International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Deep VAE AEO: Deep Variational Auto Encoder with Artificial Ecosystem Optimizer Based Cardiovascular Disease Prediction

Venkateswarlu Tata[1], Dr. R. Bhavani[2], Dr. S. V. N. Srinivasu[3], Dr. R. Priya[4]

Submitted: 16/09/2023 **Revised**: 17/11/2023 **Accepted**: 29/11/2023

Abstract: Cardiovascular complications are a major reason for mortality and morbidity in such patients. As a result, special attention must be paid to the occurrence of cardiovascular complications, particularly in high-risk populations. The underlying cause of cardiac dysfunction is the interaction of biological, autonomic, and iatrogenic factors. To make a diagnosis of heart defects, a system capable of predicting the presence of heart diseases would be required. Our main motivation in this article is to develop a reliable intelligent medical system using machine learning techniques to assist in identifying a patient's heart condition and guiding a physician in giving a precise prognosis of whether or not the patient has cardiovascular disease. Deep Variational Auto Encoder of Artificial Ecosystem Optimizer (Deep VAE AEO) is used in this paper to benefit from multiple non-linear layer upon layer without so much as an information bottleneck while not getting out of hand to the identity. Furthermore, epileptic data feature selection is carried out using the Spiral Optimization method, which utilizes an improved efficiencies rad model where In order to reach the focal point, the search process follows a lognormal spiral path. Deep VAE AEO is compared to existing methodologies in terms of different parameters, and it is discovered that Deep VAE AEO achieves 97percentage accuracy, 98percentage precision, 87 percentage recall, and 82 percentage F1-score..

Keywords: *cardiovascular disease, auto encoder, optimization, feature selection, classifier, neural network.*

1. Introduction

It is estimated that cardiac arrest and its associated illnesses cause over half of all deaths in the world [1]. The world's leading cause of death, cardiovascular diseases (CVDs) account for 30 percent of all deaths It costs the European Union €210 billion per year in the United States [2] and 45 percent in Europe [3]. Despite significant advances in medical testing over the last fifty years, doctors, primary care providers, as well as other health professionals face enormous challenges in detecting and diagnosing heart disease. A doctor's ability to diagnose CVD is also determined by a review of a patient's past medical history, a basic examination, examination findings, and biochemical markers, which are all interpreted based on the doctor's experience. Their subjective perception of scientific journals is then used to match each patient to a conventional taxonomy of physical illnesses. It is becoming increasingly inefficient and prone to errors [4]. Furthermore, physicians' jobs are becoming more complex as cardio-vascular methods become capable of collecting a greater amount of data.

It is imperative that medical interventions be simple and easy to use, quick and automated, and highly accurate in order to increase patients' well-being while lowering healthcare costs and decreasing CVD deaths. Heart failure (HF) therapies are increasingly being developed using machine learning (ML) and deep learning (DL) models. Clinical HF care, on the other hand, is faced with real-world challenges such as high readmission and death rates, inadequate patient care, and overconsumption [6]. Predictive models based on AI can address these issues; however, partnerships between data researchers and medical practitioners are required to ensure their clinical success [7]. ML is becoming increasingly popular as a trustworthy approach for merging Physician data and clinical data observations EMRs (electronic medical records) to increase the accuracy of a variety of medical procedures, owing to tremendous developments in information retrieval and warehousing capabilities [8]. Deep learning (DL) has emerged as a robust solution for medical image classification, segmentation techniques, and natural language processing (GPUs). The most popular types of machine learning are logistic regression (LogR) [9,10], support vector machines (SVM) [11], gradient boosting machines (GBM) [12-14], random forests (RF) [13, 15], artificial neural networks (ANN) [15, 16], and convolutional neural networks (CNN) are the most popular often utilized AI models in CVD [17, 18].

1Research Scholler, Department of Computer Science and Engineering, FEAT, Annamalai University, Chidambaram,
1venkat543@gmail.com.
2Professor& Head, Department of Computer Science and Engineering, FEAT, Annamalai University, Chidambaram,
2bhavaniaucse@gmail.com.
3Professor, Department of Computer Science and Engineering, Narasaraopeta Engineering College, Narasaraopeta,
3drsvnsrinivasu@gmail.com.
4Professor, Department of Computer Science and Engineering, FEAT, Annamalai University, Chidambaram,
4prykndn@yahoo.com.

The proposed effort aims to provide an accurate and dependable patient-specific CVD predictive model. The system self-learns powerful characteristics that aid in the categorization of minimal pre-processed data into 2 classifications of states. It also replaces sophisticated systems with lengthy training cycles with a much simpler, faster, and much more powerful system that takes use of the characteristics and capability of Variation auto encoders. This work makes the following contributions:

• Building a Deep Variation Auto Encoder with Artificial Ecosystem Optimizer (Deep VAE AEO) that takes use of several non-linear layers with no information bottleneck while not overfitting to the identity.

• Using the Spiral Optimization approach for feature selection, which uses a multifold sequential in this approach, the focus point is approached in a logarithmic spiral route.

This article is organized as follows: The second section discusses works that are similar to the first. Described in Section 3 are: recommended technique. Section 4 goes on performance evaluation. Section 5 concludes with the conclusion and future works.

2. Related works

[19] seeks to improve ANN performance, resulting in higher prediction accuracy. The employment of Genetic Algorithm (GA) was noticed in order to actualize. By including GA, ANN was 5.08 percent more effective when compared to the predictive performance of the lone ANN, according to simulation findings. Furthermore, in the diagnosis of cardiovascular disease, the GA-ANN predictive algorithm beat other machine learning techniques.

Through using training set data, [20] built logit model and ANN models. An evaluation of the model's predictive ability was conducted using test set data, which identified predictive indicators of coronary artery stenosis. All participants' health information was collected. IFN-, It is well known that IP-10, and MIG are strongly correlated with the extent of stenosis in patients with CHD. Despite their unrelated nature, IP-10 and MIG are major risk factors for coronary stenosis.

