

CLASSIFICATION OF CARDIAC ARRHYTHMIA

Submitted in partial fulfillment of the requirements for the award of Bachelor
of Engineering degree in Electronics and Communication Engineering

by

G. SRINIVASULU(37130136)

G. SAIRAM(37130135)



**DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
SCHOOL OF ELECTRICAL AND ELECTRONICS**

SATHYABAMA

**INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)**

**Accredited with Grade "A" by NAAC
JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI - 600 119**

APRIL - 2021



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE

www.sathyabama.ac.in

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the bonafide work of **G. SAIRAM (37130135)** and **G.SRINIVASULU (37130136)** who have done the Project work as a team who carried out the project entitled "**CLASSIFICATION OF CARDIAC ARRHYTHMIA**" under my supervision from December 2020 to April 2021.

Internal Guide

Dr.S.Poornapushpakala, Ph.D

Head of the Department

Dr. T. RAVI M.E, Ph.D.

Submitted for Viva voce Examination held on_____

Internal Examiner

External Examiner

DECLARATION

We G. SRINIVASULU and G. SAIRAM hereby declare that the Project Report entitled “**CLASSIFICATION OF CARDIAC ARRHYTHMIA**” done by me under the guidance of Dr.S.Poornapushpakala, Ph.D is submitted in partial fulfillment of the requirements for the award of Bachelor of Engineering degree in ELECTRONICS AND COMMUNICATION ENGINEERING.

DATE:

PLACE:

1. G. Srinivasulu
2. G. SaiRam

SIGNATURE OF THE CANDIDATE

ACKNOWLEDGEMENT

We pleased to acknowledge my sincere thanks to Board of Management of SATHYABAMA for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

We convey our thanks to Dr. N.M.Nanditha.,M.E.,(Ph.D)., Dean, School of ELECTRICAL AND ELECTRONICS and Dr.T.Ravi.,M.E.,Ph.D., Head of the Department, Dept. of Electronics and Communication Engineering for providing us necessary support and details at the right time during the progressive reviews.

We would like to express my sincere and deep sense of gratitude to my Project Guide Dr.S.Poornapushpakala, Ph.D. for her valuable guidance, suggestions and constant encouragement which paved way for the successful completion of our project work.

We wish to express our thanks to all Teaching and Non-teaching staff members of the Department of Electronics and Communication Engineering who were helpful in many ways for the completion of the project.

ABSTRACT

Cardiac disease is the fatal disease in the biosphere. Arrhythmia is one kind of cardiac disease which is tough to detect but by using the tedious cardiogram (ECG) recording. Due to the assortment and the noise of ECG, it consumed more time to detect it only by experts using basic eyes. Learning from the earlier study in order to help the authorities, this research develop layers CNN using MITBIH Arrhythmia Data. The dataset is firstly preprocessed. The challenge is to conquer the inequity and minor amount of data still achieve optimal accuracy. This project can be supportive for the doctors to figure out the type of arrhythmia of the patient.

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NUMBER.
1.	INTRODUCTION	9
2.	LITERATURE SURVEY	11
3.	AIM OF THE PROJECT	16
4.	DESIGN AND METHODOLOGY	17
	4.1 HARDWARE REQUIREMENTS	17
	4.2 SOFTWARE REQUIREMENTS	17
	4.3 TECHNOLOGIES USED	18
	4.3.1 Python Language	18
	4.4 APPLICATIONS OF PYTHON	
	PROGRAMMING	18
	4.4.1 Net Applications	18
	4.4.2 Scientific and numeric computing	19
	4.4.3 Creating software prototypes	19
	4.4.4 Good language to teach	
	Programming	19
	4.4.5 Opencv package	19
	4.5 FEATURES OF ANACONDA	20
	4.5.1 What is Anaconda navigator?	21
	4.5.2 Applications provide in Anaconda	
	Distribution	21
	4.5.3 New features of Anaconda 5.3	23
	4.6 PROPOSED SYSTEM	24
	4.7 ARCHITECTURE	24
	4.8 IMPLIMENTATION	24
	4.8.1 Module description	24

	4.8.2 Data collection	24
	4.8.3 Data preprocessing	25
	4.8.4 Classifier training	25
	4.8.5 CNN algorithm	25
	4.9 TESTING	25
5.	RESULT AND DISCUSSION	28
6.	CONCLUSION	36
	REFERENCES	37

LIST OF FIGURES

FIGURE NO	FIGURE TITLE	PAGE NO
4.5	ANACONDA DISTRIBUTION	20
4.5.1	ANACONDA NAVIGATOR	21
4.5.3	NEW FEATURES OF ANACONDA	23
4.7	ARCHITECTURE	24
5.1	IMPORT FILES	28
5.2	READ CSV FILES	28
5.3	UNBALANCED DATA	29
5.4	UNBALANCED FIGURE	29
5.5	RESAMPLE	30
5.6	BALANCED DATA	30
5.7	BALANCED FIGURE	31
5.8	DATAFRAME	31
5.9	SEISMIC SIGNAL	32
5.10	SEISMIC SIGNAL IN HISTOGRAM SIGNAL	32
5.11	NETWORK	33
5.12	EVALUATE	33
5.13	ACCURACY	34
5.14	CONFUSION MATRIX	35

CHAPTER 1

INTRODUCTION

1.1 OVERVIEW

Cardiovascular illnesses are the primary reason of human passing, with more than million individuals known to lose their lives every year payable to CVDs. Maximum of the total CVD deaths are among the central and low-pay parts of the the social order. Arrangement model to recognize cardiac at principal stage could efficiently decrease death. The mutual bases of cardiac arrhythmia, wherever heartbeats are to deviate from their consistent beating design. Regular heartbeats differs activity, stage and feelings. The heartbeats beats faster or slower, the condition is known as tremors. An arrhythmia does not certainly mean that the heart beats fast or slow. It means the heart beats reckless—tachycardia (>100 bpm), or slow—bradycardia (<60 bpm), skipping a beat, or in extreme cases, cardiac arrest. There are mutual types of abnormal heart rhythms includes atrial flutter, atrial fibrillation, and ventricular fibrillation. These deviations classified into several subclasses and represent different types of cardiac arrhythmia. Classification of these types helps in analyzing the treatment of heart illness patients.

The graph recordings are generally used for analyzing and foretelling internal organs for diagnosis of heart diseases. Towards this finish, experimental consultants may check up on cardiogram recordings over an extended amount of your time for detective work arrhythmia. The cardiogram could be a one-dimensional indication in place of a statistic, which may be analyzed exploitation machine learning techniques for machine- controlled recognition of bound irregularities. DL techniques are advanced, which offer vital performance in tomography copy investigation. CNN consume newly remained exposed to figure for multi-dimensional inputs however were initially developed for issues managing pictures outlined as two-dimensional inputs. For statistic knowledge, one dimensional CNNs are planned however are less versatile in comparison to two dimensional CNNs. Hence, representing the statistic knowledge during a two dimensional format may benefit bound machine learning tasks. Hence, for cardiogram signals, a two dimensional transformation should be applied to create the statistic appropriate for deep learning ways that need two dimensional pictures as input. The

short-time Fourier Transform will alter a one-dimensional signal into a two-dimensional photo and compress the period and frequency data among one matrix. The 2-D photo is analogous to hyperspectral and multi-spectral Pictures, that have numerous applications in remote sensing and clinical diagnosing, as well as haunted un-mixing, minced protection arrangement and similar, inorganic investigation, health copy arrangement, change recognition, reproduction physical documentation, target detection, activity recognition, and police investigation. The two dimensional matrix of photo coefficients might be helpful for extracting strong options for illustration of an internal organ cardiogram signal. This illustration may enable the appliance of CNN architectures (designed to work on two dimensional inputs) for the development of machine- controlled systems associated with CVDs.

CHAPTER 2

LITERATURE SURVEY

Prediction of Cardiopathy With ML

Authors: Aditi Gavhane, Gouthami Kokkula, sha Pandya ; Prof. Kailas Devadkar, 2018
Second Worldwide Meeting on Electronics, Communication and Aerospace Technology

With the rampant increase within the heart stroke rates at juvenile ages, we want to place a system in place to be able to sight the symptoms of a heart stroke at associate in nursing early stage and therefore forestall it. It's impractical for a standard man to oftentimes endure expensive tests just like the ECG and therefore there has to be a system in place that is handy and at an equivalent time reliable, in predicting the possibilities of a heart condition. Therefore we tend to propose to develop associate in the nursing application which might predict the vulnerability of a heart condition given basic symptoms like age, sex, pulse, etc. The machine learning formula neural networks has evidenced to be the foremost correct and reliable formula and therefore utilized in the planned system.

Cardiopathy Prediction using Evolutionary Rule Learning

Authors: Aakash Chauhan ; Aditya Jain ; Purushottam Sharma ; Vikas Deep, 2018
Worldwide Meeting on Computational Intelligence.

In trendy society, cardiopathy is that the noteworthy reason for a brief life. The giant population of Individuals depends on the care system so they will get the correct lead to less time. A great deal of knowledge is made and picked up by the care organization on the day-to-day. To urge intriguing information, information innovation permits extracting the information through automatization of processes. Weighted Association Rule may be a form of data processing technique accustomed to eliminate the manual task that additionally helps in extracting the information directly from the electronic records. This may facilitate in decreasing the value of services and additionally helps in saving lives. During this paper, we will realize the rule to predict a patient's risk of getting coronary unwellness. Take a look at results that have shown that overwhelming majority of the foundations helps within the best prediction of coronary malady.

