World Journal of Innovation and Modern Technology E-ISSN 2756-5491 P-ISSN 2682-5910
Vol 8. No. 6 2024 www.iiardjournals.org Online Version

# Machine Learning Technique for Determination of Normal and Abnormal ECG Heartbeat

**Idayana Alabere**
Department of Computer Science and Informatics, and Cyber Security
Federal University Otuoke, Bayelsa State
alabereii@fuotuoke.edu.ng
DOI: 10.56201/wjimt.v8.no6.2024.pg106.115

***Abstract***

*The vital role played by the heart for the sustenance of humanity is such that the heart be given proper attention and care besides the ribs. The alarming rate of heart defects and increasing death toll calls for the urgent need to improve the ECG platform with enhance machine learning technique the more accurate predictive diagnosis of heart conditions like arrhythmia, stroke and other cardio-vascular disorders. This paper proposed a machine learning approach using ANN and objective methodology to analyse and predict heartbeat condition using heart dataset of 303 patients from kaggle dataset repository and the accuracy level was 88.52 with a validation loss of 34.31 respectively which indicates that with further work with more advance machine learning algorithm the predictive capacity of the model could be optimized.*

## 1.1 Introduction

The heart is one of the most important organ of the human race little wonder it is guarded within the enclosure of the rib cage which serves as a protective covering from external attacks. Be it as it may, the heart the engine room of the humanity. It serves dual purpose of pumping oxygenated blood through the human system and carrying deoxygenated blood from at the body. The automation of the heart is one of the wonders of nature and creation which must be taken care of with utmost care. Once the heart stops beating life on earth comes to an end and the flow of oxygen around the body stops. The heart is a kind of muscular organ which pumps blood into the body and is the central part of the body's cardiovascular system which also contains lungs. Cardiovascular system also comprises a network of blood vessels, for example, veins, arteries, and capillaries (Nashif et al., 2018). These blood vessels deliver blood all over the body. Abnormalities in normal blood flow from the heart cause several types of heart diseases which are commonly known as cardiovascular diseases (CVD). Heart diseases are the main reasons for death worldwide. World Health Organization (WHO) 2017 report indicated that an estimated 17.5 million deaths occur globally per annum due to heart attacks, strokes and other heart related defects. More than 75% of deaths from cardiovascular diseases occur mostly in middle-income and low-income countries. Also, 80% of the deaths that occur due to CVDs are because of stroke and heart attack (Hazra et al., 2017). Similarly, the latest WHO data published in 2017 Coronary Heart Disease Deaths in Nigeria reached 76,410 or 3.76% of total deaths. The age adjusted Death Rate is 117.12 per 100,000 of population ranks Nigeria number 90 in the world (WHR, 2019).

Therefore, detection of cardiac abnormalities at the early stage and tools for the prediction of heart diseases can save a lot of life and help doctors to design an effective treatment plan which ultimately reduces the mortality rate due to cardiovascular diseases. Due to the development of advance healthcare systems, lots of patient data are nowadays available (i.e. Big Data in Electronic Health Record System) which can be used for designing predictive models for Cardiovascular diseases. Data mining or machine learning is a discovery method for analyzing big data from an assorted perspective and encapsulating it into useful information (Patel et al., 2016)

Electrocardiographic signal processing has a long and rich history that has greatly enhanced the diagnostic capability, mostly when signals are collected in noisy environments. An electrocardiogram is a tool that is used to graphically illustrate the heart's electrical conduction. Clinicians can detect a multitude of processes of cardiac disease by observing variations from normal on the ECG. Electrocardiography (ECG) is the processing of the heart's electrical activity recorded by an external electrode connected to the skin over time. Each of the cell membranes that form the heart cell's outer cover has an associated load that is depolarized during each heartbeat. They manifest as small electrical signals on the skin that the ECG can sense and intensify. ECG is a bio-medical signal that tracks the time-related electrical activity of the heart.

### 1.2 Statement of the problem

From the overwhelming statistics available on death heart related defects in Nigeria, there is need for a more efficient way to analyse arrhythmia to forestall further occurrence.

### 1.3 Aim and Objectives

The aim of this paper is to design a machine learning model to analyse and detect the normal and abnormal heart rate with respect to diseases like arrhythmia. The objectives include:

i. Design an artificial neural network (ANN) model for ECG analysis and classification
ii. Simulate the heartbeat dataset with python and tensorflow to detect signal patterns of normal and abnormal heart rates
iii. Ascertain how many records indicate heart disease

## 2.0 Literature review

### 2.1 Machine Learning

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. It uses the underpinning principles of statistics and computer science to create statistical models which are generally used to make prediction and inference based on available data (Dhandhania, 2018). Machine Learning models are not programmed but trained. They are presented with relevant examples and it generate a statistical structure in them which enable the system to come up with rules for automating the given task (Chollet, 2017). One example of training machine learning is the task of labelling or tagging datasets. The system would learn rule to enable it associate a data be it image or signal to a specific tag based on the already tag feature by humans.

