A Performance Based Study on Deep Learning Algorithms in the Efficient Prediction of Heart Disease

Rakibul Islam[1]
Computer Science and Engineering
Daffodil Internajtional University
Dhaka, Bangladesh
rakibul15-2388@diu.edu.bd

Abhijit Reddy Beeravolu[2]
College of Engineering, IT and Environment
Charles Darwin University
NT, Australia
abhijithreddy.b@gmail.com

Md. Al Habib Islam[3]
Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
habib15-2629@diu.edu.bd

Asif Karim[4]
College of Engineering, IT and Environment
Charles Darwin University
NT, Australia
asif.karim@cdu.edu.au
0000-0001-8532-6816

Sami Azam[5*]
College of Engineering, IT and Environment
Charles Darwin University
NT, Australia
sami.azam@cdu.edu.au
0000-0001-7572-9750

Sanzida Akter Mukti[6]
Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
sanzida15-2346@diu.edu.bd

***Abstract*— Heart disease is the leading cause of death worldwide, nearly accounting for one-third of deaths. Heart disease describes a range of conditions that affect your heart. Most of these symptoms are dependent on the type of heart disease and their risk factors, such as high blood pressure, high cholesterol, and smoking. It is important to control the conditions before they become severe, it can save countless lives. Especially in remote areas and underdeveloped countries where there's no access to necessary medical systems and medical experts at the right time. Therefore, it is important to develop a 'medical system' that can provide heart disease assessments classifications from the clinical data, so that a clinician at a faraway location can reach a decision quickly, allowing them to manage a large number of patients. To do so, collecting clinical data related to heart disease is crucial. An open source dataset that consists of 1,190 samples and multi-variate features is collected from UCI machine learning repository. A total of 14 features are selected for this research. Data normalization is performed on these features to take care of irrelevant values, so that better results can be achieved by the trained models. This research uses three deep learning algorithms, namely Radial Basis Function Network (RBFN), Convolutional Neural network (CNN) and Artificial Neural Network (ANN) to train, validate, and test them with the selected, normalized, and separated data features. Various evaluation metrics were generated to understand the performance of the classification. This research has achieved classification scores of 98.24% and 98.49% for RBFN and ANN, respectively. Overall, the CNN model has achieved higher accuracy than the other models, with 98.75%.**

Keywords*— *Classification, CNN, Deep Learning, Heart Disease, UCI Machine Learning Repository.

I. Introduction

Heart disease nearly accounts for one third of deaths worldwide every year. According to WHO data from 2018, coronary heart disease now accounts for 15.23 % of all fatalities. The mortality rate is 109.32 per 100,000 persons when adjusted for age. Other factors contributed to heart disease being one of the most well-known causes of mortality in a society where individuals have a demanding work schedule that causes restlessness and stress. Excessive smoking, alcohol use, salty foods, high blood pressure, obesity and over-weight, and a lack of physical activity are all major causes of heart disease in Bangladesh. Our objective under these inevitable conditions is to provide an early diagnosis of cardiac problems, which will undoubtedly benefit the people of the world. Most clinicians follow the American Heart Association's recommendations [1], which examines four well-known risk factors such as cholesterol, hypertension, smoking, and diabetes [2, 3]. Heart disease has been and continues to be one of the main causes of mortality in humans, killing 20 million people each year [4]. We present a technique that utilises deep learning, CNN, ANN, and other deep learning methods to tackle the challenge of predicting cardiac disease. In this subject, several machine learning approaches have been tried and failed. As a result, we will utilise these ways to address those flaws and improve the system's performance over alternative approaches. We combined five different datasets into a single one. The goal of this study is to enhance the accuracy and efficiency of the system so that it can predict the likelihood of a heart attack.

II. Literature review

Heart disease has long been a major issue all around the world. Age, sex, and an irregular pulse rate are some of the risk factors that might aid in the early detection of heart disease. Others, such as one's lifestyle, smoking and drinking habits, and so on, have an impact on heart disease. Deep learning algorithms have proven highly useful in detecting cardiac disease due to the rapid rise of machine learning algorithms being utilised throughout the world. Depending on our datasets, deep learning algorithms provide more exact findings. The article employed a variety of deep learning techniques. A CNN and a multilayer perceptron (MLP) were used to assess foetal heart rate records, with an accuracy of 85% [5]; for automated identification of irregular beating rhythm in records, Recurrent Neural Networks (RNN) has also been suggested, which showed an accuracy of 83% [6]. In [7], a long-short term memory (LSTM) network was used to classify atrial fibrillation from a variety of electrocardiographic data, and it achieved 78% accuracy and

Preprint
IEEE International Conference on Informatics and Software Engineering, 2021, Turkey

