**Heart Disease Prediction**

**An Engineering Project in Community Service – DSN3009**

**Phase – II Report**

***Submitted by***

**Team Members List**

**Sl. No. Register Number Name**

1. **20BCE10760 Uzer Khan**
2. **20BAI10240 Aditya Kumar Verma**
3. **20BCE10881 Nimish Sarathe**
4. **20BCE10506 Bharti Gattani**
5. **20BCE10726 Anoop Parashar**
6. **20BCE10023 Ayush Mishra**
7. **20BCE10773 Nimish Singh Chouhan**
8. **20BAI10350 Raj Dama**

***in partial fulfillment of the requirements for the degree of***

***Bachelor of Engineering and Technology***

****

**VIT Bhopal University**

**Bhopal(M.P.)**

**May 2023**

**Sl. No. Topic Page No.**

1. Introduction
   1. Motivation
   2. Objective
2. Existing Work
3. Proposed Work
4. System Architecture
5. Working Principle
6. Expected Result
7. Individual Contribution
8. Conclusion
9. References

****

**Bonafide Certificate**

Certified that this project report titled **“Heart Disease Prediction”** is the Bonafede work of “20BCE10760 Uzer khan, 20BAI10240 Aditya Kumar Verma, 20BCE10881 Nimish Sarathe, 20BCE10506 Bharti Gattani, 20BCG10023 Ayush Mishra, 20BCE10773 Nimish Chouhan, 20BC10726 Anoop Parashar, 20BAI10350 Raj Dama” who carried out the project work under my supervision.

This project report (Phase II) is submitted for the Project Viva-Voce examination held on 15 May 2023.

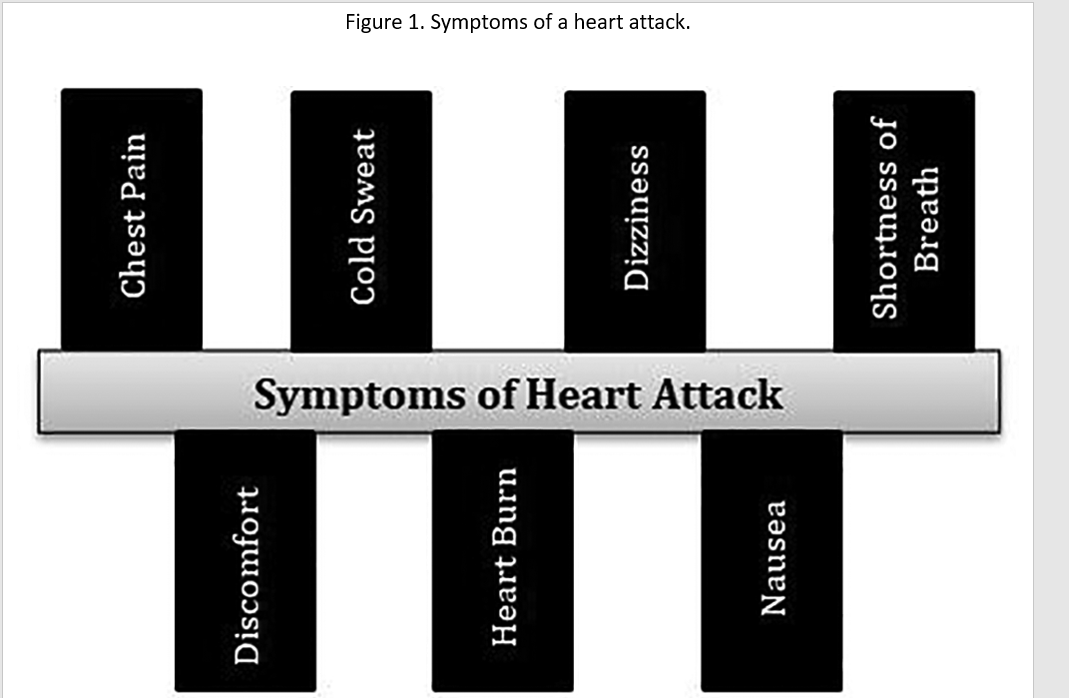
**Supervisor**

**Comments & Signature (Reviewer 1)**

**Comments & Signature (Reviewer 2)**

# 1-Introduction

A heart complaint that is similar to acute myocardial infarction (AMI) is one of the most serious conditions in the member of the cardiovascular complaint. It occurs due to the interruption of blood rotation to the muscle of the heart which damages the heart the muscle. Diagnosing heart complaints is also a pivotal task. The symptoms, physical examination, and understanding of the different signs of this complaint are needed to diagnose heart complaints. Different factors including cholesterol, inheritable heart complaint, high blood pressure, low physical exertion, rotundity, and smoking can be reasons for the circumstance of heart complaints. The major reason for heart conditions is the cessation of blood to the arrhythmia highways. The red blood cells (RBC) start getting low when blood inflow is reduced; due to this the mortal body stops getting necessary oxygen and loses knowledge. The early opinion through symptoms and signs can help cases of heart conditions if the vaticination is accurate enough. Figure 1 shows different symptoms of a heart complaint. The work presented takes 13 features/ attributes as input having number values. It has been stated that little variations in life including quitting smoking/ alcohol/ tobacco, having healthy food habits, and routine exercises can help in the forestallment of heart conditions. Any person living a healthy life with early treatment after opinion can greatly increase the positive results. still, it's delicate to identify the high threat of heart complaints where different pitfalls like diabetes, high blood pressure, and cholesterol problems are present. It's a system that's made by the use of deep literacy algorithms for guessing the possible conditions grounded on the case’s symptoms. The growth of technology has been perfecting our lives so far. It provides numerous tools that can save millions of lives, and deep literacy is one of them. Deep literacy is used to develop systems that can help us prognosticate so numerous conditions grounded on symptoms. It can suggest the croaker, and probability of the possible conditions. And opinion can be done grounded on the suggestion, therefore cost could be reduced. We're living in the age of technology and currently, humans can say that nearly anything is possible with the help of technology. moment we've so numerous tools and styles to pierce information from any region of this world and Information at this age is so important that without information we'd not survive. We've tools that can give us or suggest applicable information at our fingertips and the internet is one of those tools. moment billions of hunt queries are performed daily and occasionally their given results are applicable and occasionally they're not. In those hunt queries, thousands of quests are related to medical advice. People frequently want to know if they've any serious conditions grounded on their signs and symptoms. But there are no tools available to give them proper information. This exploration tries to give them tools so that possible complaint vaticination information can be handed to the end-stoner.