## 2.2 Artificial Neural Network (ANN)

A machine learning paradigm designed in the similitude of the human brain is called an artificial nueral network. Artificial neural network are information processing systems that has charactestics in common with the biological neural network sysems (Yadav et al, 2015). it is the mainstay of deep learning formed by the interconnection of weighted neurons (Buduma, 2017). Saxena (2016), views artificial neural netwok (ANN) as a computational model based on the structure of the brain. Individual Neuron recieves series of input via its connections and generates a single output which is associated to the state of the neuron and its activation function. Consequently, this output may get to several other neurons in the network. The inputs are the outputs thus activations of the incoming neurons multiplied by the connection weights or synaptic weights which the strength or amplitude of connection between two nodes, or in biology the amount of influence the firing of one neuron has on another. The input of a network is associated to its weight, in other words the weight of a network is equal to the input it received, therefore to compute the activation of the neuron, the activation or threshold function is applied to the weighted sum of inputs in addition to the bias (Yadav et al, 2015).

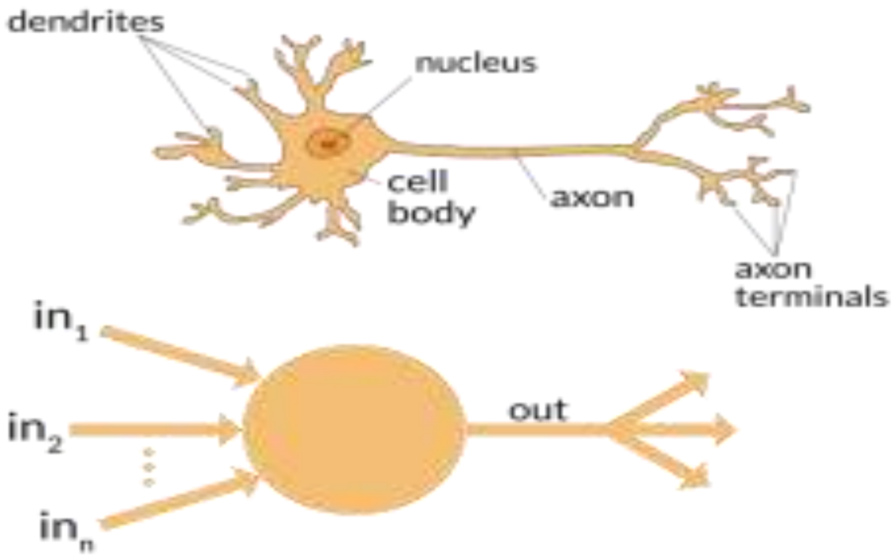


Figure 1: Brain Neuron and Neural Net. (Source: Buduma, 2017)

## 2.3 Heart Rate Variability (HRV) and Arrhythmia

Heart rate variability (HRV) is the temporal difference between sequences of consecutive heartbeat. On a standard electrocardiogram (ECG), the maximum upwards deflection of a normal QRS complex is at the peak of the R wave, and the duration between two adjacent R wave peaks is termed the R-R interval. The ECG signal requires editing before HRV analysis can be performed, a process requiring the removal of all non-sinus-node-originating beats. The resulting period between adjacent QRS complexes resulting from sinus node depolarizations is termed the N-N (normal-normal) interval (Eur, 1996). HRV is the measurement of the variability of the N-N intervals (Reed, Robertson and Addison, 2005). Although counterintuitive, it is possible that HRV confers a survival advantage. Any system exhibiting intrinsic variability is primed to respond rapidly and appropriately to demands placed upon it.