79% F1 score. J. S. Sonawane et al., [9] presented a novel ANN that can be taught with a vector quantization method and random order incremental training. They also employed 13 clinical characteristics in their training and were able to obtain an accuracy of 85.55% on the dataset. M. G. Feshki et al., [10] utilised C4.5, feed-forward backpropagation, Multilayer Perceptron and Sequential Minimal Optimization among other classification methods. On the dataset, they found that the PSO method with neural networks had the greatest accuracy of about 91.94%. R.W. Jones et al., [11] proposed utilising an ANN to forecast cardiac disease. The neural network was trained using a self-administered questionnaire. The backpropagation technique was used to train the neural network, which included three hidden layers. The Dundee rank factor score was used to validate the design, and it yielded a 98% relative operational characteristic value on the dataset. In the buried layer, R. R. Manza et al., [12] presented an ANN with a large number of Radial Basis Function neurons. On this architecture, they got approximately 97% accuracy. For feature selection, P. Ramprakash et al., [13] presented a deep neural network approach. Turay Karayilan et al., [14] looked at how ANN performed with different numbers of hidden layers. Using five hidden layers, they were able to obtain an accuracy of 95.55%. Mehmood et al., [15] used the characteristics derived from the dataset acquired from the UCI repository to forecast a probable heart attack. The authors emphasised the significance of attribute extraction approaches in data mining for prediction. Using attribute extraction approaches, they noted, different patterns may be formed to detect heart disease sooner. This research paper explains several ANN methods. The accuracy of the ANN is 94.7%, whereas the accuracy of the principal component analysis is 97.7%. A. Dizaj et al., [16] employ the data mining approach as well. The authors looked at the effectiveness of data mining algorithms and used a decision tree to forecast the probability of a stroke in patients primarily based on the risk variables that influence it. To examine the possibility of implementing the CNN based classification model, Iman et al., [17] developed Hierarchical Edge-based deep learning (HEDL) based healthcare IoT system. In addition, a scenario designed using ECG classifications is utilised to evaluate the proposed system's execution time and accuracy. Liangzhi et al., [18] presented a Fog-based Efficient Manufacture Inspection (FEMI) system for smart industry that uses deep learning to handle huge amounts of data quickly. Furthermore, the FEMI system adjusts the CNN model to the fog based computing platform, resulting in a considerable increase in processing efficiency and an improvement in only testing accuracy. The deep learning algorithm was employed. The Keras model was utilised in conjunction with dense layers and the RBFN algorithm in this study. A combination of 5 heart disease dataset.csv is the name of our dataset. Data was gathered via the University of California at Irvine's dataset repository. The Keras dense model has a 98.75% accuracy, 95.37% testing accuracy, and 93.89 % sensitivity. Again, our RBFN algorithm has a 98.24 % accuracy, a 94.11% testing accuracy, and a 90.57% sensitivity. Our ANN algorithm has a 98.49 % accuracy, a 95.76% testing accuracy, and a 95.31% sensitivity.

III. RESEARCH METHODOLOGY

The model is constructed using TensorFlow GPU (2.4.0), Keras (2.4.3) and Python (3.8.5). Five different datasets were collected and aggregated into one. Min max standard scaler approach [19] was introduced to keep the working dataset between 0 and 1 [20]. Then the dataset is pre-processed to begin with and subsequently split into training and testing phases. More than 80% of the cases are taken for training and others are for testing. Afterwards, three deep learning classifiers are trained with the training which gave us the resulting model on the testing data. However, the CNN model performs a better result in terms of accuracy. A detailed explanation is shown in Fig. 1.

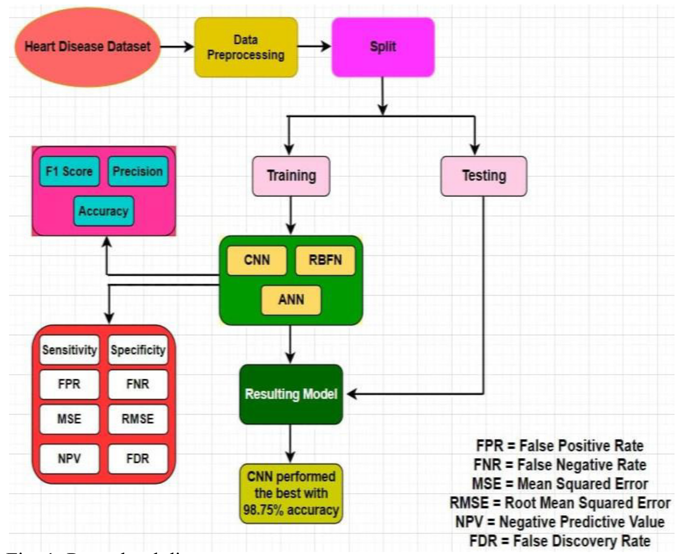

Fig. 1. Procedural diagram.

