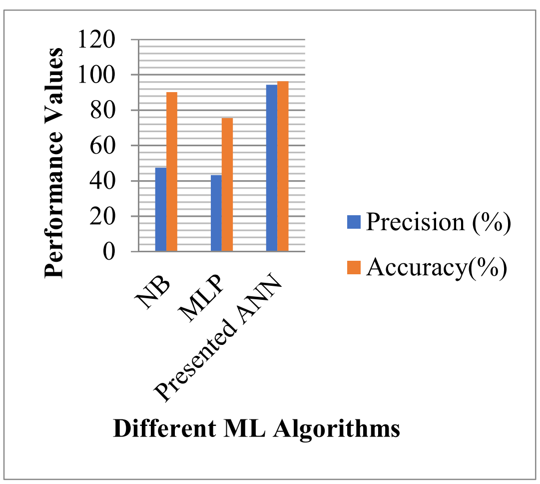

Figure 3: Comparative Graph for Accuracy and Precision

Compared to NB and MLP algorithms, presented ANN with SMO algorithm has high accuracy and precision for CVD prediction. Hence presented Artificial Neural Network with Spider Monkey Optimization Algorithm for Cardio Vascular Disease Prediction approach has effectively predicted the presence and absence of CVD.

5. CONCLUSION

In this work, Artificial Neural Network with Spider Monkey Optimization Algorithm for Cardio Vascular Disease Prediction is presented. In this analysis, the ECG reports of patients are collected for the prediction of CVD. Here CVD dataset from kaggle is used to train the ANN for CVD prediction. With the purpose of providing that the extracted features are efficient and accurate for the classification of heart rates, data preprocessing is performed to remove unwanted parts from the signal. The combination of SMO (Spider Monkey Optimization) and Bird Swarm optimization Algorithms namely: SMBS is used along with ANN to predict the CVD. The SMO utilizes the intelligent behaviour, whereas the BSO mimics the foraging, flight, and vigilance behaviours. This approach has effectively predicted the ECG reports of patient as non-CVD (healthy) or CVD. The presented approach's performance is analyzed in terms of precision and accuracy, MAE, and RMSE. Compared to Naïve Bayes and Multi Layer Perceptron algorithms, presented ANN with SMO algorithm has better results and high prediction accuracy. Therefore presented approach will be the best solution of CVD prediction, if the CVD is predicted accurately then appropriate diagnosis and treatment will be provided. As a result patient life will be saved.

REFERENCES:

[1] Ch. Anwar ul Hassan, Jawaid Iqbal, Rizwana Irfan, Saddam Hussain, Abeer D. Algarni, Syed Sabir Hussain Bukhari, Nazik Alturki and Syed Sajid Ullah, "Effectively Predicting the Presence of Coronary Heart Disease Using Machine Learning Classifiers", Sensors 2022, 22, 7227, doi:10.3390/s22197227

[2] Ali Al Bataineh and Sarah Manacek, "MLP-PSO Hybrid Algorithm for Heart Disease Prediction", Journal of Personalized Medicines, 2022, doi:10.3390/jpm12081208

[3] Jialong Zhou, "Cardiovascular Disease Diagnosis Based on Stacking Technology", 2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), DOI: 10.1109/IPEC51340.2021.9421128

[4] Harshit Jindal, Sarthak Agrawal, Rishabh Khera, Rachna Jain and Preeti Nagrath, "Heart disease prediction using machine learning algorithms", IOP Conference Series: Materials Science and Engineering, doi:10.1088/1757-899X/1022/1/012072

[5] Sudarshan Nandy, Mainak Adhikari, Venki Balasubramanian, Varun G. Menon, Xingwang Li, Muhammad Zakarya, "An intelligent heart disease prediction system based on swarm artificial neural network", Neural Computing and Applications, Springer, 2021, doi:10.1007/s00521-021-06124-1

[6] Rohit Bharti, Aditya Khamparia, Mohammad Shabaz, Gaurav Dhiman, Sagar Pande, and Parneet Singh, "Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning", Hindawi

Computational Intelligence and Neuroscience Volume 2021, Article ID 8387680, 11 pages, doi:10.1155/2021/8387680

[7] Aqsa Rahim, Yawar Rasheed, Farooque Azam, Muhammad Waseem Anwar, Muhammad Abdul Rahim, And Abdul Wahab Muzaffar, "An Integrated Machine Learning Framework for Effective Prediction of Cardiovascular Diseases", IEEE Access, 2021, volume 9, DOI: 10.1109/ACCESS.2021.3098688