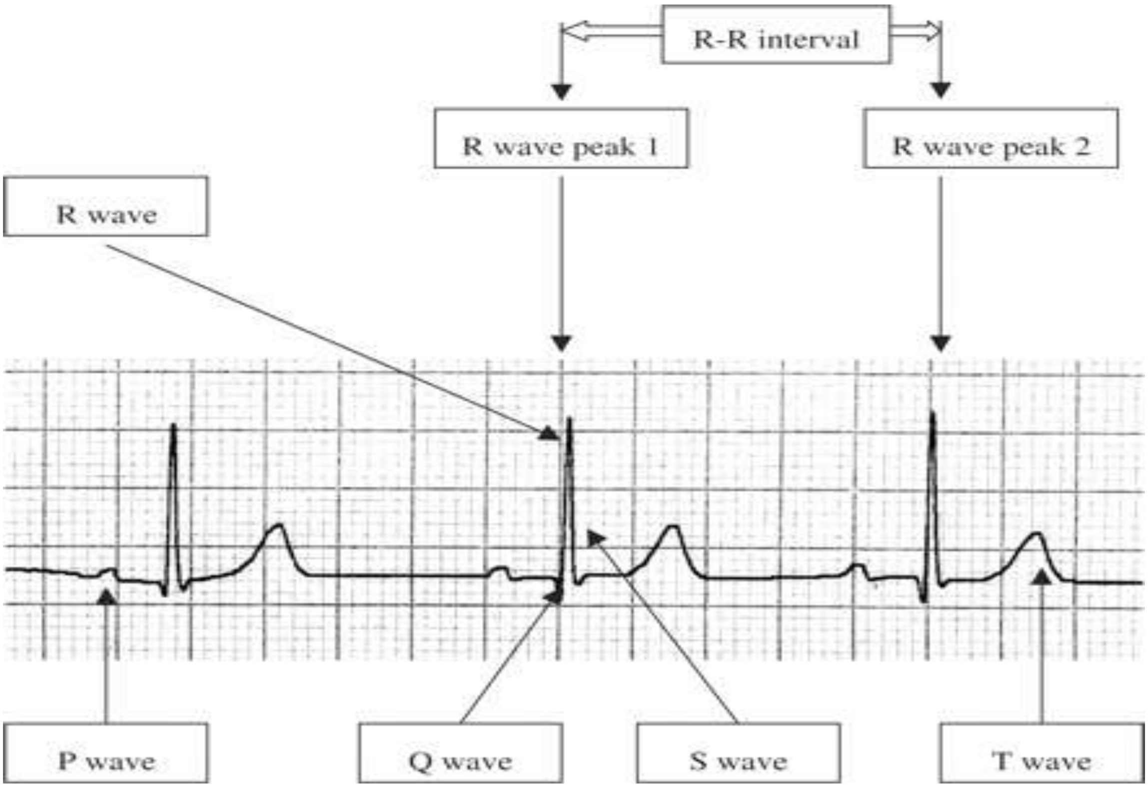


Figure 2 Heart Rate Variation (Source: Reed et al., 2005)

Reed et al., (2005), further explains that the HRV measures the balance between sympathetic mediators of heart rate (HR) which is the effect of epinephrine and norepinephrine, released from sympathetic nerve fibres, acting on the sino-atrial and atrio-ventricular nodes that result to an increase in the rate of cardiac contraction and enables conduction at the atrio-ventricular node, and parasympathetic mediators of HR which is the influence of acetylcholine, released by the parasympathetic nerve fibres, acting on the sino-atrial and atrio-ventricular nodes that leads to a decrease in the heart rate(HR) and a slowing of conduction at the atrio-ventricular node. Sympathetic mediators appear to exert their influence over longer time periods and are reflected in the low frequency power (LFP) of the HRV spectrum (between 0.04 Hz and 0.15 Hz.2,3 Vagal mediators exert their influence more quickly on the heart, and principally affect the high frequency power (HFP) of the HRV spectrum (between 0.15 Hz and 0.4 Hz).4 Thus, at any point in time, the LFP and HFP ratio is a proxy for the sympatho-vagal balance.

On the other hand Arrhythmia depicts a pattern of variability and instability in heartbeats meaning the heart beats too fast or too slow and this condition portends grave danger for the individuals suffering from it. Heart rhythm, the pattern of cardiac contraction followed by relaxation, represents the myocardial response to electrical activation of specialized cells and fibres within the atria and ventricles. Heart rhythm should be regular, with a rate of between 60-99 beats per minute at rest, but under normal conditions even a healthy heart displays slight beat-to-beat variability, as reflected in the R-R interval. (Consider sinus arrhythmia, where the R-R interval shortens during inspiration and lengthens during exhalation.) Thus, a healthy heart rhythm is not strictly regular, but varies slightly as a result of numerous factors, especially vagus nerve activity. Measuring heart rate variability (HRV) may provide useful clinical information about autonomic tone and heart function. In addition, reductions in HRV have been associated with a wide range of disorders