## 1.1 Motivation

The main motivation for doing this research is to present a heart disease prediction model for the prediction of the occurrence of heart disease. Further, this research work is aimed at identifying the best classification algorithm for identifying the possibility of heart disease in a patient. This work is justified by performing a comparative study and analysis using Convolutional neural networks used at different levels of evaluation. Although these are commonly used deep learning algorithms, the heart disease prediction is a vital task involving highest possible accuracy. Hence, the three algorithms are evaluated at numerous levels and types of evaluation strategies. This will provide researchers and medical practitioners to establish a better understanding and help them identify a solution to identify the best method for predicting the heart diseases.

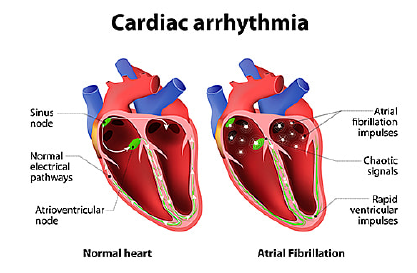
## 1.2 Objective

### **1.2.1Heart disease in the context of deep learning**

Previous works have declared that prediction can be improved with the application of feature selection and proper engineering. An experiment with different deep learning approaches and models by tuning various hyper-parameters has been performed and improved the performance with optimized accuracy. Neural networks performed well when compared to other deep learning methods Sequential, Dense, Convolutional 2D, Max Pooling, Flattening. Along with Rela and SoftMax activation functions. Other researchers worked on the reduction of cardiovascular features and extracted nonlinear features with discriminant analysis.[2](https://f1000research.com/articles/11-1126#ref2) Fisher was utilized for the experiment’s purpose to tackle overfitting problems and to improve the training speed. Results stated that 100% accuracy has been shown for the detection of arrhythmia disease.

# 2-Existing Work / Literature Review

Cardiovascular conditions (CVDs) are the number one cause of death moment. Over17.7 million people failed from CVDs in the time 2017 each over the world which is about 31 of all deaths, and over 75 of these deaths do in low and middle- income countries. Arrhythmia is a representative type of CVD that refers to any irregular change from the normal heart measures. There are several types of arrhythmias including atrial fibrillation, unseasonable compression, ventricular fibrillation, and tachycardia.

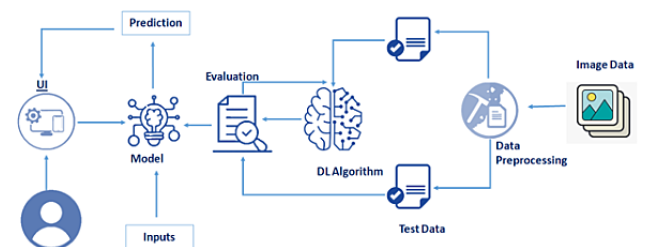


# 3-Proposed Work –

An” itinerant electrocardiogram" or an ECG) about the size of a card or digital camera that the case will be using for 1 to 2 days, or over to 2 weeks. The test measures the movement of electrical signals or swells through the heart. These signals tell the heart to contract(squeeze) and pump blood. The case will have electrodes taped to your skin. It's effortless, although some people have mild skin vexation from the tape recording used to attach the electrodes to the chest. They can do everything but shower or bathe while wearing the electrodes. After the test period, case will go back to see your croaker. They will be downloading the information

**4-System Design / Architecture-**

We will prepare the project by following the stages outlined below:

1. utilizing Sequential modeling.
2. utilizing Keras capabilities.
3. utilizing image processing techniques.
4. the Flask framework to create a web application.

**5-Working Principle**

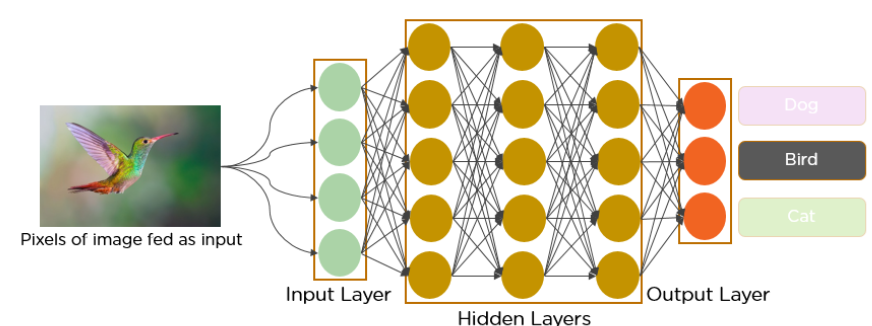
In this design, we've stationed our training model using CNN on IBM Watson plant and in our original deep. We're planting 4 types of CNN layers in a successional manner, starting from.