Based on ECNN-LSTM, [21] propose a novel computational approach for diagnosing congenital cardiac disease in live time. A multi-resolution singular value decomposed model is suggested a test for predicting congenital cardiac disease in babies and children. Using the multiresolution decomposition of singular values technique, infants and toddlers' whole-life vibrating time-domain signals are approximated and comprehensive with varying resolutions.

[22] offers a novel Using recurrent neural networks (RNNs) with gated recurrent units (GRUs), long short-term memories (LSTMs), and Adam optimization, a hybrid deep learning model can predict cardiac disease. Experiments were carried out with several customized models, and the findings revealed that the suggested model, which used RNN with multilayer GRU using the synthetic minority oversampling method (SMOTe) performed best.

[23] provides a deep learning technique for better prediction of cardiac disease. To accomplish efficient as a result of the Feature learning is performed by designing a clustering sparse auto encoder network (SSAE). The network is made up of many sparse auto - encoders as well as a soft-max classifier. To modify the settings of the small materials auto encoder, they suggest a particle swarm optimization (PSO)-based approach. The PSO optimization increases the SSAE's feature learning as well as classification performance.

[24] suggest utilizing χ^2statistical models while searching for the optimally designed deep neural network (DNN) using an exhaustive search technique. The proposed hybrid model, χ 2DNN, is assessed by contrasting its performance to traditional ANN and DNN algorithms, There are also state-of-the-art machine learning models and previously described approaches for predicting heart disease. The suggested model has a 93.33percentageprediction accuracy.

[25] proposes a two-stage approach for accurately predicting cardiac disease. In the first stage, Enhanced sparse auto encoders (SAE), unsupervised neural networks, learn the optimal representation of training data by training them. An artificial neural network (ANN) is used in the second step to forecast the health condition based on the learned records. In order to train an adequate system, the SAE was tuned.

[26] proposes a hybridization approach wherein decision tree with artificial neural network classifications are hybridized for improved prediction of cardiac disease. WEKA is used for this purpose. On a dataset of heart disease patients obtained from the UCI repository, a tenfold verification test is conducted to determine the algorithm's performance.

In this study, Using Back-propagation, a Multilayer Perceptron Neural Network was trained. The purpose of this work is to develop a diagnostic approach for predicting cardiac disease. According to medical literature, the suggested approach uses 14 key qualities for the diagnosis of heart disease. When compared with other techniques, the findings clearly show that the devised diagnosis is able to predict the overall risk of heart disease efficiently.

[28] looks at an ensemble classification approach for increasing the precision of weak algorithms by mixing different classifiers. Ensemble methods such as boosting

and bagging improve the predictive performance of weak classifiers and are effective in predicting heart disease risk. Ensemble classification provided a maximum gain in accuracy of 7% for poor classifiers.

Current classification models may be affected by a variety of factors. In real-world data, one example is the class unbalance of the training data. Models were frequently biased against the class label and failed to generalize. For ECG-based heartbeat classification, a Deep Variational Auto Encoder Using Artificial Ecosystem Optimizing framework solutions was suggested, and it demonstrated strong classification performance for unbalanced multi-category classification tasks.

3. Proposed Methodology

Framingham Heart Study-Cohort (FHS-Cohort) is utilized to predict heart disease. In order to remove null value features from a dataset, we used preprocessing techniques. Spiral Optimization is used to choose features from pre-processed data. As a result, the chosen characteristics are sent into the Deep Variational Auto Encoder Using Artificial Ecosystem Optimizer, as seen in figure-1