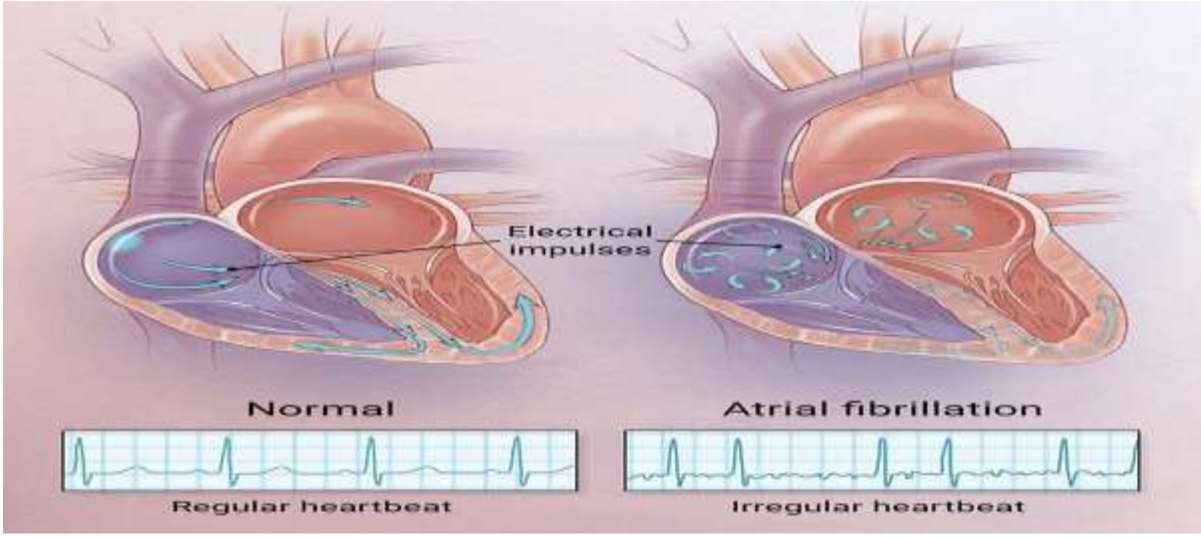


Figure 3: Regular and Irregular heartbeat (Source: Yancheng, 2016)

**Related works**

Dubey and Richariya (2013), in their paper "A Neural Network Approach for ECG Classification" asserts that ECG been an important diagnostic tool for assessing heart functions, the interpretation of ECG signal is pattern recognition function with signal pre-processing, QRS detection, feature extraction and ANN is best suitable for algorithm for classifying the feature extracted for the patterns. Different ECG feature inputs were used in the experiments to compare and find a desirable features input for ECG classification. Among different structures, it was found that a three layer network structure with 25 inputs, 5 neurons in the output layer and 5 neurons in its hidden layers possessed the best performance with highest recognition rate of 91.8% for five cardiac conditions.

**Methodology**

The methodology here deployed is the objective methodology with diagnostic information based on delineation. One advantage of the objective methodology is that is required lesser human interaction which makes is most appropriate for use in a machine learning model for the analysis and classification of heat beat rhythm. The programming tool used for the analysis is Python. The system was deployed online with Google Colab Python, Tensorflow and Jupyter notebook environment which takes advantage of the online GPU to simulate the heat beat datasets to ascertain the arrhythmia and other heart defects. The dataset was sourced from Keggle. The links are below:

