Prediction of Heart Disease with Autoencoder based ANN

N Deshai[1], S. Ramya, Dr. B V D S Sekhar, Dr. S. V. Ramana
[1]Department of Information Technology, S R K R Engineering college, Bhimavaram, AP, India

A two-stage technique is suggested in this paper to accurately predict major heart problem. The very first phase encompass of training in order to find the proper interpretation of the neural network with an enhanced sparse auto encoder (SAE), an unsupervised neural network over the cervical cancer dataset. Another phase includes the complete prediction of health status to use an artificial neural network (ANN) focused on relevant documents collected. The SAE has been configured because which is required to practice an effective model. The experimental outcome could achieve with the proposed method shows that it increases the efficiency of the ANN classifier and is more stable than that of the ANN classifier.

1. INTRODUCTION

Heart disease (HD) has also been categorized as one of the most toxic Genetic disorders and their therapy and diagnosis are very challenging. It's indeed difficult but essential to predict heart problems since the death rates if the disease is diagnosed initial and disease prevention steps are necessary, the frequency could be immensely reduced [1]. Consequently, precise estimation of the patient's the risk of heart disease is essential to decrease the associated risks of Painful situations in the heart [2]. To bring about that and save humanity there unprocessed heart information must be effectively processed to achieve suitable categorization as classifier. In Numerous scientists, to maximizing the effectiveness of high dimensional designs, using machine learning technology to identify different modes. In [2], the researchers present a model for predicting heart disease, In which they made a decrease in the features and it had an influence on the Implementation of the classification algorithms they tested on, also with support vector machine getting a precision of 88 per cent. Likewise, in the case of Ref.[3] a Technique which really produces decision-making rules in order to achieve efficient is recommended Categorize the risk of heart disease, and the experimental result has shown us that The precision of their technique was 86.7%. Even so, to go much further To enhance the advances achieved so far, other methodologies must be used. Deep learning techniques deployed in many fields, particularly Deep autoencoders in picture and visual processing, and in recent days, Such unsupervised computers had gained outstanding quality in some Activities in learning. A potential solution to the heart disease problem is Detection owing to its excellent progress in learning great qualities Interpretations of both broad and complex datasets. An autoencoder essentially contains two capacities, normally an encoder can plot the first d-dimensional information into middle and shrouded description, and then finally decoding which mapping the concealed demonstration reverse to a d-dimensional vector due to necessary like a conceivable to the first contribution encoder. This cycle is describing a reproduction error, while the distinction among the decoder result and the input of encoder is known as the recreation blunder [4]. Examination could be appeared where the order execution can be enhanced when demonstration are found out with a manner that empowers sparsity. Like inadequate autoencoders, the preparation model remembers a sparsity punishment for the code layer.

In this paper, i suggest a dynamic also well trustworthy sparse auto encoder (SAE) method smoothly handle unsupervised feature learning; efficiently predict major conditions over the heart. I centre on building up an SAE model to learn compelling highlights from the High dimensional dataset and afterwards execute characterization utilizing the learning highlights. Specifically advanced utilizing the model and adaptive moment estimation (Adam) calculation as accomplish the actively change of various boundaries, and a group standardization procedure is useful to dodge over fitting also to grow up the way of working, velocity, and security of the classifier. The advanced setting additionally guarantees reproduction blunder is essentially limited.

The improvement of the proposed technique is checked by contrasting and an independent ANN, a efficiently working algorithms, for example, k- nearest neighbor (KNN), classification regression (CART), Logistic regression (LR), and linear discriminate analysis (LDA) classifier execution. Lastly focus essential thing is outcome shows that our expected approach gets main classifier execution.

2. RELATED WORKS

Generally related working examines several earlier workings which considered also utilized an inadequate autoencoders. Also ongoing instant autoencoders are discovered huge functions in different unaccompanied learning errands in a few functions areas. In a strategy could suggested which consolidated SVM and scanty autoencoder [5]. The normal the old style Support Vector Machine has constraints for enormous scope applications; consequently, the need to utilize an inadequate autoencoder to develop the exhibition. Actually each and every creator utilized numerous autoencoder meager layers to perform highlight learning and utilized the SVM for grouping, along these lines improving the presentation of the SVM in dealing with huge scope datasets.

A comparable study, the authors are recommended a strategy to achieve highlight getting the hang of utilizing sparse auto encoders to develop the execution of the regression form on actual esteemed time series knowledge [6]. Finally the study purpose was to improve vehicular traffic stream estimating. Also, in an offer to build the precision of the sparse auto encoder, which has introduced a fall model which influences on the blend of some sort of low and more higher-level highlights, and a stochastic inclination drop calculation be

utilized like relapse method. an additional exploration be led to do abnormality identification by adapting inconsistency scores [7]. This examination view remaking mistake of foundation information is little while that of irregularity information is moderately high, the rebuilding blunder could be utilized as a peculiarity score. Consequently, which recommended a spare auto encoder manner abnormality identification strategy which employments a double concentric window.

