ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240319120727

Original scientific paper
Received: 2024-03-19, Accepted: 2024-04-15

A Deep Learning Approaches for Enhancing Clinical Solutions to Cardiovascular Prediction Using EHR

Mounika Valasapalli*, Nallagatla Raghavendra Sai

Abstract: The prediction of cardiovascular disease gained immense significance in the medical field with the alignment of increasing focus on promoting healthier lifestyle. Current methods for cardiovascular disease prediction is leading to so many miss classifications, urging the need of modern automated Deep learning approaches. The main purpose of these approaches is to detect the occurrence of cardiovascular disease (CVD) using patient information from comprehensive electronic health records (HER). Moreover, it is a complex task to choose appropriate features from Electronic Health Records data, and it is a huge confronts to attain robust and accurate results because of the incomplete data entry errors, incomplete record of the patient and patient self-reporting and data integration issues. In this paper we propose an efficient end-to-end framework known as Risk prediction with Deep Residual Neural Network (DRNN), which not only acquires the most influencing features; but also considers the time-based medical data and temporal data to help the patient disease progression, treatment effectiveness, and to check how other diseases are affecting the state of patient. The experimentation is done with the online available Kaggle dataset for cardiovascular disease (CVD) prediction. The result of DRNN demonstrate that the anticipated model significantly enhances the prediction accuracy and F-Measure, Sensitivity compared to various existing approaches. The anticipated model establishes superior trade-off among other approaches.

Keywords: cardiovascular disease; deep residual neural network; high-risk prediction; medical data; prediction accuracy

1 INTRODUCTION

One of the important reasons for death is cardiovascular diseases such as CVD in today's world [1]. Because of cardiovascular diseases, there are 17.9 million people affected out of 57.5 million people, which has been proven as a reported death in 2015 across the world. In addition, the non-negligible burden in economic which is brought by patients with cardiovascular diseases and create critical disability throughout their lifetime. However, this is estimated that around 90 % of cardiovascular diseases are prevented using relevant precautions [2]. The onset of cardiovascular diseases needs to be predicted for the person in the medical domain. There are few acknowledged metrics of pathology for detecting biomarkers for cardiovascular diseases, like angiography and electrocardiogram (ECG). The Authority method used for diagnosis is called angiography in the domain of medicine for cardiovascular diseases, and this obtains greater accuracy in the prediction and diagnosis of cardiovascular diseases. Moreover, the invasive and expensive method is angiography. Another general method for predicting and diagnosing cardiovascular diseases is called the ECG. This accuracy is highly dependent on the knowledge and experience of the experts or the medical staff in the medical domain. Then the guaranteed and considerable area of study is called computer-assisted higher prediction of risk in cardiovascular diseases. Added to the point, the conventional method for higher risk prediction depends on machine learning that focuses on achieving the automatic computer system that needs to be turned with crucial and potential features extracted from the patients' historical EHR, such as electronic health records. It is a manual vulnerable, lower-cost, and non-invasive method compared to conventional pathological measures [3].

EHR drives the higher risk prediction task as the critical difficulty in accurately achieving the patient's portrait, called learning the representation of the patient or the feature engineering. In addition, the EHR consists of different information about the patient, which helps to represent the series of timely ordered visits to hospitals. Everyone has a lot of medical variables, like the diagnosis of demographic procedures and medications of vital science and outcomes from laboratory tests. The number of distinct medical variables in EHR systems is generally very high. The previous predictive models help manage it by having a sparse representation of the features using different dimension reduction approaches. Conventional human intervention measures of feature engineering obtained poor generalization and scalability due to the high dependents on the particular EHR system and the authors' experiences. More scalable and uncomplicated methodologies are used with the help of automatic feature representation that is proposed, like as Bag-of-Words as BoW and One-Hot [4]. Moreover, every feature is considered the independent and discrete word in these methods that create the features that cannot obtain the hidden semantic information accurately among the temporal dynamics and features in the data of EHR. Hence, the process of designing an efficient method for handling the representation of the feature for high-dimension heterogeneous EHR data and sequential is the essential problem.

The important contributions are presented below as the summary session in the proposed system:

- DRNN is a robust and end-to-end model that is suggested in the proposed system for predicting the answer to the greater risk of cardiovascular diseases in accurately having diabetes, hypertension, or hyperlipidemia history for patients who are not having any medical experts' assistance.
- Different clinical data is integrated as a whole, having a close relationship in the proposed system. There is no full utilization of the difference between the heterogeneous data of EHR by the DRNN and the relationship between them.
- DRNN has the performance to be evaluated on the real medical data set in the proposed system. It makes the

DeepRisk obtain better performance on the higher dimensional, heterogeneous data and timely ordered EHR data having the integration of model and reduction in dimensionality.

The work is drafted as: section 2 gives a wider analysis on various prevailing researches on CVD prediction; section 3 elaborates methodology with DRNN. Section 4 describes the numerical outcomes. Section 5 projects the research summary with future research directions.

2 RELATED WORKS

There is difficulty even now in obtaining greater performance using the combination of temporal and heterogeneous clinical data, which is more comprehensive even though many authors have obtained the accurate representation of learning for the higher dimension of data with conventional neural networks (CNN) and recurrent neural network (RNN). Relationships and the differences among the various kinds of medical data are eradicated in the feature learning in the previous research [5]. Also, the temporal data processing of EHR is done with the models based on the CNN [6] that are often restricted due to the need for medical codes during the visits to the hospital or the timely ordered episode. The theoretical attention related to the model of RNN called DeepRisk is suggested in the proposed system for the prediction of the onset of the higher risk cardiovascular diseases and also to deal with the earlier mentioned difficulties for the patient to have hypertension, diabetes, or the history of hyperlipidemia. DeepRisk can excavate the essential information automatically between the raw data and also represent features accurately for the patient having very less dimensions when the patients are turned with the heterogeneous temporal and higher dimensional raw data of EHR.