**1-Convolutional Layer-** 2DA 2- D convolutional subcaste applies sliding convolutional pollutants to 2- D input. The subcaste convolves the input by moving the pollutants along the input vertically and horizontally and calculating the fleck product of the weights and the input, and also adding a bias term.

**2-Pooling Layer** - Subcaste Pooling layers are used to reduce the confines of the point maps. therefore, it reduces the number of parameters to learn and the quantum of calculation performed in the network. The pooling subcaste summarises the features present in a region of the point chart generated by a complication subcaste.

**3-Fully- Connected Layer**- After rooting features from multiple complication layers and pooling layers, the completely- connected subcaste is used to expand the connection of all features. Eventually, the SoftMax subcaste makes a logistic retrogression bracket. Completely- connected subcaste transfers the weighted sum of the affair of the former subcaste to the activation function.

**4-Dropout Layer** - There’s generally a powerhouse subcaste before the fully connected subcaste. The powerhouse subcaste will temporarily dissociate some neurons from the network according to the certain probability during the training of the complication neural network, which reduces the common rigidity between neuron bumps, reduces overfitting, and enhances the conception capability of the network.



**5.2-ALGORITHM FOR DISEASE PREDICTION SYSTEM-**

**5.2.1-Dataset Collection:**

The dataset contains six classes:

1. Left Bundle Branch Block
2. Normal
3. Premature Atrial Contraction
4. Premature Ventricular Contractions
5. Right Bundle Branch Block
6. Ventricular Fibrillation

Image Preprocessing: Image Pre-processing includes the following main tasks.

Import Image Data Generator Library: Image data augmentation is a technique that can be used to artificially expand the size of a training dataset by creating modified versions of images in the dataset. The Keras deep learning neural network library provides the capability to fit models

using image data augmentation via the Image Data Generator class.

**5.2.2-Configure Image Data Generator Class:**

There are five main types of data augmentation techniques for image data; specifically:

1. Image shifts via the width\_shift\_range and height\_shift\_range arguments.
2. Image flips via the horizontal\_flip and vertical\_flip arguments.
3. Image rotates via the rotation\_range argument
4. Image brightness via the brightness\_range argument.
5. Image zooms via the zoom\_range argument.

An instance of the Image Data Generator class can be constructed for train and test.

**Applying Image Data Generator functionality to the trainset and test set**:

This function will return batches of images from the subdirectories -

Left Bundle Branch Block, Normal, Premature Atrial Contraction, Premature Ventricular Contractions, Right Bundle Branch Block and Ventricular Fibrillation, together with labels 0 to 5

{'Left Bundle Branch Block': 0, 'Normal': 1, 'Premature Atrial Contraction': 2, 'Premature Ventricular Contractions': 3, 'Right Bundle Branch Block': 4, 'Ventricular Fibrillation': 5}

We can see that for training there are 15341 images belonging to 6 classes and for testing there are 6825 images belonging to 6 classes.

**Model Building**

We are ready with the augmented and pre-processed image data; we will begin our build our model by following the below steps:

1. Import the model building Libraries
2. Initializing the model
3. Keras has 2 ways to define a neural network
4. Sequential
5. Function API

The Sequential class is used to define linear initializations of network layers which then,

collectively, constitute a model. In our example below, we will use the Sequential

constructor to create a model, which will then have layers added to it using the add ()

method. Now, will initialize our model.

**Adding CNN Layers**:

We are adding a convolution layer with an activation function as “rely” and with a small filter size (3,3) and a number of filters as (32) followed by a max-pooling layer. The Max pool layer is used to down sample the input, the flatten layer flattens the input.

**Adding Hidden Layers:**

Dense layer is deeply connected neural network layer. It is most common and frequently used layer.

**Adding Output Layer:**

Understanding the model is very important phase to properly use it for training and prediction purposes. Keras provides a simple method, summary to get the full information about the model and its layers.

**Configure the Learning Process:**

The compilation is the final step in creating a model. Once the compilation is done, we

can move on to the training phase. The loss function is used to find error or deviation in

the learning processes. Keras requires loss function during the model compilation process.

**Optimization** is an important process that optimizes the input weights by comparing the prediction and the loss function. Here we are using Adam optimizer Metrics is used to evaluate the performance of your model. It is similar to loss function, but not used in the training process.

**Training the model:**

We will train our model with our image dataset. fit generator functions used to train a deep learning neural network

**Saving the model:**

The model is saved with .h5 extension as follows An H5 file is a data file saved in the Hierarchical Data Format (HDF). It contains multidimensional arrays of scientific data.

**Testing the model:**

Load necessary libraries and load the saved model using load\_model. Taking an image as input and checking the results

Note: The target size should for the image that is should be the same as the target size that you have used for training.

**7-Expected Results-**

The analysis and identification of the best classification algorithm in this research work is done and the results are provided here. For the validation of the results, several ranges of experiments are carried out using Cross validation and Percentage split methods which are described in the sections given below.

In this work, the evaluation of the performance metrices are being done with four Deep learning methods where, For the project we firstly take an ECG Image dataset of 6 classes

1. Left Bundle Branch Block
2. Normal
3. Premature Atrial Contraction
4. Premature Ventricular Contractions
5. Right Bundle Branch Block
6. Ventricular Fibrillation

The training data will have 15341 images while the testing dataset will have 6825 images. For Image Preprocessing from the Keras library, using the Deep learning modules Dense, Sequential, Convolutional, Pooling, Flattening.