Blood Thickness founded Cardiopathy Risk Forecast Technique in Edge/Fog Computer science

Authors: R Latha ; P Vetrivelan, 2019 11th Worldwide Conference on Communication Systems & Networks

In this paper, cardiopathy prediction sculpturesque victimization partly evident mathematician call method (POMDP) is projected. In an emergency, the patient is alerted through the doctor by fog computing. Ambulance sent to the place of a patient at essential things. The doctor gets the information through fog computing iFogSim. Fog computing in care may be a new space, that gains a lot off attraction in analysis community. Several researches concentrate on upset i.e. cardiopathy. The vital risk issue for cardiovascular disease is increase in blood viscousness. The extremely viscous nature of blood doesn't enable the blood to flow making a resistance within the blood flow. Cardiopathy risk factors area unit high blood pressure, obesity, diabetes, increased blood viscosity, etc. With the assistance of POMDP's states, observations, beliefs, likelihood transitions the patient health is noted. The POMDP model cardiopathy prediction computes the policy approximation victimization states and timeslots. Rewards area unit tabulated victimization policy approximations over different iterations.

The Study of Cardiopathy Prediction Models

Authors; G. Shanmugasundaram ; V. Malar Selvam ; R. Saravanan ; S. Balaji, 2018 iee international conference on system, computation, automation and networking (icscan)

Heart disease indicates the sort of condition that results in heart malfunction. Not all individuals with arteria unwellness have pain as a symbol. There area unit alternative factors like R-Blood Pressure, S-Cholesterol, F-Blood Sugar, R-ECG, Thalach, Ex-Ang, range of major Vessels blocked, metallic element Scan that additionally leads to cardiopathy. At associate initial stage the prediction of cardiopathy will save human lives. Data processing techniques like Naïve mathematician, Call Tree, K-nearest neighbor, etc., area unit utilized in the center unwellness prediction supported the parameters / factors. The target of this analysis is to analyze concerning the varied factors and its significance in identifying the heart disease. Further it also explores about the existing techniques and models used for prediction. This work concludes with the key challenges

that aren't however explored.

Newfangled Method for Primary Recognition and Inhibition of Ischemic Cardiopathy using Data Processing

Authors: Dhara B. Mehta ; Nirali C. Varnagar, 2019 Worldwide Meeting on Movements in Electronics and Informatics.

The significant advances in biotechnology and health science have LED to provide great amount of knowledge. It's offensively difficult to spot disease supported symptoms. For that we had like tons of expertise and information. Finding the precise reason behind unwellness will facilitate to patient cure at early stage of risk level. Anaemia cardiopathy is one among the most foremost common causes of death these days. During these analysis work, we have applied data processing classification techniques on the stroke dataset and planned a recommendation system during which an individual will come back to grasp their risk level of IHD. By taking correct cure and treatment will facilitate to survive a lot of years if they need high risk of IHD. We have used five data processing algorithms Logistic Regression, Decision Tree, K nearest Neighbor, Naïve Bayes and SVM on Ischemic Stroke Dataset.

Real-time machine learning for primary finding of Cardiopathy using big data method

Authors: Abderrahmane Ed-Daoudy ; Khalil Maalmi, 2019 Global Session on Radiocommunication Technologies, Embedded and Intelligent Systems.

Over the previous couple of decades, heart condition is that commonest reason for world death. therefore primary recognition of heart condition and continuous watching will rscale back the death rate. The exponential growth of knowledge from totally different sources like wearable devices utilized in Internet of Things health watching, streaming system is generating an enormous amount of knowledge on an eternal basis. The mixture of streaming big data analytics and machine learning may be a breakthrough technology which will have a major impact in aid field particularly early detection of heart condition. This technology may be additional powerful and less expensive. To beat this issue, this paper propose a real-time heart condition prediction system based on apache Spark that stand as a robust massive scale distributed computing platform that will be used with succes for streaming knowledge event against machine learning through in-

memory computations. The system consists of two main sub elements, specifically streaming process and knowledge storage and visualization. The primary issues Stimulus MLlib by Stimulus flowing and applies classification model on knowledge events to predict heart condition. The seconds uses Apache Cassandra for storing the big volume of generated knowledge.

Predictive Analysis of Rapid Spread of Heart Disease with Data Mining

Authors: Radhanath Patra ; Bonomali Khuntia, 2019 IEEE Universal Session on Electric, Computer and Communication Technologies

cardiovascular disease is that the leading reason for mortality for each sexes in worldwide. Cardiovascular disease is increasing at a speedy rate in each older and younger generation of males and females currently days. Thus in want demand of right ways, progress and application of effective health obsorvation policies ought to be stressed to combat the epidemic of heart connected diseases. Thus early detection and treatment with the utilization of each typical and innovative technique sholud be most popular during this paper we have used the machine learning repository Cleveland cardiovascular disease information having 303 instance and 76 attributes. For the planned model we have used the knowledge gain idea for selection of best attribute and processes the chosen options vicamization rail and python. This paper identifies the gap of analysis on prediction of cardiovascular disease supported python Anaconda navigator, spyder and rail platform on that we have abundant stressed. The varied techniques, processes that have used to train the model of heart datasets like feature choice, numpy, pandas library, coil tree classifier, KNN classifier, entropy, gini-index, confusion matrix. The result shows that coil tree classifier is best and applicable for prediction of UCI repository Cleveland heart dataset.

Non-Laboratory-Based Risk Factors for Automated Heart Disease Detection

Authors: H. Mai ; T. T. Pham ; D. N. Nguyen ; E. Dutkiewicz, 2018 12th International
Developing a cardiopathy detection model using easy non-laboratory risk factors plays a vital role in preventive care, particularly for prime risk subjects. The model permits physicians/epidemiologists to effectively diagnose an individual as having cardiopathy. During this work, we have a inclination aim to develop a non-invasive risk prediction

model for automatic cardiopathy detection that involves age, gender, rest force per unit area, most pulse rate, and rest diagnostic technique. We have a inclination examine four public datasets from 1071 participants UN agency were referred for a special X-ray of the heart's arteries (i.e., to check if they're narrowed or blocked). The themes conjointly undertook a physical examination and 3 non-invasive tests. To estimate the cardiopathy status, we have a inclination to apply a generalized linear model with regularization methods via coordinate descent. Even while not laboratory-based knowledge (e.g., humor steroid alcohol, fast blood sugar), we have a inclination to determined a prediction accuracy as high as seventy two percentage, compared with seventy six percentage of different comprehensive models. This observation suggests that few non-invasive factors utilizing recent advances in data analytics will replace this practices of cardiopathy risk assessment.

CHAPTER 3

AIM OF THE PROJECT

This project aims to produce associate degree algorithmic program for the recognition of upset that runs directly on cardiogram knowledge. These days CNN is cable of analyzing cardiogram dataset. Cardiovascular disease is one among the foremost hazardous sicknesses for many of the persons within the world. It's typically confessed as a coronary failure. The injury of the center muscle causes the cadiac heart disease.. The injury is irreversible, therefore it's going to result in death. Fast and precise recognition of heart disease is needed to scale back the death rate. There area unit many diagnostic tools like blood tests, cardiogram is accessible for the analysis of cardiovascular disease. Among all tools, cardiogram provides effective ends up in determinative heart disease, however the manual interpretation of the cardiogram signal might take time for the doctor to spot the symptoms of heart disease. The physical understanding might vary from person to person computer-aided diagnostic tool is needed to investigate cardiogram indications.

Cardiovascular diseases have embellished severely between Non-Communicable diseases, that have thin over the past recent years. In 2018, around seventeen million that is associate degree calculable thirty one percentage of the people have died worldwide because of CVDs. A unique ML algorithm for continuous observation, credentials and classification of internal organ from graphical record knowledge is given here.

CHAPTER 4

DESIGN AND METHODOLOGY

4.1 HARDWARE REQUIREMENTS

The most common set of needs outlined by any software system or software system application is that the physical computer resources, also known as hardware. The smallest hardware needs are as follows,

1. Processor : Pentium IV
2. RAM : 8 GB
3. Processor : 2.4 GHz
4. Main Memory : 8GB RAM
5. Hard Disk Drive : 1tb
6. Keyboard : 104 Keys

4.2 SOFTWARE REQUIREMENTS

Software needs deals with process resource needs associated conditions that must be installed on a computer to provide functioning of an application. The smallest software system needs are as follows,

1. Front end : python
2. Dataset : csv
3. IDE : anaconda
4. Operational System : Windows 10

4.3 TECHNOLOGIES USED:

4.3.1 Python Language

It is an object-oriented language created by Guido Rossum. It's ideally designed for speedy prototyping of complicated requests. It's interfaces to several OS system calls and libraries and is protractile to C or C++. Several giant companies use the Python programming language include NASA, Google, YouTube, BitTorrent, etc.

Python Programming Characteristics

- It affords wealthy information sorts and informal to browse syntax
- It may be a platform independent written language with full access to operating system API's
- Compared to different languages, it permits a lot of run-time flexibility
- It contains the essential text handling services of Perl and Awk
- Python might take single or a lot of categories and free functions
- Libraries in Pythons are compatible with uinix.
- Python are often compiled to byte-code
- Python supports practical and structured programming as well as OOP
- It supports interactive mode that permits interacting Testing
- There is no compilation step, editing, debugging and testing is quick.

4.4 Applications of Python

4.4.1 Net Applications

You can produce scalable net Apps using frameworks and CMS (Content Management System) that are designed on Python. A number of the popular platforms for making Web Apps are: Django, Flask, Pyramid, Plone, Django CMS. Sites like Mozilla, Reddit, Instagram and PBS are written in Python.

4.4.2 Scientific and Numeric Computing

There are various libraries on the market in Python for scientific and numeric computing. There are libraries like: SciPy and NumPy that are employed in general purpose computing. And, there are specific libraries like: EarthPy for natural science, AstroPy for Astronomy then on. Also, the language is heavily employed in machine learning, data processing and deep learning.

4.4.3 Creating software Prototypes

Python is slow compared to compiled languages like C++ and Java. It would not be an honest alternative if resources are restricted and potency may be a should. However, Python may be a great language for making prototypes. For example: You will be able to use Pygame (library for making games) to form your game's prototype initial. If you prefer the epitome, you will be able to use language like C++ to form the particular game.

