Grenze International Journal of Engineering and Technology, Jan Issue

Detection of Cardiovascular Diseases using Artificial Neural Networks

Vani E S[1], Ananya.S[2], Mrudula.V[3] and Abishree D.V[4]
[1-4]Nitte Meenakshi Institute of Technology, ISE Department, Bangalore, India
Email: vani.es@nmit.ac.in, ananya.s9701@gmail.com, mrudulavmithula@gmail.com, abhishreedv@gmail.com

***Abstract*— Abnormalities in normal blood flow from the heart cause several types of heart disease known as cardiovascular diseases. Over more than 17.5 million Global death occurs due to cardiovascular disease and 75% of death occurs from low- and middle-income countries. Early detection can save a lot of people's lives. ECG signals reflect the electrical activity of the heart. Thus, heart rhythm disorders or alterations in the ECG waveform are evidence of underlying cardiovascular problems, such as arrhythmias. However, manual analysis of electrocardiogram (ECG) signals is a laborious and prone-to-error task, even for a specialist with many hours of experience.**
This work focuses on identifying the five diseases of a dataset with the help of an artificial neural network. We have used a sequential model in keras with 5 dense layers. Then the ANN keras model is trained from the selected parameters of dataset and then test dataset is used. The classification report tells us how well the model has identified the diseases.

***Index Terms*— ECG signal, signal detection, artificial neural network, machine learning**

I. INTRODUCTION

The electrocardiogram (ECG) is the most widely used non-invasive technique in heart disease diagnoses. Since it reflects the electrical activity within the heart during contraction, the time it occurs, and its pattern provides much information about the state of the heart.
Each heartbeat constitutes a cardiac cycle, which is reflected in the ECG graph: a P-wave, a QRS complex and a T-wave. In addition, a small U-wave may be visible in certain ECGs. The line voltage ECG base is also known as the isoelectric line or baseline. Normally, the baseline is the portion of the stroke that follows the T-wave and precedes the next P-wave.
The paper is organized in the following format. The first part of the paper gives the need or motivation for the work to be done. It provides a brief introduction to the problems related to the heart. The Second part of the paper provides details regarding the working mechanism of the system proposed. In short, the methodology of the work is discussed here. The next part shows the implementation. The last part of the paper shows the results obtained and the thoughts on the results.

II. LITERATURE REVIEW

The dataset has 5 diseases which have been diagnosed with 95% accuracy. This review paper helped in understanding the analysis of ecg signals and how it helped in identifying the five diseases present in the dataset. [7]. We see how artificial neural networks improve the use of other existing models in the paper [1]. ANN models

Grenze ID: 01.GIJET.10.1.52

such as LSTM were used but the model didn't work properly due to abundance of numeric input in the data [2].
Neural networks are initialized randomly. So different neural networks have different results. Hence, to stabilize our artificial neural network, we have trained it with same parameters[5]. This work has 5 layers of Artificial neural network when compared to convolutional neural network of 12 layers which focuses on micro classes of heart diseases [6].
Seeing how commonly employed features such as morphological, statistical, and frequency domain features, as well as more advanced techniques such as wavelet transforms, time-frequency analysis, and nonlinear dynamics measures were chosen has helped in this paper.[8]
Understanding how different types of papers in artificial neural network work lets us know how to make this work better. NN-based classifier, The ResNet34-LSTM3 model and PNN used for classification of signals along with the ANN model were some of the types we came across in our literature survey.[9][10][11]

III. Methodology And Implementation

An ECG dataset is taken. The source is physionet MIT-BIH arrhythmia dataset . Preprocessing is done for noise removal. The ECG dataset obtained after preprocessing is split into test dataset and training dataset. Eighty percent is used for training and twenty percent is used for testing. Artificial neural network is used for detection of various heart diseases. Hardware required is a fully functional laptop and the software used is jupyter notebook.
ECG dataset will be pre-processed (noise removal). Then the model is trained using machine learning algorithm. 60% is used to train the model and 40% to test the model. Then it will give identified ECG diseases classification. The model uses classification algorithm which will identify all the 6 diseases. Fig.1 shows the framework of the system used.
The model has 5 layers:
The first layer is a dense layer with 32 units and uses the ReLU activation function. It has an input shape of (20,), which corresponds to the 20 features selected by the ANOVA F-test.
The second layer is a dense layer with 64 units and uses the ReLU activation function.
The third layer is a dense layer with 128 units and uses the ReLU activation function.
The fourth layer is a dense layer with 64 units and uses the ReLU activation function.
The fifth and final layer is a dense layer with 5 units, which corresponds to the 5 classes of ECG signals. It uses the softmax activation function to produce the class probabilities.

