An intelligent system for heart disease diagnosis using regularized deep neural network

P. Rani[a]* • R. Kumar[b]* • A. Jain[c]

[a]*MMICTBM, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala, Haryana, India*
[b]*Department of Computer Engineering, MMEC,*
Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala, Haryana, India
[c]*School of Computer Sciences, University of Petroleum and Energy Studies, Dehradun,*
Uttarakhand, India

Abstract: Heart disease is one of the deadly diseases. Timely detection of the disease can prevent mortality. In this paper, an intelligent system is proposed for the diagnosis of heart disease using clinical parameters at early stages. The system is developed using the regularized deep neural network model (Reg-DNN). Cleveland heart disease dataset has been used for training the model. Regularization has been achieved by using dropout and L2 regularization. Efficiency of Reg-DNN was evaluated by using hold-out validation method.70% data was used for training the model and 30% data was used for testing the model. Results indicate that Reg-DNN provided better performance than conventional DNN. Regularization has helped to overcome overfitting. Reg-DNN has achieved an accuracy of 94.79%. Results achieved are quite promising as compared to existing systems in the literature. Authors developed a system containing a graphical user interface. Hence, the system can be easily used by anyone to diagnose heart disease using clinical parameters.

*Corresponding author.
E-mail address: pooja.rani@mmumullana.org (P. Rani), drrajneeshgujral@mmumullana.org (R. Kumar).
Peer Review under the responsibility of Universidad Nacional Autónoma de México.

1. Introduction

Heart disease is a chronic disease that has a complex diagnosis process. It has many symptoms including chest pain, heartburn, and breathlessness, etc. It is a challenging and important task to predict heart disease at early stages because it can reduce the mortality rate (Kumar & Rani, 2020). Heart disease must be diagnosed at early stages; otherwise, it can lead to life-threatening situations. The patient does not experience severe symptoms until blockage in blood vessels exceeds a level. If the blockage in arteries increases it can also lead to a heart attack (Rajkumar et al., 2016).

Although several methods of predicting heart disease exist, they are not completely accurate. The most precise method of diagnosing heart disease is angiography. But it is a costly and invasive method having certain side effects. So, there is a requirement for non-invasive methods that can diagnose heart disease with increased accuracy. Automated diagnostic systems are an effort in this direction (Saxena & Sharma et al., 2016). Algorithms of machine learning can be used to develop such an automated diagnostic system (Lamba et al., 2021; Sachdeva et al., 2022). Different automated systems have been proposed by researchers to diagnose heart disease. These systems diagnose heart disease with different levels of accuracy. Recent research also shows that deep learning methods can be used to develop such kind of systems (Rani et al., 2021a; Mathur et al., 2020).

The principle of deep learning is based on the working of the brain that uses multiple neurons to extract complex patterns from data. Deep learning can use multiple hidden layers to learn complex patterns of data. Unlike machine learning, DNN does not have to select relevant features. It can perform feature engineering automatically. Deep learning can also work on unstructured data. Researchers have leaned towards DNN to develop automated decision support systems due to the benefits of deep learning over machine learning (Lamba et al., 2022a). In this paper, a heart disease diagnostic system is proposed using a deep neural network. The main motivation behind this work is to diagnose the disease in its initial stages and reduce the mortality rate. Another motivation is the availability of limited medical resources in developing countries. Such kind of automated systems can be useful in remote areas if heart specialists are not available.

The prime objective of this paper is to develop an optimized DNN model that performs well with both training and testing data. Different experiments were conducted on DNN to find out an optimum configuration of the DNN model. Furthermore, regularization techniques were applied to remove overfitting and increase the accuracy of the model. Proposed DNN can reduce the human error associated with diagnosing heart disease. The mortality rate can be reduced by providing appropriate treatment through timely diagnosis. It can act as an advisor to doctors to increase the reliability of diagnosis. Main research contributions of this paper are as follows:

1. A Regularized DNN has been developed for heart disease diagnosis.
2. Optimized network configuration is searched by experimenting with different combinations of layers and nodes.
3. The performance of conventional DNN has been compared to regularized DNN.

The remaining sections of this paper are organized as follows: Section 2 contains a literature review of related work. Section 3 contains a description of the dataset and methodology of the proposed system. The results and discussions are contained in section 4. Section 5 contains the conclusion and future scope.

2. Literature review

Wiharto et al. (2016) proposed a system to diagnose heart disease using K-star algorithm. System achieved sensitivity of 80.1%. Jabbar et al. (2016) used chi-square method for selecting important features. Heart disease was diagnosed using random forest using these features. System achieved 83.60% accuracy. Dwivedi et al. (2018) predicted heart disease using six classifiers artificial neural network (ANN), logistic regression (LR), support vector machine (SVM), decision tree (DT), k-nearest neighbor (KNN) and naive bayes (NB). LR achieved the highest accuracy of 85%.

