IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205294 |

A Novel Approach to Heart Disease Prediction using Xgboost

Punith kumar N, Rajashekara G N, Sahithi M R, Talari Sai Sreenivasulu, Mrs. Pavithra

U.G. Student, Department of Information Science and Engineering, SJB Institute of Technology, Bengaluru, Kengeri, India

Assistant Professor, Department of Information science and Engineering, SJB Institute of Technology, Bengaluru, Kengeri, India

ABSTRACT: Heart abnormalities are It is one of the leading causes of loss of life in the world Patients frequently have no symptoms until fatal events occur, requiring skilled personnel to capture cardiac abnormalities even in their care In recent decades electrocardiogram (ECG) [1]. Using a virtual format would provide evidence of increased machine capacity utilization detected to encounter such abnormalities Machine-getting to know (ML) has shown to be effective in predicting profitability on it in how beautiful [2]. This study aims to develop machine learning definitions of cardiovascular disease based on appropriate features. We used the UCI the benchmark data set for cardiovascular prediction in this study, which includes 14 specific parameters associated with heart failure [3] Support vector machines, logistic regression, artificial neural networks, K-nearest on special neighbors, Naïve bays, and decision trees are just a few of the scripts that this tool is based on. Trendy feature selection methods relief minimum redundancy maximum relevance, minimum Absolute shrinkage selection operator and adjacent classes were used said of the article to eliminate improprieties and eliminate inappropriate [4] . With approximately 40 thousand ECGs tagged through cardiologists in specific hospitals and countries, our fashions can cope with 7 unique alerts: normal, AF, heart failure, heart failure, heart failure, heart failure cavity, or other noise prediction models using multiple integrations and multiple statistical characteristic methods [5] . Using an XGBoost rule from the first gadget learning method, we already use the version-derived F1 rating at zero-nuety-three – zero-nine. To our knowledge, this is the first model reporting high performance across hospitals, international locations, and file types [6]

KEYWORDS: Heart disease, Machine Learning, Feature Selection, XGBOOST Algorithm.

I.INTRODUCTION

Heart disease is considered a threat to human life worldwide. According to statistical reports from the World Health Organizations, 17.9 a million individuals will perish from cardiovascular disease (32% of all global deaths) in 2019. [1] Many lives can be saved by detecting HD and appropriate interventions Early diagnosis and treatment of many heart diseases is exceptionally difficult, especially in developing countries, due to the lack of accredited medical centers and. accurate diagnosis of other factors affecting inflammatory heart disease can be reduced Atrial fibrillation (AF) and the typical doses, affecting 1-2% of the total population can lead to stroke, heart attack or myocardial infarction [4]. An issue with the diagnosis of AF is that it is very routinely asymptomatic (discovered incidentally in electrocardiograms of 30–forty-five patients for unrelated reasons [19]) and requires trained personnel them to detect problems from electrocardiograms (ECGs). Doing so could be a deadly opportunity for injury. Similarly, tachycardia (increased heart rate) and bradycardia (increased heart rate) are not rare heart conditions Although not as harmless as AF, They have the potential to seriously complicate cardiac arrest, if untreated) procedures are increasingly timeconsuming in global health care , according to guidelines for standardized diagnostic methods. In fact, the algorithm can be useful amo in many ways

II. RELATED WORK

The confusion surrounding the new aesthetic comes from its definition, as it includes specific (undefined) details of the heart, making it miles harder for the statue to determine its exact nature. which will be reclassified Figure 2 shows the essential model and its features. The CSI index is highly descriptive in classifying the signal as normal or new, and is actually a popular index used in the literature to capture anomalies in the ECG with R-R intervals greater than 50 seconds is the same as a person with atrial fibrillation has an irregular heartbeat surrounded by non-current devices.