4.4.4 Good Language to Teach Programming

Python is employed by several companies to show programming to children and newbies. It's an honest language with a lot of options and capabilities. Yet, it's one among easiest language to learn because of it has simple syntax.

4.4.5 Opencv Package

Python might be a universally useful language began by Guido van Rossum, that became highly regarded briefly time principally owing to its simplicity and code readability. It allows the coder to precise his concepts in fewer lines of code while not reducing any readability.

Compared to different languages like C/C++, Python is slower. However another necessary feature of Python is that it are often simply extended with C/C++. This feature helps us to put in writing computationally intensive codes in C/C++ and make a Python wrapper for it so we are able to use these wrappers as Python modules. This offers us two advantages: initial, our code is as quick as original C/C++ code (since it is the particular C++ code working in background) and second, it is terribly straight forward to code in Python. it's a Python wrapping everywhere original C++ operation.

And the support of Numpy makes the task a lot of easier. Numpy may be a extremely optimized library for numerical operations. It offers a MATLAB-style syntax. All the OpenCV array structures are regenerate to-and-from Numpy arrays. Thus no matter

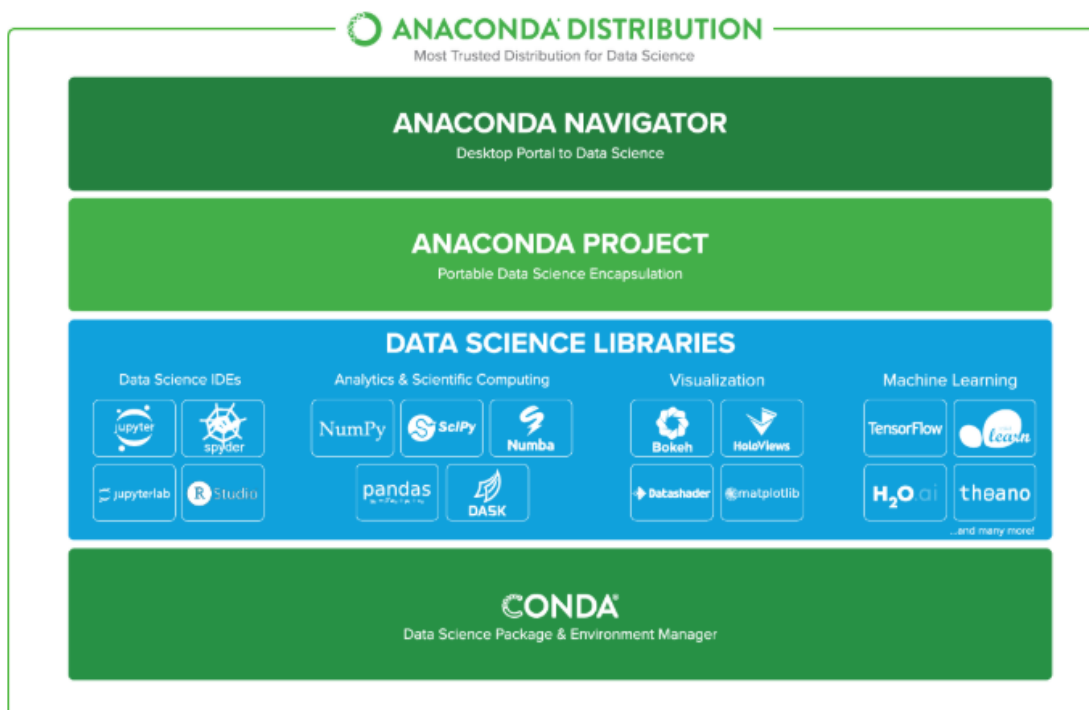
operations you will be able to knock of Numpy, you will be able to mix it with OpenCV, that will increases variety of weapons in your arsenal. Besides that, many different libraries like SciPy, Matplotlib that supports Numpy can be used with this.

As OpenCV-Python is an appropriate tool for quick prototyping of computer vision problems.

4.5 FEATURES OF ANACONDA

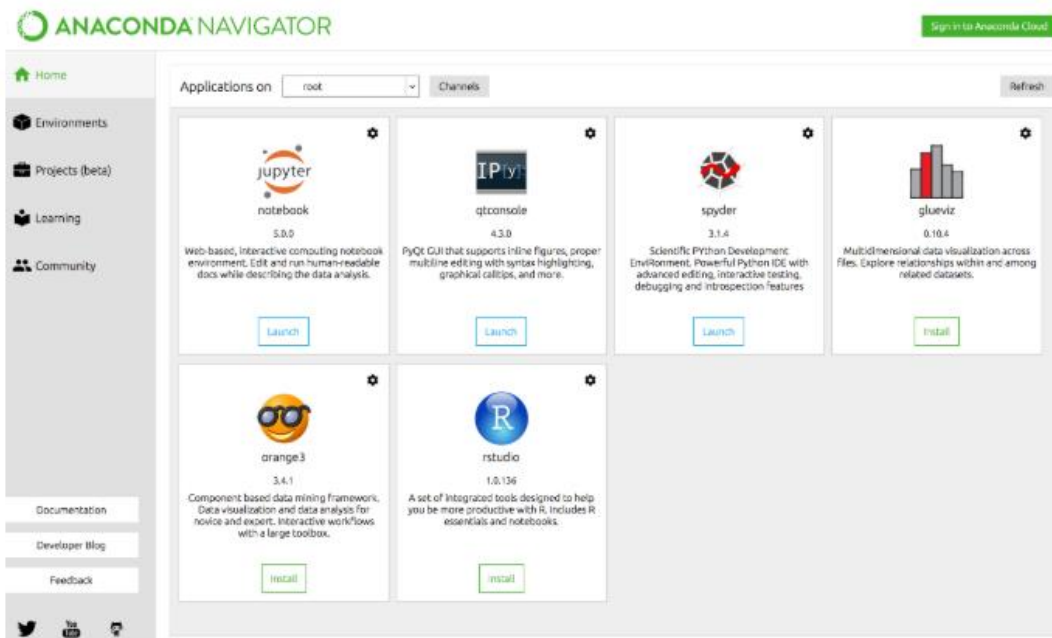
Anaconda could be a free source, easy to install distribution of R, Python programming. It provides a operating setting that is employed for scientific computing, data science, applied math analysis.

The latest version is Anaconda 5.3 and is released in 2018. It's contains conda package, environment manager and a collection of open source packages with free community.



4.5.1 What is Anaconda Navigator?

Anaconda Navigator could be a desktop graphical user interface enclosed within the distribution. It permits us to launch requests provided within the Anaconda distribution and simply manage conda packages, environments and channels while not the employment of command-line arguments. It's out there for Windows and Unix.



Anaconda Navigator

4.5.2 Applications Provided In Anaconda Distribution

The Anaconda distribution comes with the subsequent applications along with Anaconda Navigator.

- JupyterLab
- Jupyter Notebook
- Qt Console

- Spyder
- Glueviz
- Orange3
- RStudio
- Visual Studio Code
- JupyterLab: It is an extendable employed atmosphere for cooperative and reproducible computing,supported the Jupyter Notebook and design.
- Jupyter Notebook: It is often a Net-based, cooperating computing notebook setting. We are able to edit and run human-readable docs whereas describing the information analysis.
- Qt Console: It's the PyQt GUI that supports level figures, correct multiline editing with syntax highlighting, graphical calltips and a lot of.

4.5.3 New Features of Anaconda software



Accumulated with Modern Python release: Anaconda 5.3 is accumulated with Python 3.7, taking benefit of Python's speed and have enhancement.

- **Better Reliability:** The dependableness of Anaconda has been improved within the latest release by capturing and storing the package information for connected packages.
- **Enhanced CPU Performance:** The Information Mathematics Kernel Library 2019 for Deep Neural Networks has been announced in Anaconda 5.3 distribution. Users deploying Tensorflow will build use of MKL 2019 for DNN. These Python binary packages are provided to achieve high CPU performance.
- **New packages are added:** There are over 230 packages which has been updated and additional in the new release.

4.6 PROPOSED SYSTEM

- We propose an ECG cardiac arrhythmia classification technique using a deep convolutional neural network (CNN). The classifier achieved the very best accuracy and also the lowest loss.
- The main advantage of CNN compared to its predecessors is that it mechanically detects the vital options without any human supervision.
- For example, given several photos of ECG it learns distinctive features for every category by itself. CNN is additionally computationally efficient.

4.7 ARCHITECTURE

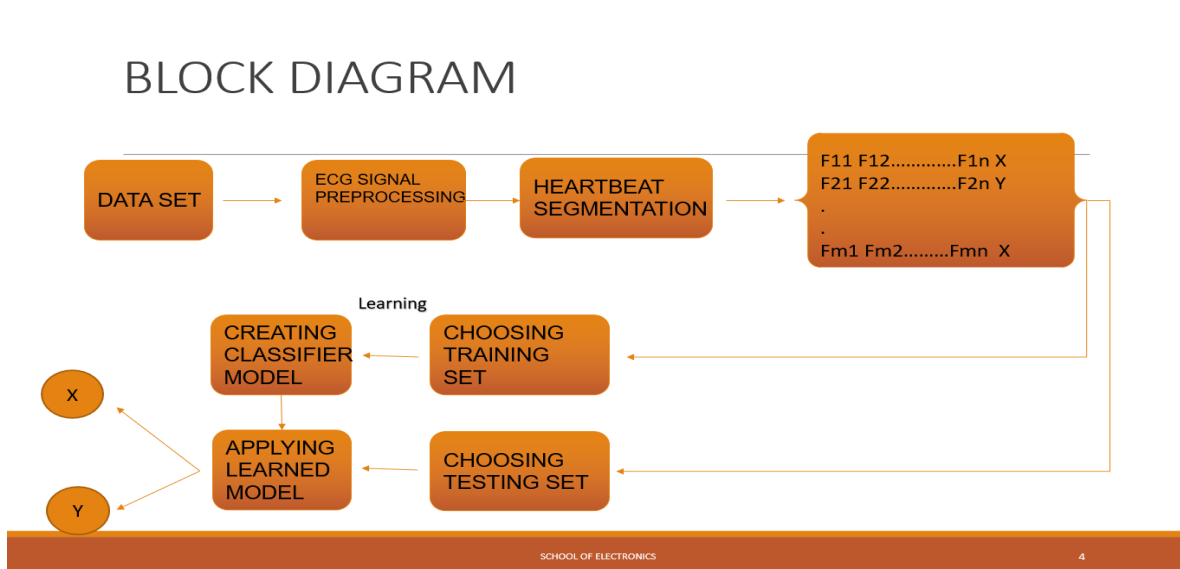


Fig.1 System Architecture

4.8 IMPLEMENTATION

4.8.1 MODEULE DESCRIPTION

The proposed work is to detect and to classify the Cardiac Arrhythmia using a convolutional neural network.