Samuel et al. (2017) developed ANN model to predict the heart failure. Weights of the network were optimized by using the fuzzy approach to calculate the contribution of the attributes in the prediction. Accuracy of 91.10% was achieved. Acharya et al. (2017) proposed a convolutional neural network-based system to diagnose heart disease using ECG signals and achieved 95.22% accuracy. Liu et al. (2017) developed a hybrid system to diagnose heart disease using C4.5, multilayer perceptron and naive bayes classifiers. Boosting had increased the performance of the system. Relief and Rough set feature selection methods were used. The system achieved 92.59% accuracy.

Haq et al. (2018) used different combinations of feature selection algorithms and classifiers to develop a system for disease prediction. Best accuracy of 89% was provided with LR. David and Belcy (2018) proposed a heart disease diagnosis system using random forest (RF), DT and NB. RF provided the accuracy of 81%. Authors used Statlog heart disease dataset. Malav and Kadam (2018) used ANN (artificial neural network) to predict heart disease. Clustering was performed using K-Means to provide input to ANN. System provided 93.52%

accuracy. Miao and Miao (2018) developed a deep neural network based upon multilayer perceptron architecture and diagnosed heart disease with 83.67% accuracy.

Latha and Jeeva (2019) used majority voting with NB, multilayer perceptron (MLP) and RF to propose an ensemble model to predict heart disease with 85.48% accuracy. Ali et al. (2019) proposed a system to diagnose heart disease using a deep neural network. Dimensionality reduction was done using the chi-square method. Grid search method was used to optimize network configuration. The system achieved 93.33% accuracy.

Dutta et al. (2020) identified important features using LASSO method and then performed heart disease prediction using a convolutional neural network. The system achieved 79.5% accuracy. Gárate-Escamila, et al. (2020) performed the dimensionality reduction using chi-square and Principal Component Analysis. An optimized set of features was classified by RF for heart disease diagnosis with 98.7% accuracy.

Tama et al. (2020) developed a system based upon stacked architecture. Relevant features were chosen using particle swarm optimization. Ensemble model was constructed using gradient boosting (GB) and RF classifiers. The system achieved 85.71% accuracy. Verma and Mathur (2019) used correlation and cuckoo search method to selected important features and developed a DNN. Severity of the disease was informed using case-based reasoning. An accuracy of 85.48% was achieved using this approach.

Terrada et al. (2020) used decision tree, ANN, and Adaboost classifiers to diagnose heart disease and performed experiments on three datasets. ANN provided the best accuracy of 91.47% on Cleveland dataset, 90% on Hungarian dataset and 94% on Z-alizadeh sani dataset. Verma (2020) developed an ensemble model using J48, CART and random forest classifiers. System achieved 84.82% accuracy.

Javid et al. (2020) developed an ensemble model using voting mechanism for heart disease prediction. SVM, KNN, RF, gated-recurrent unit (GRU) and long short term memory (LSTM) classifiers were used to develop the ensemble model. System achieved 85.71% accuracy. Mienye et al. (2020a) partitioned the dataset into different segments. Different models were developed on the partitioned datasets using classification and regression tree (CART). An ensemble model was developed by combining different CART models. System achieved 93% accuracy on Cleveland dataset and 91% accuracy on Framingham dataset.

Sarkar (2020) developed a two-tier optimized prediction system for heart disease. The first level identified the optimal ratio of training and testing data. Perfect rule induction by sequential method (PRISM) was used to generate rules with the best training set. In the second level, the rule set was optimized with parallel genetic algorithm. The system was tested on several datasets. System achieved 89.81% accuracy on Cleveland dataset, 86.21% on Hungarian dataset and

88.34% on Statlog dataset. Mienye et al. (2020b) proposed a two-stage model for diagnosis of heart disease combining sparse autoencoder with ANN. Sparse autoencoder learned the low dimensional features and provided input to ANN. ANN performed the classification based upon the representation learned by sparse autoencoder. System achieved 90% accuracy. Rani et al. (2021b) proposed a hybrid system for heart disease diagnosis. Important features were selected using hybrid genetic and recursive feature elimination algorithm. SVM, NB, LR, RF and Adaboost classifiers were used to develop the system. System achieved maximum accuracy of 86.60% with RF classifier. Rani et al. (2021c) obtained 82.69% accuracy with SVM classifier. Salhi et al. (2021) selected important features using correlation matrix. Authors used ANN, KNN and SVM to diagnose heart disease using selected features. Maximum accuracy of 93% was achieved with ANN.