Long short-term memory (LSTM) and attractive technique in the predictive model is used in the proposed system to utilize the sequential data of EHR more comprehensively [7]. An independent module is used in the proposed system to obtain the relevant characteristics to consider the differences among various types of medical data [8]. On the other hand, the combination of these various data is given to another module in the proposed system to achieve the data regarding the relationship that is potentially used [9]. At last, the modules give all the features learned to integrate for obtaining the greater risk prediction task in the proposal system. The outcomes from the experiments depend on the real data set of the clinic of the cardiovascular diseases that establishes the DeepRisk, which is more robust and accurate when compared with the modern methodologies used in the prediction task based on the HER [10].

3 METHODOLOGY

3.1 Proposed Model

The origination of attention technique originated from the perception of human vision. The global images are scanned first when the human perceives an object and then concentrates on the particular region to obtain more detailed data and also tries to suppress other unwanted data. Google machine translation research team proposes the self-attention technique that helps to obtain extensive attraction having the upcoming study of the attention technique due to the learning process of relationship among the other positions and the particular position to obtain the dependence of context. The self-attention network is used to broad the task. The visual attention technique is utilized to enhance the accuracy. Self-attention technique is first suggested in the proposal system for predicting seizure research to attend to the information of the HER globally. These processes will be elaborated in the upcoming session.

Fig. 2 depicts the process of feeding to the conventional layer for producing two feature maps, Y and Z, provided a local feature $X \in R^{C*H*W}$. Here, Y and Z are the dimensions for both R^{C*H*W}. In addition, the process of reshaping is done for R^{C*N}. Here, $N = H \times W$. Thus, the matrix multiplication is performed among the transportation of Y and Z in the proposed system and then achieves the weight matrix having the dimension with the help of the softmax layer.

$$S_{ji} = \frac{\mathrm{e}^{Y_i \cdot Z_j}}{\sum_{i=1}^{N} \mathrm{e}^{Y_i \cdot Z_j}} \tag{1}$$

Here, j^{th} position influences on i^{th} position is presented by S_{ji}. More same features give representations of the channels higher relevance among them. In addition, a new representation of feature $T \in R^{C*H*W}$ is created using the process of feeding X and is reshaped to R^{C*N} and the convolution layers simultaneously. However, the matrix multiplication of T and S is performed in the proposed system, and outcome is reshaped to R^{C*H*W}. At last, the scaling parameter α is used to multiply the outcome. The element-wise sum operation having X is executed to obtain the last outcome as $E \in R^{C*H*W}$ is presented in Eq. (2). Here, 0 is assigned to α, which is allocated little by little more weight. Hence, E is the last feature, the original feature, and the sum of weighted feature which includes the perspective related to context and gathering the information globally, which depends on the attention map optionally.

$$E_j = \alpha \sum_{i=1}^{N} (S_{ji} * T_i) + X_j \tag{2}$$

The various features of the channel represent various semantics of EHR signals. The interdependence among the mapping of the channel is mined by using the channel attention technique, and various representations of semantics are relevant to one another. In the first step, the process of reshaping X to R^{C*N} is done using the transpose of X and the matrix multiplication among X. At last, the soft wax layer is used in the proposed system for achieving the map of channel attention $P \in RC \times C$ presented in Eq. (3). Here, i^{th} channel influences on the j^{th} channel is measured by P_{ji}.

$$P_{ji} = \frac{\mathrm{e}^{X_i \cdot X_j}}{\sum_{i=1}^{C} \mathrm{e}^{X_i \cdot X_j}} \tag{3}$$

The matrix multiplication is performed on X and P in the proposed system, and result is reshaped to R^{C*H*W}. The scaling parameter β is used to multiply the result in the proposed system. The element-wise sum operation having X is executed to obtain the last output as $E \in R^{C*H*W}$ is presented in Eq. (4). Hence, 0 is initialized to β. The integration of two attention models using the element-wise sum in the proposed system to completely utilize the spectrum and channel information of context. After fusing original features, last feature map needs to be achieved using the average pooling in the proposed system.

$$E_j = \beta \sum_{i=1}^{C} (P_{ji} * X_i) + X_j \tag{4}$$

Here, k-fold cross-validation (CV) is chosen for every patient in the proposed system to obtain outcomes that are the same as the real conditions. The complete interictal recordings are classified into n parts. Every part includes the t/n hours gathered randomly, having any patient's recordings when the patient has t hours and n recordings. In addition, feature pairs are provided for testing, this is performed for n times. On the other hand, balance $n - 1$ pair or utilized for the training stage. A few researchers commonly randomly classified 20 % testing data and 80 % training data. These sets are utilized as the validation to focus on overfitting. Moreover, the anticipated methodology is relevant for HER data independency. In addition, HER data is time-dependent and samples are chosen in the proposed system at various periods to observe whether the model started to overfit during the training period. The proposed system chooses 25 % samples from the recordings of patients as the set of validation to monitor in the training set, and the balance 75 % samples are utilized as training set. The problem of overfitting is still present to the number of alterations that upsurges the training accuracy in training process and the early stop method is used to solve this issue. The training has to be stopped and storing the network parameters at the low loss of validation is done immediately during the detection of the loss on the set of validation which is started to maximize.

Tab. 2 depicts the parameter of DRNN model. The model has the input, which is 1 × 22 × 9 × 114. Here, the number of data is presented as 22, and the dimension of the features is represented as 9 × 114. The batch normalization, ReLu activation function, and dropout are used to follow every conventional layer. The earlier feature map helps feed the matrix to obtain 64 × 7 × 28, which leads to the convolutional layer, followed by the reshaping operation. There are 4 ResBlock layers used in the proposed system for extracting the deep features of the EHR subsequently as in Fig 1. The global features are fused using the dual self-attention, fully connected layers with activation function. Loss function is chosen in the proposed system as the function of cost. However, 32 are considered the size of the batch. Accordingly, 0.0005 and 0.5 are fixed as the rate of learning and rate of dropout. The proposed new model has been accomplished with MATLAB 2020a.