4.8.2 Data Collection

In this stage the information is collected to investigate the best projected method.

MITBIH dataset is used in research for identified the performance of projected methods.

4.8.3 Data Preprocessing

The collected data is preprocessed in this stage. It contains of Feature Selection, Data Transformation and Data Normalization.

- 1) Data Transformation is transferring the symbolic data into numerical values.
- 2) Data Normalization is converting the uneven values into a standard choice [0-1].
- 3) Feature Selection is method of finding the important options and removing the unwanted options in order to rise the speediness of the process.

4.8.4 Classifier Training

The supervised classifier technique is used in our study. After selecting the related options the proposed classifiers are implemented to identify the anomalies.

4.8.5 CNN Algorithm

A convolutional neural network is a definite kind of ANN that uses perceptrons, a ML unit procedure program, for supervised learning, to research knowledge. CNN is used to apply image processing, NLP and other kinds of cognitive tasks.

CNN image classifications takes input image, process and classify it under certain categories.

4.9 TESTING

General

In a generalized way, we can say that the system testing is a type of testing in which the main aim is to make sure that system performs efficiently and seamlessly. The process of testing is applied to a program with the main aim to discover an unprecedented error, an error which otherwise could have damaged the future of the software. Test cases which brings up a high possibility of discovering and error is considered successful.

TEST CASE

Testing, as already explained earlier, it is the method of discovering all possible weak-points in the finalized software product. Testing helps to counter the working of sub-assemblies, components, assembly and the complete result. The software is taken

through different exercises with the main aim of making sure that software meets the business requirement and user-expectations and doesn't fails abruptly. Several types of tests are used today. Each test type addresses a specific testing requirement.

Testing Techniques

A test strategy is a file which describes approach, its scope, its resources and the schedule of aimed testing exercises. It helps to identify almost other test item, the features which are to be tested, its tasks, how will everyone do each task, how much the tester is independent, the environment in which the test is taking place, its technique of design plus the both the end criteria which is used, also rational of choice of theirs, and whatever kind of risk which requires emergency planning. It can be also referred to as the record of the process of test planning. Test plans are usually prepared with signification input from test engineers.

UNIT TESTING

In unit testing, the design of the test cases is involved that helps in the validation of the internal program logic. The validation of all the decision branches and internal code takes place. After the individual unit is completed it takes place. Plus it is taken into account after the individual unit is completed before integration. The unit test thus performs the basic level test at its component stage and test the particular business process, system configurations etc.

FUNCTIONAL TESTING

The functional tests help in providing the systematic representation that purposes verified are available and specified by technical requirement, documentation of the system and the user manual.

SYSTEM TESTING

System testing, as the term advises, is the type of testing in which ensure that the software system meet the business requirements and aim. Testing of the configuration is taken place here to ensure predictable result and thus analysis of it. System testing is relied on the description of process and its flow, stressing on pre driven process and the points of integration.

INTEGRATION TESTING

These trials are intended to trial the united software program items to regulate whether they really execute as a single package or application. The challenging is event determined and thus is anxious with the elementary outcome of field. The Integration tests demonstrate that the components were individually satisfaction, as already represented by successful piece challenging, the apparatuses are apt and fine. This type of testing is specially aimed to expose the issues that come-up by the components combination.

WHITE BOX TESTING

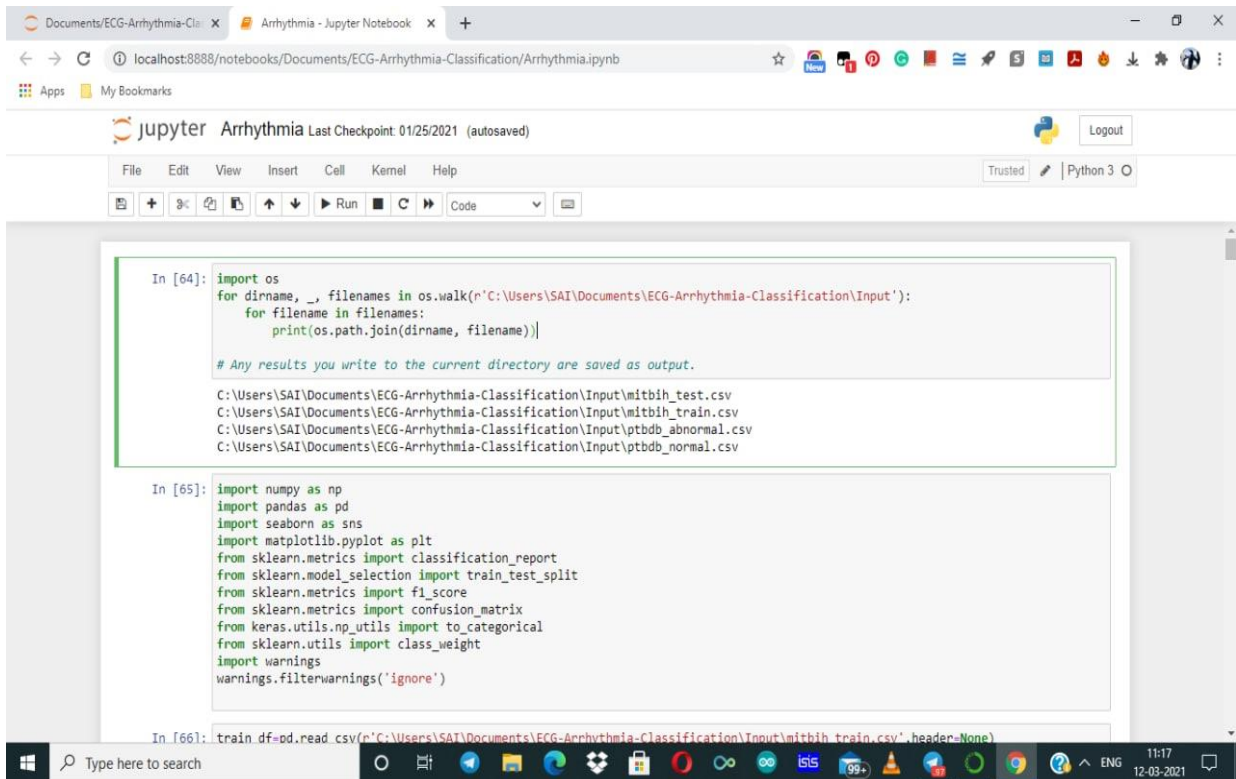
It is the type of testing in which the internal components of the system software is open and can be processed by the tester. It is therefore a complex type of testing process. All the data structure, components etc. are tested by the tester himself to find out a possible bug or error. It is used in situation in which the black box is incapable of finding out a bug. It is a complex type of testing which takes more time to get applied.

BLACK BOX TESTING

It is the testing is the type of testing in which the internal components of the software is hidden and only the input and output of the system is the key for the tester to find out a bug. It is therefore a simple type of testing. A programmer with basic knowledge can also process this type of testing. It is less time consuming as compared to the white box testing. It is very successful for software which are less complex are straight-forward in nature. It is also less costly than white box testing.

CHAPTER 5

RESULT AND DISCUSSION



The screenshot shows a Jupyter Notebook interface with the following code in cell [64]:

```
In [64]: import os
for dirname, _, filenames in os.walk(r'C:\Users\SAI\Documents\ECG-Arrhythmia-Classification\Input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))

# Any results you write to the current directory are saved as output.

C:\Users\SAI\Documents\ECG-Arrhythmia-Classification\Input\mitbih_test.csv
C:\Users\SAI\Documents\ECG-Arrhythmia-Classification\Input\mitbih_train.csv
C:\Users\SAI\Documents\ECG-Arrhythmia-Classification\Input\ptbdb_abnormal.csv
C:\Users\SAI\Documents\ECG-Arrhythmia-Classification\Input\ptbdb_normal.csv
```