https://archive.ics.uci.edu/ml/datasets/Heart+Disease

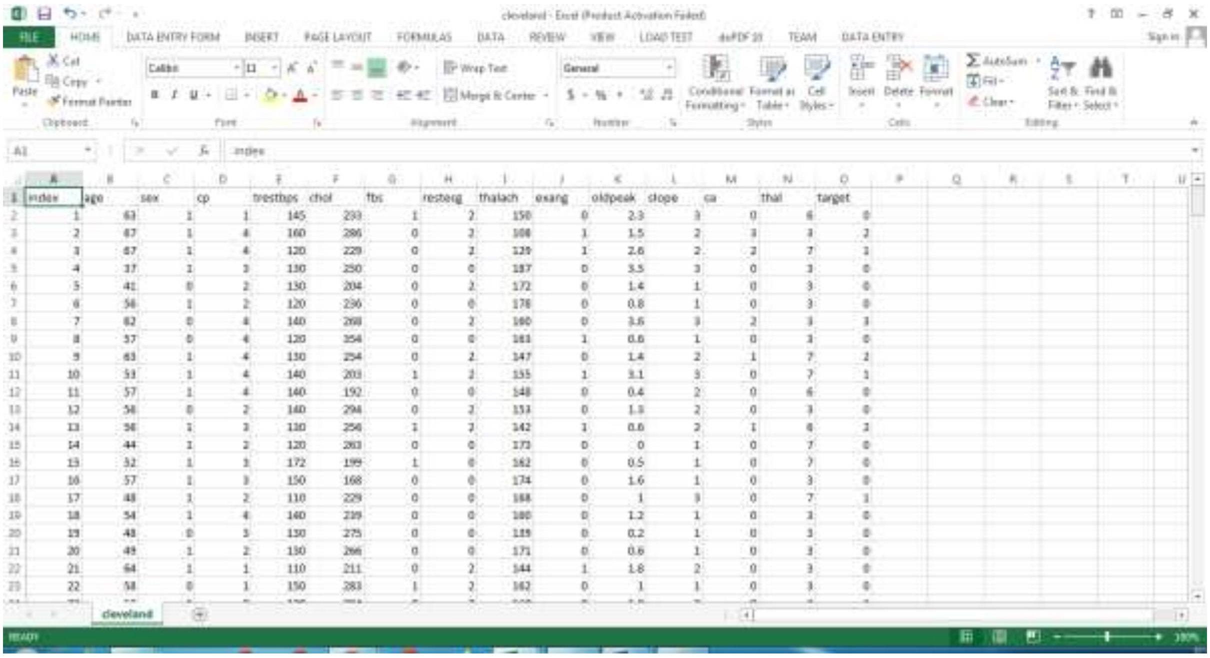

Figure 4: Heart Dataset (Source: https://archive.ics.uci.edu/ml/datasets/Heart+Disease)

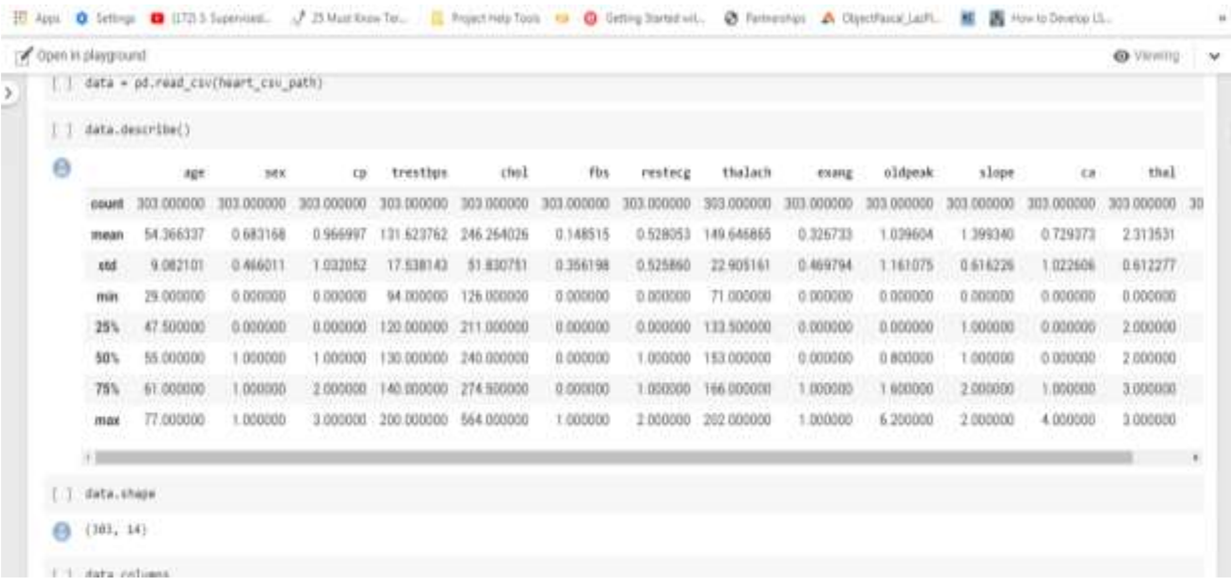

Figure 5: Pre-processed Heart Dataset

## Results and discussion

The dataset sourced from kaggle containing 303 patient records. Each record contains attributes which includes the age, sex, cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak slope, ca, thal and target respectively. These variables where used to analyse the ECG signals and predict the following results. The ANN model developed was deployed and trained using supervised learning algorithm which entails the using labelled dataset in training the model. The training dataset was 80% and the test dataset 20% respectively, the training accuracy of the model after 100 epochs was 85.67 and a validation accuracy of 88.52 respectively. The training and loss is 28.28 and 34.31 percent making the prediction accuracy high. Figure 6 and 7 show the model training and prediction as the model was able to classify and predict the rate of disease with the dataset.