3. Dataset, methodology and evaluation metrics

3.1. Dataset

Authors performed experiments on Cleveland heart disease dataset available on (University of California, Irvine) repository (https://archive.ics.uci.edu/ml/datasets/heart+disease). Authors selected this dataset because much of the existing research in the literature had used this dataset. This dataset is having records of 303 persons. Seven instances are having missing values. The dataset has 14 attributes.13 attributes in the dataset are features used for the prediction of heart disease. Age, chest pain type, resting blood pressure, fasting blood sugar, sex, electrocardiographic results, serum cholesterol, exercise-induced angina, maximum heart rate, ST depression, slope of the peak exercise, thallium and number of major vessels (0-3) colored by fluoroscopy are prediction features in the dataset (Rani et al, 2022). The distribution of features is shown in Figure 1.
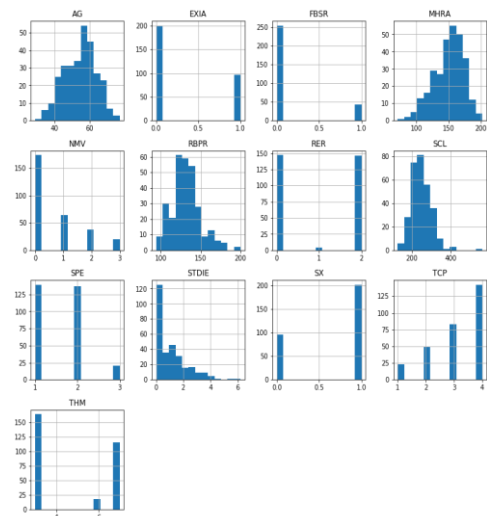

Figure 1. Distribution of features in dataset.

3.2. Methodology

Methodology used in implementing a regularized DNN is discussed in this section. Instances having missing values were removed from the dataset and data was normalized by applying scaling as a part of preprocessing.

DNN gives better performance when trained with a larger dataset, so class balancing was applied to increase dataset size. Hold-out validation method was used with 70% data for training the model and 30% data for testing the model. DNN model for heart disease diagnosis is shown in Figure 2.
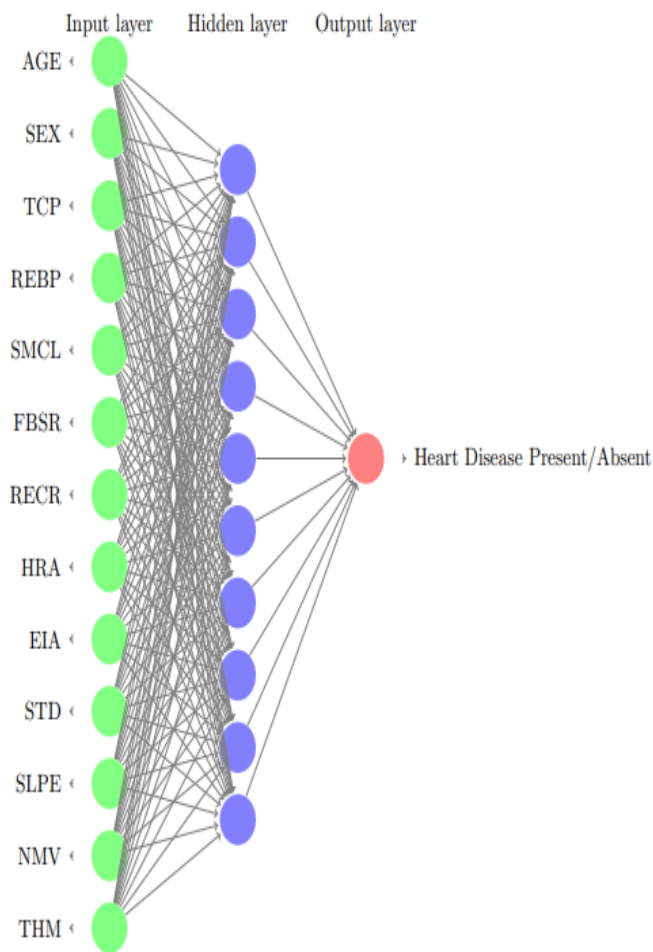

Figure 2. DNN model for heart disease diagnosis.

There are 3 layers in DNN i.e. input layer, hidden layer, and output layer. Input layer has 13 nodes. All the prediction attributes of the dataset are passed to the input layer. The hidden layer has 10 nodes and the output layer has 1 node. Authors have experimented with the different number of layers and nodes to find out the optimum combination. Authors have used dropout layers to avoid overfitting. Parameters of Reg-DNN method are shown in Table 1.

Table 1. Parameters of Reg-DNN method.