Table 1 Network architecture of the Proposed Model DRNN

Layers	Output size	Structural description
Input	1 * 22 * 9	--
Conv_1	64 * 1 * 7 * 55	22 * 3 * 5 conv Stride 1 * 1 * 2
Pooling	64 * 1 * 7 * 28	1 * 1 * 2 max pooling
RB1	64 * 7 * 28	$\begin{pmatrix} 3*3 & 64 \\ 3*3 & 64 \end{pmatrix} * 2$
RB2	128 * 4 * 14	$\begin{pmatrix} 3*3 & 128 \\ 3*3 & 128 \end{pmatrix} * 2$
RB3	256 * 2 * 7	$\begin{pmatrix} 3*3 & 256 \\ 3*3 & 256 \end{pmatrix} * 2$
RB4	512 * 1 * 4	$\begin{pmatrix} 3*3 & 512 \\ 3*3 & 512 \end{pmatrix} * 2$
Attention model	512 * 1 * 4	Attention
Avg. pooling	512 * 1	--
Classification	2 * 1	Softmax

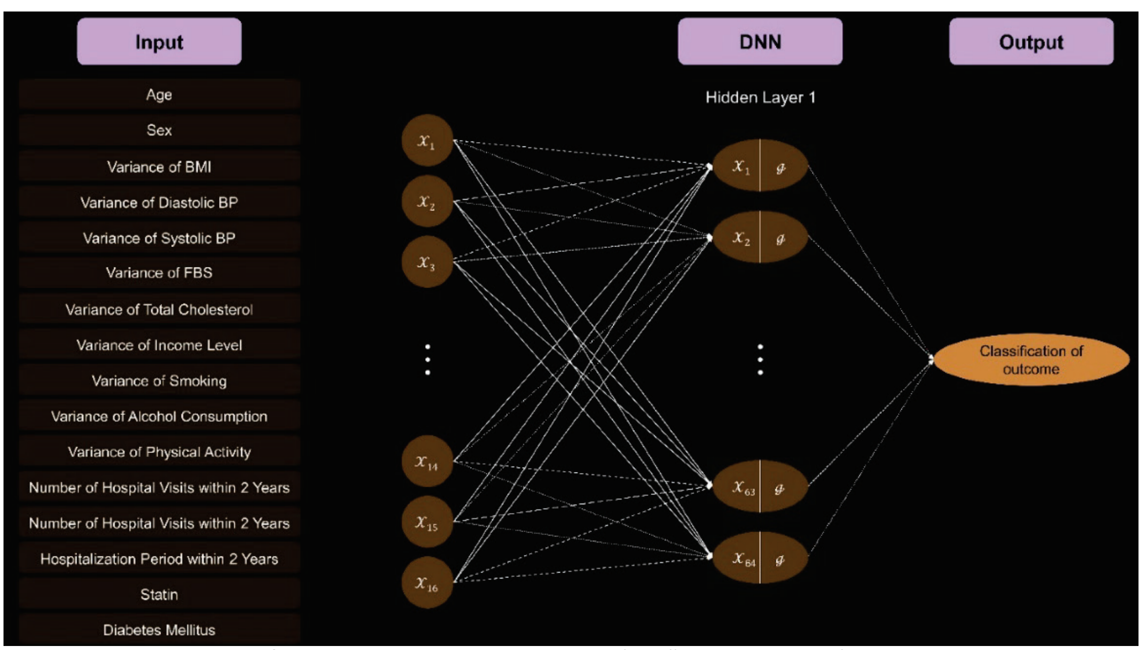

Figure 1 Showing the proposed DRNN Model with the five different compositions of the Datasets

3.2 Model Description

Residual network (ResNet) and dual self-attention techniques are presented in the proposed model. Fig. 2 depicts attraction of the internal module; the EHR data is used in the proposed system as the input to network and heart disease features are hauled out via ResNet potentially. Thus, the input is given to the attention module using the below 3 steps for creating the new global features. An attention matrix is created in the first step to present the special relationship among any features. In the second step, the multiplication matrix on the original features and the attention matrix is executed in the proposed system. In the third step, the final global feature is performed in the proposed system with the help of execution of element-wise sum among original feature and matrix outcome. The channel attention process is the same as the attention model. At last, features (ResNet and attention) are merged and combined with original features to acquire EHR's better characteristics.

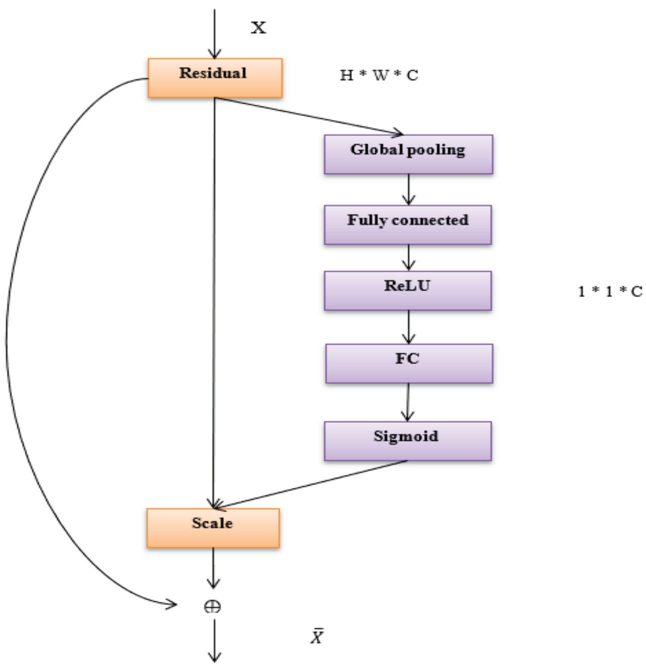

Figure 2 Showing the interaction of the extraction model for the ResNet

CNN is used broadly in natural language processing, computer vision, and so on [30]. The Residual function is used to alleviate the vanishing gradient problem and used to training of very deep neural networks. Features are processed through residual connections, global pooling, fully connected layers with ReLU activations, and finally a sigmoid activation function for binary prediction. It is efficient for depending on total network layers for obtaining the ability of the expression and fitting the mapping relationship better and potentially. There is an issue of the disappearance of radiance, and gradient descent optimization becomes very challenging as the total layers depend on the network model. In recent years, ResNet has been used for combating the mentioned issue when the training has happened for the very deep conventional network consisting of different restaurant blocks and was proposed. Fig 1 presents the addition of identity mapping by the residual block with the network via connections of shortcut that is not based on the calculation and not adding the extra parameters to solve the model degradation. Every residual block consists of 3 × 3 convolution layer, batch normalization and ReLU. The model includes identity map as x and the residual path as $F(x)$. This research has four residual blocks to use that are connected to the proposed system.