i. Less Accuracy

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205294 |

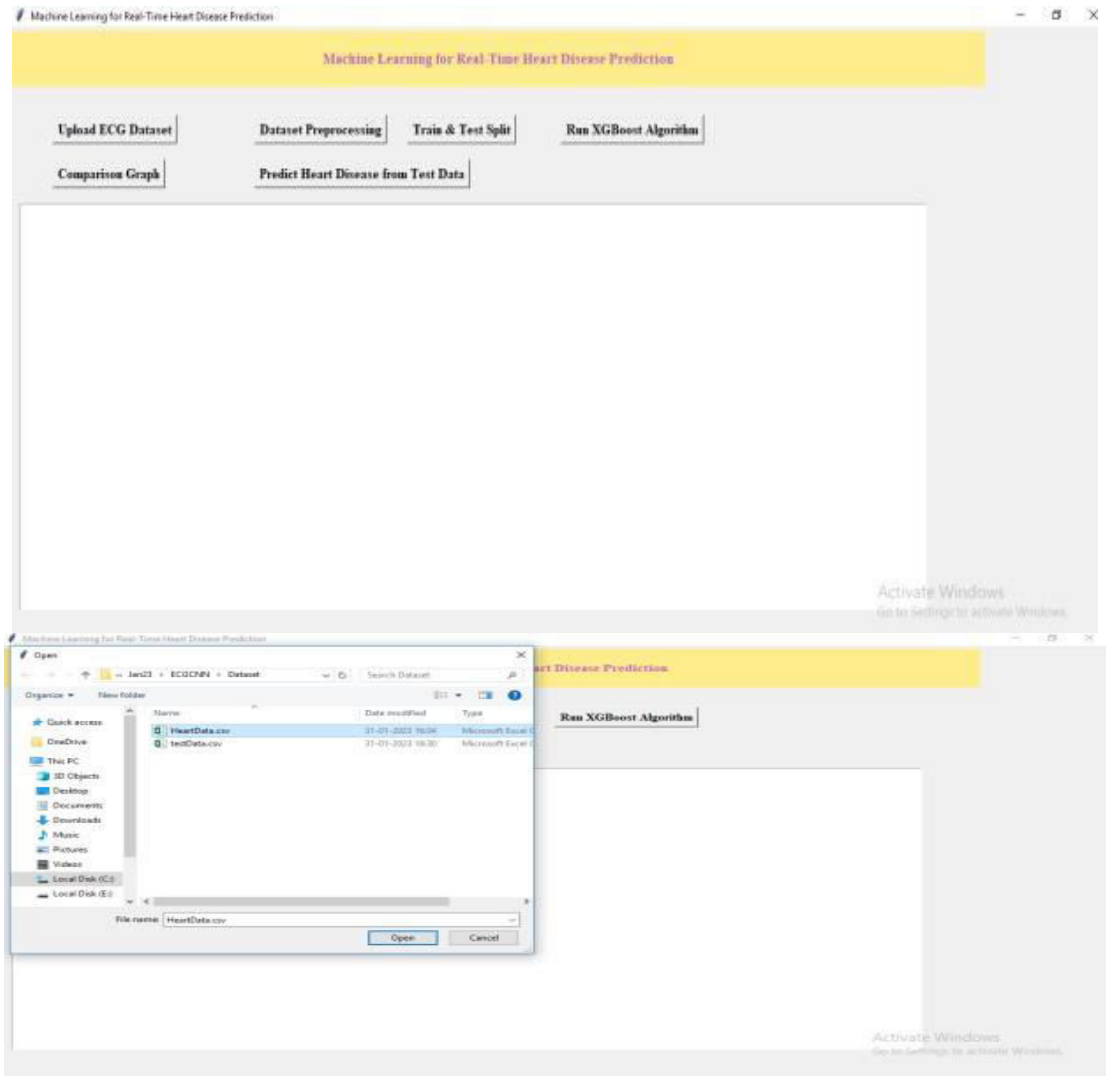
ii. Low Efficiency

III. METHODOLOGY

1) Upload ECG Dataset: This module allows us to upload datasets into the program.
2) Dataset Preprocessing:Using in this module, we'll learn all the dataset values then remove the missing values then normalize and shuffle the values
3) Train and Test Split: The processed values can be split into train and test with 80% used for training and 20% used for testing the data set.
4) Run XGBoost Algorithm: In order to determine the prediction accuracy, we will use this module to train a set of XGBOOST rules on 80% of school data and then test them on the remaining 20% of the data.
5) Comparison Graph: We will use this module to create equality and equality graph
6) Predict Heart Disease from Test Data: This module will be used to upload the test records after which the utility will check the facts and assign an expected label as regular or coronary heart attack.To run the mission can be found below by double clicking on the run.Bat file

IV. EXPERIMENTAL RESULT

To initiate the project, double-click on run.Bat file to access the screen below
In above display click on 'Upload ECG Dataset' button to upload dataset and get under outpu