Parameter	Value
Number of layers	3
Nodes in input layer	13
Nodes in output layer	1
Nodes in hidden layer	10
Dropout	0.3
Batch Size	30
Epochs	100

Output of a DNN is defined by using the equation 1:

$$Output = f(\sum(w * x) + b) \qquad (1)$$

Here f is activation function, w is weight of nodes, x is input and b is bias. Activation function is responsible for deciding which neurons will be activated in DNN. Relu (rectified linear activation function) is used in the input layer and hidden layer (Jabbar et al., 2016). Relu function is a non-linear function that can easily learn the complex patterns in data (Shankar et al., 2020). It returns the positive value as it is and converts a negative value into a positive value as defined by the equation 2:

$$R(x) = \max(0, x) \qquad (2)$$

Sigmoid activation function is used in the output layer (Verma & Mathur, 2020). Sigmoid function is defined by the equation 3:

$$f(x) = \frac{1}{1+e^{-x}} \qquad (3)$$

Network was initialized with random weights and trained with 100 epochs and batch size 30. Heart disease diagnosis is a binary classification problem, so, binary cross-entropy loss function was used which is defined by the equation 4:

$$B(x) = \sum_{i=1}^{n} A(x) log B(x) \qquad (4)$$

Where A is actual output and B is predicted output.
Model was regularized to prevent overfitting. Regularization penalizes weights assigned to nodes to remove overfitting. Proposed Reg-DNN is developed by applying Dropout and L2 regularization.

Dropout method of regularization randomly deactivates some nodes in the selected layer at each iteration. Dropout of 0.3 has been used in Reg-DNN.

L2 regularization updates the cost function of DNN by adding a regularization term. Cost function calculates the error between actual and predicted values (Rahangdale & Raut, 2019). L2 regularization updates the cost function as given by the equation 5:

$$Cost\ Function = \text{Loss} + \frac{\lambda}{2n} \quad (5)$$

$$* \sum ||w||^2$$
$$\lambda = Regularization\ Parameter$$
$$n = number\ of\ samples$$
$$w = Weight\ matrix$$

Adding the regularization term reduces the value of weight matrices and hence reducing overfitting. L2 regularization leads the weight values towards zero.

3.3. Evaluation metrics
System was evaluated using following performance metrics:

Accuracy: Accuracy measures the ability of a system to make true predictions as given in equation 6. We always strive to increase accuracy.

$$\text{Accuracy} = \left(\frac{\text{TP+TN}}{\text{TP+FP+TN+FN}}\right) * 100 \quad (6)$$

Where TP (true positive) and TN (true negative) are correct predictions whereas FP(false positive) and FN (false negative) are incorrect predictions done by the system.

Loss: In DNN we always strive to minimize error. This error can be calculated using loss function.

$$Loss = -\frac{1}{M}\sum_{i=1}^{M} x_i * \log(p(x_i)) + (1 - x_i) * \log(1 - p(x_i)) \quad (7)$$

Binarycross-entropy loss function is used for calculating loss in binary classification problems. Loss gives a measurement of the difference between actual and predicted probability distributions. Loss can be calculated as given in the equation 7 (Sachdeva & Bathla, 2022).

Where M is total number of samples, x_i is output class label for sample i, $\log(p(x_i))$ is log probability of class label of sample i.

In addition to the above evaluation metrics, receiver operating characteristic curve (ROC) of the model is drawn to demonstrate the ability of the system in performing correct diagnosis. In ROC chart, false positive rate (FPR) is taken on the x-axis and true positive rate (TPR) is taken on the y-axis. ROC chart having TPR of 1 and FPR of 0 is an ideal ROC curve. The more area under the curve, the more accurate diagnosis will be performed (Lamba et al., 2022b).

4. Results and discussion

The efficiency of Reg-DNN was evaluated by using hold-out validation method. 70% data was used for training the model and 30% data was used for testing the model. Two types of models of DNN were created: Conventional DNN and Regularized DNN.

Experiments were performed with both types of models to analyze the effect of regularization on model accuracy. Regularized DNN was developed by applying Dropout and L2 regularization. Model accuracy and loss with conventional and regularized DNN are shown in Table 2.

Table 2. Performance of deep neural network models on heart disease dataset.

DNN Model	Accuracy	Loss
Regularized DNN (Test Data)	94.79	31.34
Regularized DNN (Train Data)	83.03	37.46
Conventional DNN (Test Data)	90.62	34.69
Conventional DNN (Train Data)	83.03	27.57

In the development of DNN, it is necessary to ensure that overfitting does not occur. Inappropriate network configuration leads to overfitting of the model. When overfitting occurs network is highly fitted with training data and gives good accuracy with training data but gives a poor performance with testing data.

In our experiments, the problem of overfitting was overcome by regularization. Accuracy and loss of conventional DNN and regularized DNN over number of epochs is shown in Figure 3 and 4. Figure 3 and 4 indicate that regularized DNN increased the accuracy and reduced loss for test data. ROC chart for conventional DNN and regularized DNN is shown in Figure 5. AUC of Regularized DNN is 0.963 whereas AUC of conventional DNN is 0.954. ROC chart indicates that Regularized DNN has shown better performance than conventional DNN.

Proposed Reg-DNN performed better than many systems already existing in the literature for diagnosis of heart disease. Performance of the proposed method was compared to the performance of the existing systems as shown in Table 3. A comparison of the accuracy of the proposed Reg-DNN with existing systems is shown in Figure 6.