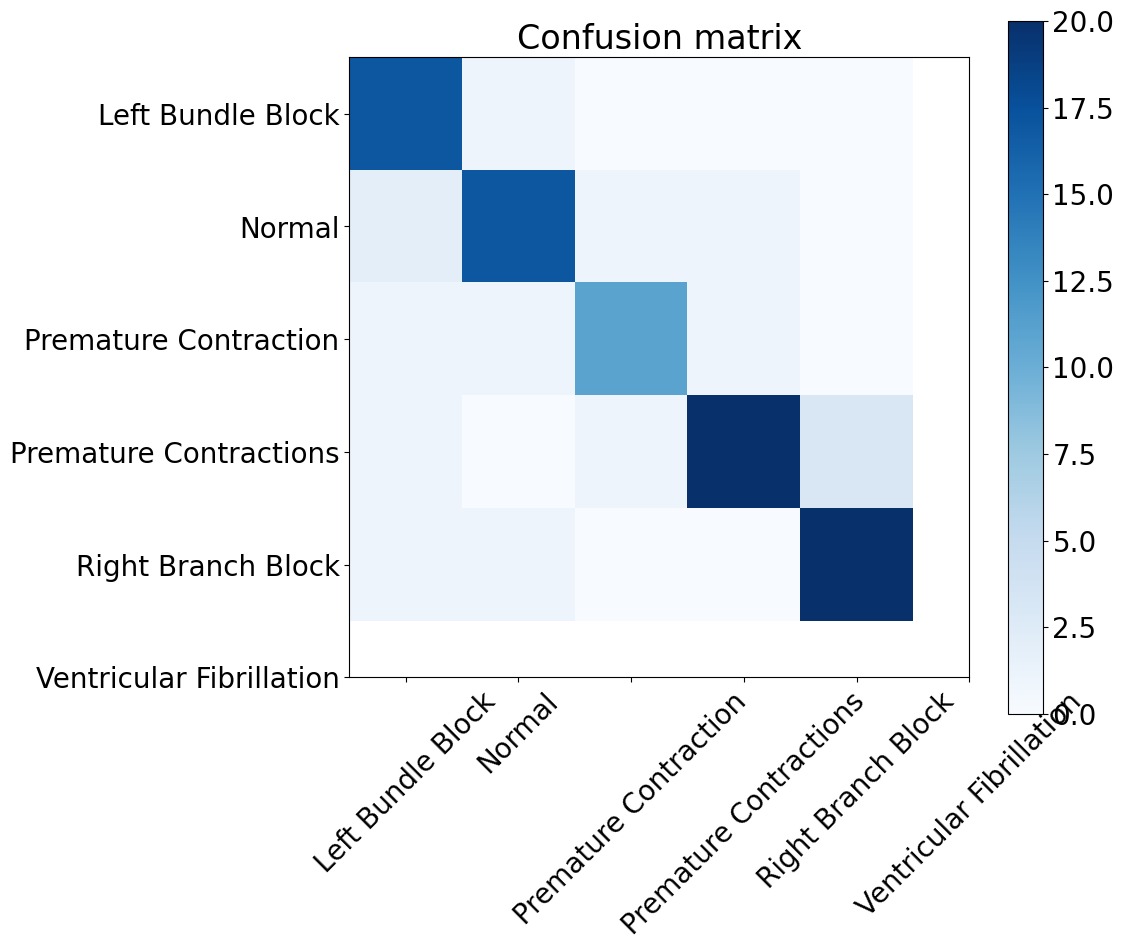
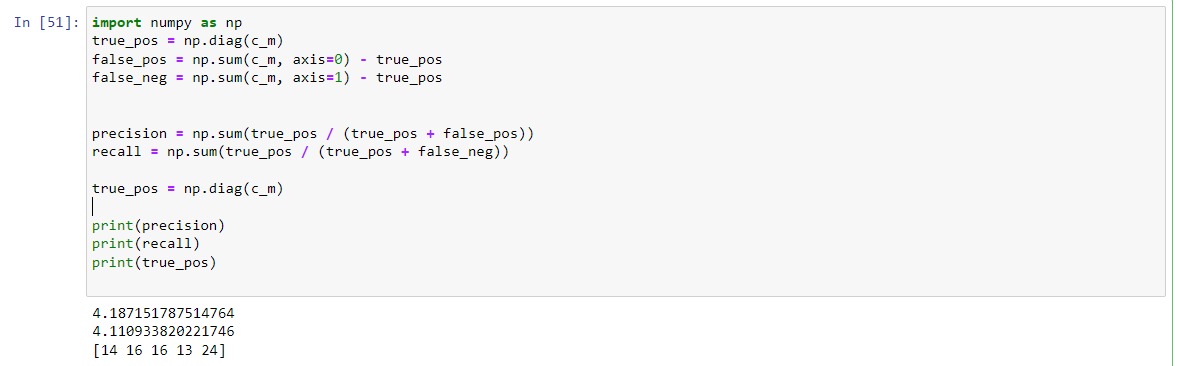
Convolution filters of 32x32 will be introduced with image kernel being 64x64 size with Re-Lu activation applied onto the hidden layers. The Convolutional Neural Network would result to a 9248 (2D Layer parameters.