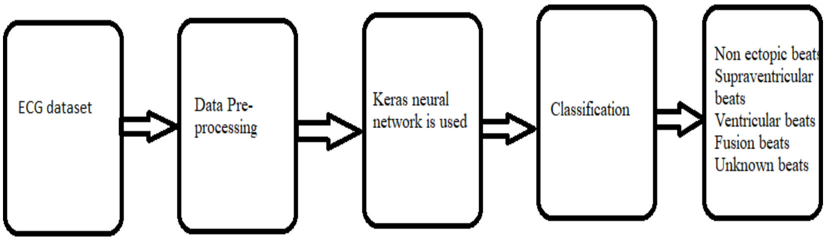

Fig 1: System Framework

Our paper focuses on detecting the five types of heart diseases based on the ecg signals from the dataset. Non ectopic beats, Supraventricular ectopic beats, Ventricular ectopic beats, Fusion beats and Unknown beats.
Non ectopic beats mean normal heartbeats. The rhythmic contraction and relaxation of heart muscles are known as heartbeat. Your pulse rate, also known as your heart rate, is the number of times your heart beats per minute. A normal resting heart rate should be between 60 to 100 beats per minute.
Generally, a lower heart rate at rest implies more efficient heart function and better cardiovascular fitness. Although there's a wide range of normal, an unusually high or low heart rate may indicate an underlying problem. Fig. 2 shows a normal heartbeat.
Supraventricular Ectopic Beats indicates atrial irritability. Isolated Supraventricular Ectopic Beats are generally not significant in nature, but a high frequency can represent more risk. An increasing trend in Supraventricular Ectopic Beats may be an indicator or sign for atrial fibrillation. Atrial Fibrillation is considered to be significant as it can lead to heart attack or stroke. Fig 3 shows supraventricular ectopic beats.
Ventricular ectopic are a type of arrhythmia or abnormal heart rhythm. It is caused by the electric signals in the heart starting in a different place and travelling a different way through the heart. The ECG features of these beats include Broad QRS complex ($\geq$ 120 ms) with abnormal morphology and Discordant ST segment and T wave changes. Fig 4 shows ventricular ectopic heartbeat.

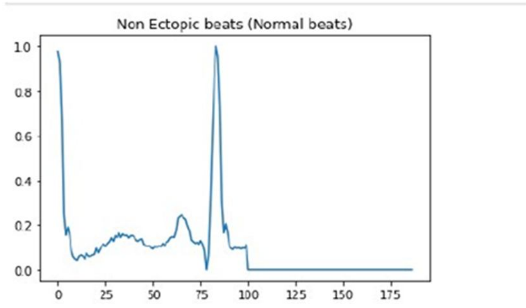

Fig 2: Non-Ectopic beats

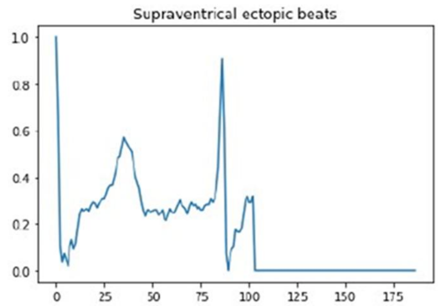

Fig 3: Supraventricular Ectopic beats

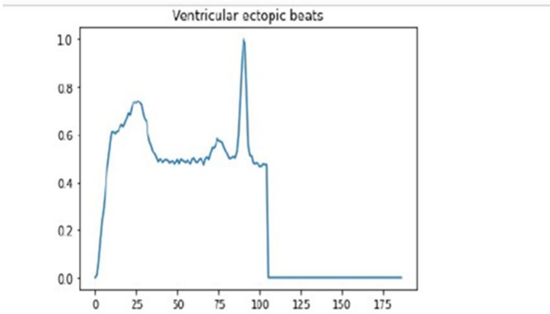

Fig 4: Ventricular Ectopic beats

A fusion beat occurs when electrical impulses from different sources act upon the same region of the heart at the same time~~.~~. If it acts upon the ventricular chambers, it is called a ventricular fusion beat, whereas colliding currents in the atrial chambers produce atrial fusion beats. The fusion beats are of intermediate width and morphology to the supraventricular and ventricular complexes. Fusion beats are seen with Ventricular Tachycardia and Accelerated idio ventricular rhythm (AIVR). Figure 5 shows fusion beats.

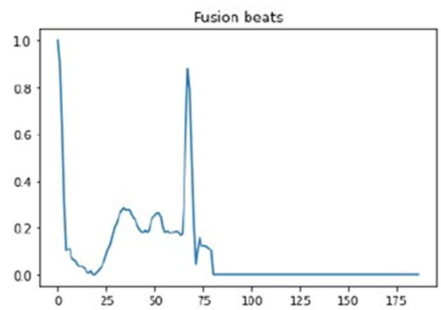

Fig 5: Fusion beats

Unknown beats are seen in myocardial infarction. Myocardial infarction (MI), colloquially known as "heart attack," is caused by decreased or complete cessation of blood flow to a portion of the myocardium. Myocardial infarction may be “silent," and go undetected, or it could be a catastrophic event leading to hemodynamic deterioration and sudden death. Hyper acute T wave changes are seen with increased T wave amplitude and width; ST elevation in ECG signals can be seen. Figure 6 shows unknown beats.

