EMITTER International Journal of Engineering Technology
Vol. 13, No. 1, June 2025, pp. 124~138
DOI : 10.24003/emitter.v13i1.936

Utilizing Evolutionary Mating Algorithm Optimized Deep Learning to Assess Cardiovascular Diseases Risk

Ahmed Mohammed Ahmed Alsarori[1], Mohd Herwan Sulaiman[1]

[1]Faculty of Electrical & Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA), Pekan Pahang, Malaysia.
Corresponding Author: herwan@umpsa.edu.my

Received January 2, 2025; Revised March 6, 2025; Accepted May 2, 2025

Abstract

Cardiovascular Diseases (CVD) continue to be a primary cause of death worldwide, underscoring the critical importance of early and accurate risk prediction. However, traditional predictive models struggle with the complexity and interdependencies in medical data. This study addresses this gap by proposing a deep learning-based risk assessment model optimized with the Evolutionary Mating Algorithm (EMA) to enhance prediction accuracy and efficiency. Our contributions include developing a dedicated risk variable for machine learning applications and benchmarking the EMA-optimized model against ADAM and Particle Swarm Optimization (PSO). The proposed method was evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Coefficient of Determination (R^2), and Standard Deviation (STD). Experimental results demonstrate that the EMA-optimized model outperforms traditional optimization methods, achieving an MAE of 0.037, RMSE of 0.0464, and an R^2 of approximately 0.91. These results highlight the effectiveness of EMA in enhancing cardiovascular risk assessment models, providing a more reliable tool for early diagnosis and clinical decision-making.

Keywords: Heart Disease, Deep Learning, Artificial Neural Network, Risk Prediction, Evolutionary Mating Algorithm.

1. INTRODUCTION

Chronic diseases are a major global health concern, contributing to 74% of all deaths, with cardiovascular diseases (CVDs) being a leading cause [1]. In 2023 alone, 41 million deaths were attributed to non-communicable diseases, and 88% of Americans over 65 suffer from at least one chronic condition [2]. CVDs, often resulting from arterial plaque buildup (atherosclerosis) and blood clot risks, remain a significant cause of mortality worldwide, particularly in developing nations. Given the complexity of CVD risk factors, early prediction models are essential for timely intervention and treatment [2–4].

Machine learning (ML) has transformed medical decision-making, enabling disease classification and risk prediction [5,6]. Deep learning models have been widely applied in cardiovascular risk assessment, leveraging factors

like blood pressure, cholesterol, and diabetes [4,7–9]. However, traditional optimizers like ADAM often struggle with global optimization, impacting predictive accuracy [3,7,10,11].

Evolutionary algorithms (EAs), inspired by natural selection, have proven effective in optimizing various domains, including healthcare [12–14]. In this study, we introduce the Evolutionary Mating Algorithm (EMA) to enhance deep learning models for cardiovascular risk prediction. EMA efficiently explores the solution space, improving accuracy compared to conventional optimizers [15,16].

This research evaluates the effectiveness of a Feedforward Neural Network (FFNN) optimized using EMA for CVD risk assessment. Performance is compared with ADAM and PSO to determine the most suitable optimization approach. Our findings aim to advance predictive modeling in cardiovascular healthcare by improving predictive reliability in medical applications.

2. RELATED WORKS

Mean Arterial Pressure's correlation with significant cardiovascular disease (CVD) events has been substantiated by the ADVANCE study [2,17]. Notably, research involving individuals with type 2 diabetes has revealed a direct association, showing that for every 13 mmHg increase in MAP, there is a corresponding 13% rise in the risk of CVD. Furthermore, an elevation in MAP is anticipated to result in a higher incidence of CVD-related hospitalizations among individuals with type 2 diabetes [17]. These findings underscore a clear and direct link between MAP and the risk of cardiovascular disease.

Feedforward neural networks have been successfully used for battery SOC estimation, while evolutionary machine learning has shown promise in CVD risk prediction. EMA, as a recent advancement, has demonstrated superior optimization capabilities, suggesting its potential for further exploration in both energy and possibly medical applications.

The study titled "An evolutionary machine learning algorithm for cardiovascular disease risk prediction" applied an evolutionary machine learning algorithm to predict cardiovascular disease (CVD) risk, leveraging evolutionary principles to optimize model performance. The study utilized a population-based search approach to refine predictive features and improve classification accuracy. Results showed that the evolutionary model achieved higher predictive performance compared to traditional machine learning methods, demonstrating the potential of evolutionary optimization in healthcare applications [18].

In the study titled "Using the evolutionary mating algorithm for optimizing deep learning parameters for battery state of charge estimation of electric vehicle", the Evolutionary Mating Algorithm (EMA) was explored to optimize deep learning parameters for battery SOC estimation. EMA, inspired by evolutionary mating strategies, was employed to fine-tune hyperparameters, leading to enhanced model accuracy and convergence speed. Their findings showed that EMA outperformed conventional

optimization methods, achieving MAE of 3.458% and RMSE of 4.7%, reducing error rates and improving the overall predictive capability of the deep learning model [19].