A. Radial Basis Function Network (RBFN) Architecture

A radial basis function network is a supervised ANN that uses 'radial bias functions' as activation functions to operate as a nonlinear classifier. Nonlinear classifiers use more sophisticated functions to analyze data than simple linear classifiers that work on lower-dimensional vectors. RBFN networks provide various advantages such as ease of design, good generalization, high input noise tolerance. The RBF networks have qualities that make them ideal for designing flexible control systems. Fig 2. Showcases the layers in the RBFN model where layers are from Keras backend. Weighted layer provides weights to the layers to feed forward the neurons. Batch size is 8 here then the output is shown.

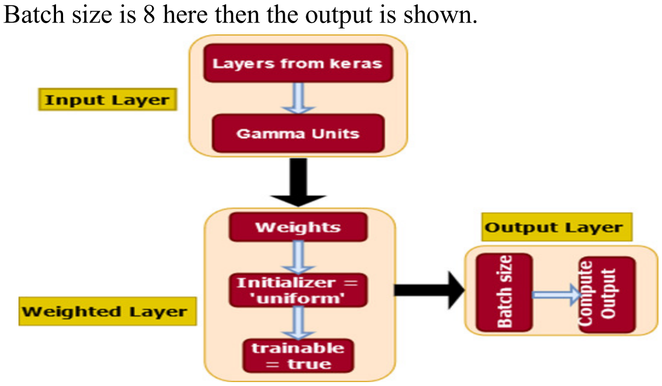

Fig. 2. RBFN model architecture.

B. Convolutional Neural Network (CNN) Architecture

CNN architecture consists of a stack of distinct layers that helps in extracting complex features from input data and providing an output. As seen in Fig. 3, the constructed CNN takes in the input data and assigns learnable weights and biases [21, 22] so that the model can easily differentiate between the features extracted from various classes. The Adaptive Moment Estimation (ADAM) loss function is used to come up with an efficient algorithm by optimizing the model to achieve better accuracy. Optimizers are methods utilized to alter the weights and learning rates, to reduce the training and testing losses. First 3 dense layer is 256 and the last one is 64. Optimizer is Adam with a learning rate of 0.09. Kernel initializers are normal. Activation layer functions are ReLU.

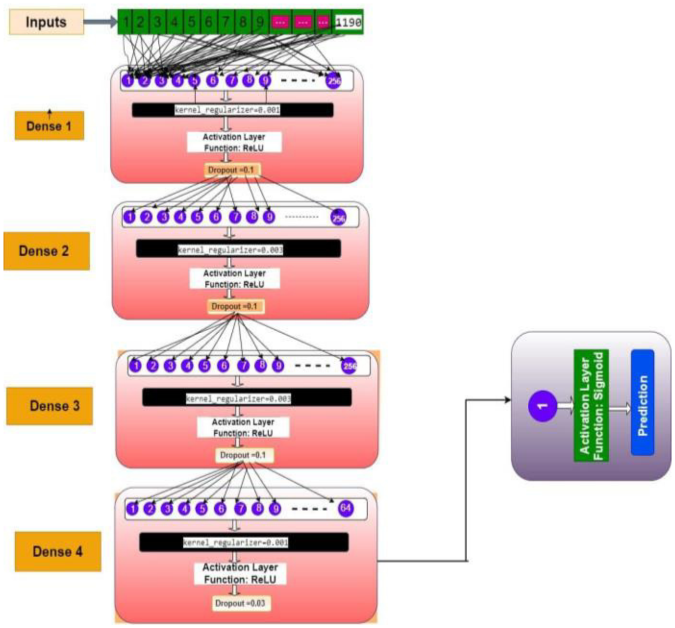

Fig. 3. CNN model architecture.

C. Artificial Neural Network (ANN) Architecture

ANNs are used to model non-linear problems and predict the output values from given input parameters during the training and testing process. ANN is a computing framework planned to recreate how the human brain analyzes and forms data. The input signal from the outside environment is received as a vector by the ANN. The notations x (n) for every n number of inputs are used to designate these inputs mathematically. The activation function then receives the sum of the weighted inputs. Dense layer have been used here with units to 256,128 and 1 respectively. Optimizer is Adam with a learning rate of 0.09. Kernel initializers are normal. The corresponding figure has been added in Fig. 4.