The sample data is castoff to calculate the limits of the model. The probability equation is nonlinear for this distribution, therefore analytic solutions are not available. As a result, for getting parameter estimations, we use an iterative approach such as the ResNet Model.

For the ResNet algorithm, the limitations of the given model revised equations are evaluated. The probability of the present model, is

$$Q(\theta,\,\theta^{(0)}) = E_{\theta^{(0)}}\left[\frac{\log L(\theta)}{\bar{x}}\right] \tag{5}$$

This implies

$$\log L(\theta) = \sum\nolimits_{s=1}^{N} \log\left(\sum\nolimits_{i=1}^{k} \alpha_i^{(l)} f_i(x_s, \theta^{(l)})\right) \tag{6}$$

The conditional probability region k is

$$P_k(x_s,\,\theta^{(l)}) = \left[\frac{\alpha_k^{(l)} f_k(x_s,\,\theta^{(l)})}{p_i(x_s,\,\theta^{(l)})}\right] \tag{7}$$

$$p_k(x_s,\,\theta^{(l)}) = \left[\frac{\alpha_k^{(l)} f_k(x_s,\,\theta^{(l)})}{\sum\nolimits_{i=1}^{k} \alpha_i^{(l)} f_i(x_s,\,\theta^{(l)})}\right] \tag{8}$$

Therefore, the ResNet distribution:

$$f_i(x_s,\,\theta^{(l)}) = \frac{\left[\dfrac{3}{(3p+\pi^2)}\right]\cdot\left[p+\left(\dfrac{x_s-\mu_i^{(l)}}{\sigma^{(l)}}\right)^2\right]\cdot \mathrm{e}^{-\left(\frac{x_s-\mu_i^{(l)}}{\sigma^{(l)}}\right)}}{\sigma_i^{(l)}\left[1+\mathrm{e}^{-\left(\frac{x_s-\mu_i^{(l)}}{\sigma^{(l)}}\right)}\right]^2} \tag{9}$$

The likelihood function of example comments must be estimated as the first step of ResNet Algorithm.

E-step: The following expectation (E) step, $\log L(\theta)$ is the expectation value, with respect to the preliminary parameter vector $\theta^{(0)}$ is

$$Q(\theta,\,\theta^{(0)}) = E_{\theta^{(0)}}\left[\frac{\log L(\theta)}{\bar{x}}\right] \tag{10}$$

Given the preliminary parameter $\theta^{(0)}$.

M-step: Parameters estimation,

$$F=\left[E\left(\log L(\theta^{(l)})\right)+\beta\left(1-\sum_{i=1}^{k}\alpha_i^{(l)}\right)\right] \tag{11}$$

The Updated equations of α_i: The appearance for α_i, we solve the following calculation

$$\frac{\partial F}{\partial \alpha_i}=0;\ \sum_{i=1}^{N}\frac{1}{\alpha_i}P_i(x_s,\ \theta^{(l)})+\beta=0;\ \beta=-N \tag{12}$$

Consequently,

$$\alpha_i=\frac{1}{N}\sum_{s=1}^{k}P_i(x_s,\ \theta^{(l)}) \tag{13}$$

The updated equations of α_i for $(l+1)^{\text{th}}$ iteration is

$$\alpha_i^{(l+1)}=\frac{1}{N}\sum_{s=1}^{k}P_i(x_s,\ \theta^{(l)}) \tag{14}$$

This implies

$$\alpha_l^{(l+1)}=\frac{1}{N}\sum_{s=1}^{N}\left[\frac{\alpha_l^{(l)}f_l(x_s,\ \theta^{(l)})}{\sum_{i=1}^{k}\alpha_i^{(l)}f_i(x_s,\ \theta^{(l)})}\right] \tag{15}$$

The Rationalized equations of μ_i: For limitation logistic type dispersal - By put on the derived with respect to μ_i, we have

$$\frac{\partial}{\partial\beta_i}\left[\sum_{s=1}^{N}\sum_{i=1}^{k}P_i(x_s,\theta^{(l)})\log\alpha_i\frac{\left[\frac{3}{(12+\pi^2)}\right]\cdot\left[4+\left(\frac{x_s-\beta_i}{\sigma_i}\right)^2\right]\cdot \mathrm{e}^{-\left(\frac{x_s-\beta_i}{\sigma_i}\right)}}{\sigma_i\left[1+\mathrm{e}^{-\left(\frac{x_s-\beta_i}{\sigma_i}\right)}\right]^2}\right]=0 \tag{16}$$

$$\frac{\partial}{\partial\beta_i}\left[\sum_{s=1}^{N}\sum_{i=1}^{k}P_i(z_s,\theta^{(l)})\log\alpha_i\frac{\left[\frac{3}{(12+\pi^2)}\right]\cdot\left[4+\left(\frac{z_s-\beta_i}{\sigma_i}\right)^2\right]\cdot \mathrm{e}^{-\left(\frac{z_s-\beta_i}{\sigma_i}\right)}}{\sigma_i\left[1+\mathrm{e}^{-\left(\frac{xz-\beta_i}{\sigma_i}\right)}\right]^2}\right]=0 \tag{17}$$