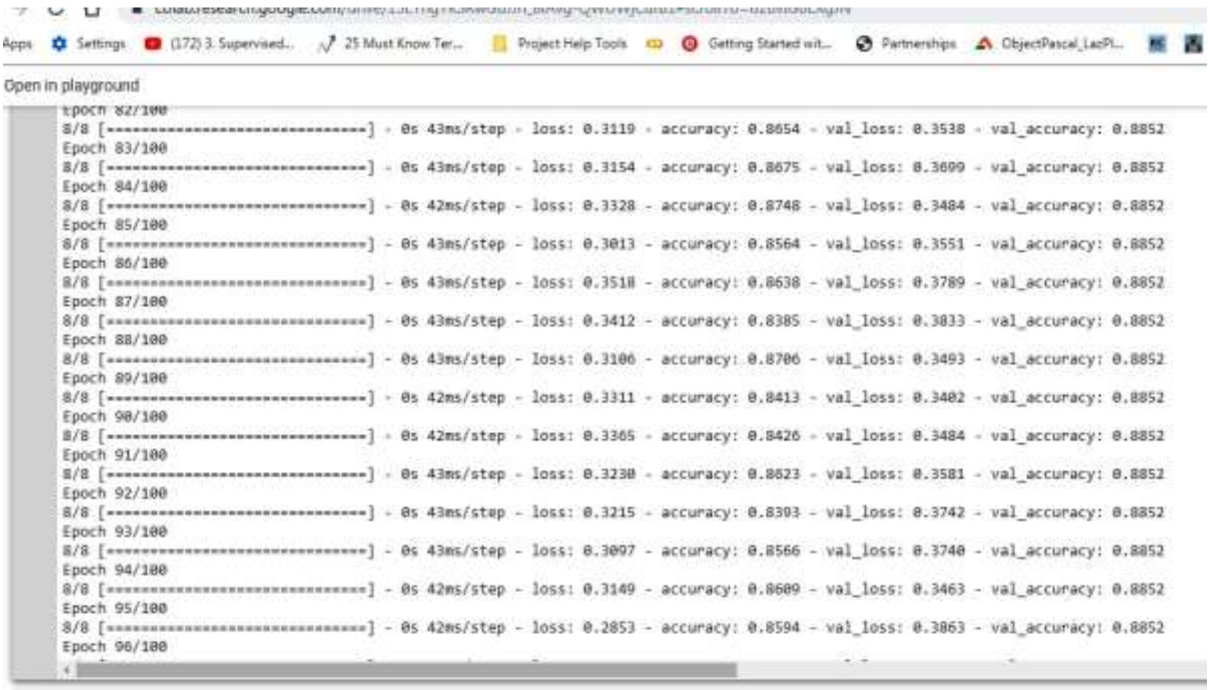


Figure 6: Model training iterations

The training process in figure 6 shows a training accuracy of 85.94% and a validation accuracy of 88.52% with a training loss of 28.53 and validation loss of 38.63% respectively. This implies that with more training the predictive potential of the model will be optimized.