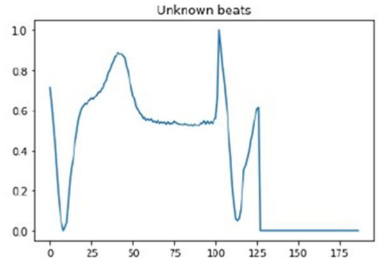

Fig 6: Unknown beats

As the type of diseases are discussed as above, the signals are plotted. Then, label encoding is done to handle the categorical variables in the data. Finally, our model is initialized. Our code creates a sequential neural network model. The model has 5 layers including the input layer. The first four layers are dense layers with ReLU activation function. The last layer has 5 neurons and a softmax activation function to output a probability distribution over 5 classes. The input shape for the first layer is (187,), meaning that the input has 187 features.
Our code compiles the previously defined neural network model in Keras. The model is compiled with Adam optimizer. The optimizer is the algorithm used to update the model's weights during training, "Adam" is a popular choice for its efficiency The loss function is used to evaluate the model's performance during training. Categorical_crossentropy is a common loss function for multi-class classification problems. The evaluation metric, "accuracy", is used to monitor the model's performance during training and testing.

IV. RESULTS

We see here that the model does as well as it did on the validation set, this indicates that the model has learned to generalize. We have a total accuracy of 95% on the test data. Also, we have to pay attention to precision and recall. We don't have the numbers we had with the validation set but still it does very well. This classification report shows several metrics for each class, including precision, recall, f1-score, and support. Precision is the ratio of true positives to the total number of predicted positives. In other words, it measures how many of the predicted positive instances are positive. For example, the precision for "non-ectopic beats" is 0.96, which means that 96% of the instances predicted as "non-ectopic beats" are actually "non-ectopic beats”. Recall is the ratio of true positives to the total number of actual positives. In other words, it measures how many of the actual positive instances are correctly identified by the model. For example, the recall for "Supraventricular ectopic beats" is 0.54, which means that only 54% of the actual "Supraventricular ectopic beats" were correctly identified by the model. F1-score is the harmonic mean of precision and recall. It is a single metric that combines both precision and recall into one value. For example, the f1-score for "non-ectopic beats" is 0.98, which is a high value and indicates good performance of the model in identifying "non-ectopic beats". Fig. 7 shows the Confusion matrix.
From the confusion matrix we can identify the diseases that, 81.550% of “non-ectopic beats”, 1.378% of “Supraventricular Ectopic beats”, 5.363% of “Ventricular Ectopic beats”, 0.402% of "Fusion beats" and 6.792% of “Unknown beats” are present in the given dataset.
In summary, the model performs very well in identifying "non-ectopic beats" and "Unknown beats", but not as well in identifying "Supraventricular ectopic beats" and "Fusion beats".

V. CONCLUSIONS

The prediction of heart disease at an early stage can prevent many mishaps. The use of an efficient algorithm can help physicians in detecting the possible presence of heart disease before it manifests. The overall accuracy of the model is 95% after 20 epochs. The model accuracy doesn’t change after 20 epochs hence the model accuracy is 95%.

Two graphs are plotted to track the model's performance. Loss graph and accuracy graph are the graphs plotted. The loss graph measures how bad the model performs. Therefore, the lower the loss, the better the model performance is. Fig. 8 shows the loss is less, proving that our model has performed well.

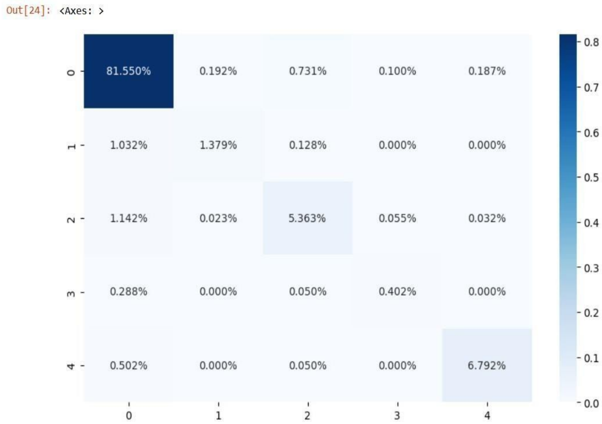

Fig 7: Confusion Matrix

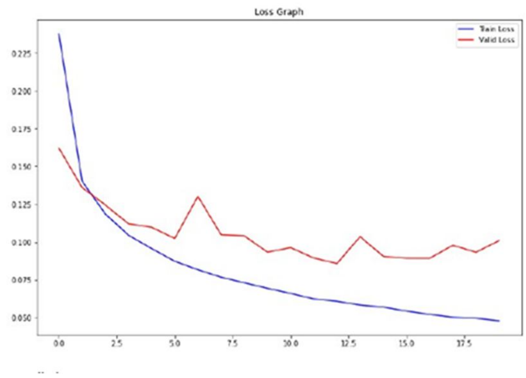

Fig 8. Loss graph

Accuracy graph has both training and validation accuracy of the model. Fig. 9 shows the accuracy graph.