This implies

$$\sum_{s=1}^{N}P_i(z_s,\theta^{(l)})\left[\left[\frac{2\left(\frac{z_s-\beta_i}{\sigma_i}\right)\left(-\frac{1}{\sigma_i}\right)}{\left[4+\left(\frac{z_s-\beta_i}{\sigma_i}\right)^2\right]}\right]+\left[\frac{1}{\sigma_i}\right]-\left[\frac{\mathrm{e}^{-\left(\frac{z_s-\beta_i}{\sigma_i}\right)^2}}{\sigma_i\left(1+\mathrm{e}^{-\left(\frac{z_s-\beta_i}{\sigma_i}\right)^2}\right)}\right]\right]=0 \tag{18}$$

$$\mu_i^{(l+1)}=\frac{\sum_{s=1}^{n}\frac{P_i(z_s,\ \theta^{(l)})(2y_s)}{(\sigma_i^{(l)})^2\left[p+\left(\frac{z_s-\beta_i^{(l)}}{\sigma_i^{(l)}}\right)^2\right]}\sum_{s=1}^{n}\frac{P_i(z_s,\ \theta^{(l)})}{\sigma_i^{(l)}}+\sum_{s=1}^{n}\frac{2P_i(z_s,\ \theta^{(l)})}{\sigma_i^{(l)}\left[1+\mathrm{e}^{\left(\frac{z_s-\beta_i^{(l)}}{\sigma_i^{(l)}}\right)}\right]}}{2\sum_{s=1}^{n}\frac{P_i(z_s,\ \theta^{(l)})}{(\sigma_i^{(l)})^2\left[p+\left(\frac{z_s-\beta_i^{(l)}}{\sigma_i^{(l)}}\right)^2\right]}} \tag{19}$$

3.3 Dataset Description

Therefore, the use of a repository gives the data, which are all available online. Data pre-processing is performed, and then different records of patients are gathered. There are 303 records of the patients presented in that data. On the other hand, the missing values are presented from the six records. Missing values for eliminated and remaining data are maintained, such as 297 records. In addition, the introduction of the multi-class variables and the binary classification are done for the dataset attributes. Some variables are employed to measure disease occurrence or not. Values are fixed as 'zero' or 'one', and no symptoms of the disease are represented by 0. Hence, the occurrence of the disease is shown as 137, and the non-occurrence of the disease is shown as 161 accordingly. Dataset description is provided in Tab. 2.

Table 2 Dataset attributes in Kaggle Dataset

Attributes	Descriptions	Type
Age	CVD patients age	N
Sex	CVD Patients' gender	N
CP	CVD type 1. Typical/atypical angina, 3. Non-anginal, 4. Asymptomatic	Nm
Tresbps	BP state	Nm
Chol	Serum cholesterol	N
FBS	Blood sugar level	N
Resting	ECG outcomes	Nm
Thali	Max. heart rate	N
Exchange	Angina owing to exercise	Nm
OldPeak	Exercise owing to depression	N
Slope	peak exercise segment	Nm
Ca	Fluoroscopy coloured vessels	N
Thal	CVD status	N
Num	CVD diagnosis value	N

N - Nominal; Nm – Numerical

4 RESULTS AND DISCUSSIONS

The proposed model needs to compute the performance using statistical analysis like precision, sensitivity, F-measure, recall, Matthews' correlation coefficient (MCC), true negative rate (TNR), false positive rate (FPR) and false negative rate (FNR). Existing techniques have the evaluation such as the Bayesian model (BM), regression model (RM), feed-forward neural network (FFNM), linear support vector machine (L-SVM), random forest (RF), decision tree (C4.5), deep auto-encoder (DAE) and ensemble classifier techniques are used to obtain the significance of the model. Eq. (5) to Eq. (12) expresses the mathematical representation of these measures.

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \quad (20)$$

$$Precision = \frac{TP}{TP+FP} \quad (21)$$

$$Recall / Sensitivity / TPR = \frac{TP}{TP+FN} \quad (22)$$

$$F\text{-}measure = \frac{2*Precision}{p+r} \quad (23)$$

$$MCC = \frac{(TP*TN)-(FP*FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (24)$$

$$FPR = \frac{FP}{FP+TN} \quad (25)$$

$$FNR = \frac{FN}{FN+TP} \quad (26)$$

$$TNR = \frac{TN}{TN+FP} \quad (27)$$

Deep Auto-encoder (DAE), Ensemble, and DRNN have achieved notably higher performance across various metrics compared to other models (Tab. 3).

The comparison of the anticipated in the simple model is presented in Tab. 3, having the prevailing techniques. In addition, a few statistical metrics include precision, accuracy, F-measure, sensitivity, FPR, MCC, FNR, and TNR. These measures have the results across the anticipated methodology: precision at 98.8%, accuracy at 98%, sensitivity at 97.5%, F-measure at 97.5%, MCC at 95%, FPR at 4.54%, FNR at 3.15% and TNR as 98% accordingly. Then, the existing model shows an accuracy of 8%, 13%, 10%, 26%, 22%, 14%, 1% and 0.8%, which are superior to the anticipated model. The existing model shows precision of 8.8%, 10.8%, 13.8%, 22.8%, 22.8%, 11.8%, 0.8% and 0.6%, which are superior to the prevailing model. The existing model shows a sensitivity of 12.5%, 11.5%, 16.5%, 44.5%, 25.5%, 22.5%, 1.5% and 1.3%, which are superior to the prevailing model. The existing model shows an F-measure of 12.5%, 12.5%, 11.5%, 35.5%, 24.5%, 16.5%, 1.5% and 1.3%, which are better than the proposed model. The existing model shows an MCC of 20%, 20%, 25%, 50%, 43%, 25%, 1% and 0.8%, which are better than the proposed model. The existing model shows FPR of 10.46%, 9.46%, 8.46%, 12.46%, 0.19%, 9.46%, 1.46% and 1% higher than the proposed model. The existing model shows an FNR of 16.85%, 18.85%, 16.85%, 48.85%, 27.85%, 21.85%, 1.2% and 1%, which are higher than the proposed model. The existing model shows TNR of 8%, 8%, 6%, 12%, 20%, 8%, 1% and 0.7%, which are higher than the anticipated model. The complete performance is considerably compared with other approaches using the comparative analysis, and the anticipated ensemble model is proved that the model provides better performance for prediction and the complexity is reduced in the learning techniques.