3. ORIGINALITY

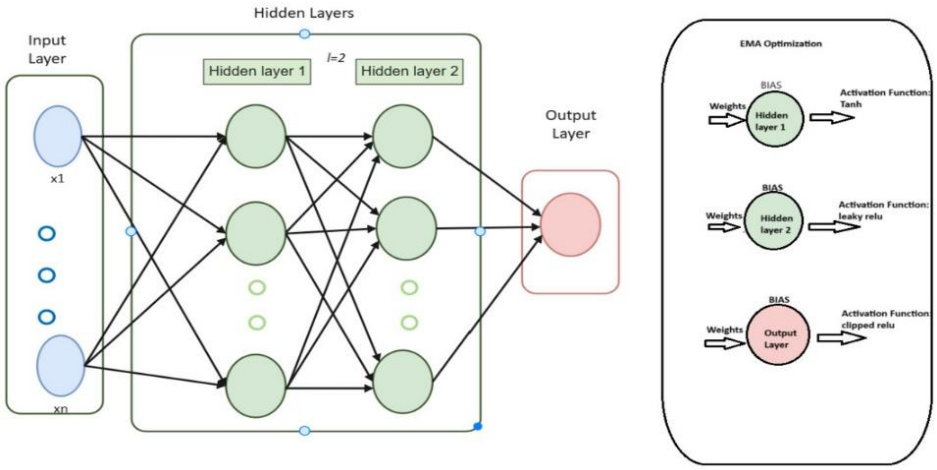

Figure 1. FFNN model

The core objective of this paper is to assess a patient's risk of developing cardiovascular diseases by employing machine learning techniques for prediction. The approach used would involve feature engineering a risk factor variable, which would indicate the risk level a patient had. This approach offers valuable support to medical practitioners, enabling them to input basic patient data and preliminary test results to estimate the patient's risk level for cardiovascular diseases accurately. To achieve this objective, the study concentrates on optimizing two key parameters of the deep learning model—weights and biases—by minimizing the Mean Absolute Error (MAE). MAE is employed as the objective function to improve predictive accuracy and is formulated as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \overline{y}_i\right| \quad (1)$$

In equation (1), n stands for the number of observations, y_i stands for the actual values, and $\overline{y}_i$ stands for predicted values.

This study employs the Evolutionary Mating Algorithm (EMA) to optimize and fine-tune the weights and biases of the proposed deep learning (DL) model, with the Feedforward Neural Network (FFNN) serving as the foundational model. The optimization process involves integrating the EMA function into the FFNN framework and executing it until the predefined maximum number of iterations is reached. Upon completion of training using the training dataset, the EMA-optimized DL model is then evaluated on previously unseen data from the testing dataset. Figure 1 provides a visual representation of the proposed EMA-DL framework. Further elaboration and thought processes can be seen from the flowchart of Figure 2.

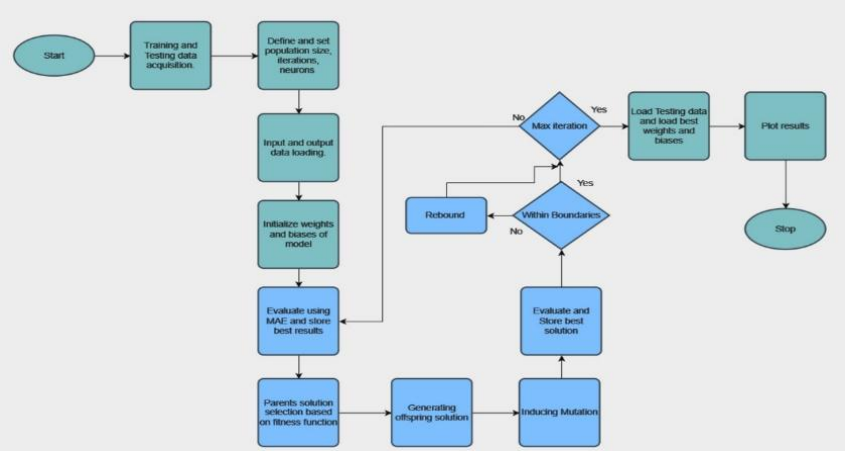

Figure 2. EMA-DL flowchart

4. SYSTEM DESIGN

4.1 Dataset

This study employs a cardiovascular disease dataset, with all pertinent details regarding its structure and usage outlined in this chapter. Table 2 presents the data configuration specifications used for both training and testing phases.