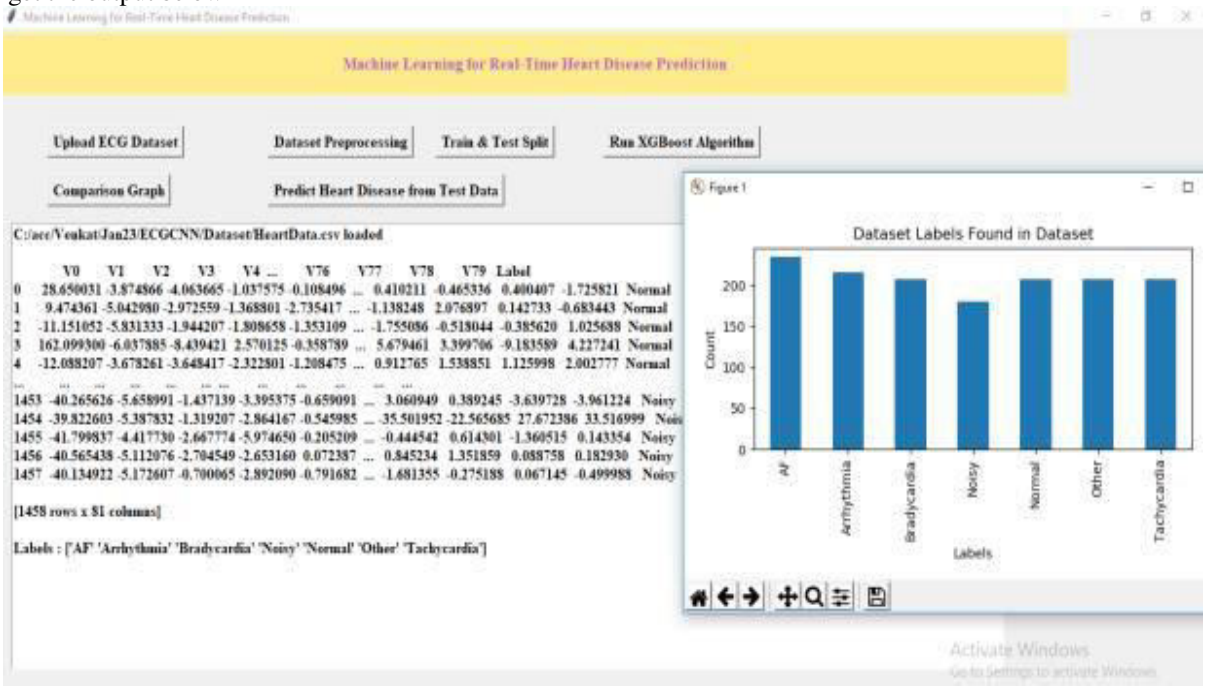
Select and upload the 'HeartData.csv' file to the screen above and after that, press "Open" to load the dataset.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

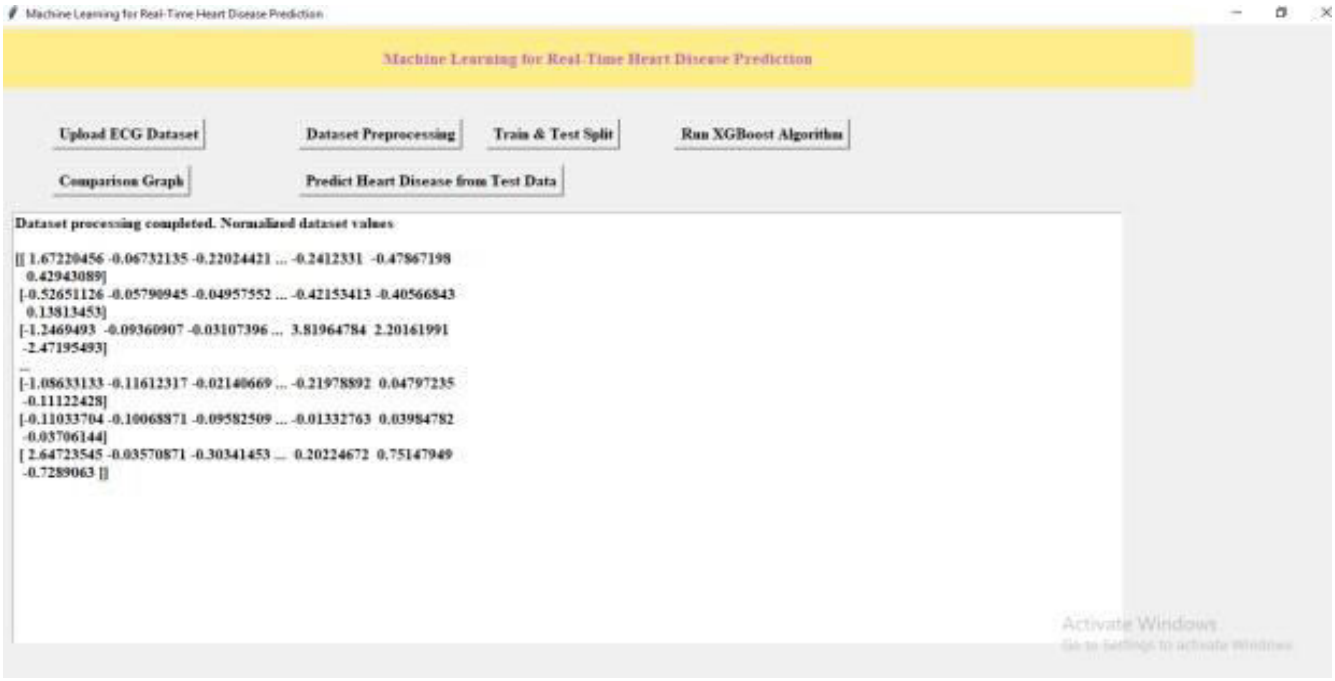
|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205294 |