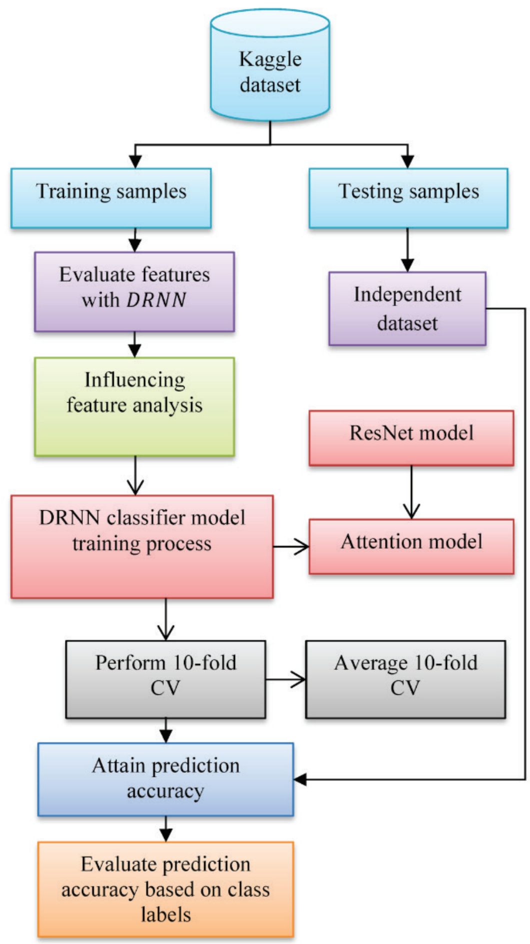

Figure 3 Flow chart of the Proposed Methodology for the different composition of the datasets

Table 3 Overall comparisons of the results with proposed model

Approaches	Accuracy	Precision	Sensitivity	F-measure	MCC	FPR	FNR	TNR
Bayesian model (BM)	90%	90%	85%	85%	75%	15%	20%	90%
Regression model (RM)	85%	88%	86%	85%	75%	14%	22%	90%
Feed Forward network model (FFNM)	88%	85%	81%	86%	70%	13%	20%	92%
Linear Support Vector Machine (L-SVM)	72%	76%	53%	62%	45%	17%	52%	86%
Decision tree (C4.5)	76%	76%	72%	73%	52%	4.35%	31%	78%
Random Forest (RF)	84%	87%	75%	81%	70%	14%	25%	90%
Deep Auto-encoder (DAE)	97%	98%	96%	96%	94%	6%	4.35%	97%
Ensemble	97.2%	98.2%	96.2%	97%	94.2%	5.54%	4.15%	97.3%
DRNN	98%	98.8%	97.5%	97.5%	95%	4.54%	3.15%	98%

The comparison of baseline classified measures is presented in Tab. 3 without a CV and with a 10-fold CV. In addition, the anticipated model's average precision, accuracy, sensitivity and F-measure are 98.5%, 99.5%, 99% and 99%, which is substantially higher than other approaches. The average accuracy of the existing approaches is 6.5%, 11.5%, 8.5%, 24.5%, 20.5%, 12.5%, 1.5%, and 1% higher than BM, RM, FFNM, L-SVM, C4.5, RF, DAE and Ensemble classifier model. The average precision of the existing approaches is 7.5%, 9.5%, 12.5%, 21.5%, 21.5%, 10.5%, 1% and 0.7% higher than other approaches like BM, RM, FFNM, L-SVM, C4.5, RF, DAE and Ensemble classifier model. The average sensitivity of the existing approaches is 12%, 12%, 16%, 44%, 25%, 22%, 2% and 1.5% higher than BM, RM, FFNM, L-SVM, C4.5, RF, DAE and Ensemble classifier model. The average F-measure of the existing approaches are 12%, 12%, 11%, 35%, 24%, 16%, 2% and 1.5% higher than BM, RM, FFNM, L-SVM, C4.5, RF, DAE and Ensemble classifier model (See Fig. 4 and Fig. 5). The re-substitution loss depends on the learning cycle presented in Fig. 7. The consideration of k - fold CV is done in machine learning as in Fig 6. Here, 10 is fixed to the value of k. Hence, the 10 recursions are needed to perform and consider the average accuracy in the prediction state. The average prediction rate is 98.5%, which is presented in Tab. 2 accordingly.

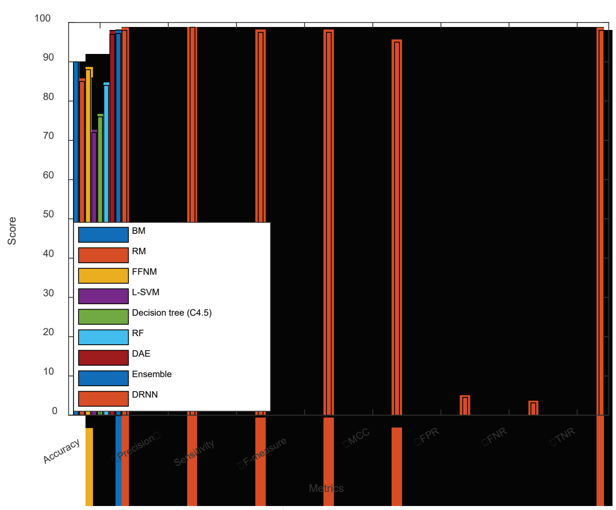

Figure 4 Comparison of various prevailing approaches

Table 4 Average value after 10-fold CV

Model	Accuracy	Precision	Sensitivity	F-measure
Bayesian model (BM)	92%	92%	87%	87%
Regression model (RM)	87%	90%	87%	87%
Feed Forward network model (FFNM)	90%	87%	83%	88%
Linear Support Vector Machine (L-SVM)	74%	78%	55%	64%
Decision tree (C4.5)	78%	78%	74%	75%
Random Forest (RF)	86%	89%	77%	83%
Deep Auto-encoder (DAE)	97%	98.5%	97%	97%
Ensemble	97.5%	98.8%	97.5%	97.5%
DRNN	98.5%	99.5%	99%	99%

The complete performance measure of the classifier is presented in Tab. 4. The baseline DRNN classifier has the complete performance that needs to be determined. The average precision of the anticipated model is 99.5%, the average sensitivity is 99%, and the average F-measure is 99% which is comparatively superior to other techniques. The p-value analysis ranges from 0.001. This analysis helps to prove that the answerable technique provides better accuracy in prediction using the under-fitting and over-fitting problems. The problems are eliminated one after the other and model across DRNN classifier gives better functionality.