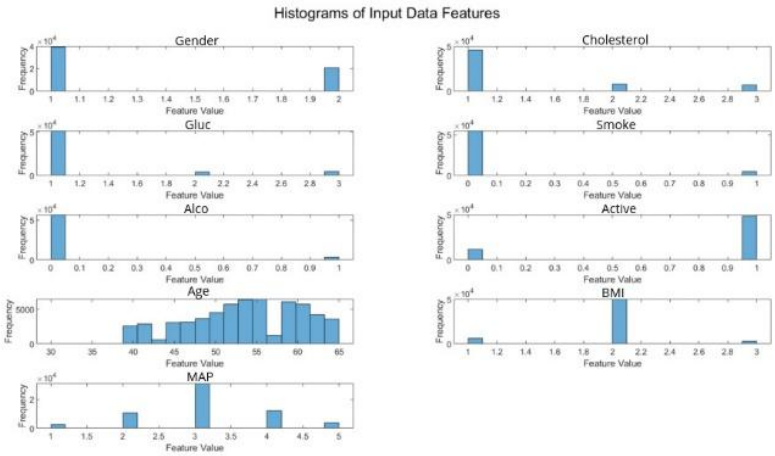

Figure 3. Input Variables

It is important to note that the risk assessment variable is continuous in nature, expressed as a percentile. While it theoretically ranges from 0 to 100, the observed values fall within the range of 50 to 97.8. The interpretation and application of this variable are intended to be at the discretion of the medical practitioner.

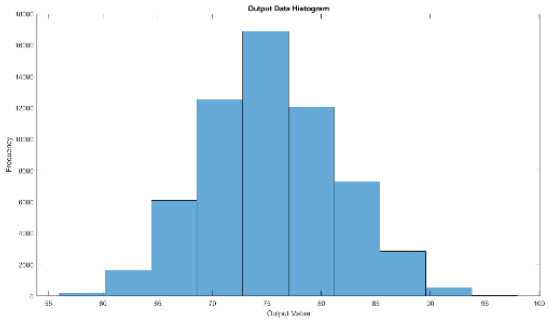

Figure 4. Output Variable (Risk)

The dataset utilized for this research comprises 70,000 patient records, as documented in the reference [20] along with their corresponding features outlined in the provided Table 1. This rich dataset is instrumental in supporting the investigation and drawing meaningful insights related to the study's objectives.

Table 1. Dataset variables

Features	*Variables*	*Missing Values*
ID	id	Non
Age	Age	Non
Height	Height	Non
Weight	Weight	Non
Gender	Gender	Non
Systolic blood pressure	ap_hi	Non
Diastolic blood pressure	ap_lo	Non
Cholesterol	Cholesterol	Non
Glucose	Gluc	Non
Smoking	Smoke	Non
Alcohol intake	Alco	Non
Physical activity	Active	Non
Cardiovascular disease	Cardio	Non

The table presented outlines the 13 original features or variables included in the dataset before preprocessing. Although most of these features are self-explanatory, it is important to highlight that the "ID" feature, associated with patient identification, will be excluded from all models in this study. Additionally, to prepare the dataset for input into the deep learning algorithm—a feedforward neural network—the data must be normalized and transposed. These steps will be performed following the preprocessing stage and the incorporation of any feature-engineered variables.

The data processing began with a comprehensive assessment to identify and address any missing or duplicate entries within the dataset. Following the completion of this initial phase, the identification and treatment of outliers were undertaken.

4.2 Features

By segmenting continuous inputs into discrete groups or bins, the algorithm was better equipped to capture nuanced differences across data classes. For instance, the age variable—originally recorded in days—was transformed into the standard age format in years [17].

Additionally, new variables were introduced for use in the study, with Mean Arterial Pressure (MAP) and Body Mass Index (BMI) identified as essential features across all risk assessment models. The risk variable itself functioned as the primary target output for these models.

In the medical field, Mean Arterial Pressure (MAP) represents the average blood pressure a person experiences throughout a single cardiac cycle and is calculated as follows:

$$Mean\ Arterial\ Pressure\ (MAP) = \frac{2\ ap\_lo + ap\_hi}{3} \quad (2)$$

This feature serves as a crucial indicator of both peripheral resistance and cardiac output. These findings underscore a clear and direct link between MAP and the risk of cardiovascular disease.

$$Body\ Mass\ Index\ (BMI) = \frac{weight}{height^2} \quad (3)$$

A comprehensive comparative examination revealed that individuals classified as obese experienced an earlier onset of incident Cardiovascular Disease (CVD), a greater proportion of their lifespan marked by CVD-related morbidity, and a reduced overall survival rate.

What sets this study apart is its emphasis on the risk variable, which is designed to estimate a patient's likelihood of developing Cardiovascular Disease (CVD). This variable incorporates critical factors including “Age,” “MAP,” “Cholesterol,” and “BMI,” and is calculated using the following formula:

$$Risk = \frac{Age + MAP + cholesterol + BMI}{oldest\ age + highest\ MAP + highest\ cholesterol + highest\ BMI} \times 100 \quad (4)$$

The selection of these components is grounded in established medical research [21], which highlights the significant roles that age and cholesterol play in the early detection and assessment of CVD risk. Indeed, age has consistently emerged as a significant factor, with older individuals being more susceptible to the onset of CVD [22].

Table 2 presents the complete set of variables utilized in this study, following the completion of data analysis and preprocessing procedures. As stated earlier the formulated features, MAP, and BMI are necessary features for any model that uses a medical dataset (as long as the initial variables are within the dataset) while the variable RISK identifies how much risk a patient is in to get afflicted with a cardiovascular disease.