and get the output below



Above display screen dataset is loaded and It is visible to us dataset of numeric and non-numeric data but gadget earning algorithm will take numeric data most efficient so we have to first process the facts to convert into numeric and x-axis in the above graph represents the name of coronary heart failure and the Y-axis represents the dependence on the statistics for that disorder observed within the dataset. Click the upper graph now, and then click the 'Dataset Preprocessing' button to execute the dataset technique and get the output



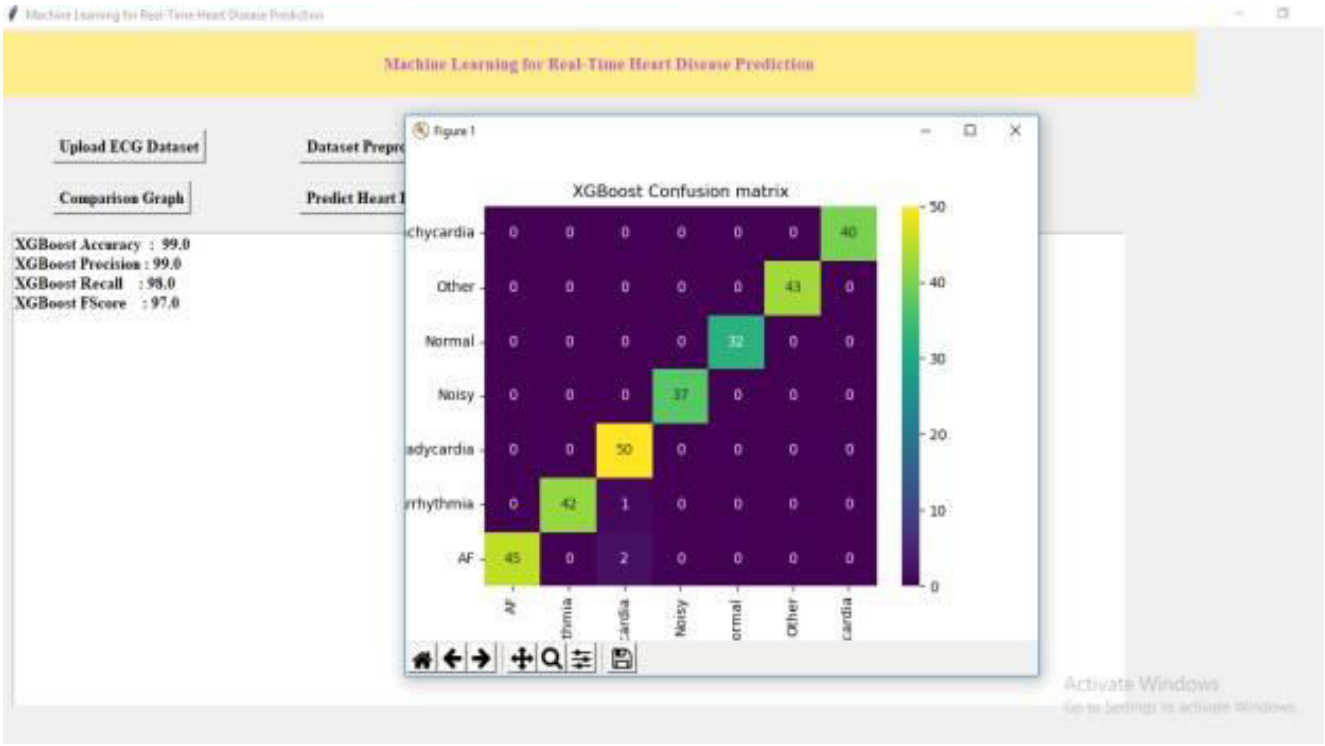
All the values on the aforementioned screen have been transformed into calculations and you can now click on the 'Train & Test Split' button to split the dataset into trains and test to get the output below