IV. Dataset And Preprocessing

A. Dataset Description

The collected Dataset is an aggregation of five distinct datasets (Switzerland, Cleveland, Hungary, and VA Long Beach [23] and Statlog [24]). Addressing missing values is critical [25]. However, our dataset has no missing values. We have 1190 instances in the whole dataset and 14 features based on which we predicted whether a person has heart disease or not. 13 features were selected as input features and the remaining "num" attribute was selected as an output class. All of them contain integer values. Class has 5 different classes (0 to 4).

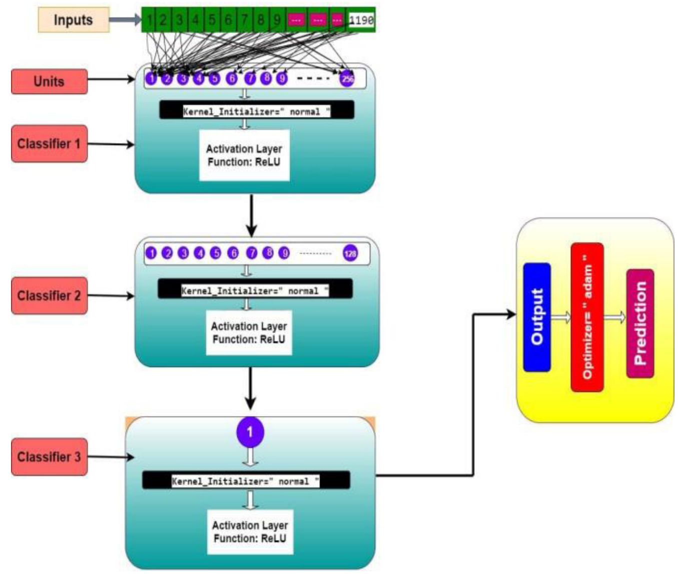

Fig. 4. ANN model architecture.

B. Preprocessing

Fig. 5 showcases the amount of data that is available for each 'num' attribute in the dataset. This research combines the data of attributes that are in the severity range of 1 to 4 into one class so that the neural network models can classify the data into two classes (0 (no disease) & 1 (disease)), as seen in Fig. 6.

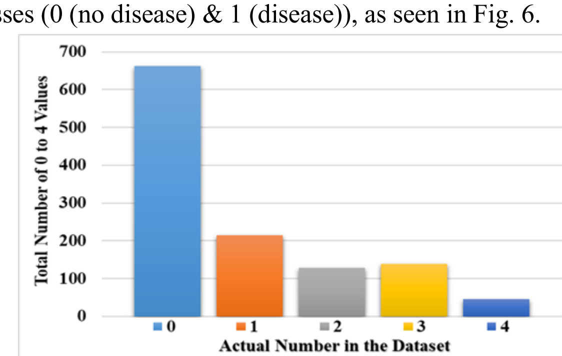

Fig. 5. Actual values vs. total values.

Fig 6. Showcases the converted 1 - 4 classes into 1 defining it to be classified as heart disease patients.

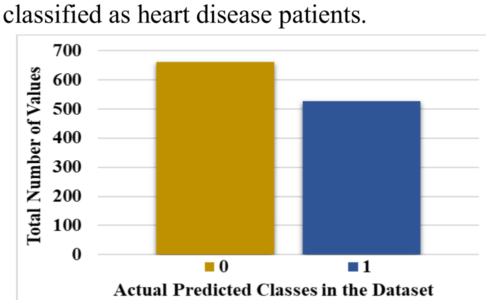

Fig. 6. Value conversion.

C. Justification of Our Proposed Method

Fig. 7 illustrates the correlation between the selected attributes of the dataset (14) in the form of a heatmap. The bar shows how strong or weak the correlation is between the dataset according to colour." ca" has the highest correlation with "thal" which is 0.85. As the red colour here defines a weak relationship, it is evident that "thalach" has the weakest relationship with "exang" which is -0.43. The white, yellow colour describes that it is not so strong or weak correlation between the data points.

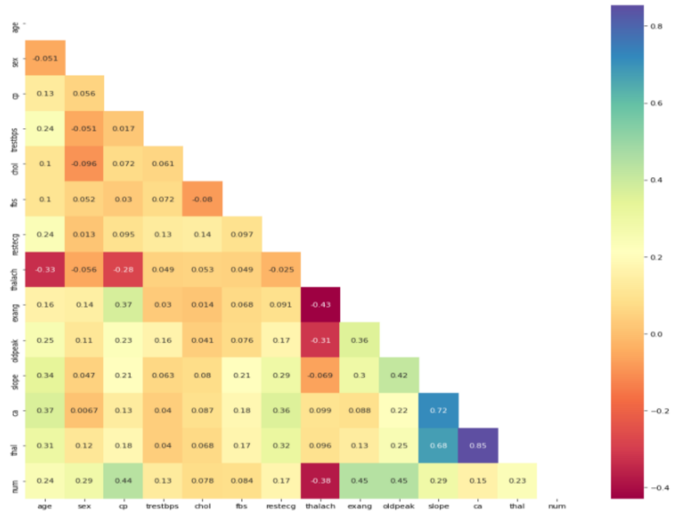

Fig. 7. Heatmap describing the correlation between data points.