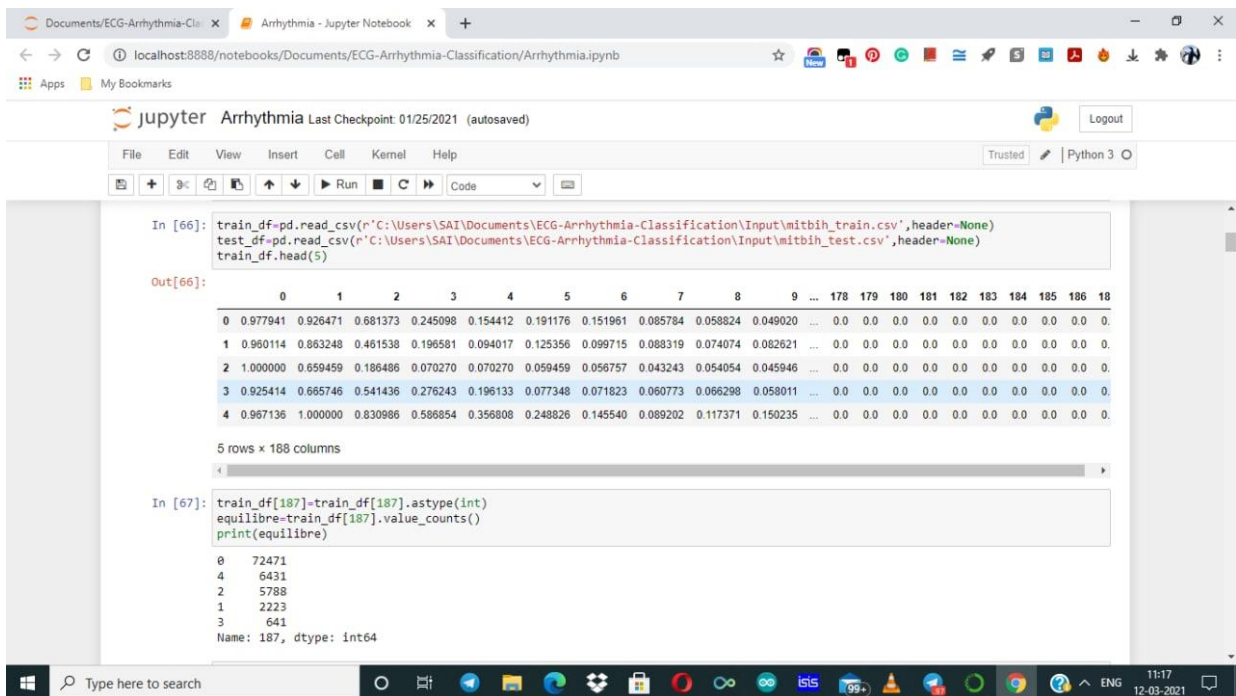
Cell [65] contains the following imports:

```
In [65]: import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.metrics import classification_report
from sklearn.model_selection import train_test_split
from sklearn.metrics import f1_score
from sklearn.metrics import confusion_matrix
from keras.utils.np_utils import to_categorical
from sklearn.utils import class_weight
import warnings
warnings.filterwarnings('ignore')
```

Cell [66] contains the following code:

```
In [66]: train_df=pd.read_csv(r'C:\Users\SAI\Documents\ECG-Arrhythmia-Classification\Input\mitbih_train.csv',header=None)
```

Fig 5.1



The screenshot shows a Jupyter Notebook interface with the following code in cell [66]:

```
In [66]: train_df=pd.read_csv(r'C:\Users\SAI\Documents\ECG-Arrhythmia-Classification\Input\mitbih_train.csv',header=None)
test_df=pd.read_csv(r'C:\Users\SAI\Documents\ECG-Arrhythmia-Classification\Input\mitbih_test.csv',header=None)
train_df.head(5)
```

The output of cell [66] is a 5x188 DataFrame:

	0	1	2	3	4	5	6	7	8	9	...	178	179	180	181	182	183	184	185	186	18
0	0.977941	0.926471	0.681373	0.245098	0.154412	0.191176	0.151961	0.085784	0.058824	0.049020	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
1	0.960114	0.863248	0.461538	0.196581	0.094017	0.125356	0.099715	0.088319	0.074074	0.082621	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
2	1.000000	0.659459	0.186486	0.070270	0.070270	0.059459	0.056757	0.043243	0.054054	0.045946	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
3	0.925414	0.665746	0.541436	0.276243	0.196133	0.077348	0.071823	0.060773	0.066298	0.058011	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
4	0.967136	1.000000	0.830986	0.586854	0.356808	0.248826	0.145540	0.089202	0.117371	0.150235	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.

5 rows x 188 columns

Cell [67] contains the following code:

```
In [67]: train_df[187]=train_df[187].astype(int)
equilibre=train_df[187].value_counts()
print(equilibre)
```

The output of cell [67] is:

```
0    72471
4    6431
2    5788
1    2223
3     641
Name: 187, dtype: int64
```