Figure 3 shows the figures of all the inputs, while Figure 4 shows the output prior to transposing them. It clearly shows the distribution of data within the dataset. It illustrates the selected input variables, how they are provided to the model, and the nature of the output, specifically, the predicted values the model is designed to generate.

Table 2. Input and Output Variables

Features	*Variables*	*Missing Values*	*Preview*	*Input/Output*
Age (years old)	Age	Non	55, 76, 62, 42	Input
Body Mass Index	BMI	Non	1 – Normal 2 – Overweight 3 – Obese	Input
Mean Arterial Pressure	MAP	Non	1 – Normal 2 –Elevated 3 – Elevated (within early stage 1 hypertension) 4 – Stage 1 Hypertension 5 – Stage 2 hypertension	Input
Gender	Gender	Non	1 – Male 2 - Female	Input
Cholesterol	Cholesterol	Non	1 - normal 2 - high 3 – very high	Input
Glucose	Gluc	Non	1 - normal 2 – high 3 – very high	Input
Smoking	Smoke	Non	0 – non-smoker 1 - smoker	Input
Alcohol intake	Alco	Non	0 – doesn't drink 1 - drinks	Input
Physical activity	Active	Non	0 – not active 1 - active	Input
Risk factor (percentiles)	Risk	Non	64.32, 88.15, 72.98 (percentiles)	Output

4.3 Feed Forward Neural Network (FFNN)

Feed Forward Neural Network is a type of deep neural network where only a forward pass of data occurs, used to map the non-linearities governed by data, due to a multi-layer perceptron [23]. A deep network is typically characterized by an increase in the number of layers, featuring, at a minimum, 2 hidden layers [24]. The number of hidden layers, neurons, and activations are best selected after various experiments so that optimum results can be acquired. The activation function of each layer was based on ref. [25]. The equations of each activation utilized in this model are represented as follows:

$$\text{Input layer}: y = u \text{ (linear function)} \tag{5}$$

$$\text{Hidden layer 1: hyperbolic tangent (tanh) } y = \frac{e^{u} - e^{-u}}{e^{u} + e^{-u}} \tag{6}$$

$$\text{Hidden layer 2}: \text{leaky rectified linear unit (ReLU) } y = \max(0.3 * u, u) \tag{7}$$

$$\text{Output Layer}: \text{clipped ReLU } y = \begin{cases} 0, & u < 0 \\ u, & 0 \leq u \leq 1 \\ 1, & u > 1 \end{cases} \tag{8}$$

In the context of these equations, y represents the output generated by each neuron, while u denotes the total input to the neuron. This total input is calculated as the weighted sum of all inputs combined with the bias. Represented as:

$$u = \sum_i w_{ij} x_i + b_j \tag{9}$$

Here in equation (9), x_i represents the output from the i^{th} neuron or node of a previous layer, w_{ij} denotes the weight of the connection between layers i-j, and b_j refers to the bias of the current layer.

4.4 Evolutionary Mating Algorithm (EMA)

Evolutionary Mating Algorithm (EMA) represents a specialized variant of EAs that emphasizes recombination or mating-like operations among candidate solutions. By combining genetic information from parent solutions, EMA enhances solution diversity and exploration capabilities, facilitating the discovery of novel and potentially superior solutions within complex and dynamic solution spaces. EMA operates through iterative cycles of selection, crossover, and mutation among candidate solutions represented as chromosomes or parameter vectors encoding treatment strategies [19]. EMA combines elements of genetic algorithms and natural mating processes to function. It is a recently established metaheuristic algorithm that stems from Hardy-Weinberg (HW) principles and is inspired by the mating process in organisms.

$$x_m = \begin{bmatrix} x_1^1 & \cdots & x_1^d \\ \vdots & \ddots & \vdots \\ x_{n/2}^1 & \cdots & x_{n/2}^d \end{bmatrix} \tag{10}$$

$$x_f = \begin{bmatrix} x_{\frac{n}{2}+1}^1 & \cdots & x_{\frac{n}{2}+1}^d \\ \vdots & \ddots & \vdots \\ x_n^1 & \cdots & x_n^d \end{bmatrix} \tag{11}$$

$$I_{mates} = 1 + \left[var\left(x_{m,*}^T\right) - var\left(X_{f,*}^T\right)\right] \tag{12}$$

$$X_{child}^T = \begin{cases} p. * X_{m,*}^T + q. * X_{f,*}^T & for\ I_{mates} \geq 0 \\ p. * X_{f,*}^T + q. * X_{m,*}^T & for\ I_{mates} < 0 \end{cases} \tag{13}$$

$$p = randn(1, d) \tag{14}$$

$$q = (1 - p) \tag{15}$$

The algorithm begins by initializing a population of potential solutions, typically derived from the available dataset. According to equation (10), during the initialization phase, the candidate solution x is divided into two distinct groups: x_m, representing the male population, and x_f, representing the female population. Here, n denotes the population size, and d indicates the dimensionality of the problem.