V. RESULTS & DISCUSSION

The generated results are presented and discussed in Table 1.

TABLE I. USED PARAMETERS IN ALGORITHMS

Models	Activation	Optimizer	Learning rate	Epoch	Batch size
CNN	RELU	Adam	0.09	20	8
RBFN	-	-	-	20	8
ANN	RELU	Adam	0.09	20	8

The accuracy is expressed as a ratio between the total numbers of correctly classified examples to total examples. It can be seen from the Fig. 8 that CNN is has higher accuracy whereas, RBFN has the lowest. On the other hand, the highest specificity score was received by the RBFN algorithm as compared to others.

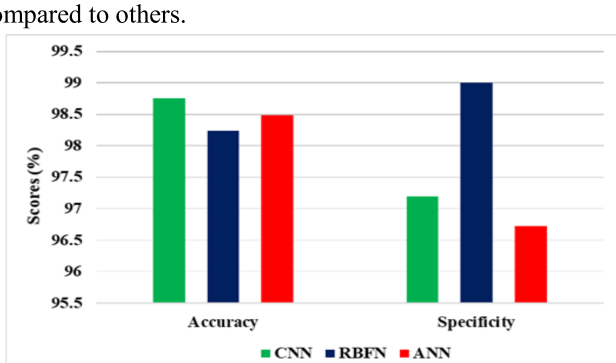

Fig. 8. Model accuracy.

The fraction of data that is incorrectly labelled as true is referred to as the false discovery rate. As seen in Fig. 8, compared to others CNN labelled more data incorrectly true, even though the accuracy is high. The fraction of real negative that is predicted as negative is described by negative predictive value. CNN model outperforms the other two by a good margin. The false negative rate, also known as the Type-II error, is the % age of results. ANN model has less false negative rate than CNN (6.10%) and RBFN (9.42%). A false-positive rate is the proportion of positive outcomes that the model incorrectly predicted. Again, RBFN has less false-positive rate than the other two, as seen in Fig. 10.

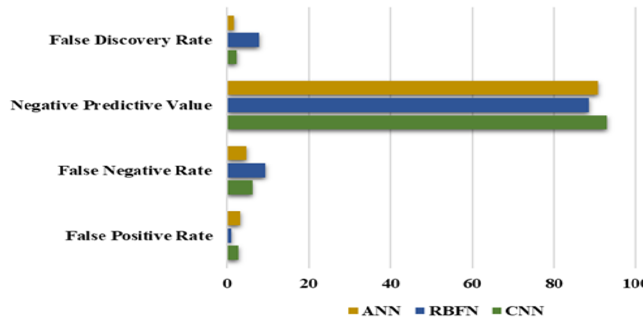

Fig. 9. FDR, NPV, FNR and FPR.

***FDR: False Discovery Rate**
***NPV: Negative Predictive Value**
***FNR: False Negative Rate**
***FPR: False Positive Rate**

Thus, generating parameters such as precision, recall, and F1-score is crucial. A model with high precision and high recall suggests that the model is returning accurate results. Whereas, a model with low precision and high recall has incorrect results. All the models have high precision and low recall, as seen in Fig. 10, which suggests that most of the predicted labels are true. Thus, it is important to calculate additional metrics such as sensitivity and specificity. Using these values recall, precision and F1-scores are calculated.

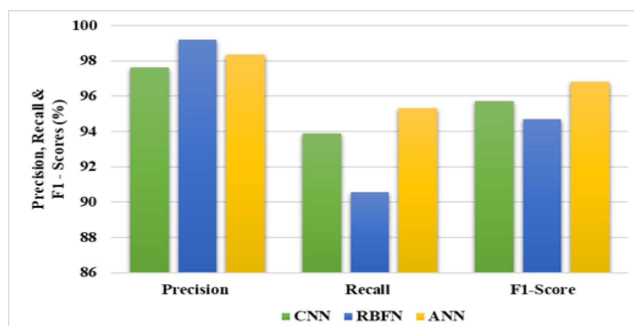

Fig. 10. Precision, Recall and F1-Score.

The total of error terms is more directly represented by Mean Absolute Error. The square root of Mean Square Error (MSE) is Root Mean Square Error (RMSE), which is an absolute measure of the quality for the fit. It is more often employed to avoid large mistakes than MSE. The R Squared method is a metric for how well a model fits its dependent variables. ANN has the highest mean squared error in Fig. 11. Showcasing that how well it fits to the model.