Authors also developed a graphical user interface for an intelligent system for heart disease diagnosis so that system can be easily used by anyone. Graphical user interface of the system is shown in Figure 7. Clinical parameters of the patient are accepted through this GUI and input is provided to the developed model. Model predicts whether patient is suffering from heart disease or not.

Finally, the user is informed whether the person should consult a cardiologist. Proposed system can be very helpful for diagnosing heart disease because of its increased accuracy. It can help health professionals to take the right decision at the right time.

Although the proposed system has given a good performance, it also has some limitations. The severity of the disease can not be identified using this system. It cannot identify in which artery the blockage is present and it is not recommending any treatment to the patient.
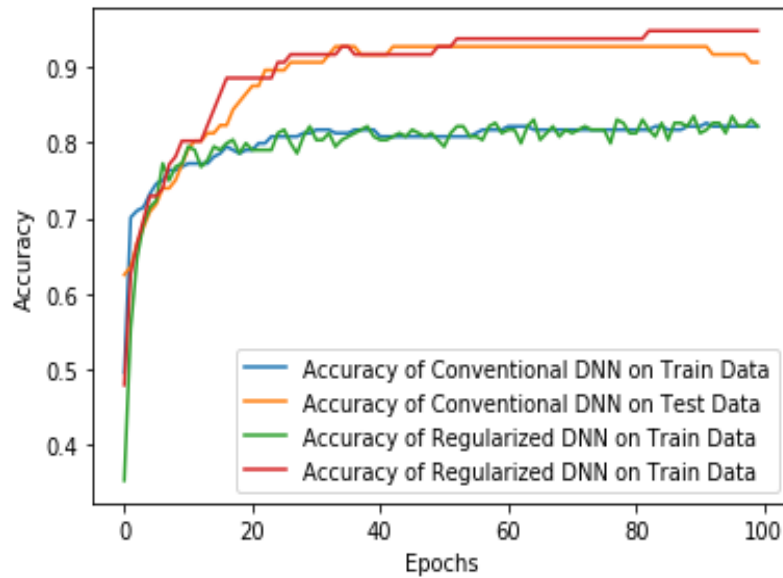
Figure 3. Accuracy of conventional DNN and regularized DNN over number of epochs.
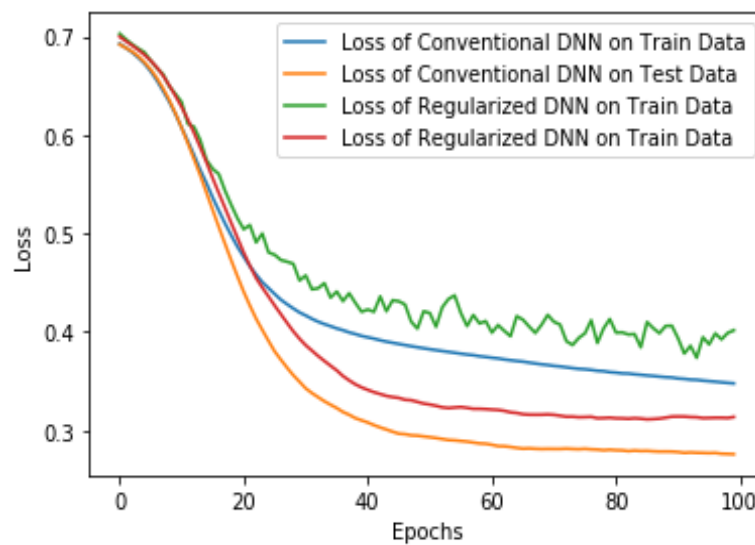

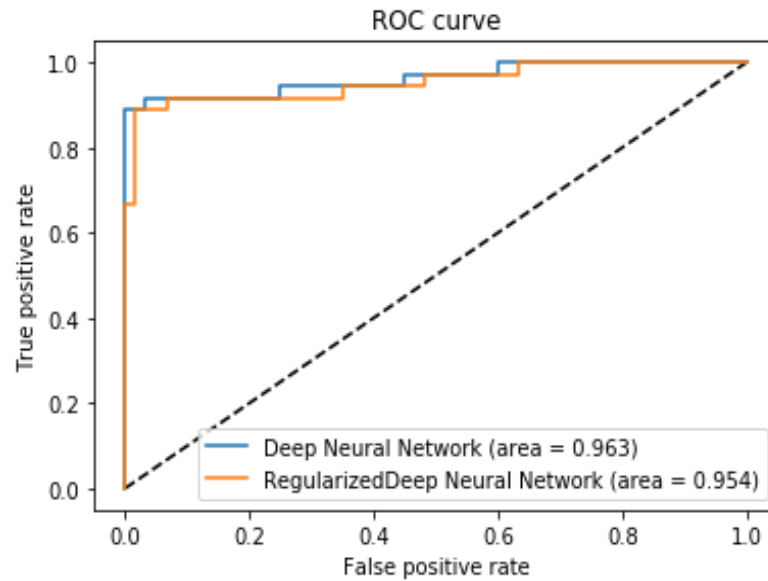Figure 4. Loss of conventional DNN and regularized DNN over number of epochs.