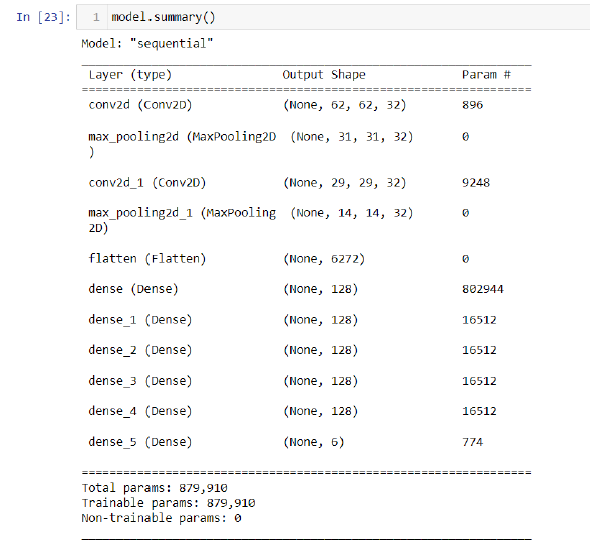
Alongside a total 879910 parameters to be trained inside the CNN. After, using the Adam optimizer we would result to a Desired Accuracy of 85.26 % to an Actual accuracy being 85.26. While the model’s expected loss is 0.1% and its actual loss is 0.61%. At the end of the successful model training, we obtain a model that can accurately predict and distinguish between different type of chronic heart conditions based on the image given as the input format. We, then integrated the model to a live webapp so that the prediction of the disease can be predicted and can be seen visually using a React.JS hosted website with its backend built on flask, alongside useful information cards present on the home page and contact page as a query-receiving response page. In this section, we will be building a web application that is integrated into the model we built. A UI is provided for the uses where he has uploaded an image. The uploaded image is given to the saved model and prediction is showcased on the UI.