***MAE: Mean Absolute Error**
***MSE: Mean Squared Error**
***RMSE: Root Mean Squared Error**

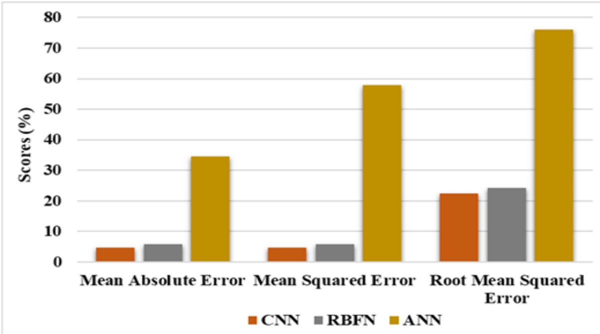

Fig. 11. MAE, MSE and RMSE.

A. *Learning Curve for the CNN Architecture*

Fig. 12 showcases the model accuracy where green depicts train and blue depicts validation. Train is always upward to almost 1 from more than 0.6. Validation is also upward.

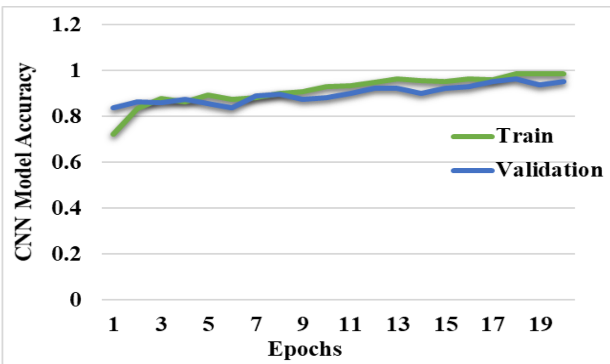

Fig. 12. Model accuracy for CNN.

Fig. 13 shows the Loss vs Epochs where training is in orange and validation is in yellow. Train is gradually decreasing from more than 0.3 to almost 0.1. Validation decreases to 0.2.

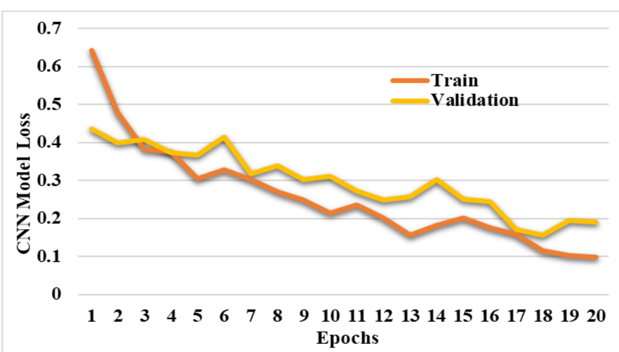

Fig. 13. Model loss for CNN

B. *Learning Curve for the RBFN Architecture*

Fig. 14 illustrates model accuracy where training is defined by orange and validation is defined yellow. Both the line nearly starts from less than 0.8 and ended around 1.

Fig. 15 showcases model loss. Loss vs. Epochs where training gradually decreases from around 0.6 to 0.1. Validation also decreases around 0.5 to in-between 0.2 and 0.1

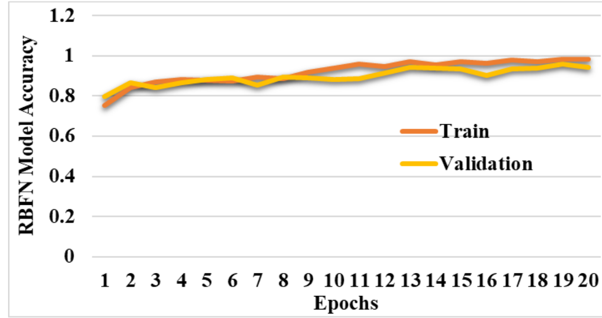

Fig. 14. Model accuracy for RBFN.

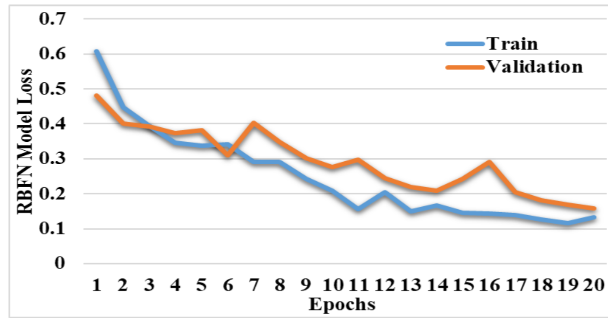

Fig. 15. Model loss for RBF.