Fig 5.2

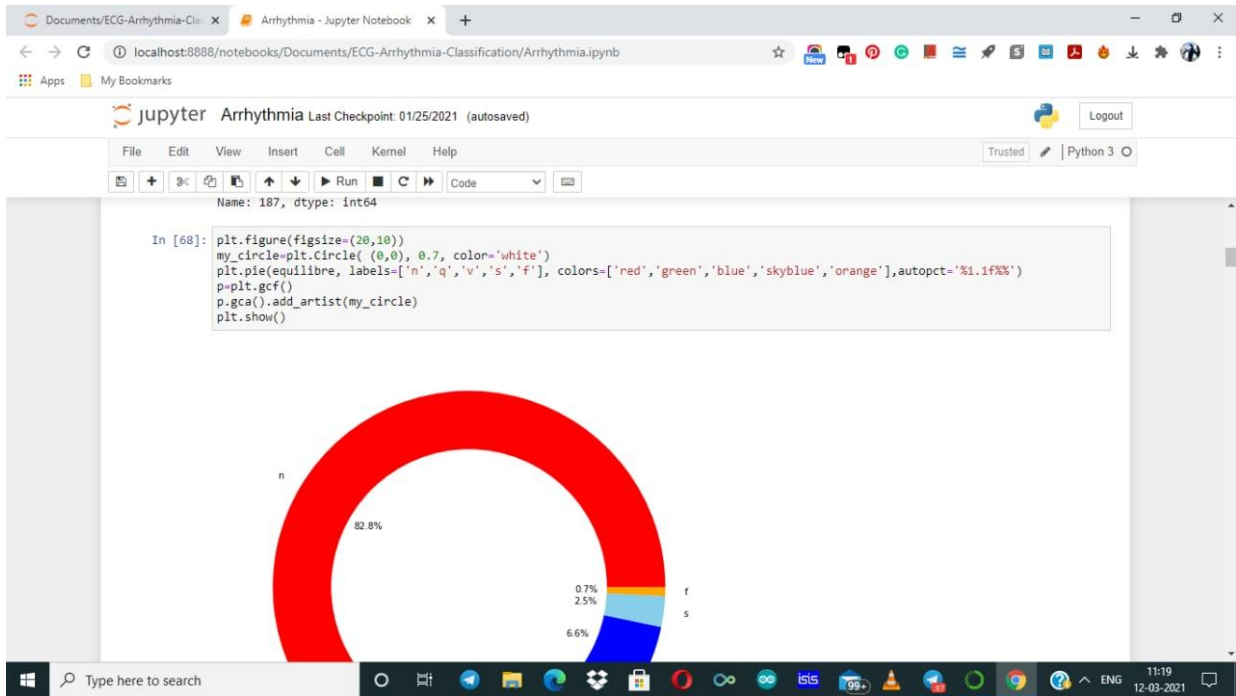


Fig 5.3

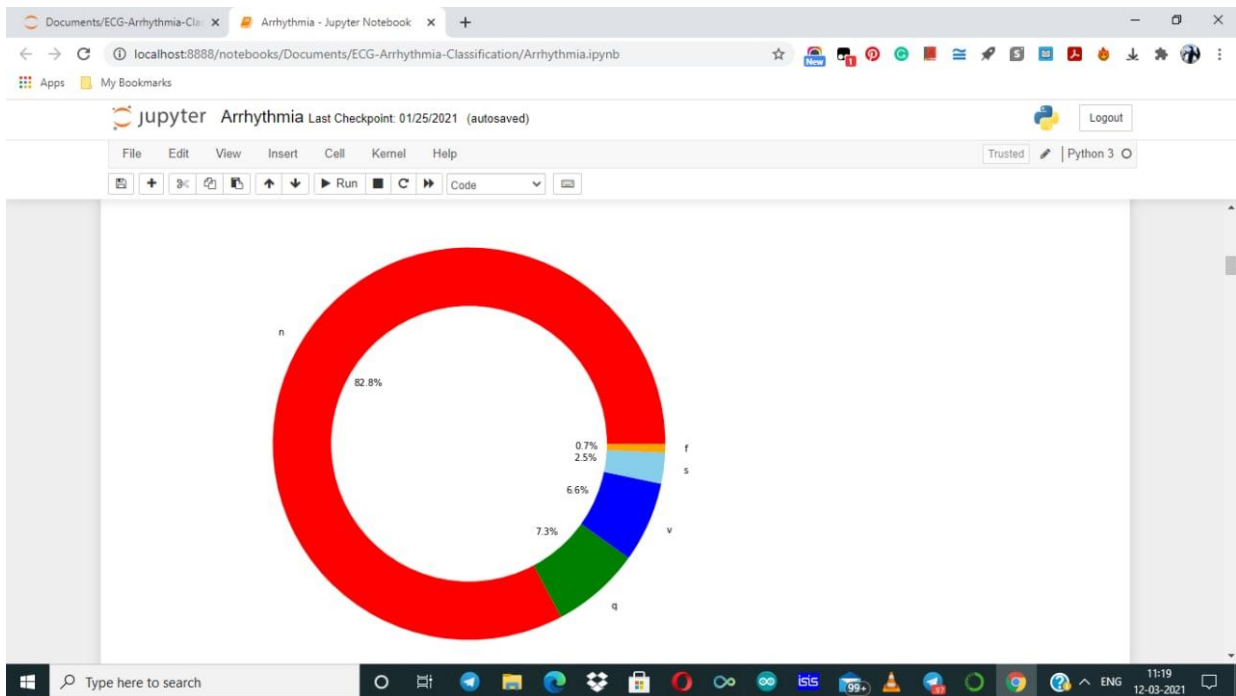


Fig 5.4

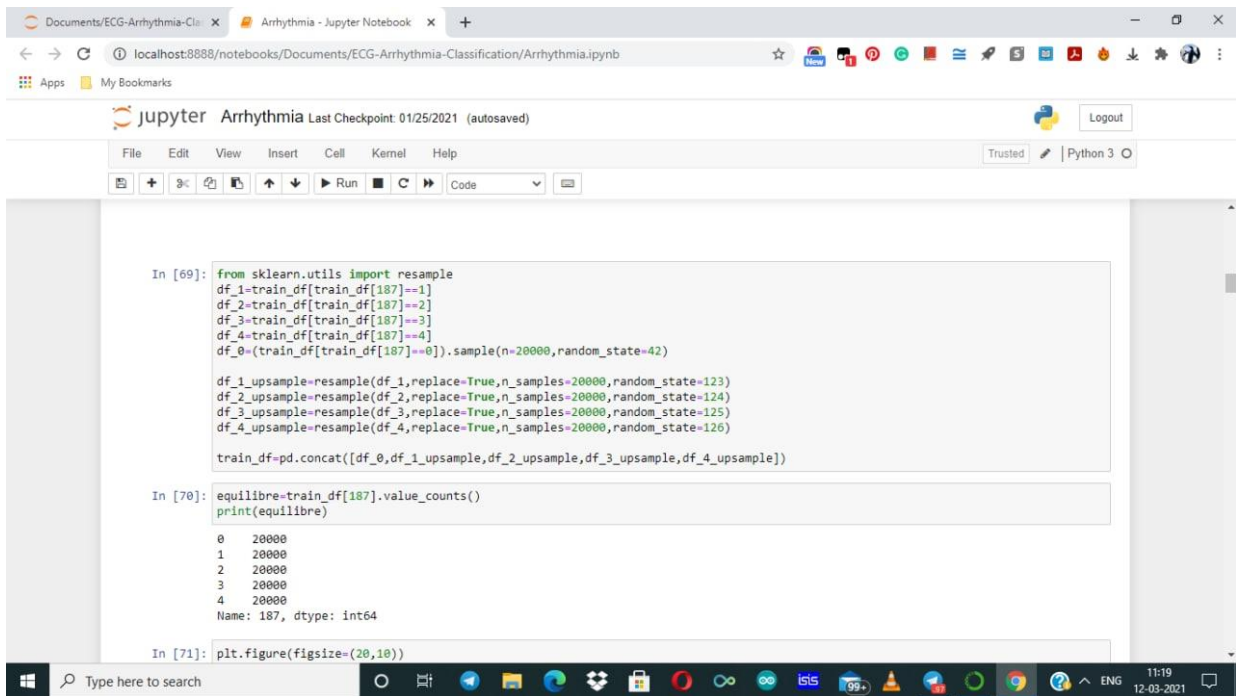


Fig 5.5

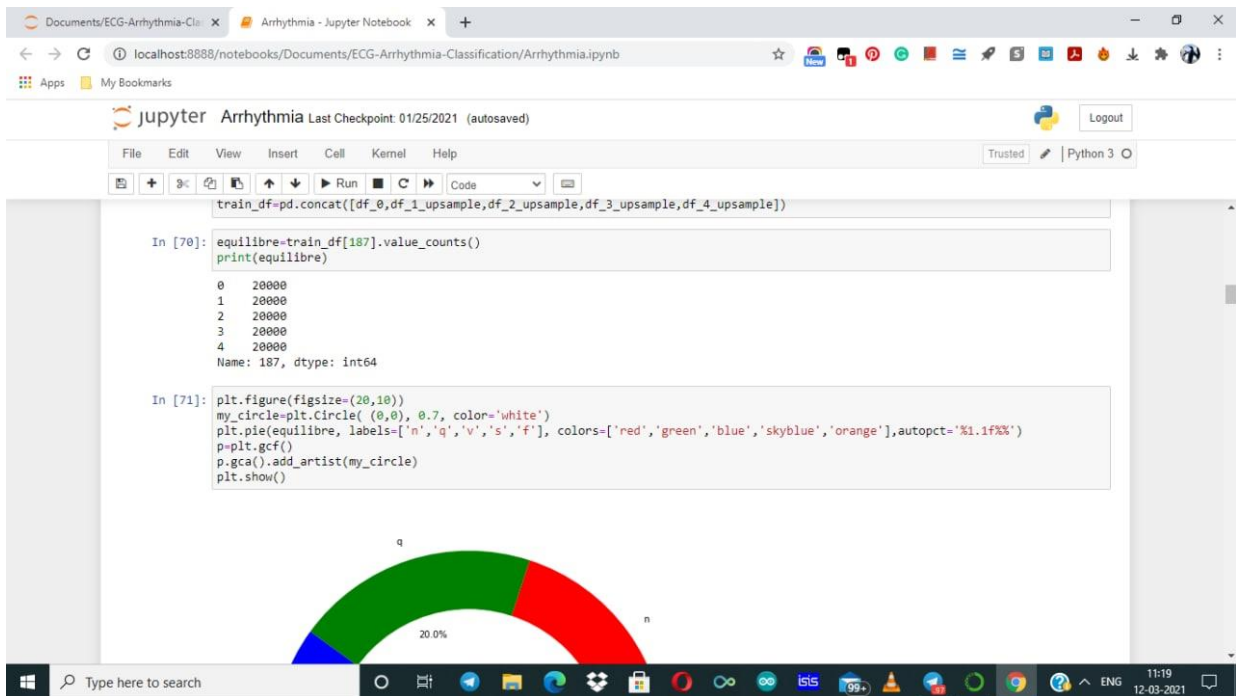


Fig 5.6

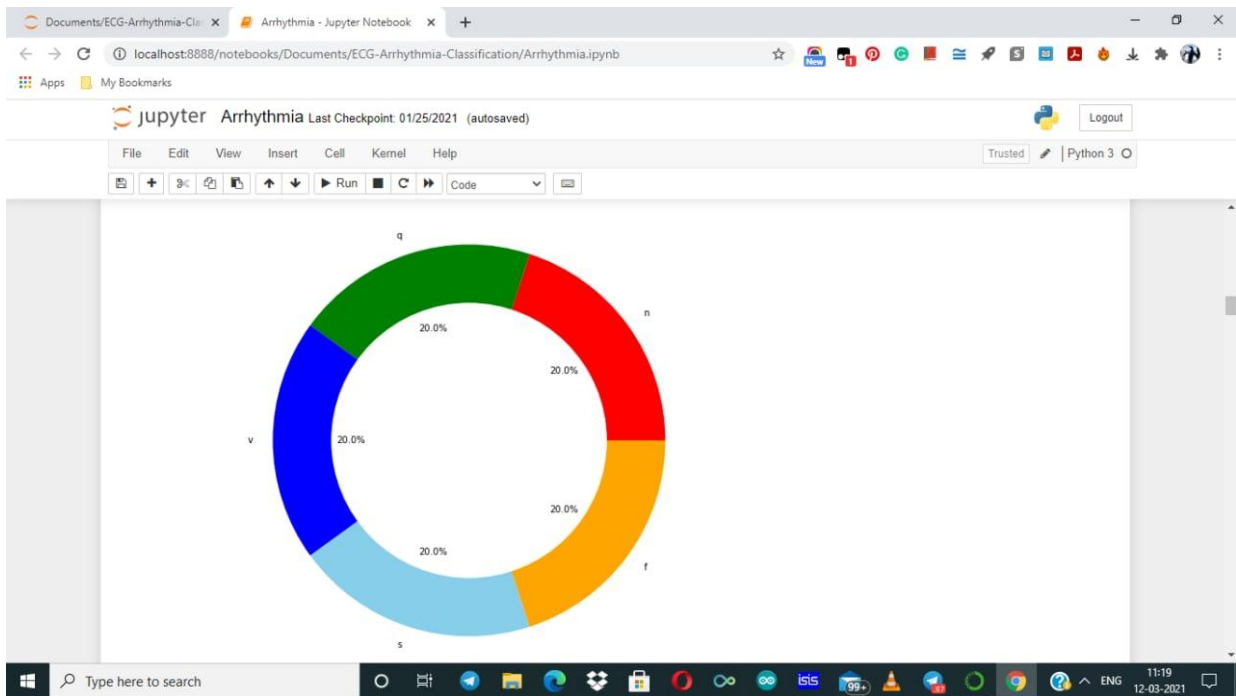


Fig 5.7

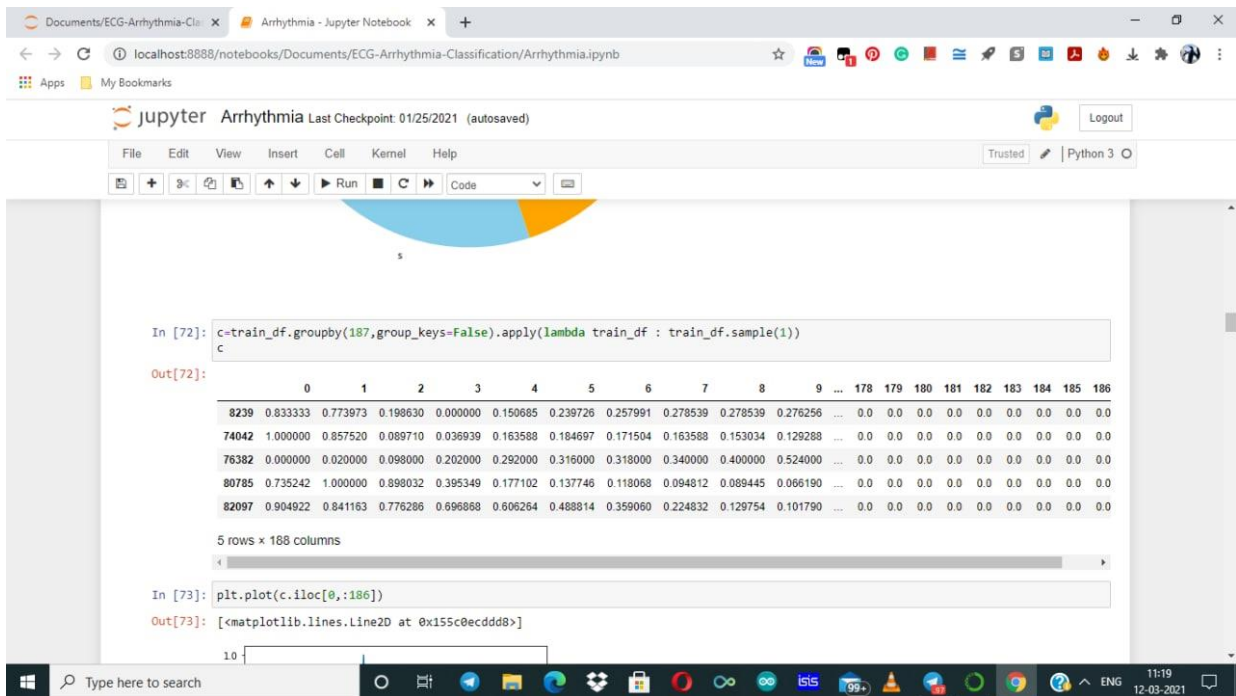


Fig 5.8

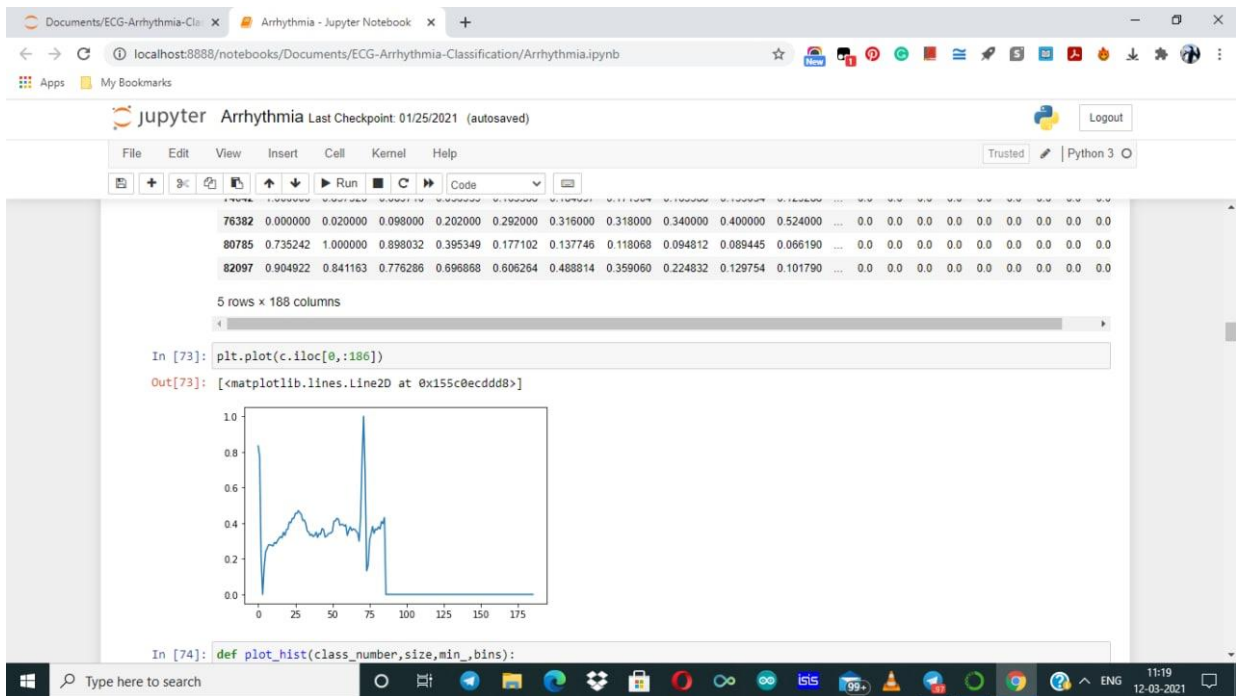


Fig 5.9

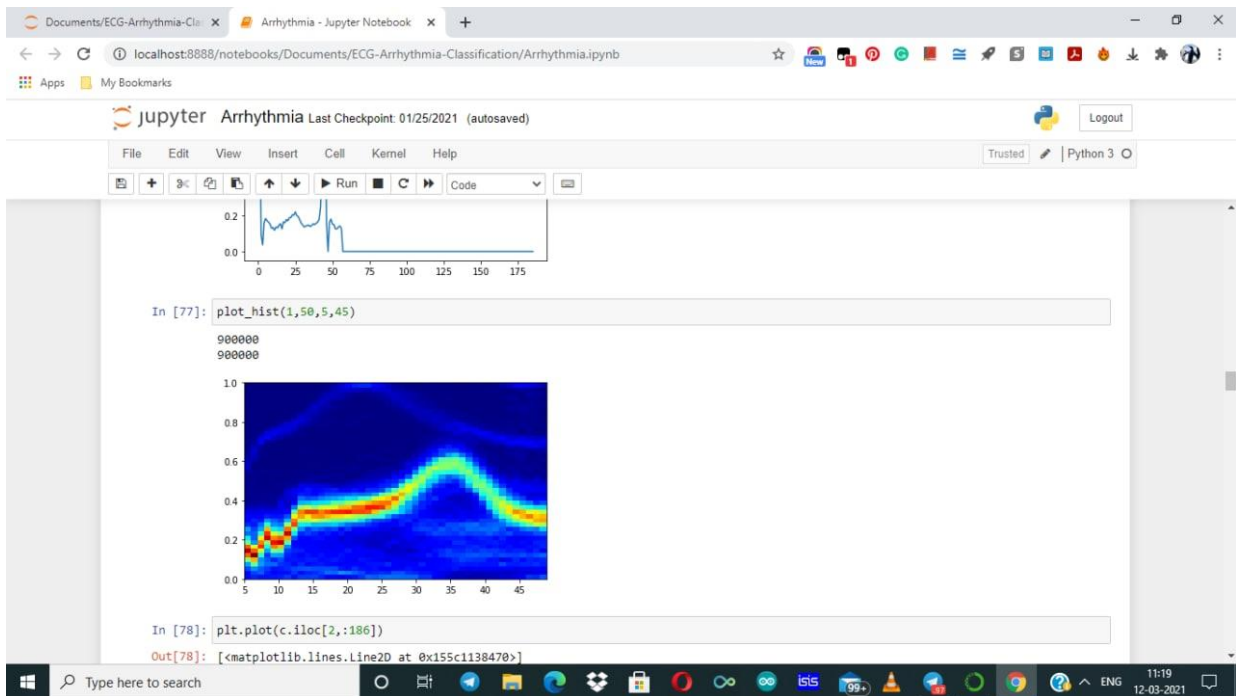


Fig 5.10

The screenshot shows a Jupyter Notebook titled 'Arrhythmia' with a last checkpoint of 01/25/2021. The code is as follows:

```

y_test=y_test_categorical(target_test)

In [87]: X_train=train_df.iloc[:,1:186].values
         X_test=test_df.iloc[:,1:186].values
         for i in range(len(X_train)):
             X_train[i,186]= add_gaussian_noise(X_train[i,1:186])
         X_train = X_train.reshape(len(X_train), X_train.shape[1],1)
         X_test = X_test.reshape(len(X_test), X_test.shape[1],1)

In [88]: def network(X_train,y_train,X_test,y_test):

         im_shape=(X_train.shape[1],1)
         inputs_cnn=Input(shape=(im_shape), name='inputs_cnn')
         conv1_1=Convolution1D(64, (6), activation='relu', input_shape=im_shape)(inputs_cnn)
         conv1_1=BatchNormalization()(conv1_1)
         pool1=MaxPool1D(pool_size=(3), strides=(2), padding='same')(conv1_1)
         conv2_1=Convolution1D(64, (3), activation='relu', input_shape=im_shape)(pool1)
         conv2_1=BatchNormalization()(conv2_1)
         pool2=MaxPool1D(pool_size=(2), strides=(2), padding='same')(conv2_1)
         conv3_1=Convolution1D(64, (3), activation='relu', input_shape=im_shape)(pool2)
         conv3_1=BatchNormalization()(conv3_1)
         pool3=MaxPool1D(pool_size=(2), strides=(2), padding='same')(conv3_1)
         flatten=Flatten()(pool3)
         dense_end1 = Dense(64, activation='relu')(flatten)
         dense_end2 = Dense(32, activation='relu')(dense_end1)