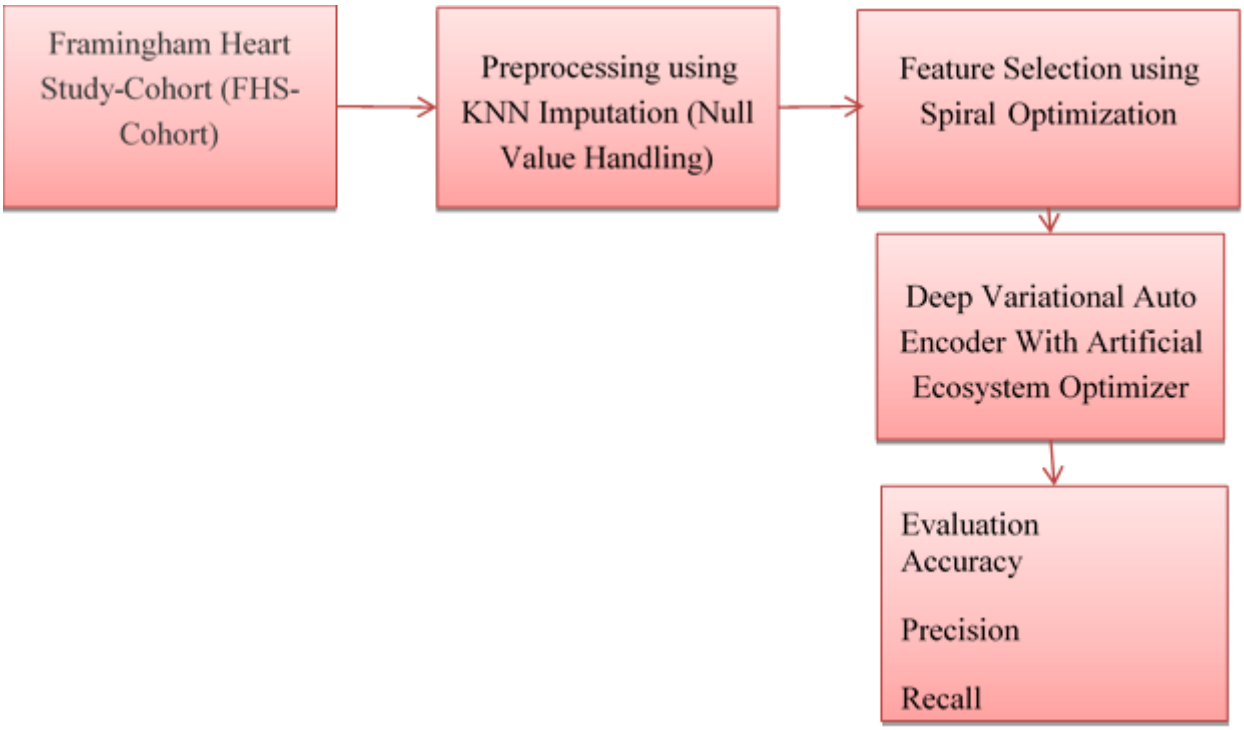

Fig-1 Block diagram for optimized neural network-based CVD prediction

4. Dataset description

The dataset was obtained from such Kaggle provides access to ongoing cardiovascular research of Framingham, Massachusetts residents [29]. The classification is used to evaluate if a patient has a 10% probability of getting In the next ten years, there will be an increase in coronary heart disease (CHD). The collection provides

patient information and includes 4,240 entries and 15 characteristics. Each feature might be a danger factor. Risk factors include concerns about demographic, behavioral, and medical aspects.

1.The gender of the subject: male(0) or female(1); (Nominal)

2.Age: The patient's age; (Continuous - Although recorded ages are truncated to whole numbers, the concept of age is continuous).

3.Smoker: whether the patient smokes now (Nominal)

4.CigarettesPerDay: number of cigarettes smoked on average in a day. (can be considered continuous since one can have a cigarette or even half a cigarette every day.)

5.BPMeds: Whether patient takes blood pressure medications (Nominal)

6. Stroke prevalence: whether the patient has had a stroke before (Nominal)

7.PresenceHyp: whether the patient was hypertensive (nominal)

8. Diabetes: patient health status (nominal)

9. totChol: total cholesterol (constant)

10. SBP: systolic blood pressure

11. DiaBP is Diastolic Blood Pressure

12. The BMI (Continuous) is your body mass index

13. HeartRate: heart rate (Continuous) - In medical research, variables like heart rate can be monitored are considered continuous because there are many possible values.

14. Glucose: Continuous level of glucose

15. CHD risk over 10 years (binary: "1" means "Yes", "0" means "No") - Target Variable

Fig-2: Description of variables (Framingham dataset)

5. **Preprocessing of data**

To impute missing values, the k-NN-based technique finds data with expression patterns comparable to the information of interest. Based on data A, which One value is missing in experiment 1 this approach would uncover K additional data in experiments 2-N with values present in experiment 1 and expressions most similar to A. (where N represents the total number of experiments). The value of data A is estimated using a modified average of the results from laboratory experiment from the K nearest data. The contributions of each information is weighted in the weighted sum by the similarity of its interpretation with that of data A. After studying a variety of data similarity metrics (correlations, Euclidean distances, and variance minimization), it was Euclidean distance was indeed an accurate norm. This discovery is rather surprising considering that perhaps the Euclidean distance metric is frequently susceptible to outliers, which may exist in microarray data. However, log-transforming the data appears to lessen the influence of anomalies on data resemblance determination considerably. The KNN imputation procedure is seen here:

- Choose data A with the expression that contains a missing value.
- Enter the value k for neighbor data with the complete affirmation to find the k nearest neighbors.
- Determine Euclidean distance among data A with incomplete data and all training expressions with absolute gene values.
- Organize distance and count the number of nearest neighbors depending here on kth shortest distance
- Sort those neighbors into categories. Replace the missing value with the corresponding gene value calculated from weighting of most related comprehensive data expression.

An outlier in statistics is an information that is quantitatively distinct from the remainder of the data. Outliers are explored and deleted in the experiment using the Interquartile filter in Weka. Outlier cases are deleted after the imputation procedure before the classification step begins. This process guarantees that the data is accurate and reduces data size, which is useful when a data collection has a significant number of attributes.

Human cardiac signals are made up of linear data, mainly period, bandwidth, and analytical method, A focus will be placed on ECG sequence signal and amplitude statistics.

ECG sequence amplitude is measured by wave coefficient:

$$F_i(n) = \frac{1}{M-1}\sum_{j=1}^{M-1}|a_n(j+1) -| - a_n \quad \text{----- (1)}$$

Here, a_n is amplitude of n-th ECG data after wavelet transform; M is signal length. NECG data with similarity tolerance N. Approximate entropy $AE(m,r)$indicates likelihood that 2 series of adjacent m points remain relatively close after tracing to m +1dimensional space mostly on basis of r in initial N sample point series.

$$AE(m,r) = \varphi^m(r) - \varphi^{m+1}(r) \quad \text{----- (2)}$$

$$\begin{cases} \varphi^m(r) = \frac{1}{N-m}\sum_{i=1}^{N-m} lnC_i^m\,(r) \\ C_i^m(r) = \frac{Q}{N-m-1} \\ d_{ij} = max\,[|x(i+k) - x(j+k)|] \end{cases} \quad \text{-----(3)}$$

Here, $\varphi^m(r)$, $C_i^m(r)$, $d_{ij}$are intermediate variables. Approximate entropy represents signal similarity, limits capacity to analyze problems. As a result, feature selection is added.