C. *Learning Curve for the ANN Architecture*

Fig. 16 depicts ANN model accuracy where train and validation are blue and orange respectively. Train starts around 0.8 and finished almost 1. Where validation starts from little higher than 0.8 and finished in between 1 and 0.8

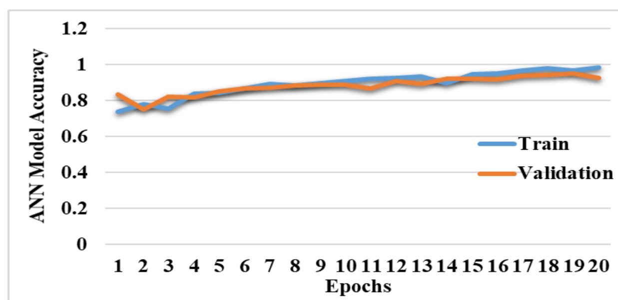

Fig. 16. Model accuracy for ANN.

Fig. 17 Showcases major ups and down between trainings which is blue and validation which is orange. Loss started from around 0.6 and finished in just down to 0.5. Validation ended and started at the same point.

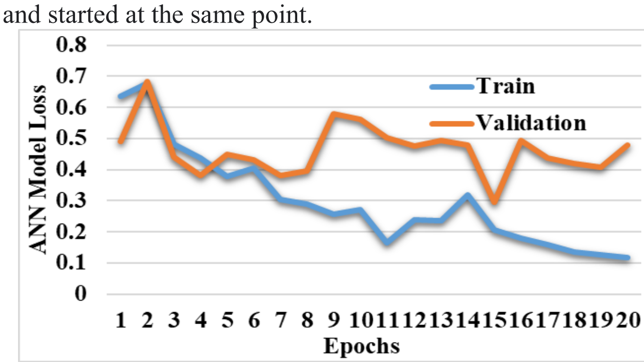

Fig. 17. Model loss for ANN.

D. *AUC Accuracy for the Used Algorithms*

Fig. 18 showcases the AUC scores where see that RBFN has the highest AUC score of 98% and CNN and ANN has the same AUC score of 95%.

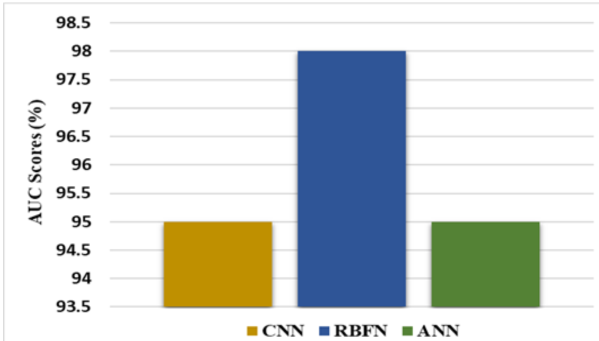

Fig. 18. AUC accuracy.

VI. CONCLUSION

Early diagnosis of heart disease is the primary aim of this research. Three Trained Neural Network models are created using CNN, RBFN, and ANN. All the models have high precision and low recall suggesting that most of the predicted labels are true. Of the three CNN performed better in classifying the clinical data into two classes, namely, no disease (0) and disease (1). CNN achieved 98.75% accuracy, whereas, RBFN and ANN achieved 98.24% and 98.49%, respectively. Adam optimizer is used to improve the performance of the models by helping with the overfitting issues. We intend to expand our research in the future by deploying more neural networks to predict and classify the disease.

REFERENCES

[1] Fisher, Edward. ”Coronary Artery Disease - coronary heart disease.” American Heart Association. 26 Apr, 2017

[2] ”FHS Research Policies.” Framingham Heart Study: A project of the national heart, blood, and lung institute and Boston University.

[3] Goff DC, Lloyd-Jones DM, Bennett G, Coady, et al. ”2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines.” National Institute of Health.

[4] World Health Organization. Global Status Report on Noncommunicable Diseases. Switzerland: WHO, 2014.

[5] Li, J.; Chen, Z.Z.; Huang, L.; Fang, M.; Li, B.; Fu, X.; Zhao, Q. Automatic classification of fetal heart rate based on convolutional neural network. IEEE Internet Things J. 2018, 6, 1394–1401.

[6] Golgooni, Z.; Mirsadeghi, S.; Baghshah, M.S.; Ataee, P.; Baharvand, H.; Pahlavan, S. Deep Learning-Based Proarrhythmia Analysis Using Field Potentials Recorded From Human Pluripotent Stem Cells Derived Cardiomyocytes. IEEE J. Transl. Eng. Health Med. 2019, 7, 1–9.