Figure 5. ROC chart of conventional and regularized DNN (Reg-DNN).

Table 3. Comparison of the proposed system with existing systems.

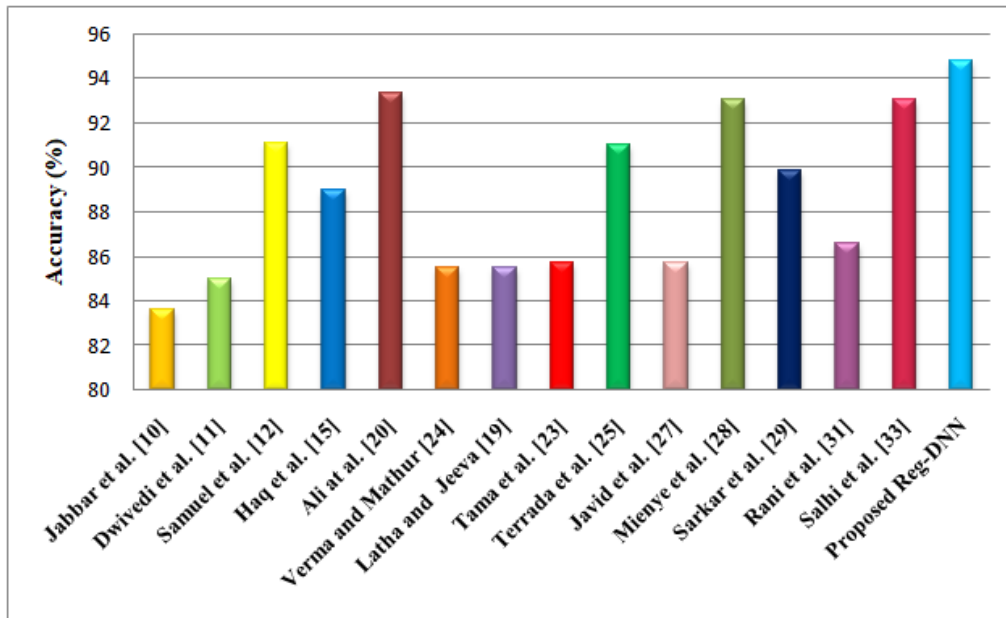| S.No.	Study	Year	Dataset used	Method	Accuracy(%)
1	Jabbar et al.	2016	Cleveland	Random forest classifier.	83.60
2	Dwivedi et al.	2016	Cleveland	Logistic regression	85
3	Samuel et al.	2017	Cleveland	Artificial neural network optimized by fuzzy approach	91.10
4	Haq et al.	2018	Cleveland	Logistic regression	89
5	Ali at al.	2019	Cleveland	DNN optimized by chi-square	93.33
6	Verma and Mathur	2019	Cleveland	Deep learning with features optimized by correlation and cuckoo search method	85.48
7	Latha and Jeeva	2019	Cleveland	Ensembled model using NB, MLP and RF	85.48
8	Tama et al.	2020	Cleveland	Two tier ensemble method	85.71
9	Terrada et al.	2020	Cleveland	Artificial neural network	91
10	Javid et al.	2020	Cleveland	Ensmebled model using majority voting	85.71
11	Mienye et al.	2020	Cleveland	CART	93
12	Sarkar et al.	2020	Cleveland	Two-tier PRISM optimized with parallel genetic algorithm	89.81
13	Rani et al.	2021	Cleveland	Random forest	86.60
14	Salhi et al.	2021	Cleveland	Artificial neural network	93
15	Proposed Reg-DNN	2021	Cleveland	Regularized deep neural network	94.79

Figure 6. Comparison of proposed Reg-DNN with existing systems.

Figure 7. Graphical user interface of an intelligent system for heart disease diagnosis.

6. Conclusions

In this paper, an intelligent system using a regularized deep neural network (Reg-DNN) has been developed for heart disease diagnosis. L2 regularization and dropout methods are used to prevent overfitting. The performance of the proposed system has been evaluated using accuracy, loss, AUC and ROC. Reg-DNN has given 94.79% accuracy and 29.25% loss. The proposed system can be used as an assistant by doctors to diagnose the disease. It can reduce the human error associated with the diagnosis process and hence reduce the mortality rate.

In future, authors have planned to collect the real-time data through the interface shown in Figure 7 and assist the doctor. Further, authors have also planned to increase the accuracy of the system by searching for optimum configuration of the network by using more sophisticated algorithms like genetic algorithm.