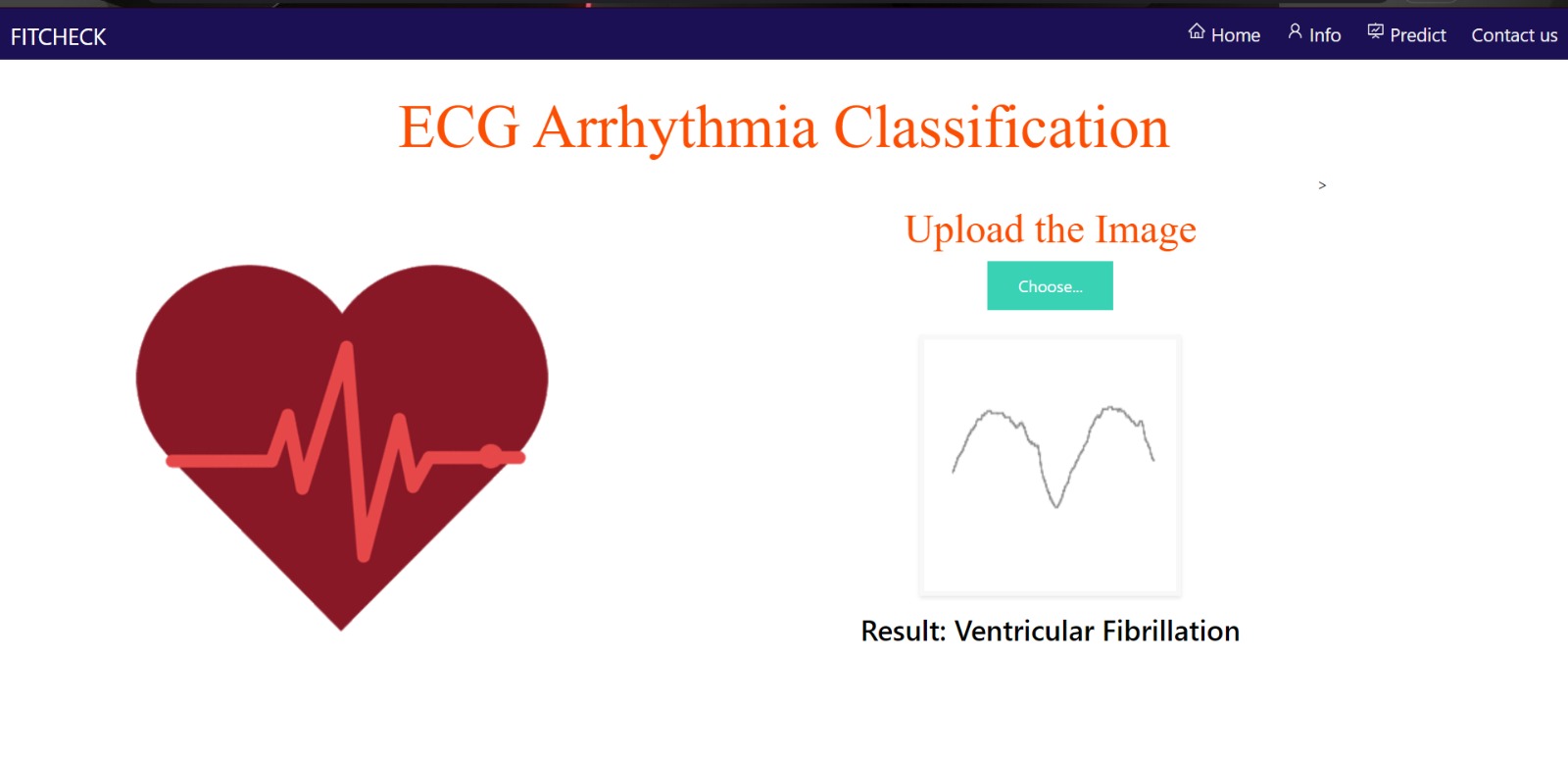
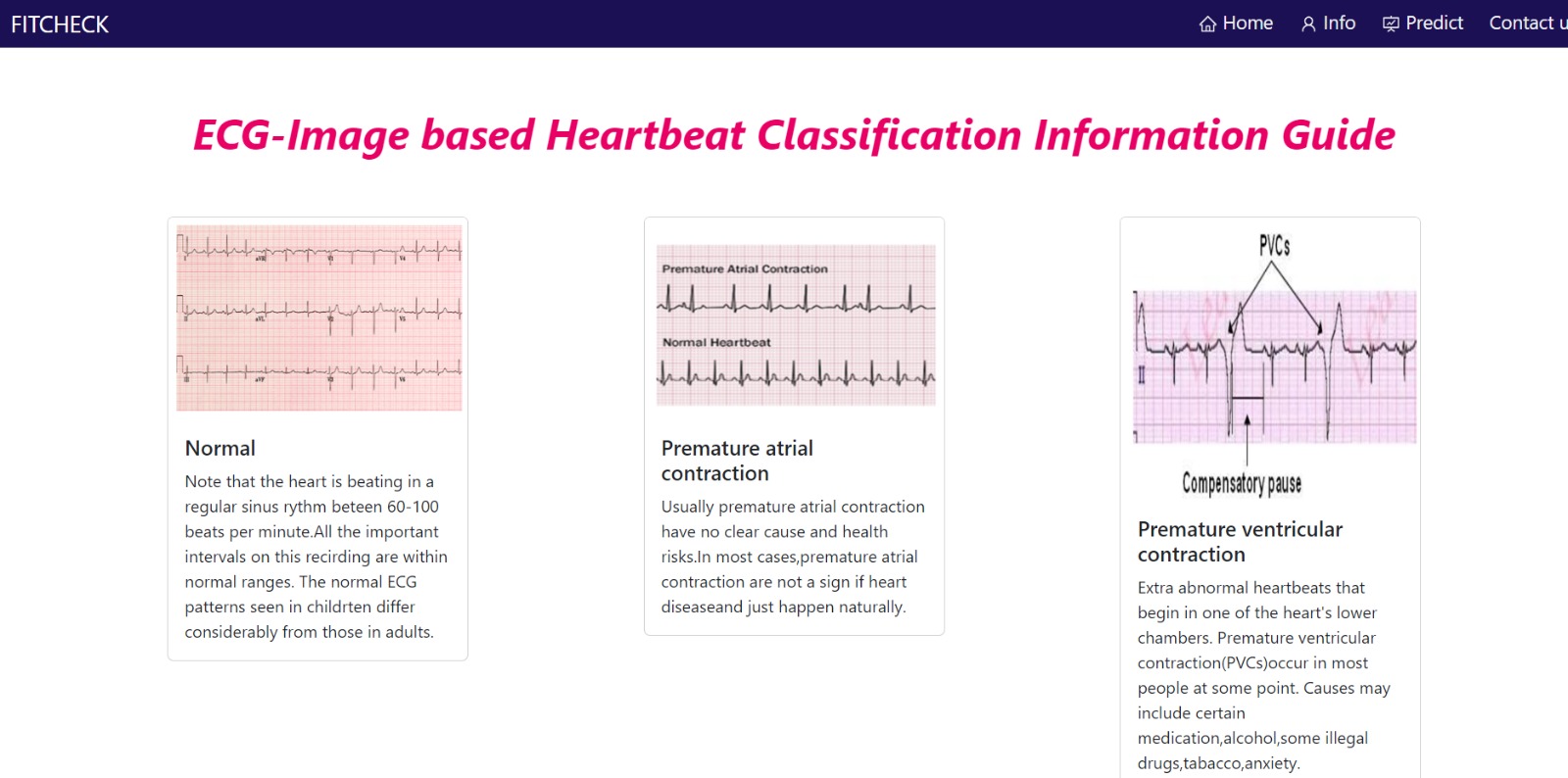
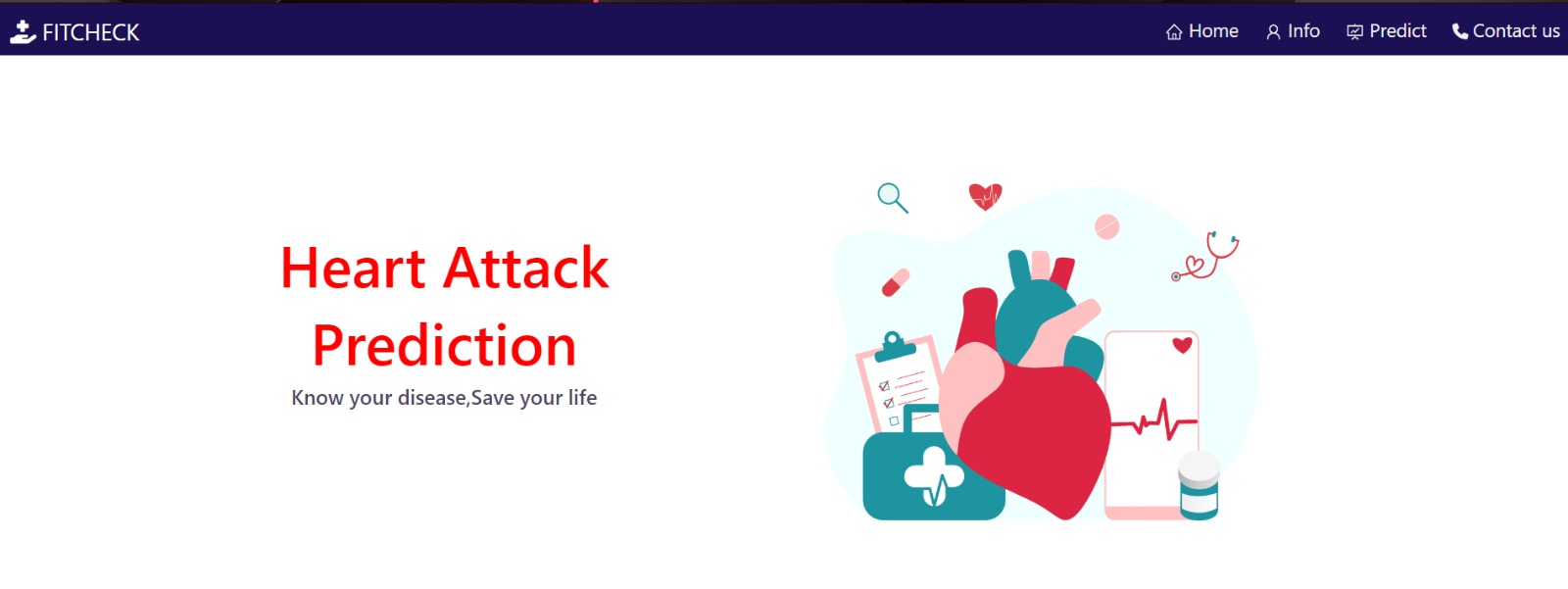
After the result we form a confusion matrix which provides us with an overall accuracy of 98.35 ,where the Precision is 4.18 and Recall bieng 4.11 of our deeep learning model.