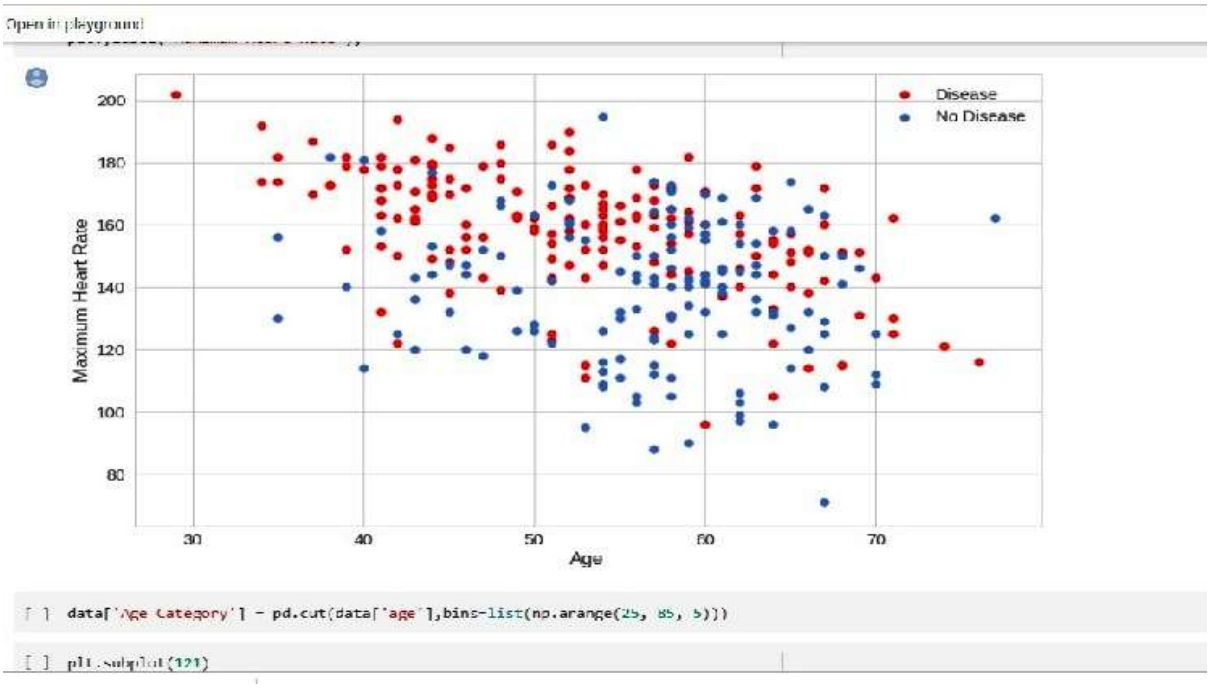


Figure 7. Heart Showing Diseased and No Disease

In figure 7, the model prediction shows the rate and possibility of heart disease as it related with the age of the patient. The red spot indicate the possible presence of one form of heart defect or another while the blue spot show no heart defect. The model indicates the increase and possibility of patient to have any form of heart defect from the age of 40 and above as the red spot which indicate the presence of heart disease appears more from age 40 and above. This also indicated that the older the patient the more they are prone to heart defect therefore the need for regular medical checks and the adoption of a healthy lifestyle at this age bracket.

Figure 8 and 9 shows the accuracy and the loss of the model. The model accuracy determines the model prediction. The prediction accuracy of the model is 88.52% with a loss of 38.63% which can be optimized with more data and training as machine learning model's optimization is enhanced with big data and more training epochs.