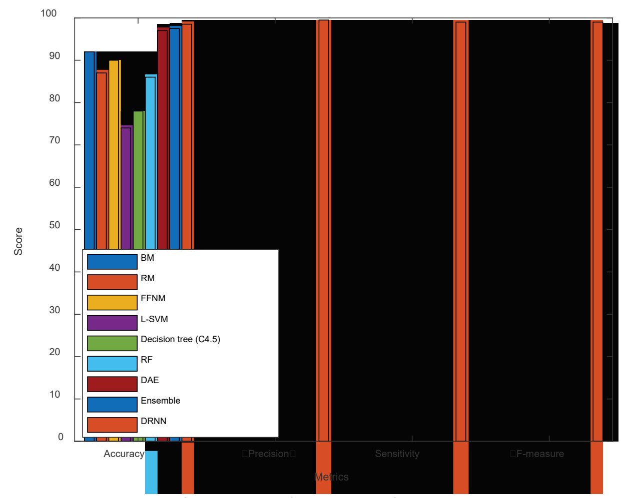

Figure 5 Comparison of average performance metrics

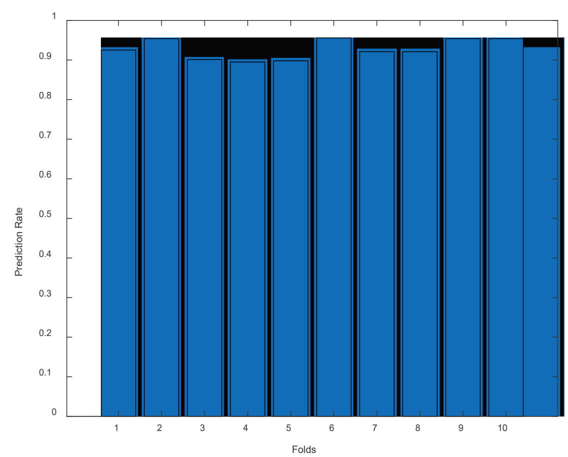

Figure 6 *k*-fold cross validation (k = 10)

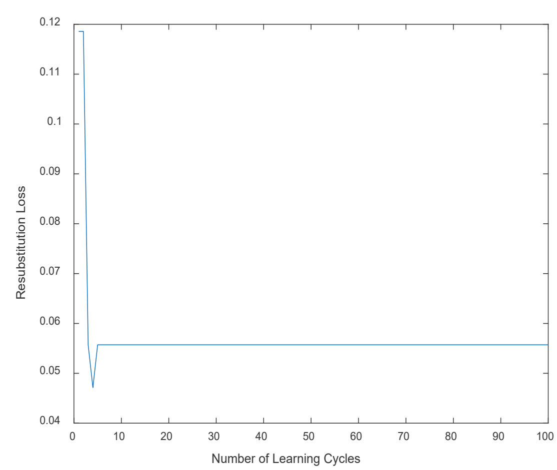

Figure 7 Learning vs. loss analysis

5 CONCLUSION

The process of predicting CVD is a novel approach presented in the studies. The proposed model uses the deep network model for identifying global feature set and is utilized for identifying the optimal configuration. In addition, three main problems are resolved with the help of predicting heart disease using the proposed DRNN. The optimal network configurations need to be identified, and the important issues are underfitting and overfitting. The anticipated model is evaluated with traditional DNN using optimization approaches. However, other network configurations are analyzed to compare the optimized network configuration outcomes. The proposed model performed well than other approaches, and the prevailing techniques are presented to the machine learning using the comparative analysis. The features selection technique improves the performance and obtains good performance over other approaches compared with the different feature selection methods. In addition, the proposed DRNN improves the prediction of heart diseases with prediction accuracy of 94% from comparative and experimental outcomes in the proposed system. Also, the clinical experts helped to generate effective decisions. It is essential for future studies that the complexity of time in the proposed cabinet approach is required to be investigated as the most important feature in the healthcare field. Also, the proposed approach is suggested to test some data sets to evaluate the technique's effectiveness with more instances.

6 REFERENCES

[1] Bharti, R., Khamparia, A., Shabaz, M., Dhiman, G., Pande, S. & Singh, P. (2021). Prediction of heart disease using a combination of machine learning and deep learning. *Comput Intell Neurosci*. https://doi.org/10.1155/2021/8387680

[2] Ramalingam, V. V., Dandapath, A. & Raja, M. K. (2018). Heart disease prediction using machine learning techniques: a survey. *International Journal of Engineering & Technology, 7*(2.8), 684-687. https://doi.org/10.14419/ijet.v7i2.8.10557

[3] Fatima, M. & Pasha, M. (2017). Survey of machine learning algorithms for disease diagnostic. *J Intell Learn Syst Appl., 9*(1), 1-16. https://doi.org/10.4236/jilsa.2017.91001

[4] Garg, A., Sharma, B. & Khan, R. (2021). Heart disease prediction using machine learning techniques. *IOP Conf. Ser.: Mater. Sci. Eng.* 1022 012046. https://doi.org/10.1088/1757-899X/1022/1/012046

[5] Raut, S., Magar, R., Memane, R. & Rupna, V. S. (2020). Heart disease prediction using machine learning. *Journal of Emerging Technologies and Innovative Research, 7*(6), 2081-2085. https://www.jetir.org/papers/JETIR2006301.pdf

[6] Vembandasamy, K., Sasipriya, R. & Deepa, E. (2015). Heart disease detection using Naive Bayes algorithm. *Int J Innov Sci Eng Technol. 2*(9), 441-444.