Following initialization, the fitness function is evaluated for each individual in the population. Based on these evaluations, the most optimal solutions from both x_m and x_f are identified and retained, with fitter solutions being favored for the next steps.

Next, a crossover operation is performed to generate new offspring solutions by combining selected male and female solutions. This mating process, described in equation (12), is guided by sexual selection principles. The terms var $(x_{m,*}^T)$ and var $(x_{f,*}^T)$ represent the variances of the chosen male and female individuals, respectively, at iteration T, and I_{mates} denotes the EMA's mating mechanism.

Once mating concludes, the generation of offspring (X_{child}^{T}) proceeds, as shown in equation (13). The variables p and q, derived from equation (14) and (15), are based on normal random distributions and guide the variation in offspring creation. These new solutions are then evaluated using the fitness function, and the best-performing individuals are selected to serve as parents in subsequent generations. This cycle is repeated iteratively until the algorithm converges.

Additionally, two parameters—Cr and r—are introduced to simulate environmental pressures such as predator encounters. Cr represents the crossover probability, while r denotes the likelihood of encountering predators, both of which influence the evolutionary dynamics of the algorithm. With environmental changes, the best solution's characteristics would significantly be altered, since the offspring could be assumed to be dead or alive.

5. EXPERIMENT AND ANALYSIS

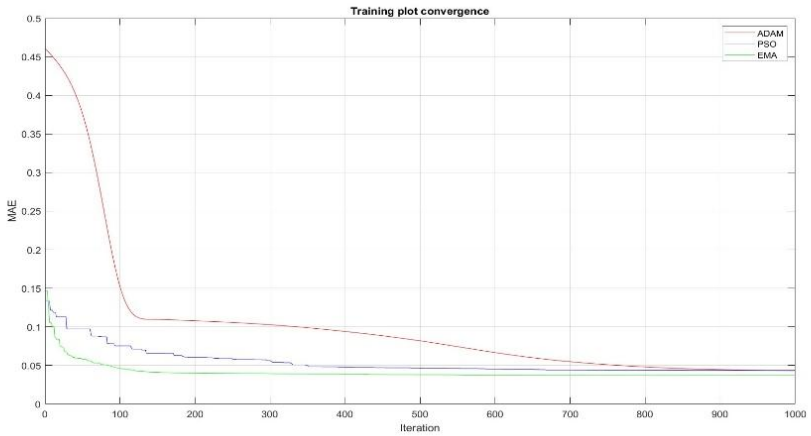

Figure 5. Convergence Plot

The study was carried out using MATLAB on a laptop equipped with an Intel Core i7 processor, Intel Iris Xe graphics, and 16GB of RAM. MATLAB served as the main platform for dataset processing and model development. The dataset underwent a division into groups of two subsets: a training dataset and a testing dataset, with 80-20, 70-30, and 60-40 splitting ratios. Subsequently, each model was trained using the training dataset, and its efficacy was evaluated through the testing dataset. The initial dataset comprised 70,000 rows and 13 attributes. However, after undergoing preprocessing and data cleaning, the dataset was refined to approximately 60,000 rows and 10 attributes, specifically tailored for the risk assessment task. Normalization was done to standardize and scale the data. Finally, it was transposed and then provided to the models.

This study employs the performance metrics mean absolute error, mean squared error, root mean squared error, standard deviation, coefficient of determination, and maximum error to evaluate risk assessment. Models

executed under each data partition were compared to determine the most suitable partition for this research.

The efficiency and convergence of each model's training process are illustrated in Figure 5, which highlights the 60:40 data partition, identified as the optimal split ratio following multiple trials. Table 3 presents the outcomes associated with this partition, reflecting the effectiveness of each model's optimizer during training. Notably, all models were trained a minimum of ten times to ensure consistency and reliability in the results.

Table 3. Optimized DL algorithms

(%)		R2	MAE	MSE	RMSE	MAX Error	Precision	Standard of Deviation	Computational time
Adam-dl	60:40	87.9	4.3	.28	5.3	24	53	5.27	20 sec
Pso-dl		86.8	4.3	.30	5.5	34	77	5.51	10 min
Ema-dl		90.6	3.7	.22	4.6	19	87	4.6	220 sec

Table 3 presents a comparative analysis of three optimization techniques—Adam-dl, PSO-dl, and Ema-dl—evaluated based on their performance using key metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Coefficient of Determination (R^2), Standard Deviation (STD), precision, and computational time at the optimal data partition ratio. Adam-dl stands out as the most time-efficient optimizer, requiring only 20 seconds on average for training across all splits, whereas Ema-dl takes significantly longer (220 – 300 seconds on average) across all splits. Despite its slower computation, Ema-dl consistently demonstrates superior performance, achieving the best metric values in all partitions. This highlights its capability to deliver high-quality results, making it a strong contender in applications where predictive performance outweighs the importance of computational speed.