6. Feature Selection using Spiral Optimization

Following preprocessing, the Spiral Optimization approach A method of identifying epileptic features is used. According to The spiral optimization (SPO) algorithm described by Tamura and Yasuda [30] is a simple metaheuristic optimization approach inspired by spiral occurrences. The SPO method uses a multifold spiral model where the search process follows a logarithmic spiral path towards the focus point. The resulting focal points are linked to local optimal solutions, which are regularly updated when a significant improvement is discovered. A spiral optimization technique is created for a difficult n-Dimensional issue that involves a hybrid model made up of a rotation matrix as well as a step rate. The suggested method aims to generate a multi-objective spiral optimization with optimal rotation matrix and step rate settings that converges to a stationary location as an optimized design candidate. The spiral optimization model$x(k) \in R^n$is from initial position, which converges to x^*as a center with a logarithmic spiral trajectory.

$$x(i+1) = x^* + r\Omega(\alpha)(x(i) - x^*)(i = 0,1,2,\dots) \quad \text{----- (4)}$$

here $\alpha \in [-\pi, \pi]$ and step rate of distance between x^* and $x(i)$ is specified by r and $\Omega(\alpha)$ is rotation matrix

$$\Omega(\alpha) = (-1)^{\beta}\varphi_{i1j1}(\alpha) \times \dots \times \varphi_{i\tau j\tau}(\alpha), \beta \in \{0,1\} \text{-----(5)}$$

$$\varphi_{i\tau j\tau}(\alpha).\,l = \begin{matrix} 1 & 0 & 0 \\ \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 1 \end{matrix} \text{------(6)}$$

Where $l = \{1,2,\dots\tau\}$

The spiral optimization method emphasizes two entities, namely intensification and diversification. Diversification is the factor which conducts the worldwide search, whilst intensification represents for exploring and executing deep searches. The hybrid rotational matrix is created with a critical factor in mind.$\varphi(\alpha)$ and step rate is given by $r(k)$where$k$ is maximum iteration in spiral algorithm.

The performance of the search is determined by the composites rotation matrix R the step rate r(k), and the beginning coordinates x i (0). The specified spiral optimization technique is not mathematically that it will converge to an equilibrium position. The SPO method is developed from direct search method, which has been shown to convergence to a constant location when there is a continuous differentiable objective function. To improve the outcome, the finite searching vector are incorporated into the existing best position. With this method, at least 1 vector direction is decreasing, and the length of the targeted search vectors is lowered when the search fails. A multi-objective spiral optimization technique is developed to attain the best exergy efficiency and product cost. Multi-objective method have several objectives (m) and decision variable (n) thatcan begiven by mathematical terms

$$minF(X) = [f_1(X), f_2(X), \dots . f_m(X)]^T \text{-----(7)}$$

Subjects to

$$g_j(X) \leq 0, j = 1,2 \dots . J \text{----(8)}$$

$$h_j(X) \leq 0, k = 1,2 \dots . K \text{---(9)}$$

$$x_{i,min} \leq x_1 \leq x_{i,max} \text{------(10)}$$

here $F(X)$is fitness function for mentioned problems, X is decision variable vector. Inequality constraint is given by $\boldsymbol{g_j(X)}$ and equality constraint is given by$\boldsymbol{h_j(X)}$**.** Although non-dominated sorting plays an important role in specifying improved solutions, recently summation-based sorting as well as diversified selection methods have been employed for multi-objective spiral optimization due to the balance of searching among distinct objectives.

The extracted features are Male, age, education, ex-smoker, cigsPerDay, BPMeds, prevalent Stroke, prevalent Hypertension, diabetes, totalChol, sysBP, diaBP, BMI, heart rate, glucose, and ten-year CHD. We selected the following features: male, current smoker, cigarettes per day, prevalent stroke.

7. Classification of signals using Deep Variational Autoencoder (DVAE)With Artificial Ecosystem Optimizer

Because it has a comparable structure (an encoder coupled to a decoder), DVAE is regarded a generalization version of a conventional auto encoder (AE), but it accomplishes its work differently, as seen in figure-2. Whereas an AE learns the compressed representation of an input and then recreates that inputs from that representations, a DVAE learns all

characteristics of a probability density P of the input's latent space representation X. is important generation system having encoder network $Q\varphi(Z|X)$ and decoder network $P\theta\ (X|Z)$. The gradient descent approach may be used to train DVAE to learn approximation inference. The parameterized encoder network φ learns an effective compressing of the data in with this lower-dimensional space, mapping data X into a continuously latent variable Z. The parameterized decoder network generates data by using θ latent variable, which translates Z to reconstructed data ˆ. Deep neural networks are used for building encoder and decoder with parameters θ and φ.

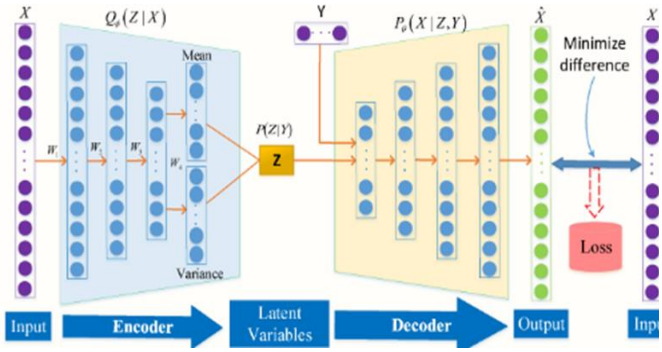

Fig-3 diagrammatic representation of Deep Variational Auto Encoder

The DVAE's encoding section consists of 4 2D convolutional layers and 4 max-pooling layers that are utilized alternately. 3 2x2 filters and 1 2x3 filter are used for down sampling with max-pooling layers. The DVAE returns two vectors of the same length: the mean and the standard deviation. The latent transfer function z is sampled using both vectors. The restoration loss of every training sample is calculated depending on the class, and the highest restoration loss for every class is used as the screening criterion. We can assume that the decoder $P((x_i|z,y),\ here\ i=i,2,\dots n)$ obeys Bernoulli distribution, i.e.,

$$P(x=1|z,y)=\propto_{z,y}\text{----------(11)}$$

$$P(x=0|z,y)=1-\propto_{z,y}\text{-------(12)}$$

For observation, likelihood is:

$$L=\alpha_{z,y}^{x}\left(1-\propto_{z,y}\right)^{1}-x\ \text{---------(13)}$$

Decoder's output is given as parameter of Bernoulli distribution, i.e.