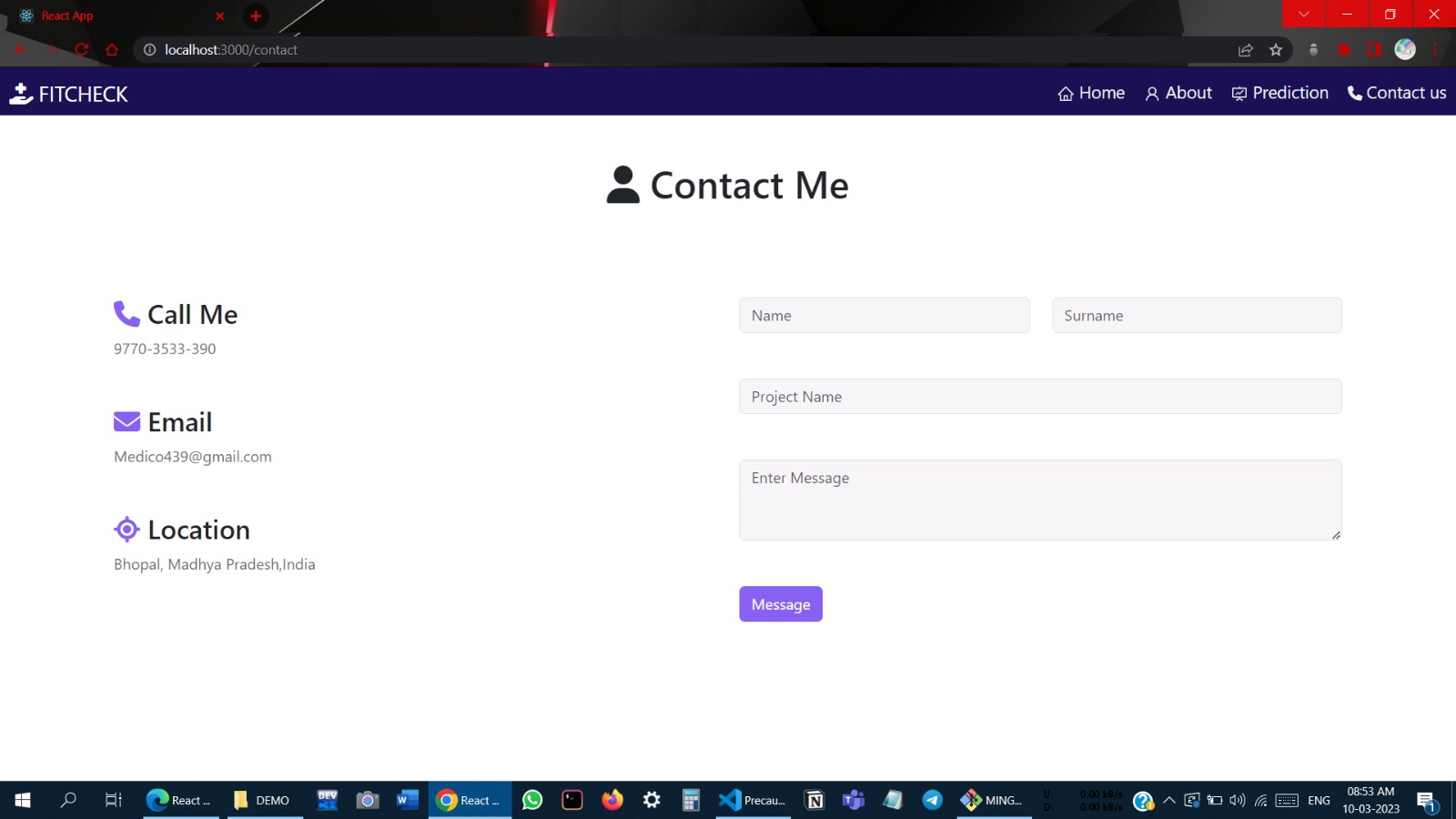
This section has the following tasks

1. Building HTML Pages:
2. We use HTML to create the front end part of the web page.
3. Here, we created 4 html pages- home.html, predict\_base.html, predict.html,
4. home.html displays the home page.
5. Info.html displays all important details to be known about ECG.
6. Predict-base.html and predict.html accept input from the user and predicts the values.
7. Contact.html – responsive contact query page









**7-INDIVIDUAL CONTRIBUTION**

**ADITYA KUMAR VERMA – 20BAI10240**

****

Developed The Deep Learning Model and other training part of the architecture of the project.

Data Searching and Idea formulation. He also different types of machine learning things in our project. Cleaning and Modification of Dataset: Responsible for modification of the dataset used in the AI model from Kaggle and modified it according to the project's needs and goals. This involved cleaning the data, removing any irrelevant information, and adding new data points that were necessary for the project. This helped in ensuring that the AI model could provide accurate and relevant predictions to the users.