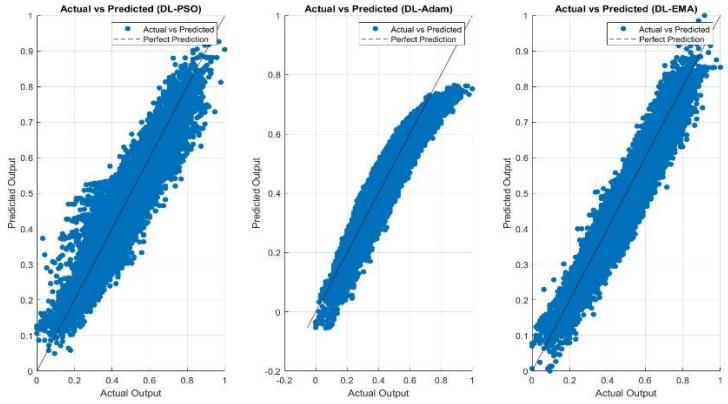

Figure 6. Scatter plot of Actual vs predicted (60/40)

Ema-dl's superior performance stems from its evolutionary approach, which optimally balances the exploration and exploitation of the solution space. For instance, at the 60:40 data partition, Ema-dl achieves an MSE of

0.22% and an MAE of 3.7%, outperforming both Adam-dl and PSO-dl. On the other hand, Adam-dl, though computationally efficient, exhibits slightly worse evaluation results, indicating a trade-off between speed and precision. PSO-dl's results fall between Adam-dl and Ema-dl, but it fails to outperform Ema-dl in any specific aspect.

The choice of data partition also plays a critical role in the model's performance. The 60:40 split, where 60% of the data is used for training and 40% for testing, produces the most favourable results across all optimizers. This balance allows the model to generalize well by leveraging sufficient training data while maintaining diversity in the testing dataset. As the proportion of training data increases (70:30 and 80:20), the performance slightly declines, likely due to reduced testing data, which affects the model's ability to evaluate generalization effectively.

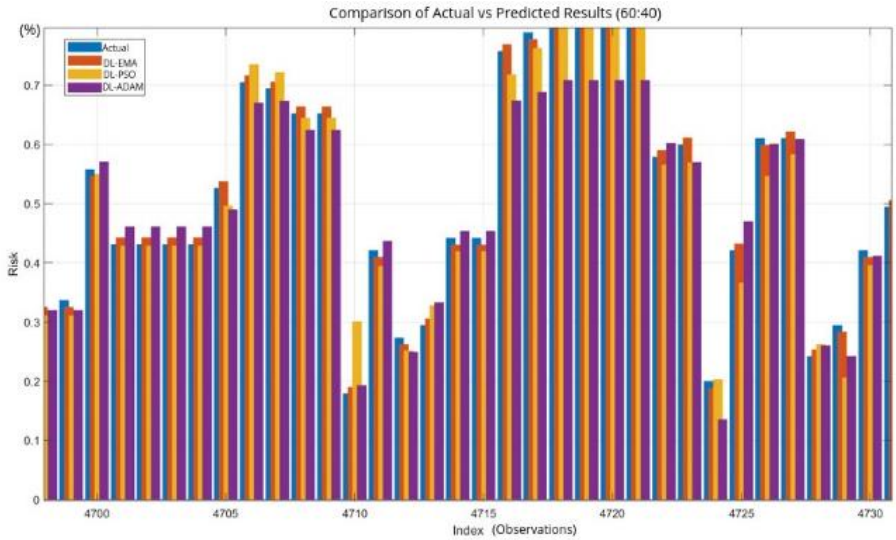

Figure 7. Risk assessment Actual vs Predicted

The plot in Figure 5 is the training convergence plot showing the Mean Absolute Error (MAE) over its run (the iterations/epochs) for different optimization methods at the 60:40 splitting ratio. Among the convergence shown by each model, models that had been optimized by EMA appear to have the fastest initial convergence rate, as it was shown to reach lower MAE values quicker than models optimized by ADAM and PSO. Models optimized by ADAM tend to have a smooth convergence rate and generally the best convergent rate but not the lowest MAE overall.

The Evolutionary Mating Algorithm (EMA) has proven highly effective in optimizing deep learning models, especially by enhancing the distribution and accuracy of predictions. As illustrated in Figure 7, EMA yields results that more closely match actual outcomes compared to other optimization techniques. This superior distribution suggests a balanced and consistent performance across various data points, which is crucial for robust and reliable models, which could be inferred from Table 3's r^2 and precision results of ema-dl, being 90.6% and 87% respectively.