$$\propto_{z,y}=decoder\ 9z,y)=x' \qquad \text{----(14)}$$

$$-\log L=-[x.\log(x')+(1-x).\log(1-x')\text{-----(15)}$$

Cross entropy is clearly represented by negative log - likelihood function in Equation (9). This cross entropy is used as the decoder's reconstruction loss. Following each training sample (xi,yi) is fed into trained DVAE, reconstruction loss $li(xi,yi)$ is calculated as follows:

$$li(xi,yi)=-[xi.\log(x')+(2-xi)\log(1-x')\text{ ------ (16)}$$

Decoder network is made up of 4 2D convolutional layers and 4 up-sampling layers that work together to reverse the encoding layers' work and produce a rebuilt version of a input signal. Except for the last convolutional layer (which uses sigmoid activation), all other convolutional layers use rectified linear unit (ReLU) relu activation. Here between convolutional and pooling layers, batch normalization layers are utilized to regularize and accelerate training. What distinguishes our system out of a standard DVAE architecture is that the hidden representations z vector is entirely coupled to a single neuron with such an activated sigmoid function to conduct preictal as well as interictal signal categorization. The DVAE training is performed only once in order to minimize the normal DVAE losses while simultaneously minimizing the supervised classification loss. The size of latent dimensional space is a parameter that may be changed. After testing with several numbers, we decided on a value of 16. In our DVAE, the Artificial Ecosystem method is employed as the optimizer, and binary bridge is used as the cost function for both the restoration and classification losses.

8. Artificial Ecosystem Optimizer

The behavior of the flow of energy in the earth system of life inspired the Artificial Ecosystem Optimizer. This process mimics some of the behaviors of live creatures, such as breakdown, consumption, and creation. The process of updating may be stated as follows:

i. Production Method: As per procedure, producer location is chosen at random, and the associated manufacturer is worst. For best solution, as represented by decomposer, is characterized as follows:

$$X\_1\ (t+1)=(1-d)\ X\_n\ (t)+d.X\_rand\ \ (t)\text{-----(17)}$$

$$d=(1-t/T\_max\ )\ 〖rand〗\_1\text{-------(18)}$$

$$X\_rand\ \ (t)=〖rand〗\_2\ \ ().(ub-lb)+lb\text{--------(19)}$$

here t and T_maare Current iteration and number of iterations. The upper and lower boundaries of the search space are defined by ub and lb. 〖rand〗_1and 〖rand〗_2gives A variable in the interval [0,1] and a weight parameter d. X_rand (t) donates a solution generated randomly in the search space.

ii. Consumption procedure: in this procedure, the first user feeds the lower energy user or the producer.

The locations of herbivores can be modernized only with respect to producers:

X_i (t+1)=X_i (t)+K.(X_i (t)-X_1 (t))-------(20)

Here X_1 gives location of producer and Kgives parameter for consumption, it can be calculated by levy flight by using the following equations:

K=1/2 u/v,u∈norm(0,1),v∈norm(0,1)------(21)

Here norm(0,1) Variable generated using normal distribution with "0" mean and unit variance.

Carnivores can be updated through arbitrary customers with different levels of energy. The procedure can be modeled as follows:

X_i (t+1)=X_i (t)+K.(X_i (t)-X_l (t))

l=randi([2i-1]),i=1,2,….N-------(22)

Here randi[a,b] In this function, random integer numbers are generated [a,b]] C. Omnivores' positions are updated based on both producer and randomly chosen consumer with high energy index levels(l) as framed follows:

X_i (t+1)=X_i (t)+K.(〖rand〗_3.X_i (t)-X_1 (t)))+(1-〖rand〗_3).(X_i (t)-X_l (t),l=randi----------(23)

iii. Decomposition process: This is the final stage of living process where each operator passes onto remaining components are separated. This step is referred to as AEO exploitation and is written as

X_i (t+1)=X_i (t)+D.(e.X_n (t)-h.X_i (t))i=0,1,2..N

h=2.〖rand〗_4-1--------(24)

As shown in Equation (7), parameter D gives the decomposition factor, while parameters h and e give the weight parameters. 〖rand〗_4- A random number is generated between [0,1] AEO has several benefits above all other MH methods, including the fact that no parameters must be established during optimization. Furthermore, Consequently, it achieves better convergence and avoids becoming stuck at local optima due to improved exploration and exploitation capabilities.

Bilinear procedure and pooling produce a bilinear vector with size 64 × 64 × 1. Predictions are derived using the final layer, a soft - max layer, with every node representing the likelihood that the control signal corresponds to a certain seizure type. To address the issue of class imbalance, class weights are added into the classifier's training, giving larger weights to minority classes with lower weights to organization are generally. To avoid over-fitting throughout the training phase, early halting is used as a regularization strategy. The approach measures validation loss and stops the training phase if it does not improve after 10 epochs.

9. Evaluation of Performance

Based on the experimental results, There are five parameters: accuracy, precision, recall, specificity, and F1-score are analysed using the PYTHON software. The suggested Deep Variational Auto Encoder Including Artificial Ecosystem Optimizer (Deep VAE AEO) is compared to the current Ensemble machine learning (Ens ML), XGBoost Random Forest (XGBRF) Classifier, and Support Vector Classifier (SVC) utilizing the Framingham Heart Study-Cohort (FHS-Cohort).