         main_output = Dense(5, activation='softmax', name='main_output')(dense_end2)

         model = Model(inputs= inputs_cnn, outputs=main_output)

```

Fig 5.11

The screenshot shows the continuation of the Jupyter Notebook. The code is as follows:

```

         im_shape=(X_train.shape[1],1)
         inputs_cnn=Input(shape=(im_shape), name='inputs_cnn')
         conv1_1=Convolution1D(64, (6), activation='relu', input_shape=im_shape)(inputs_cnn)
         conv1_1=BatchNormalization()(conv1_1)
         pool1=MaxPool1D(pool_size=(3), strides=(2), padding='same')(conv1_1)
         conv2_1=Convolution1D(64, (3), activation='relu', input_shape=im_shape)(pool1)
         conv2_1=BatchNormalization()(conv2_1)
         pool2=MaxPool1D(pool_size=(2), strides=(2), padding='same')(conv2_1)
         conv3_1=Convolution1D(64, (3), activation='relu', input_shape=im_shape)(pool2)
         conv3_1=BatchNormalization()(conv3_1)
         pool3=MaxPool1D(pool_size=(2), strides=(2), padding='same')(conv3_1)
         flatten=Flatten()(pool3)
         dense_end1 = Dense(64, activation='relu')(flatten)
         dense_end2 = Dense(32, activation='relu')(dense_end1)
         main_output = Dense(5, activation='softmax', name='main_output')(dense_end2)

         model = Model(inputs= inputs_cnn, outputs=main_output)
         model.compile(optimizer='adam', loss='categorical_crossentropy',metrics = ['accuracy'])

         callbacks = [EarlyStopping(monitor='val_loss', patience=8),
                      ModelCheckpoint(filepath='best_model.h5', monitor='val_loss', save_best_only=True)]

         history=model.fit(X_train, y_train,epochs=10,callbacks=callbacks, batch_size=32,validation_data=(X_test,y_test))
         model.load_weights('best_model.h5')
         return(model,history)

In [89]: def evaluate_model(history,X_test,y_test,model):
         scores = model.evaluate(X_test,y_test, verbose=0)