Sparse demonstration manner classification strategy could be recommended utilizing deep learning manner working transductive. The network contains a completely associated layer and a convolution autoencoder. The completely associated layer is put linking the encoder furthermore, decoder, and it can locate the scanty portrayal, though the autoencoder network learns successful profound highlights for characterization. At the point when the assessed inadequate codes are utilized for classification of some datasets, the recommended technique demonstrated improved execution.

3. PROPOSED METHODOLOGY

The encoder mapping the input to a novel demonstration. So novel demonstration is then decoded at the output to restructure the input ***x***' according to Equations where x is the input and z the novel demonstration.

$$Z=h\,(Wx+b)$$
$$X'=g\,(W'z+b')$$

Here h is the activation function to neurons of hidden layer and g is neurons for the output layer, are weight matrices are W and W', and b and b' are the encoder and decoder bias vectors, respectively.

$$\mathbf{h} = g = \frac{\mathbf{1}}{\mathbf{1} + e^{-x}}$$

In this paper, the sigmoid activation function is make use of the others such as Relu, Tanh

The re-enactment error method E linking the input x and reconstructed input x' uses the mean squared error (MSE)

$$\mathrm{E} = \frac{\mathbf{1}}{\mathbf{N}} \sum_{i=1}^{N} xi + x'i^2$$

N denotes the fraction of input units. Though, in this study, a sparse autoencoder is employed to achieve a useful low-level description of the input data supporting sparse constraints. Therefore, sparsity is included by adding regularization to the cost function. Let pi be the normal activation of neurons in the hidden layer.

$$\mathrm{pi}\hat{} = \frac{\mathbf{1}}{\mathbf{n}} \sum_{j=1}^{n} \mathrm{zi(xj)}$$

The HD dataset is taken from the Kaggle website [7]. The dataset was acquired later a cardiovascular examination on citizens of Framingham, Massachusetts. The purpose for doing the Kaggle Framingham heart dataset is since it produces a tremendous quantity of things (4238) compared to the Cleveland, Hungarian, and Long Beach heart datasets which have 303, 294, and 200 instances correspondingly. The dataset includes patient knowledge and it proposes at predicting their 10-year risk of coming coronary heart disease (CHD). It has of 4238 individuals and 16 characteristics. Each feature is a feasible risk factor, and they include behavioural, demographic, and pharmaceutical risk circumstances. The dataset includes absent properties. After filtering rows with missing characteristics, 3656 records were moved; between which 3099 held negative and 557 positives. We employed 70-30% train-test data partitioning strategy.

4. EXPERIMENTAL SETUP

The recommended strategy includes two actions. Initially, the dataset is preprocessed to compose it becoming for making our model. Later pre-processing the original data, the dataset is then divided into train and test parts. The sparse autoencoder was trained to utilize the negative situations in the training set. The reason here is that if the model could learn the proper representation of these negative individuals, while achieved with hidden patterns both positive and negative that should recognize them. Our concern is to get the potential representation of the input received by the sparse autoencoder pattern and practice it to train the artificial neural network. Hence, once the sparse autoencoder design is prepared, the encoding section is employed to generate extra network. This interface is utilized to convert the train and test sets, including both positive and negative samples. This softly changes the dataset to a low-dimensional description dataset.

To demonstrate the effectiveness and achievement of the features learned by our recommended sparse autoencoder, primly, we applied an ANN utilizing the new data and further working the learned features as shown in Table 1

Table 1. Different classifier results

Algorithm	Accuracy	Precision	Recall	F1 Score
K-Nearest Neighbor	80	74	82	77
Classification & Regression	76	74	75	76
Linear Regression	84	83	82	75
Bayesian Classifier	83	77	83	80
Linear discriminant analysis	84	80	84	78
sparse auto encoder + artificial neural network	91	88	92	91

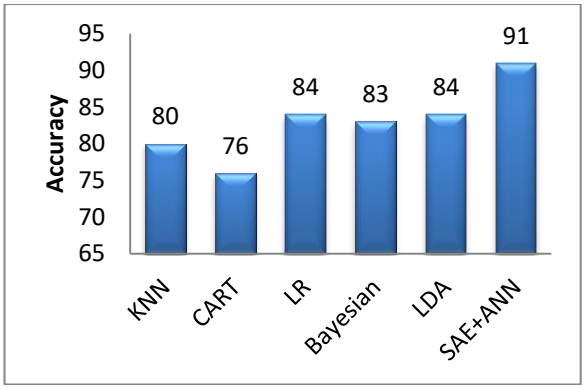

Fig.1. Different classifier results on Accuracy

From the outcomes achieved so far, it could be observed that the suggested approach gives meaningful growth associated to the other classifications in terms of classification enforcement. And the recommended sparse autoencoder raises the precision of the ANN related to a situation where the ANN individual was employed to perform forecasts. The outcomes also note that upgraded performance can be accomplished not only by growing the structure of the neural network or making hyper-parameter tuning of algorithms but further by increasing the preprocessing stage of the classification procedure.