These are the performance metrics:

Accuracy: An accuracy check involves comparing the total expected values with the number of predictors used to categorize successfully. Mathematically, it is expressed.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \text{--------(25)}$$

Precision: It determines the proportion of positive attributes to total anticipated values. Mathematically, it is expressed.

$$\text{Precision} = \frac{TP}{TP+FP} \text{----------(26)}$$

Recall: It calculates the percentage of correct predictions to total predictions. It is expressed mathematically.

$$Recall = \frac{TP}{TP+FN} \text{----------(27)}$$

F1-Score: Used to calculate the mean average accuracy to recall ratio. Equation provides the F1-score.

$$\text{F1} - \text{Score} = 2 * \frac{Precision*recall}{precision+recall} \text{---------(28)}$$

Table 1 - Train and test for the FHS-Cohort dataset

Parameters	Values for Testing	Values for Training
accuracy	0.9796	0.9761
Precision	0.9667	0.9857
Recall	0.8735	0.864
F1-score	0.9177	0.9208

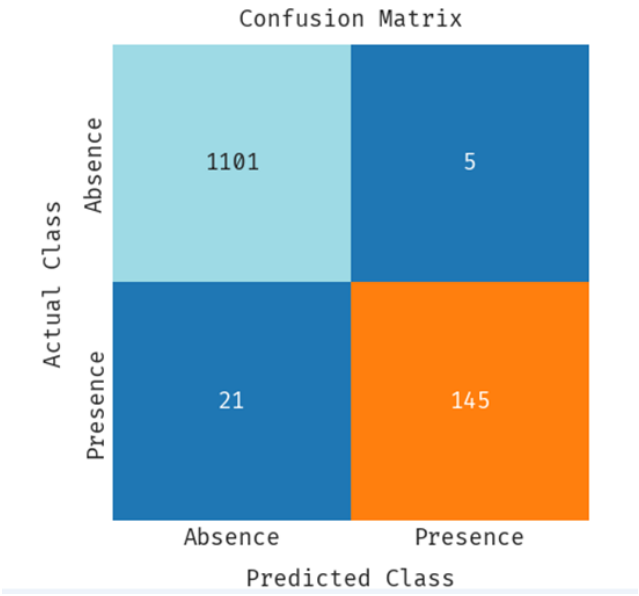

Fig-4 A confusion matrix for the validation of CVD tests

Figure 3 shows the confusion matrix represents the expected class and the actual class in CVD test validation. Diagonally colored cells represent correctly or incorrectly classified systems. Each class is represented by a column on the right, and its actual performance is reflected in a row at the bottom.

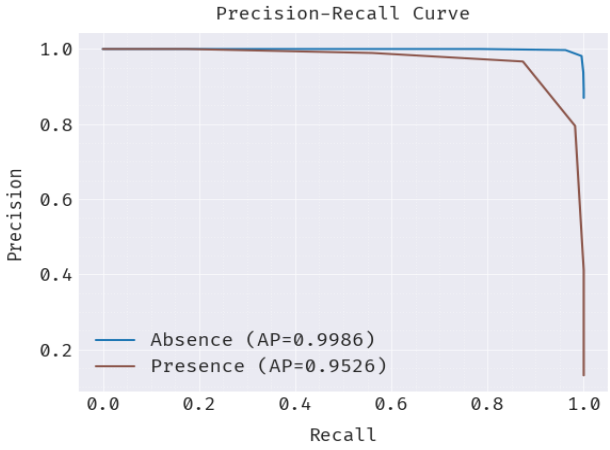

Fig-5 precision-recall curve CVD testing validation

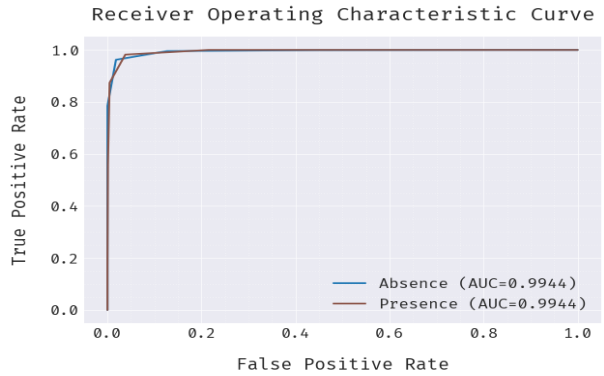

Fig-6 A ROC curve for validating CVD testing

A precision-recall curve is shown in **figure 4,** where recall is represented on the x-axis and accuracy is represented on the y-axis. It has been discovered that an average precision (AP) of 0.9986 denotes absence and 0.9526 suggests presence. In Figure 5, the ROC curve for CVD testing validation is shown with the x-axis representing false positives and the y-axis representing true positives. There is an AUC of 0.994 for absences and 0.9944 for present cases.

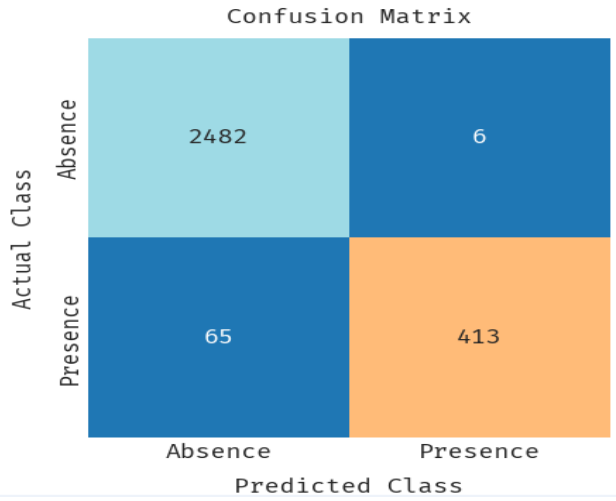

Fig-7 confusion matrix for CVD Training validation

A clustering algorithm for CVD Trained validation is shown in Figure 6, The rows represent the class value and the columns represent the actual information type. Diagonally colored cells indicate correctly classified electric networks and erroneously classified electric networks. On the right side, you can see the anticipated performance of each class, and on the bottom row you can see the actual performance of each class.