Table 4. Sample table of observations (Actual vs predicted)

Sample	Inputs	Output (Risk - Actual)	EMA - DL	PSO - DL	ADAM - DL
1	*Gender (2)* *Cholesterol (1)* *Glucose level (1)* *Smoking (0)* *Alcoholic (0)* *Active (0)* *Age (62)* *BMI (1)* *MAP (3)*	24.21%	22.78%	21.27%	23.75%
2	*Gender (1)* *Cholesterol (1)* *Glucose level (1)* *Smoking (0)* *Alcoholic (0)* *Active (0)* *Age (57)* *BMI (1)* *MAP (2)*	30.53	30.14	29.57	31.34%
3	*Gender (2)* *Cholesterol (1)* *Glucose level (1)* *Smoking (0)* *Alcoholic (0)* *Active (1)* *Age (40)* *BMI (1)* *MAP (2)*	11.58%	13.35%	14.3%	8.13%
4	*Gender (1)* *Cholesterol (1)* *Glucose level (1)* *Smoking (0)* *Alcoholic (1)* *Active (1)* *Age (44)* *BMI (2)* *MAP (3)*	33.68%	33.83%	32.16%	37.08%
5	*Gender (1)* *Cholesterol (1)* *Glucose level (1)* *Smoking (0)* *Alcoholic (0)* *Active (1)* *Age (62)* *BMI (1)* *MAP (3)*	46.32%	46.08%	41.24%	48.02%

One of the key strengths of EMA lies in its ability to enhance precision, which reflects the consistency and reliability of predictions. Table 4 shows a sample of the model's testing results. It includes the actual risk values and the predicted risk values of all optimized models. The table contains a random sampling of 5 entries.

In terms of performance metrics, EMA-optimized models achieve remarkable results, boasting an MAE of 0.037 and an RMSE of 0.0464. These metrics indicate that EMA minimizes prediction errors more effectively than other optimization methods. These metrics highlight EMA's overall robustness in producing dependable models that can consistently minimize prediction inaccuracies.

While the ADAM optimizer offers a faster and more straightforward training process, EMA outshines it in producing more precise and better-distributed predictions. Figure 6 further enhances this point by showing the scatterplot of real vs predicted results of all three optimized models. From these plots, it could be seen that the Adam-optimized model, though it strives to deliver precise results, is neither as consistent as the other 2 optimized models nor as precise. Ema-optimized models show the most consistent and precise results.

5.1 Application and limitations

A practical application of this algorithm is its integration into clinical decision support systems for early cardiovascular disease risk assessment. Given that the model outputs a percentile risk score (e.g., 0.44 indicating a 44% risk), it can be used as a preliminary screening tool in hospitals and telemedicine platforms. Physicians could input basic patient data (blood pressure, cholesterol levels, BMI, age) into the system to compute a personalized risk score, and help prioritize high-risk patients for further evaluation. However, this study has some limitations. First, the model was only designed and tested within a controlled dataset, meaning real-world clinical validation is still required. Additionally, the algorithm does not incorporate feature selection, which might impact model interpretability. Lastly, while EMA outperformed ADAM and PSO in accuracy, it has a higher computational cost, which could affect scalability in large-scale healthcare applications. Future work should focus on optimizing the algorithm for real-time deployment and evaluating its effectiveness in diverse patient populations.

6. CONCLUSION

In this conducted study, an innovative approach was employed to assess the risk factor associated with the likelihood of each patient developing cardiovascular disease. The primary objective of this research was to serve as a preliminary or early diagnostic tool for medical practitioners, aiding them in evaluating the potential risk of cardiovascular diseases for individual patients.

In the course of this investigation, the risk assessment process was executed mainly through the application of a feed forward neural network optimized by the EMA (ema-dl) model. The results acquired from this investigation were better than the other optimization techniques, specifically ADAM and PSO. Notably, the results acquired indicated that the ema-dl model algorithms consistently exhibited better performance, with the 60:40 data split ema-dl model algorithm achieving the best MAE and RMSE, reaching approximately 0.037 and 0.0464 respectively. The only downside is its computational time, which is generally higher than the other optimized model, adam-dl. This could be mitigated by running the model on a better-performing device or improving the algorithm itself.

Therefore, this study demonstrated the feasibility of incorporating a preliminary risk assessment as a potential solution within the medical sector.

With further refinement, this approach could potentially be employed to promptly address and treat high-risk patients, or at the very least, alleviate symptoms, thereby enhancing the prospects of recovery and survivability. The findings suggest that integrating such proactive risk assessment measures holds promise for optimizing medical interventions and improving patient outcomes in the context of high-risk conditions.

ACKNOWLEDGEMENT

This work was supported by UMPSA under Distinguished Research Grant RDU223003.

REFERENCES

[1] World Health Organization, **Noncommunicable Diseases**, https://www.who.int/health-topics/noncommunicable-diseases, 2023.

[2] R. Sigit, R. Rokhana, Setiawardhana, T. Hidayat, Anwar, and J. Jaenputra**, Implementation of Portable Ultrasound for Heart Disease Detection Using Cloud Computing-Based Machine Learning**, *EMITTER International Journal of Engineering Technology.*, Vol. 12, No. 2, pp. 196–212, 2024.

[3] R. Waigi, S. Choudhary, P. Fulzele, and G. Mishra, **Predicting the risk of heart disease using advanced machine learning approach**, *European Journal of Molecular and Clinical Medicine*, Vol. 7, pp. 1638–1645, 2020.