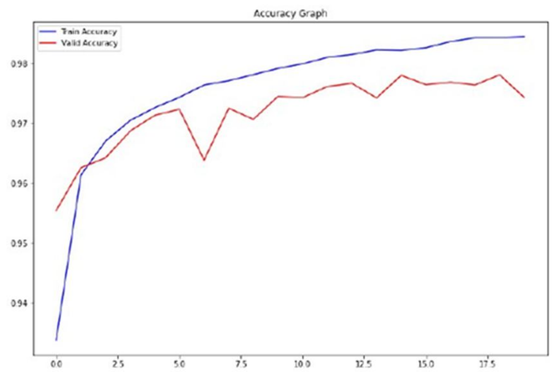

Fig. 9: Accuracy graph

The gap between the training and validation accuracy indicates the amount of overfitting. The gap between train accuracy and valid accuracy here is little indicating less over fitting. Less over fitting is recommended as it means the model can generalize the new data well.

The code at the end selects 10 samples from the test set and plots the corresponding ECG signal along with its predicted and actual labels. We can see the comparison of actual versus predicted signal graphs here. The below sample is an example. Fig 8 is one random sample. The actual class is supraventricular ectopic beats which the model has predicted correctly. The random index is the index of the sample which has been selected randomly.

REFERENCES

[1] Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring Nashif, Md.R., Islam, Md.R. and Imam, M.H. (2018) Heart Disease Detection by Using System. World Journal of Engineering and Technology, 6, 854-873. S., Raihan,

[2] Abdullah Alqahtani ,1 Shtwai Alsubai ,1 Mohemmed Sha , Lucia Vilcekova, Cardiovascular Disease Detection using Ensemble Learning, August 2022

[3] Awais Mehmood1 · Munwar Iqbal1 · Zahid Mehmood2 · Aun Irtaza1 · Marriam Nawaz1 · Tahira Nazir1 ·Momina Masood1, Prediction of Heart Disease Using Deep Convolutional Neural Networks, November 2020

[4] Sanchez, FA Rivera, and JA Gonzalez Cervera. "ECG classification using artificial neural networks." In Journal of Physics: Conference Series, vol.: 1221, no. 1,p.012062. IOP Publishing, 2019.

[5] Ribeiro, Antônio H., Manoel Horta Ribeiro, Gabriela MM Paixão, Derick M. Oliveira, Paulo R. Gomes, Jéssica A. Canazart, Milton PS Ferreira et al. "Automatic diagnosis of the 12-lead ECG using a deep neural network." Nature communications 11, no. 1 (2020): 1760.

[6] Wu M, Lu Y, Yang W and Wong SY (2021) A Study on Arrhythmia via ECG ,Signal Classification Using theConvolutional Neural Network.,Front. Comput. Neurosci. 14:564015.doi: 10.3389/fncom.2020.564015Y.

[7] JOSHI, ANAND & TOMAR, ARUN & TOMAR, MANGESH. (2014). A Review Paper on Analysis of Electrocardiograph (ECG) Signal for the Detection of Arrhythmia Abnormalities. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering. 03. 12466-12475. 10.15662/ijareeie.2014.0310028.

[8] Houssein, Essam & Kilany, Moataz & Hassanien, Aboul Ella. (2017). ECG signals classification: a review. International Journal of Medical Engineering and Informatics. 5. 376-396. 10.1504/IJIEI.2017.10008807.

[9] El-Khafif, Sahar H., and Mohamed A. El-Brawany. "Artificial neural network-based automated ECG signal classifier." International Scholarly Research Notices 2013 (2013).

[10] Xie, Qiyang, Xingrui Wang, Hongyu Sun, Yongtao Zhang, and Xiang Lu. "ECG signal detection and classification of heart rhythm diseases based on ResNet and LSTM." Advances in Mathematical Physics 2021 (2021): 1-1

[11] Kshirsagar, PRAVIN R., SUDHIR G. Akojwar, and R. A. M. K. U. M. A. R. Dhanoriya. "Classification of ECG-signals using artificial neural networks." In Proceedings of International Conference on Intelligent Technologies and Engineering Systems, Lecture Notes in Electrical Engineering, vol. 345. 2017.