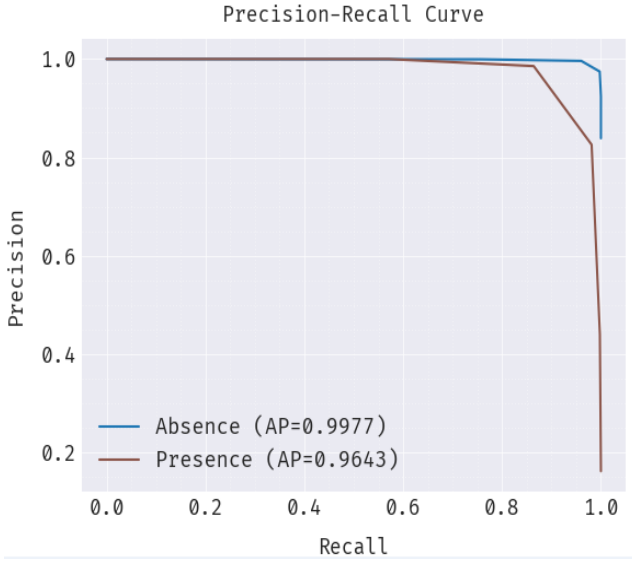

Fig-8 precision-recall curve CVD training validation

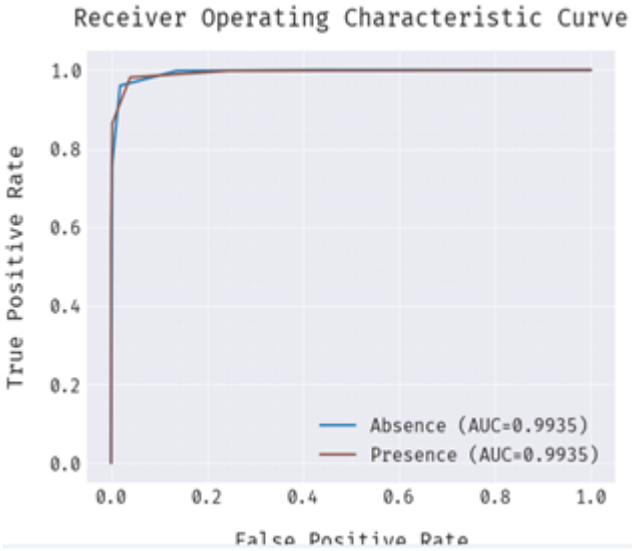

Fig-9 ROC curve CVD training validation

In Figure 7, The x axis represents recall, and the y axis represents accuracy CVD training validation. Average accuracy (AP) is found to be 0.9977 for absence and 0.9643 for presence. ROC curve for CVD training verification is shown in Figure 5, The x-axis represents false positives, while the y-axis represents true positives. Absence has an AUC of 0.9935 and presence has an AUC of 0.9935.

Table-2 overall comparative analysis between existing and proposed **methods**

Parameters	Ens_ML	XGBRF	SVC	Deep_VAE_AEO (proposed)
Accuracy (%)	**90.24**	**75.63**	**64.99**	**97**
Precision (%)	**92**	**73.13**	**65.64**	**98**
Recall(%)	**88**	**68.25**	**38.58**	**87**
F1-score (%)	**90**	**70.61**	**48.60**	**82**

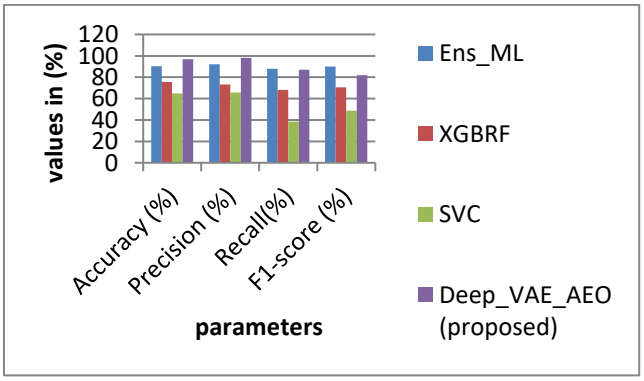

Fig-10 overall comparative analysis

The above figure 8 depicts the total comparison analysis, and the x-axis represents the parameters On the y-axis, we see percentage values. Deep VAE AEO outperforms Ens ML in terms of accuracy by 7.24%, 6%, and 1%8%. Better than XGBRF by 22.63%, 25.13%, 21.25%, and 12.61%. Better than SVC in 33.99%, 33.64%, 49.58%, and 34.60%.

10. Conclusion

We discovered in this work that classification technique of CVD datasets is difficult due to the size, dispersion of records, and amount of significant factors. We proved that feature selection strategies were effective in reducing a significant number of characteristics. The goal of this effort is to categorize 4 forms of CVD. The distribution of each category is extremely imbalanced, however, because a significant portion of the records suggest patients were not identified with any form of disease. To solve the categorization challenge, we created a neural network model including an optimizer. Despite the poor quality of our structured dataset, the suggested Deep Variation Auto Encoder Including Artificial Ecosystem Optimizer (Deep VAE AEO) obtained a good classification performance. We compared the suggested technique to three baseline models, demonstrating its poor classification performance versus the proposed neural network models. Other neural network architectures, like Generative Adversarial Network (GAN) or Attention-based Recurrent Neural Network, may be implemented in future research. This will provide a better understanding of the behavior of the forecast of various forms of heart disease. Additionally, changing each output target's binary classification into a multi-label categorization job may be explored for our dataset in order to correct its imbalance and maintain its structure.

Reference

[1] M. McClellan, N. Brown, R. M. Califf, and J. J. Warner, "Call to action: urgent challenges in cardiovascular disease: a presidential advisory from the American Heart Association," *Circulation*, vol. 139, no. 9, pp. e44–e54, 2019.

[2] E. J. Benjamin, M. J. Blaha, S. E. Chiuve et al., "Heart disease and stroke statistics—2017 update: a report from the American Heart Association," *Circulation*, vol. 135, no. 10, pp. e146–e603, 2017.

[3] S. Allender, P. Scarborough, V. Peto et al., "European cardiovascular disease statistics," *European Heart Network*, vol. 3, pp. 11–35, 2008.

[4] P. Bizopoulos and D. Koutsouris, "Deep learning in cardiology," *IEEE reviews in biomedical engineering*, vol. 12, pp. 168–193, 2018.

[5] N. Luo, C. M. O'Connor, L. B. Cooper et al., "Relationship between changing patient-reported

outcomes and subsequent clinical events in patients with chronic heart failure: insights from HF-ACTION," *European Journal of Heart Failure*, vol. 21, no. 1, pp. 63–70, 2019.