[7] Louridi, N., Douzi, S. & El Ouahidi, B. (2021). Machine learning-based identification of patients with a cardiovascular defect. *J Big Data, 8*, 133. https://doi.org/10.1186/s40537-021-00524-9

[8] Benjamin, E. J., Muntner, P., Alonso, A., et al. (2019). Heart Disease and Stroke Statistics–2019 Update: A Report from the American Heart Association. *Circulation, 139*(10), e56-e528. https://doi.org/10.1161/CIR.0000000000000659

[9] Ismaeel, S., Miri, A. & Chourishi, D. (2015). Using the Extreme Learning Machine (ELM) technique for heart disease diagnosis. *IEEE Canada International Humanitarian Technology Conference (IHTC 2015)*, 1-3. https://doi.org/10.1109/IHTC.2015.7238043

[10] Kanika, P. & Shah, D. R. (2016). Prediction of cardiovascular diseases using support vector machine and Bayesian classification. *International Journal of Computer Applications, 156*(2). https://doi.org/10.5120/ijca2016912368

[11] Siontis, K. C., Noseworthy, P. A., Attia, Z. I., et al. (2021). Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nat Rev Cardiol*. https://doi.org/10.1038/s41569-020-00503-2

[12] Fathima, K. & Vimina, E. R. (2022). Heart Disease Prediction Using Deep Neural Networks: A Novel Approach. In: Raj, J. S., Palanisamy, R., Perikos, I., Shi, Y. (eds) *Intelligent Sustainable Systems. Lecture Notes in Networks and Systems, vol 213*. Springer, Singapore. https://doi.org/10.1007/978-981-16-2422-3_56

[13] Dwivedi, A. K., Imtiaz, S. A. & Villegas, E. R. (2019). Algorithms for automatic analysis and classification of heart sounds - a systematic review. *IEEE Access, 7*. https://doi.org/10.1109/ACCESS.2018.2889437

[14] Kiran, J., Debbarma, N. & Ganjala, S. (2023). Heart Disease Prediction Using Machine Learning. In: Bhateja, V., Yang, X. S., Lin, J. C. W., Das, R. (eds) *Evolution in Computational Intelligence, FICTA 2022. Smart Innovation, Systems and Technologies, vol 326*. Springer, Singapore. https://doi.org/10.1007/978-981-19-7513-4_24

[15] Ansarullah, S. I. & Kumar, P. (2019). A systematic literature review on cardiovascular disorder identification using knowledge mining and machine learning method. *International Journal of Recent Technology and Engineering, 7*(6S), 1009-1015.

[16] Nazir, S., Shahzad, S., Mahfooz, S. & Nazir, M. (2018). Fuzzy logic-based decision support system for component security evaluation. *The International Arab Journal of Information Technology, 15*(2), 224-231.

[17] Amin, M. S., Chiam, Y. K. & Varathan, K. D. (2019). Identification of significant features and data mining techniques in predicting heart disease. *Telematics and Informatics, 36*, 82-93. https://doi.org/10.1016/j.tele.2018.11.007

[18] Mienye, I. D., Sun, Y. & Wang, Z. (2020). An improved ensemble learning approach for the prediction of heart disease risk. *Informatics in Medicine Unlocked, 20*, Article ID 100402. https://doi.org/10.1016/j.imu.2020.100402

[19] Amin, M. S., Chiam, Y. K. & Varathan, K. D. (2019). Identification of significant features and data mining techniques in predicting heart disease. *Telematics and Informatics, 36*, 82-93. https://doi.org/10.1016/j.tele.2018.11.007

[20] Enriko, I. K. A., Suryanegara, M. & Gunawan, D. A. (2016). Heart Disease Prediction System Using K-Nearest Neighbor Algorithm with Simplified Patient's Health Parameters. Springer, Berlin, Germany, 59-65.

[21] Kompella, S. & Boddu, V. (2019). Neural network based intelligent system for predicting heart disease. *International Journal of Innovative Technology and Exploring Engineering, 8*. 484-487.

[22] Tarawneh, M. & Embarak, O. (2019). Hybrid approach for heart disease prediction using data mining techniques. *Acta Scientific Nutritional Health, 3*(7), 147-151. https://doi.org/10.1007/978-3-030-12839-5_41

[23] Jagtap, Malewadkar, P., Baswat, O. & Rambade, H. (2019). Heart disease prediction using machine learning. *International*

Journal of Research in Engineering, Science and Management, 2(2), 352-355.

[24] Kim, J. K. & Kang, S. (2017). Neural network-based coronary heart disease risk prediction using feature correlation analysis. *Journal of Healthcare Engineering*, Article ID 2780501. https://doi.org/10.1155/2017/2780501

[25] Yin, L. W., Gregory, P. A., Amanda, Z. & Margaux, G. (2021). Development of a novel clinical decision support system for exercise prescription among patients with multiple cardiovascular disease risk factors. *Mayo Clinic Proceedings: Innovations, Quality & Outcomes, 5*(1), 193-203. https://doi.org/10.1016/j.mayocpiqo.2020.08.005

[26] Pe, Rubini, Subasini, C., Katharine, A., Kumaresan, V., Gowdhamkumar, S. & Nithya, T. (2021). A Cardiovascular Disease Prediction using Machine Learning Algorithms. *Annals of the Romanian Society for Cell Biology, 25*(2), 904-912.

[27] Sherubha, P., Sasirekha, S. P., Manikandan, V. et al. (2020) Graph based event measurement for analyzing distributed anomalies in sensor networks. *Sādhanā, 45*, 212. https://doi.org/10.1007/s12046-020-01451-w

[28] Sherubha, P. & Mohanasundaram, N. (2019). An Efficient Network Threat Detection and Classification Method using ANP-MVPS Algorithm in Wireless Sensor Networks. *International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8*(11), 1597-1606. https://doi.org/10.35940/ijitee.K3958.0981119

[29] Sherubha, P. & Mohanasundaram, N. (2020). An Efficient Intrusion Detection and Authentication Mechanism for Detecting Clone Attack in Wireless Sensor Networks. *Journal of Advanced Research in Dynamical and Control Systems (JARDCS), 11*(5), 55-68.

[30] Singh, A. & Kumar, R. (2020). *Heart disease prediction using machine learning algorithms. International Conference on Electrical & Electronics Engineering, (ICE3)*, Gorakhpur, India, 452-457. https://doi.org/10.1109/ICE348803.2020.9122958

Authors' contacts:

Mounika Valasapalli, Research Scholar
(Corresponding author)
Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Guntur-522302, India
mounikatkrcse@gmail.com

Nallagatla Raghavendra Sai, Associate Professor
Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Guntur-522302, India
nallagatlaraghavendra@kluniversity.in