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

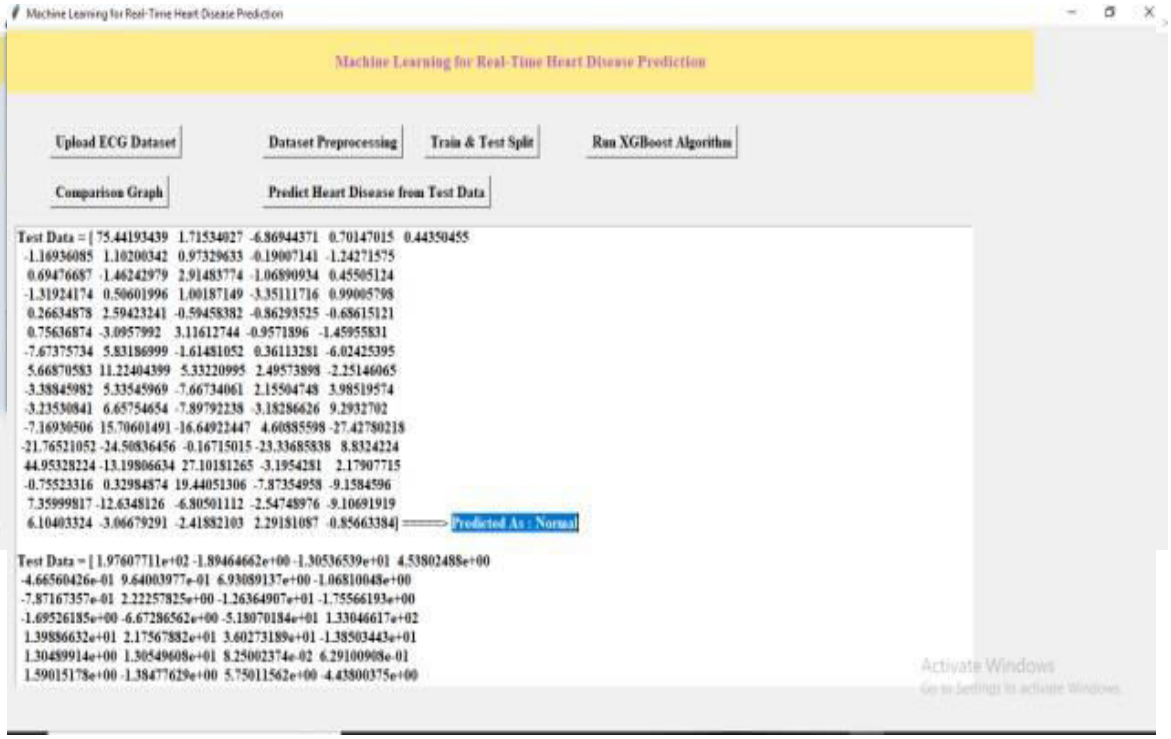
|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205294 |



In the screen above where XGBOOST we achieved 99% accuracy as well as in the confusion matrix graph the x-axis represents Predicted Labels and the y-axis represents True Labels and all colored boxes represent correct prediction estimates and blue boxes have incorrect prediction too conservative approximations.Now close the top graph and then click on the 'Comparison Graph' button to get the bottom graph

In above screen selecting and uploading 'testData.csv' file and then click on 'Open' button to load dataset and get below output