[8] B Padmajaa, Chintala Srinidhib, Kotha Sindhuc, Kalali Vanajad, N M Deepikae, E Krishna Rao Patrof, "Early and Accurate Prediction of Heart Disease Using Machine Learning Model", Turkish Journal of Computer and Mathematics Education 4516 Research Article Vol.12 No.6 (2021), 4516-4528

[9] Karna Vishnu Vardhana Reddy, Irraivan Elamvazuthi, Azrina Abd Aziz, Sivajothi Paramasivam, Hui Na Chua and S. Pranavanand, "Heart Disease Risk Prediction Using Machine Learning Classifiers with Attribute Evaluators", Appl. Sci. 2021, 11, 8352, doi:10.3390/app11188352

[10] Gihun Joo, Yeongjin Song, Hyeonseung Im, And Junbeom Park, "Clinical Implication of Machine Learning in Predicting the Occurrence of Cardiovascular Disease Using Big Data (Nationwide Cohort Data in Korea)", 2020 IEEE Access, volume 8, DOI: 10.1109/ACCESS.2020.3015757

[11] Rahul Katarya, Sunit Kumar Meena, "Machine Learning Techniques for Heart Disease Prediction: A Comparative Study and Analysis", Health and Technology, 2020, Springer, doi:10.1007/s12553-020-00505-7

[12] Yar Muhammad, Muhammad Tahir, Maqsood Hayat & Kil To Chong, "Early and accurate detection and diagnosis of heart disease using intelligent computational model", Scientific Reports, 2020, doi: 10.1038/s41598-020-76635-9

[13] Muktevi Srivenkatesh, "Prediction of Cardiovascular Disease using Machine Learning Algorithms", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249-8958 (Online), Volume-9 Issue-3, February 2020

[14] Youness Khourdifi1, Mohamed Bahaj, "The Hybrid Machine Learning Model Based on Random Forest Optimized by PSO and ACO for Predicting Heart Disease", ICCWCS 2019, April 24-25, Kenitra, Morocco, 2019 EAI
DOI 10.4108/eai.24-4-2019.2284088

[15] Abderrahmane Ed-daoudy, Khalil Maalmi, "Real-time machine learning for early detection of heart disease using big data approach", 2019 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS), 978-1-5386-7850-3/19, 2019 IEEE

[16] Shadman Nashif, Md. Rakib Raihan, Md. Rasedul Islam, Mohammad Hasan Imam, "Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System", World Journal of Engineering and Technology, 2018, 6, 854-873, DOI: 10.4236/wjet.2018.64057 Nov. 22, 2018

[17] Jae Kwon Kim and Sanggil Kang, "Neural Network-Based Coronary Heart Disease Risk Prediction Using Feature Correlation Analysis", Hindawi Journal of Healthcare Engineering, Volume 2017, Article ID 2780501, 13 pages, doi:10.1155/2017/2780501

[18] Cheryl Ann Alexander and Lidong Wang, "Big Data Analytics in Heart Attack Prediction", Journal of Nursing and Care, 6;2, 2017, DOI: 10.4172/2167-1168.1000393

[19] Chun-yan Zhu, Sheng-qiang Chi, Run-ze Li, Dan-yang Tong, Yu Tian, Jing-song Li, "Design and Development of a Readmission Risk Assessment System for Patients with Cardiovascular Disease", 2016 8th International Conference on Information Technology in Medicine and Education

[20] Jaymin Patel, Prof.TejalUpadhyay, Dr. Samir Patel, "Heart Disease Prediction Using Machine learning and Data Mining Technique", IJCSC,

2016, Volume 7, doi:10.090592/IJCSC.2016.018.

[21] Kumar, G. K. L., Asif, S., & Veeresh, U., "A study on heart disease prediction using supervised machine learning models". In AIP Conference Proceedings (Vol. 2358, No. 1, p. 100018), July 2021.

[22] Lakshmi Anusha Kothamasu, Dr E. Kannan, "Sentiment analysis on twitter data based on spider monkey optimization and deep learning for future prediction of the brands", Concurrency and Computation: Practice and Experience, Vol 34, Issue 21, June 2022. DOI: 10.1002/cpe.7104.

[23] Kumar, S., Sharma, B., Sharma, V. K., & Poonia, R. C, "Automated soil prediction using bag-of-features and chaotic spider monkey optimization algorithm". Evolutionary intelligence, Vol 14, 2021.

[24] G. Krishna Lava Kumar, A. Robert Singh, "Impactful study on the heart disease based on machine learning techniques". In AIP Conference Proceedings (Vol. 2463, No. 1, p. 020039), May 2022.