**RAJ DAMA – 20BAI10350**

****

Developed The Deep Learning Model and other training part of the architecture of the project.

Data Validation and Pre-Processing: Responsible for validating the data collected from various sources. This involved ensuring that the data was accurate, complete, and relevant. Made sure that AI model is trained on a valid unbiased training data and the application could provide accurate and relevant information to the users. Data Validation and Pre-Processing: Responsible for validating the data collected from various sources. This involved ensuring that the data was accurate, complete, and relevant. Made sure that AI model is trained on a valid unbiased training data and the application could provide accurate and relevant information to the users

**UZER KHAN – 20BCE10760**

****

Integrate the AI model to the frontend by using flask. I am not familiar with the machine learning before this project but after doing this project I get the idea of the AI. The main problem is faced during the integration of this I am getting error all the time and I resolve all the bugs one by one and by installing different things like tenser flow, keras etc. This is really a very tough part. It takes a lot of time. Manage the whole team and plays the important role in design the frontend of the project and always contact to the ml developers which problem is coming in which part .and gives the resources and idea of how the problem will resolve

**NIMISH SARATHE – 20BCE10881**



Add Functionality and design the information page in which all four types of heart problem is there.Read and learn about the different types of heart problem occur in human being like normal, atrial contraction, etc. and guide us what type of things we have to for different things.

**BHARTI GATTANI – 20BCE10506**

****

Design the front page and in this, she added the functionality of the card. In the Cards we have mention the precaution we need to take so that one can save away from attacks like quit smoking, daily exercise eats fresh things She collects all the information and implements in our website.

UI Design: Took the lead in designing the UI of the website and ensured that the interface was visually appealing, intuitive, and easy to navigate for fusers. carefully selected appropriate colour schemes, typography, and layout to enhance the overall user experience.

**Nimish Singh Chouhan – 20BCE10773**

****

Icons in UI Design: Selected and imported high quality and animated icons and images for UI design of the heart disease application. This helped in enhancing the user experience of the application by providing helped in providing a better user experience for the user.Read and Studied about the Cardiac arrythmia and studied about the different things in heart like Sinus nondermal electrical pathway, atrioventricular node, atrial fibrate node, etc and tell us the developer to do things. The concept which we did not know he will guide us to do these things in project.

**ANOOP PARASHAR-20BCE10726**



Design and add functionality to the contact page. In this we have added different things like mail, phone number for contact. He also added the query section in which if the user has any query related to the issue he is facing, we will resolve that issue by this feature. Custom Layouts and Widgets: Developed custom layouts and widgets for the website, which helped in enhancing the user experience of the application. By creating custom layouts and widgets, we were able to provide a unique look and feel to the application that was both aesthetically pleasing and functional

**AYUSH MISHRA – 20BCG10023**

****

Software Testing and Debugging: Was responsible for ensuring that our website worked as expected. Tested the application and identified any bugs or errors that needed to be fixed. Further debugged by identifying the cause of issues and developed possible fix to resolve all the issues to ensure smooth running of the application.

# 8-CONCLUSION:

Cardiovascular disease is a major health problem in today's world. The early diagnosis of cardiac arrhythmia highly relies on the ECG. Unfortunately, the expert level of medical resources is rare, visually identify the ECG signal is challenging and time-consuming.

The advantages of the proposed CNN network have been put to evidence. It is endowed with an ability to effectively process the non-filtered dataset with its potential anti-noise features. Besides that, ten-fold cross-validation is implemented in this work to further demonstrate

the robustness of the network. With the help of a disease prediction system, it was possible to diagnose people based on symptoms. Disease prediction system provides only possible outcomes it does not guarantee that it will predict the disease correctly. But it has significantly higher accuracy for predicting possible diseases. In our research, we have analyzed the accuracy of this system for heart diseases and our accuracy can go up to 98.35%.

**FUTURE SCOPE**

It would be fascinating to investigate the usage of optimization techniques in the future to develop a feasible design and solution. The limitation of our study is that we have yet to use any optimization techniques to optimize the model parameters, and we expect that with the application of the optimization, the performance of the suggested solution will be able to be elevated to the next level.

**Tech Stack Used:** HTML, CSS, JAVASCRIPT, REACT.JS, PYTHON (SCIKIT LEARN, NUMPY, PANDAS, Flask), JUPYTER NOTEBOOK, GOOGLE COLLAB & VS CODE

# 9-Reference:

**OUR PROJECT LINK - https://github.com/Uzerkhan786/HEART\_ATTACK**

1- M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, “Disease prediction by deep learning over big data from healthcare communities” IEEE Access, vol. 5, no. 1, pp. 8869–8879, 2017.

**2-S**ayali Ambedkar, Rashmi Panikkar, “Disease Risk Prediction by Using Convolutional Neural Network” IEEE, 978-1-5386-5257-2/18, 2018.

3-Naganna Chetty, Kunwar Singh Vaisala and Nagma Patil, “An Improved Method for Disease Prediction using Fuzzy Approach” IEEE, DOI 10.1109/ICACCE.2015.67, pp. 569-572, 2015.

4- Dhiraj Dah wade, Gajanan Patel and Ekata Mesh ram, “Designing Disease Prediction Model Using Deep Learning Approach” IEEE Xplore Part Number: CFP19K25-ART; ISBN: 978-1-5386-7808-4, pp. 1211-1215, 2019.

5-Lambodar Jena and Ramakrishna Swain, “Chronic Disease Risk Prediction using Distributed Deep Learning Classifiers” IEEE, 978-1-5386-2924-6/17, pp. 170-173, 2017.