[4] J. Jagriti, N. Sharma, and S. Aggarwal, **Impact of Principal Component Analysis on the Performance of Machine Learning Models for the Prediction of Length of Stay of Patients**, *EMITTER International Journal of Engineering Technology*, Vol. 12, No. 2, pp. 128–149, 2024.

[5] P. Natarajan, J. Frenzel, and D. Smaltz, **Demystifying Big Data and Machine Learning for Healthcare**, *CRC Press*, Ed. 1, 2021.

[6] D. Shah, S. Patel, and S. K. Bharti, **Heart Disease Prediction using Machine Learning Techniques**, *SN Comput. Sci.*, Vol. 1, No. 6, 2020.

[7] S. Mohan, C. Thirumalai, and G. Srivastava, **Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques**, *IEEE Access*, Vol. 7, pp. 81542–81554, 2019.

[8] S. Sengupta, R. Das, and S. Chakrabarti, **A Deep Dive into a Groundbreaking Approach to Machine Learning-Powered E-Learning**, *EMITTER International Journal of Engineering Technology*, Vol. 12, No. 2, pp. 213–236, 2024.

[9] P. Balasundaram, P. Ganesh, K. P, and R. K. Mukesh, **A Novel Technology Stack for Automated Road Quality Assessment Framework using Deep Learning Techniques**, *EMITTER International Journal of Engineering Technology*, Vol. 12, No. 1, pp. 62–89, 2024.

[10] Y. Li, J. Zhao, and Z. Lv, J. Li, **Medical image fusion method by deep learning**, *International Journal of Cognitive Computing in Engineering*, Vol. 2, pp. 21–29, 2021.

[11] T. Zhou, Q. R. Cheng, H. L. Lu, Q. Li, X. X. Zhang, and S. Qiu, **Deep learning methods for medical image fusion: A review**, *Comput. Biol. Med.*, Vol. 160, pp. 106959, 2023.
[12] A. Ghaheri, S. Shoar, M. Naderan, and S. S. Hoseini, **The Applications of Genetic Algorithms in Medicine**, *Oman Med. J.*, Vol. 30, No. 6, pp. 406–416, 2015.
[13] A. E. Eiben, and J. E. Smith, **Introduction to Evolutionary Computing**, *Springer*, 2024.
[14] Y. Clapper, J. Berkhout, R. Bekker, and D. Moeke, **A model-based evolutionary algorithm for home health care scheduling**, *Comput. Oper. Res.*, Vol. 150, pp. 106081, 2023.
[15] A. Slowik, and H. Kwasnicka, **Evolutionary algorithms and their applications to engineering problems**, *Neural Comput. Appl.*, 2024.
[16] A. Suresh, R. Kumar, R. Varatharajan, B. A. Suresh, R. Kumar, and R. Varatharajan, **Health care data analysis using evolutionary algorithm**, *J. Supercomput*, Vol. 76, pp. 4262–4271, 2020.
[17] D. Yu, Z. Zhao, and D. Simmons, **Interaction between Mean Arterial Pressure and HbA1c in Prediction of Cardiovascular Disease Hospitalisation: A Population-Based Case-Control Study**, *J. Diabetes Res*, pp. 1–7, 2016.
[18] M. Ordikhani, M. S. Abadeh, C. Prugger, R. Hassannejad, N. Mohammadifard, and N. Sarrafzadegan, **An evolutionary machine learning algorithm for cardiovascular disease risk prediction**, *PLoS One*, Vol. 17, No. 7, 2022.
[19] M. H. Sulaiman, Z. Mustaffa, F. Zakaria, and M. Saari, **Using the evolutionary mating algorithm for optimizing deep learning parameters for battery state of charge estimation of electric vehicle**, *Energy*, Vol. 279, pp. 128094, 2023.
[20] S. Ulianova, **Cardiovascular disease dataset**, https://www.kaggle.com/datasets/sulianova/cardiovascular-disease-dataset, 2019.
[21] National Center for Chronic Disease Prevention and Health Promotion, **Heart Disease and Stroke**, https://www.cdc.gov/chronicdisease/resources/publications/factsheets/heart-disease-stroke.htm, 2022.
[22] National Institute on Aging, **Heart Health and Aging**, www.nia.nih.gov/health/heart-health/heart-health-and-aging, 2018.
[23] C. Vidal, P. Malysz, M. Naguib, A. Emadi, and P. Kollmeyer, **Estimating battery state of charge using recurrent and non-recurrent neural networks**, *J. Energy Storage*, Vol. 47, pp. 103660, 2022.
[24] S. A. Ali Shah, I. Uddin, F. Aziz, S. Ahmad, M. A. Al-Khasawneh, and M. Sharaf, **An Enhanced Deep Neural Network for Predicting Workplace Absenteeism**, *Complexity*, 2020.
[25] Vidal C, Kollmeyer P, Naguib M, Malysz P, Gross O, Emadi A., **Robust xEV Battery State-of-Charge Estimator Design Using a Feedforward Deep Neural Network**, *SAE Int. J. Adv. Curr. Pract. Mobil*., Vol. 2, No. 5, pp. 2872–2880, 2020.