Authors have also planned to collect clinical parameters of patients automatically by combining the system with IoT. Authors have planned to find out optimum prediction features for heart disease diagnosis by consulting a cardiologist. Besides, the authors have planned to enhance the system by suggesting some precautionary measures and treatments for the patients.

Conflict of interest

The author(s) has(have) no conflict of interest to declare.

Funding

References

Acharya, U. R., Fujita, H., Oh, S. L., Hagiwara, Y., Tan, J. H., & Adam, M. (2017). Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. *Information Sciences*, *415*, 190-198. https://doi.org/10.1016/j.ins.2017.06.027

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 statistical model and optimally configured deep neural network. *IEEE Access*, *7*, 34938-34945. https://doi.org/10.1109/ACCESS.2019.2904800

David, H. B. F., & Belcy, S. A. (2018). Heart disease prediction using data mining techniques, ICTACT, *Journal on Soft Computing*, *9*(1), 1824-1830.

Dwivedi, A. K. (2018). Performance evaluation of different machine learning techniques for prediction of heart disease. *Neural Computing and Applications*, *29*, 685-693. https://doi.org/10.1007/s00521-016-2604-1

Dua, D. & Graff, C. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

Dutta, A., Batabyal, T., Basu, M., & Acton, S. T. (2020). An efficient convolutional neural network for coronary heart disease prediction. *Expert Systems with Applications*, *159*, 113408. https://doi.org/10.1016/j.eswa.2020.113408

Gárate-Escamila, A. K., El Hassani, A. H., & Andrès, E. (2020). Classification models for heart disease prediction using feature selection and PCA. *Informatics in Medicine Unlocked*, *19*, 100330. https://doi.org/10.1016/j.imu.2020.100330

Haq, A. U., Li, J. P., Memon, M. H., Nazir, S., & Sun, R. (2018). A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms. *Mobile information systems*, *2018*, 1-21. https://doi.org/10.1155/2018/3860146

Jabbar, M. A., Deekshatulu, B. L., & Chandra, P. (2016). Prediction of heart disease using random forest and feature subset selection. In *Innovations in Bio-Inspired Computing and Applications: Proceedings of the 6th International Conference on Innovations in Bio-Inspired Computing and Applications (IBICA 2015) held in Kochi, India during December 16-18, 2015* (pp. 187-196). Springer International Publishing. https://doi.org/10.1007/978-3-319-28031-8_16

Javid, I., Alsaedi, A. K. Z., & Ghazali, R. (2020). Enhanced accuracy of heart disease prediction using machine learning and recurrent neural networks ensemble majority voting method. *International Journal of Advanced Computer Science and Applications*, *11*(3). http://dx.doi.org/10.14569/IJACSA.2020.0110369

Kumar, R. & Rani, P. (2020). Comparative analysis of decision support system for heart disease. *Advances in Mathematics: Scientific Journal, 9*(6), 3349-3356. https://doi.org/10.37418/amsj.9.6.15

Lamba, R., Gulati, T., Al-Dhlan, K. A., & Jain, A. (2021). A systematic approach to diagnose Parkinson's disease through kinematic features extracted from handwritten drawings. Journal of Reliable Intelligent Environments 7, 253–262 (2021). https://doi.org/10.1007/s40860-021-00130-9

Lamba, R., Gulati, T., & Jain, A. (2022a). Automated Parkinson's Disease Diagnosis System Using Transfer Learning Techniques. In *Emergent Converging Technologies and Biomedical Systems: Select Proceedings of ETBS 2021* (pp. 183-196). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-16-8774-7_16

Lamba, R., Gulati, T., & Jain, A. (2022b). A hybrid feature selection approach for parkinson's detection based on mutual information gain and recursive feature elimination. *Arabian Journal for Science and Engineering*, *47*(8), 10263-10276. https://doi.org/10.1007/s13369-021-06544-0

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://doi.org/10.1016/j.imu.2019.100203

Liu, X., Wang, X., Su, Q., Zhang, M., Zhu, Y., Wang, Q., & Wang, Q. (2017). A hybrid classification system for heart disease diagnosis based on the RFRS method. *Computational and mathematical methods in medicine*, *2017*. https://doi.org/10.1155/2017/8272091

Malav, A., & Kadam, K. (2018). A hybrid approach for heart disease prediction using artificial neural network and K-means. *International Journal of Pure and Applied Mathematics*, *118*(8), 103-10.