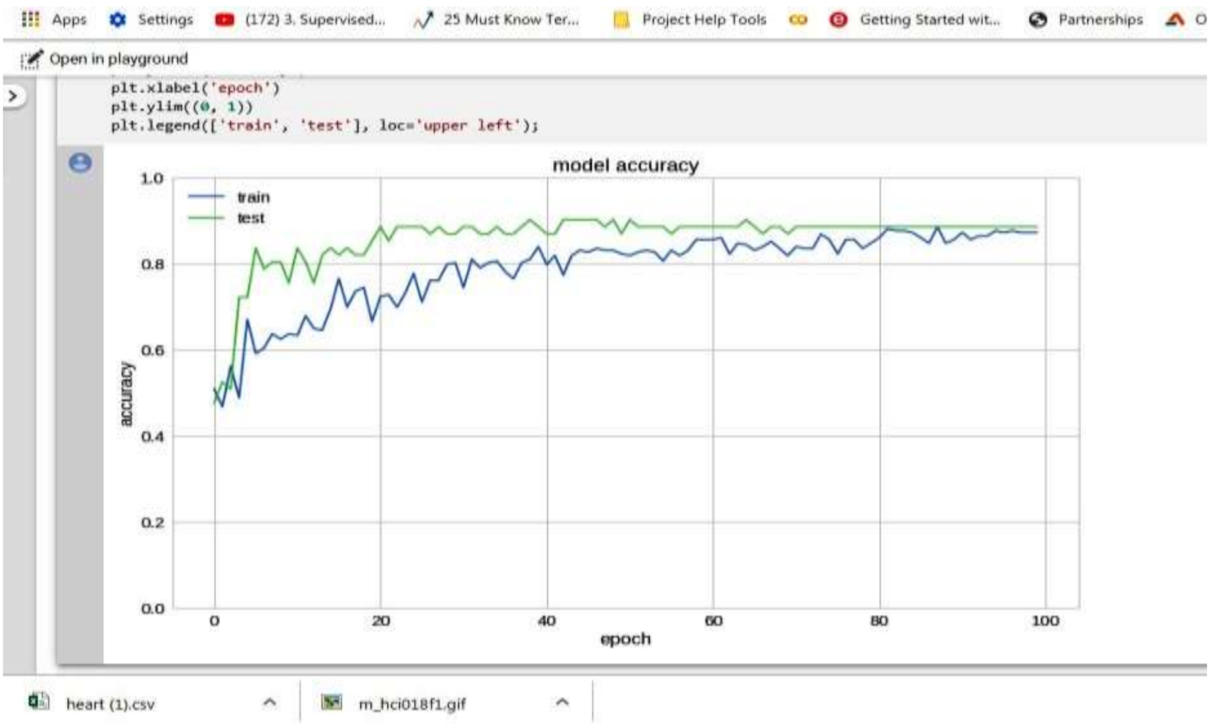


Figure 8: Model Training Accuracy

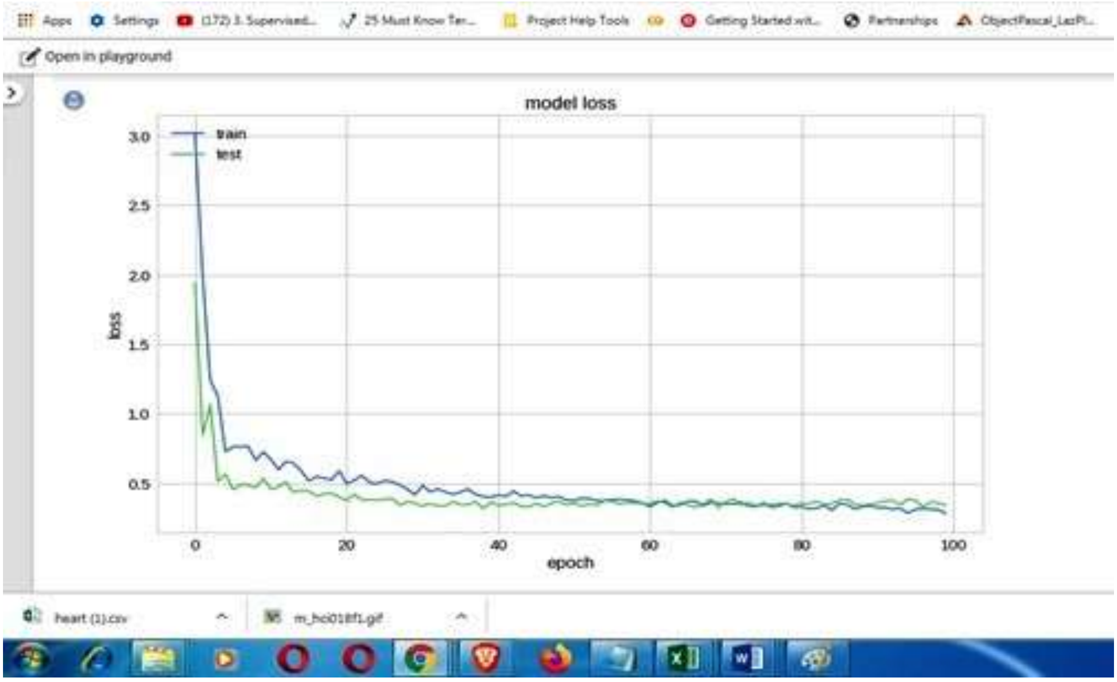


Figure 9: Model Loss

## Conclusion

Using machine learning algorithm for ECG analysis and prediction is more efficient than other methods as the accuracy level in respect to the above experiment indicated 88.52 accuracy level which implies that error margin is low. The place of machine learning in predictive analysis cannot be overemphasised especially in bioinformatics and medical science research as it has proven to be a veritable tool that over time has enhance medical diagnostic services. ECG is not left out in this revolution as the heart continues to beat and generate life sustaining signals so the need for more research to ensure longevity and wellbeing of the individual and the society.

## REFERENCES

Buduma, N. (2017). *Deep Learning in a Nutshell*. Retrieved October 15, 2019, from Common Lounge: https://www.commonlounge.com/discussion/20b7692fa77b418e9956e963d29646ad

Chollet, F. (2017). *Deep Learning with Python.* New York: Manning Publishers. Retrieved November 2, 2019, from Manning Publication: https://livebook.manning.com/#!/book/deep-learning-with-python/chapter-1/12

Dhandhania, K. (2018). Retrieved November 14, 2019, from Common Lounge: https://www.commonlounge.com/discussion/040066e852e04889810c0e82910ad10f/history

Dubey,V and Richariya, V. (2013). A Neural Network Approach for ECG Classification. *International Journal of Emerging Technology and Advanced Engineering, 3*(10).

Eur, J. (1996). *Heart rate variability. Standards of measurement, physiological interpretation, and clinical use.* Task for of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology.

Hazra, A., Mandal, S., Gupta, A. and Mukherjee, A. (2017). Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review. *Advances in Computational Sciences and Technology, 10*, 2137-2159.

Nashif, S., Raihan, R., Islam, R and Imam, M.H. (2018). Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System. *WJET, 6*(4).

Patel, J., Upadhyay, P. and Patel, D. (2016). Heart Disease Prediction Using Machine learning and Data Mining Technique. *Journals of Computer Science & Electronics, 7,* 129-137.

Reed, M.J., Robertson, C.E and Addison, P.S. (2005). Heart rate variability measurements and the prediction of ventricular arrhythmias. *An International Journal of Medicine, 85*(2), 87–95.

Saxena, A. (2016). *Convolutional Neural Networks (CNNs): An Illustrated Explanation*. Retrieved October 26, 2019, from Crossroads: The ACM MAgazine for Students: https://blog.xrds.acm.org/2016/06/convolutional-neural-networks-cnns-illustrated-explanation/

WHR. (2019). *NIGERIA : CORONARY HEART DISEASE*. Retrieved November 12, 2019, from World Health Ranking: https://www.worldlifeexpectancy.com/nigeria-coronary-heart-disease

Yadav, N; Yadav, A and Kumar, M. (2015). *An Introduction to Neural Netwok Methods for Differential Equations.* New York: Springer.