[7] Maknickas, V.; Maknickas, A. Atrial fibrillation classification using qrs complex features and lstm. In Proceedings of the 2017 Computing in Cardiology (CinC), Rennes, France, 24–27 September 2017; pp. 1–4.

[8] Grzegorczyk, I.; Soli ´nski, M.; Łepek, M.; Perka, A.; Rosi ´nski, J.; Rymko, J.; Gierałtowski, J. PCG classification using a neural network approach. In Proceedings of the 2016 Computing in Cardiology Conference (CinC), Canada, 11–14 September 2016; pp. 1129–1132.

[9] J. S. Sonawane and D. R. Patil, 2014, March. Prediction of heart disease using learning vector quantization algorithm. In 2014 Conference on IT in Business, Industry, and Government (CSIBIG) (pp. 1-5). IEEE.

[10] Feshki, M.G. and Shijani, O.S., 2016, April. Improving the heart disease diagnosis by evolutionary algorithm of PSO and Feed Forward Neural Network. In 2016 Artificial Intelligence and Robotics (IRANOPEN) (pp. 48-53). IEEE.

[11] Shen, Z., Clarke, M., Jones, R.W. and Alberti, T., 1993, October. Detecting the risk factors of coronary heart disease by use of neural networks. In Proceedings of the 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Societ (pp. 277-278).

[12] Hannan, S.A., Mane, A.V., Manza, R.R. and Ramteke, R.J., 2010, December. Prediction of heart disease medical prescription using radial basis function. In 2010 IEEE International Conference on Computational Intelligence and Computing Research (pp. 1-6). IEEE.

[13] “New initiative launched to tackle cardiovascular disease, the world's number one killer,” World Health Organization. [Online]. Available:http://www.who.int/cardiovascular_diseases/globalhearts/Global_hearts_initiative/en/.

[14] Ramprakash, P., Sarumathi, R., Mowriya, R. and Nithyavishnupriya, S., 2020, February. Heart Disease Prediction Using Deep Neural Network. In 2020 International Conference on Inventive Computation Technologies (ICICT) (pp. 666-670). IEEE.

[15] Awan, S.M.; Riaz, M.U.; Khan, A.G.: Prediction of heart disease using artificial neural network. VFAST Trans. Softw. Eng. 13, 102–112, 2018

[16] Alizadeh-dizaj, G.: Risk Prediction and Stratification of Patients with Stroke Using Data Mining Techniques. Tabriz University of Medical Sciences, School of Management and Medical Informatics (2018).

[17] Azimi, I., Janne T.M., Arman A., Amir M. R., Juha-Pekka, S and Pasi, L. “Empowering healthcare IoT systems with hierarchical edge-based deep learning.” In 2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), pp. 63-68. IEEE, 2018.

[18] Li, Liangzhi, Kaoru Ota, and Mianxiong Dong. “deep learning for smart industry: efficient manufacture inspection system with fog computing.” IEEE Transactions on Industrial Informatics 14, no. 10 (2018).

[19] P. Ghosh, A. Karim, S. T. Atik, S. Afrin, and M. Saifuzzaman, “Expert cancer model using supervised algorithms with a lasso selection approach,” International Journal of Electrical and Computer Engineering (IJECE), vol. 11, no. 3, p. 2631, 2021.

[20] P. Ghosh, F. M. Javed Mehedi Shamrat, S. Shultana, S. Afrin, A. A. Anjum and A. A. Khan, "Optimization of Prediction Method of Chronic Kidney Disease Using Machine Learning Algorithm," 15th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP), Bangkok, Thailand, 2020.

[21] A. Karim et. al., “A Comparative Study of Different Deep Learning Model for Recognition of Handwriting Digits”, ICICNIS 2020, January 2021, Available at SSRN: https://ssrn.com/abstract=3769231 or http://dx.doi.org/10.2139/ssrn.3769231

[22] P. Ghosh, S. Azam, K. M. Hasib, A. Karim, M. Jonkman, A. Anwar, “A Performance Based Study on Deep Learning Algorithms in the Effective Prediction of Breast Cancer,” International Joint Conference on Neural Networks (IJCNN 2021), 2021.

[23] “Heart disease datasets from UCI machine learning repository,” [online]. Available: https://archive.ics.uci.edu/ml/datasets/ Heart+ Disease, (Accessed on 31/11/2021).

[24] “Heart disease Statlog dataset of UCI machine learning repository,” [online]. Available: http://archive.ics.uci.edu/ml/datasets/ statlog+ (heart), (Accessed on 31/11/2021).

[25] P. Ghosh, S. Azam, A. Karim, M. Hassan, K. Roy, and M. Jonkman, “A comparative study of different machine learning tools in detecting diabetes,” Procedia Computer Science, vol. 192, pp. 467–477, 2021.