```

Fig 5.12

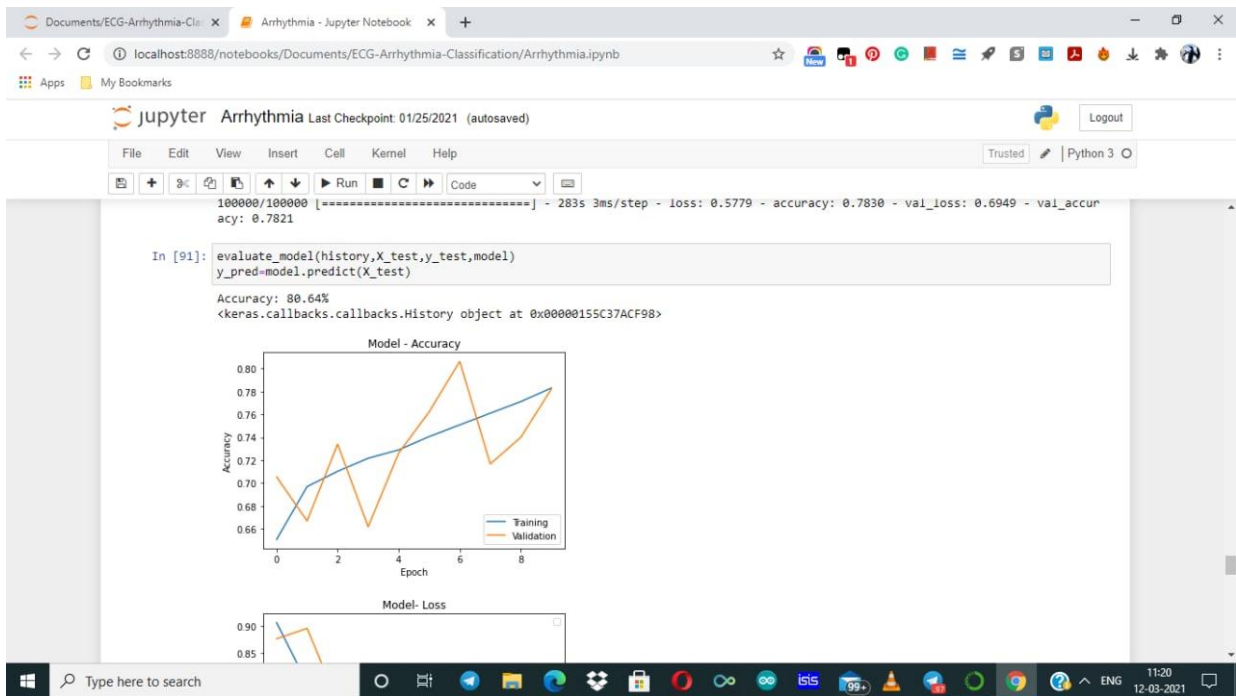


Fig 5.13

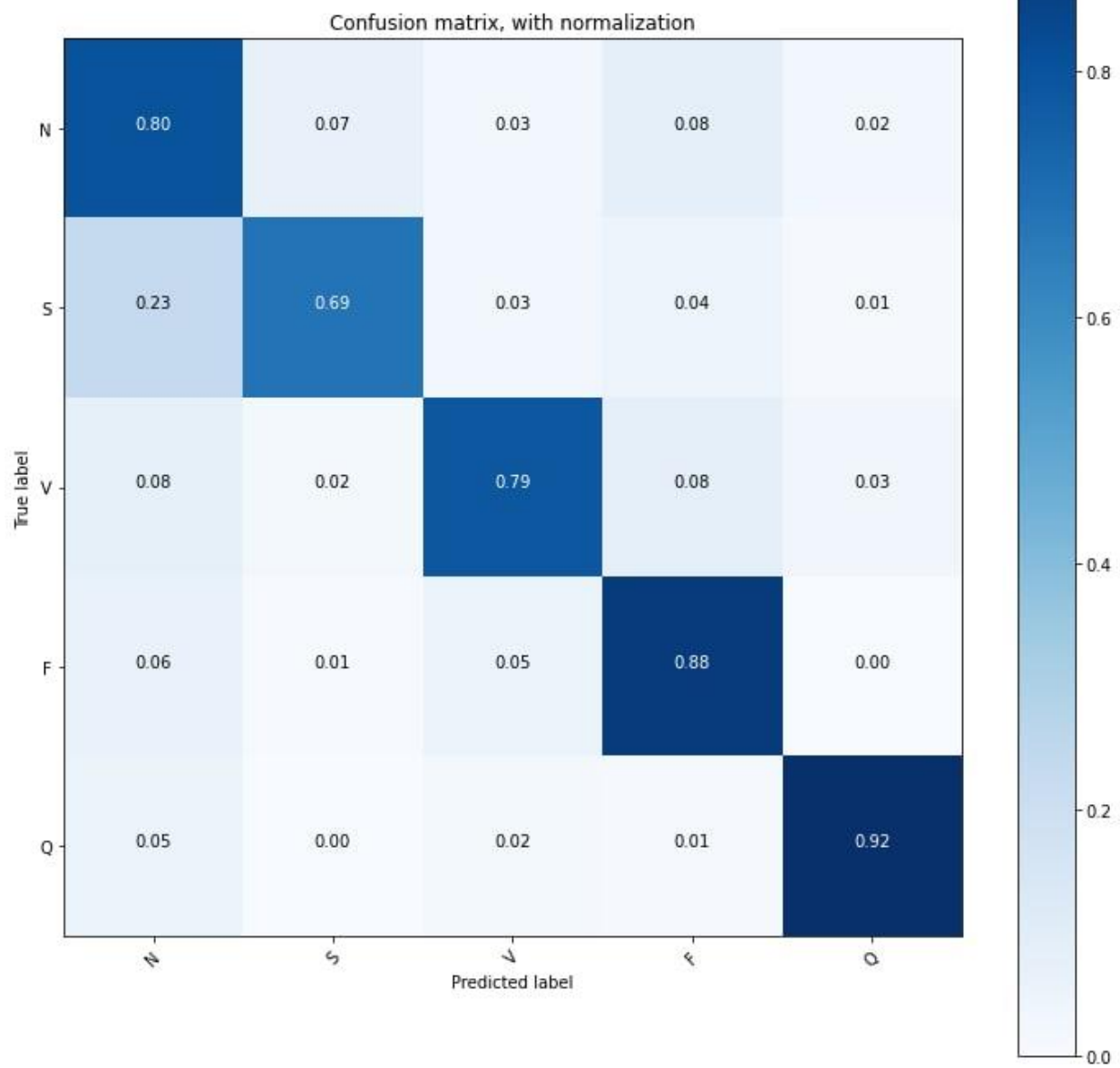


Fig 5.14

CHAPTER 6

CONCLUSION

In this project, a CNN-based organization procedure for programmed organization of cardiovascular utilizing cardiogram signals is proposed. A specific arrangement of cardiogram signals is useful in the inhibition and diagnosis of CVDs. Deep CNN has demonstrated beneficial in enhancing the accuracy of diagnosis algorithms in the fusion of medicine and modern machine learning technologies. The projected CNN-based classification algorithm, by means of dataset, will be classify eight types of Arrhythmia and it accomplished ninety seven percentage average sensitivity, ninety nine percentage specificity, ninety nine percentage average accuracy, and ninety eight percentage positive predictive value (precision). These outcomes specify that the forecast and organization of arrhythmia with cardiogram representation as spectrograms and the CNN prototypical is a dependable functioning model in the diagnosis. The planned scheme can help experts diagnose CVDs by referring to the automated classification of cardiogram signals. The project helpful only a single-lead cardiogram indication.

REFERENCES

- 1 Prediction of cardiopathy Using Machine Learning Authors: Aditi Gavhane ; Gouthami Kokkula ; Isha Pandya ; Prof. Kailas Devadkar,2018 Second Worldwide Meeting on Electronics, Communication and Aerospace Technology, 2019.
- 2 Cardiopathy Prediction using Evolutionary Rule Learning Authors:Aakash Chauhan ; Aditya Jain ; Purushottam Sharma ; Vikas Deep, 2018 Worldwide Meeting on Computational Intelligence & Communication Technology, 2018.
- 3 Blood Thickness based Cardiopathy Risk Prediction technique in Edge/Fog Computing Authors: R Latha ; P Vetrivelan,2019 11th Worldwide Meeting on Communication Systems & Networks, 2017

127_37130136_37130135(3)

ORIGINALITY REPORT

34%

SIMILARITY INDEX

17%

INTERNET SOURCES

23%

PUBLICATIONS

13%

STUDENT PAPERS

PRIMARY SOURCES

1

Dhara B. Mehta, Nirali C. Varnagar.

6%

"Newfangled Approach for Early Detection and Prevention of Ischemic Heart Disease using Data Mining", 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019

Publication

2

R Latha, P Vetrivelan. "Blood Viscosity based Heart Disease Risk Prediction Model in Edge/Fog Computing", 2019 11th International Conference on Communication Systems & Networks (COMSNETS), 2019

3%

Publication