Mathur, P., Srivastava, S., Xu, X., & Mehta, J. L. (2020). Artificial intelligence, machine learning, and cardiovascular disease. *Clinical Medicine Insights: Cardiology*, *14*, 1179546820927404. https://doi.org/10.1177/1179546820927404

Miao, K. H., & Miao, J. H. (2018). Coronary heart disease diagnosis using deep neural networks. *International Journal of Advanced Computer Science and Applications*, *9*(10). https://doi.org/10.14569/IJACSA.2018.091001

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

Mienye, I. D., Sun, Y., & Wang, Z. (2020b). Improved sparse autoencoder based artificial neural network approach for prediction of heart disease. *Informatics in Medicine Unlocked*, *18*, 100307. https://doi.org/10.1016/j.imu.2020.100307

Rahangdale, A., & Raut, S. (2019). Deep neural network regularization for feature selection in learning-to-rank. *IEEE Access*, *7*, 53988-54006. https://doi.org/10.1109/ACCESS.2019.2902640

Rajkumar, R., Anandakumar, K., & Bharathi, A. (2006). Coronary artery disease (CAD) prediction and classification-a survey. *ARPN*, *Journal of Engineering and Applied Sciences*, *11*(9), 5749-5754.

Rani, P., Kumar, R., & Jain, A. (2021a). Coronary artery disease diagnosis using extra tree-support vector machine: ET-SVMRBF. *International Journal of Computer Applications in Technology*, *66*(2), 209-218. https://doi.org/10.1504/IJCAT.2021.119772

Rani, P., Kumar, R., Ahmed, N. M. S., & Jain, A. (2021b). A decision support system for heart disease prediction based upon machine learning. *Journal of Reliable Intelligent Environments*, *7*(3), 263-275. https://doi.org/10.1007/s40860-021-00133-6

Rani, P., Kumar, R., Jain, A., & Chawla, S. K. (2021c). A hybrid approach for feature selection based on genetic algorithm and recursive feature elimination. *International Journal of Information System Modeling and Design (IJISMD)*, *12*(2), 17-38. https://doi.org/10.4018/IJISMD.2021040102

Rani, P., Kumar, R., & Jain, A. (2022). A Hybrid Approach for Feature Selection Based on Correlation Feature Selection and Genetic Algorithm. *International Journal of Software Innovation (IJSI)*, *10*(1), 1-17. http://doi.org/10.4018/IJSI.292028

Sachdeva, R. K., & Bathla, P. (2022). A machine learning-based framework for diagnosis of breast cancer. *International Journal of Software Innovation (IJSI)*, *10*(1), 1-11. https://doi.org/10.4018/IJSI.301221

Sachdeva, R. K., Bathla, P., Rani, P., Kukreja, V., & Ahuja, R. (2022). A systematic method for breast cancer classification using RFE feature selection. In *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 1673-1676). IEEE. https://doi.org/10.1109/ICACITE53722.2022.9823464

Shankar, V., Kumar, V., Devagade, U., Karanth, V., & Rohitaksha, K. (2020). Heart disease prediction using CNN algorithm. *SN Computer Science*, *1*(3), 170.
https://doi.org/10.1007/s42979-020-0097-6

Salhi, D. E., Tari, A., & Kechadi, M. T. (2021). Using machine learning for heart disease prediction. In *Advances in Computing Systems and Applications: Proceedings of the 4th Conference on Computing Systems and Applications* (pp. 70-81). Springer International Publishing.
https://doi.org/10.1007/978-3-030-69418-0_7

Samuel, O. W., Asogbon, G. M., Sangaiah, A. K., Fang, P., & Li, G. (2017). An integrated decision support system based on ANN and Fuzzy_AHP for heart failure risk prediction. *Expert Systems with Applications*, *68*, 163-172.
https://doi.org/10.1016/j.eswa.2016.10.020

Sarkar, B. K. (2020). Hybrid model for prediction of heart disease. *Soft Computing*, *24*, 1903-1925.
https://doi.org/10.1007/s00500-019-04022-2

Saxena, K., & Sharma, R. (2016). Efficient heart disease prediction system. *Procedia Computer Science*, *85*, 962-969.
https://doi.org/10.1016/j.procs.2016.05.288

Tama, B. A., Im, S., & Lee, S. (2020). Improving an intelligent detection system for coronary heart disease using a two-tier classifier ensemble. *BioMed Research International*, *2020*.
https://doi.org/10.1155/2020/9816142

Terrada, O., Hamida, S., Cherradi, B., Raihani, A., & Bouattane, O. (2020). Supervised machine learning based medical diagnosis support system for prediction of patients with heart disease. *Advances in Science, Technology and Engineering Systems Journal*, *5*(5), 269-277.
https://dx.doi.org/10.25046/aj050533

Verma, L., & Mathur, M. K. (2019). Deep learning based model for decision support with case based reasoning. *Int J Innov Technol Explor Eng*, *8*(6C), 149-153.

Verma, P. (2020). Ensemble Models for Classification of Coronary Artery Disease using Decision Trees. *International Journal of Recent Technology and Engineering*, *8*(6), 940-944.
http://dx.doi.org/10.35940/ijrte.F7250.038620

Wiharto, W., Kusnanto, H., & Herianto, H. (2016). Interpretation of clinical data based on C4. 5 algorithm for the diagnosis of coronary heart disease. *Healthcare informatics research*, *22*(3), 186-195.
https://doi.org/10.4258/hir.2016.22.3.186