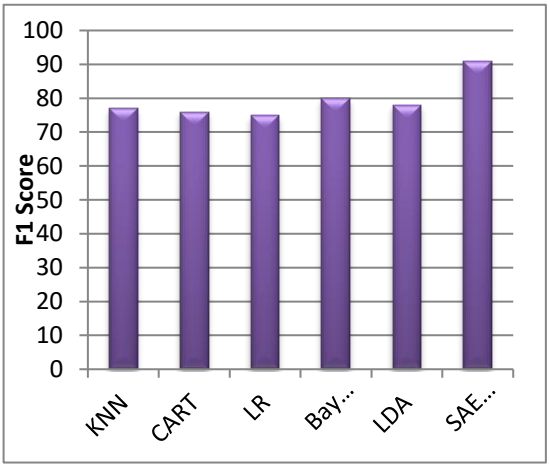

Fig.4. Different classifier results on F1 score

5. CONCLUSION

In this paper, an enhanced manner ANN that is sparse autoencoder is projected to assist the efficient prediction of major heart disease as most reliable and efficient manner. The sparse autoencoder has been used to find out the most excellent demonstration of the dataset whereas the ANN uses to compose predictions from learned aspects. The SAE could be optimized with Adam approach also batch normalization deployed. The precision of the classifier on tested dataset has 91%. Match up to to a few conventional machine learning strategies and ANN, but here the recommended technique demonstrate enhanced result.

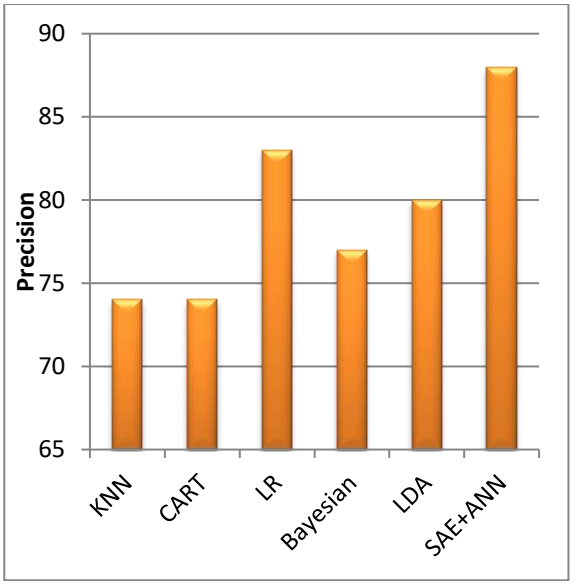

Fig.2. Different classifier results on Precision

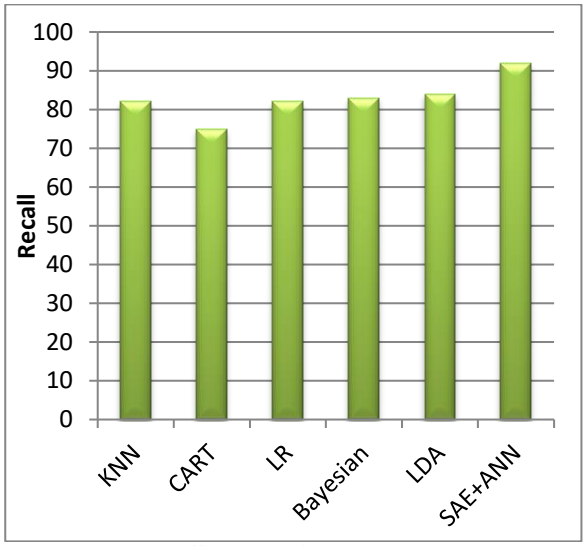

Fig.3. Different classifier results on Recall

6. REFRENCES:

[1] D.ShinyIreneaT.SethukarasibN.Vadivelanc,"Heart disease prediction using hybrid fuzzy K-medoids attribute weighting method with DBN-KELM based regression model"Medical Hypotheses 143 (2020) 110072.

[2] FarmanAlia, ShakerEl-Sappagh, S.M. RiazulIslam, DaehanKwak, AmjadAli,MuhammadImran, Kyung-SupKwak,"A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion"Information Fusion, Volume 63, November 2020, Pages 208-222

[3] EktaMainia,BonduVenkateswarlu,BaljeetMaini,DheerajMarwaha,"Machine learning–based heart disease prediction system for Indian population: An exploratory study done in South India

[4] AniruddhaDutta,TamalBatabyal, MeheliBasu,Scott T.Acton"An efficient convolutional neural network for coronary heart disease prediction",Expert Systems With Applications 159 (2020) 113408 pp.1-16

[5] Shaodong,Zhenga,Jinsong,Zhao,"A new unsupervised data mining method based on the stacked autoencoder for chemical process fault diagnosis"Computers and Chemical Engineering 135 (2020) 106755 pp.1-17

[6] YongmingLi,YanLei,PinWang,MingfengJiang,YuchuanLiua," A deep learning method combined sparse autoencoder with SVM",

[7] PelinGörgel,AhmetSimsek,"Face recognition via Deep Stacked Denoising Sparse Autoencoders (DSDSA)"Applied Mathematics and Computation 355 (2019) 325–342