[6] H. G. C. Van Spall, S. F. Lee, F. Xie et al., "Effect of patient-c transitional care services on clinical outcomes in patients hospitalized for heart failure," *JAMA*, vol. 321, no. 8, pp. 753–761, 2019.

[7] V. L. Patel, E. H. Shortliffe, M. Stefanelli et al., "The coming of age of artificial intelligence in medicine," *Artificial Intelligence in Medicine*, vol. 46, no. 1, pp. 5–17, 2009.

[8] I. Kononenko, "Machine learning for medical diagnosis: history, state of the art and perspective," *Artificial Intelligence in Medicine*, vol. 23, no. 1, pp. 89–109, 2001.

[9] A. A. Voors, W. Ouwerkerk, F. Zannad et al., "Development and validation of multivariable models to predict mortality and hospitalization in patients with heart failure," *European Journal of Heart Failure*, vol. 19, no. 5, pp. 627–634, 2017.

[10] G. S. Bowen, M. S. Diop, L. Jiang, W. C. Wu, and J. L. Rudolph, "A multivariable prediction model for mortality in individuals admitted for heart failure," *Journal of the American Geriatrics Society*, vol. 66, no. 5, pp. 902–908, 2018.

[11] G. Lorenzoni, S. S. Sabato, C. Lanera et al., "Comparison of machine learning techniques for prediction of hospitalization in heart failure patients," *Journal of Clinical Medicine*, vol. 8, no. 9, p. 1298, 2019.

[12] E. D. Adler, A. A. Voors, L. Klein et al., "Improving risk prediction in heart failure using machine learning," *European Journal of Heart Failure*, vol. 22, no. 1, pp. 139–147, 2020.

[13] R. Chen, A. Lu, J. Wang et al., "Using machine learning to predict one-year cardiovascular events in patients with severe dilated cardiomyopathy," *European Journal of Radiology*, vol. 117, pp. 178–183, 2019.

[14] S. F. Weng, J. Reps, J. Kai, J. M. Garibaldi, and N. Qureshi, "Can machine-learning improve cardiovascular risk prediction using routine clinical data?" *PloS one*, vol. 12, no. 4, Article ID e0174944, 2017.

[15] S. E. Awan, M. Bennamoun, F. Sohel, F. M. Sanfilippo, and G. Dwivedi, "Machine learning-based prediction of heart failure readmission or death: implications of choosing the right model and the right metrics," *ESC heart failure*, vol. 6, no. 2, pp. 428–435, 2019.

[16] D. Chicco and G. Jurman, "Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone," *BMC Medical Informatics and Decision Making*, vol. 20, no. 1, p. 16, 2020.

[17] M. Porumb, E. Iadanza, S. Massaro, and L. Pecchia, "A convolutional neural network approach to detect congestive heart failure," *Biomedical Signal Processing and Control*, vol. 55, Article ID 101597, 2020.

[18] U. R. Acharya, H. Fujita, S. L. Oh et al., "Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals," *Applied Intelligence*, vol. 49, no. 1, pp. 16–27, 2019.

[19] Arroyo, J. C. T., &Delima, A. J. P. (2022). An optimized neural network using genetic algorithm for cardiovascular disease prediction. *Journal of Advances in Information Technology*, *13*(1).

[20] Cheng, X., Han, W., Liang, Y., Lin, X., Luo, J., Zhong, W., & Chen, D. (2022). Risk Prediction of Coronary Artery Stenosis in Patients with Coronary Heart Disease Based on Logistic Regression and Artificial Neural Network. *Computational and Mathematical Methods in Medicine*, *2022*.

[21] Bai, J., Fu, J., Du, S., Chen, X., & Zhang, C. (2022). Prediction of Heart Failure in Children with Congenital Heart Disease Based on Multichannel LSTM. *Mobile Information Systems*, *2022*.

[22] Krishnan, S., Magalingam, P., & Ibrahim, R. (2021). Hybrid deep learning model using recurrent neural network and gated recurrent unit for heart disease prediction. *International Journal of Electrical & Computer Engineering (2088-8708)*, *11*(6).

[23] Mienye, I. D., & Sun, Y. (2021). Improved heart disease prediction using particle swarm optimization based stacked sparse autoencoder. *Electronics*, *10*(19), 2347.

[24] Ali, L., Rahman, A., Khan, A., Zhou, M., Javeed, A., & Khan, J. A. (2019). An automated diagnostic system for heart disease prediction based on ${\chi^{2}} $ statistical model and optimally configured deep neural network. *IEEE Access*, *7*, 34938-34945.

[25] Mienye, I. D., Sun, Y., & Wang, Z. (2020). Improved sparse autoencoder based artificial neural network approach for prediction of heart disease. *Informatics in Medicine Unlocked*, *18*, 100307.

[26] Maji, S., & Arora, S. (2019). Decision tree algorithms for prediction of heart disease. In *Information and communication technology for competitive strategies* (pp. 447-454). Springer, Singapore.

[27] Subhadra, K., & Vikas, B. (2019). Neural network based intelligent system for predicting heart disease. *International Journal of Innovative Technology and Exploring Engineering*, *8*(5), 484-487.

[28] Latha, C. B. C., & Jeeva, S. C. (2019). Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques. *Informatics in Medicine Unlocked*, *16*, 100203.

https://www.kaggle.com/eeshanpaul/framingham.

[29] Tamura K, Yasuda K. The Spiral Optimization Algorithm: Convergence Conditions and Settings. IEEE Trans Syst Man, Cybern Syst 2020.

[30] Anand, D., Tata, V., Samriya, J.K., Kumar, M. (2023). A Review on Deep Learning-Enabled Healthcare Prediction Technique: An Emerging Digital Governance Approach. Soft Computing: Theories and Applications. Lecture Notes in Networks and Systems, vol 627. Springer, Singapore. https://doi.org/10.1007/978-981-19-9858-4_22