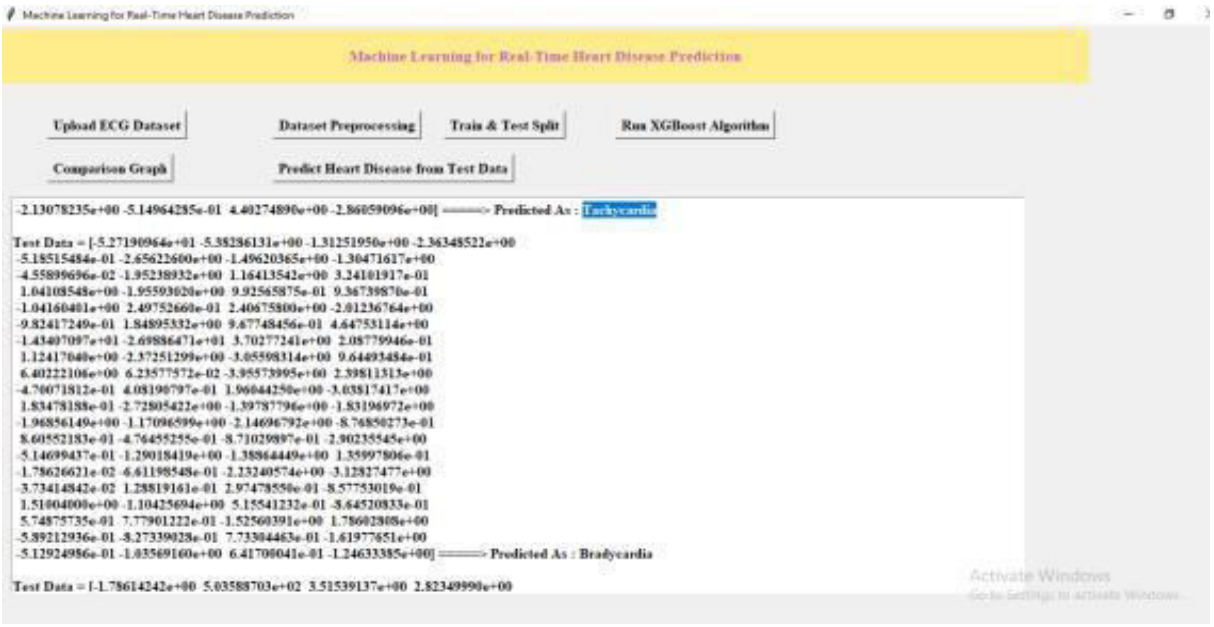


| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205294 |

In the square cell of the top screen we can see the ECH signal test data and after the arrow symbol =∟ we can see the predicted output and scroll down the output screen to view the new prediction output



V. CONCLUSION

In this paper, we suggest a unique approach to discover cardiac abnormalities from freshly recorded ECGs. The prediction techniques can be summed up like this:
signal preprocessing, feature filtering, version schooling, size and evaluation. We formulate a function extraction pipeline of 110 capabilities, with which we educate 5 unmarried fashions on 3 collected information sets. Our models show overall overall performance predictions which might be very strong with admire to unobservable facts, however can be generalized to facts sets of ECGs recorded in precise contexts, and to populations with precise inherent characteristics. Bradycardia, different (non-exceptional), arrhythmia and noise. To correspondingly boom the reliability of our models, we use Temperature Scaling to lessen the anticipated size errors in our put on measurements. Our fashions confirm that on the spot search for symptoms of the QRS complicated gives excessive prediction accuracy. It needs to be expansive functionality effect on the lives of people tormented by heart illnesses. In truth, we expected our paintings to be implemented in a real time setting, with a wearable tool that can continuously show the coronary heart beat of the sufferers at danger.

REFERENCES

1. Chapman University and shaoxing people's hospital. https://figshare.com/bibliography/ChapmanECG/4560497/1,
2. Tianchihefei High technology cup ecg man-machine intelligence competition. http://hangzhou.oss-cn-hangzhou. aliyuncs.com/231754/round2/hf_route2_train.zip, .
3. Alivecor, Inc. https://www.alivecor.com/#, https://www.alivecor.com/#.
4. T.S., eds. Akiba, S. Sano, T.S. Yanase , T. Ohta , M. Koyama . Optuna: A next generation hyperparameter optimization framework. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,
5. Z. D. G. Ari L. Goldberger A. Schwilkin Editor-in- Chief. Goldberger's Treatment in Electrocardiology. https://www.sciencedirect.com/topics/ medicine-and- dentistry/qrs-complex, 2017.
6. ZI Attia, PA Nakworthy, F Lopez-Jimenez, SJ Asirwatham, AJ Deshmukh, BJ Gersh, RE Carter, X Yao

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details