# Classification of Arrhythmia by Using Deep Learning With 2-D ECG spectral imageRepresentation

Done by

Team Leader: Prasanth A

Team Member: Rithan S

Team Member: Naveensridhar K

Team Member: Bharath S

Team Member: Prasanth R

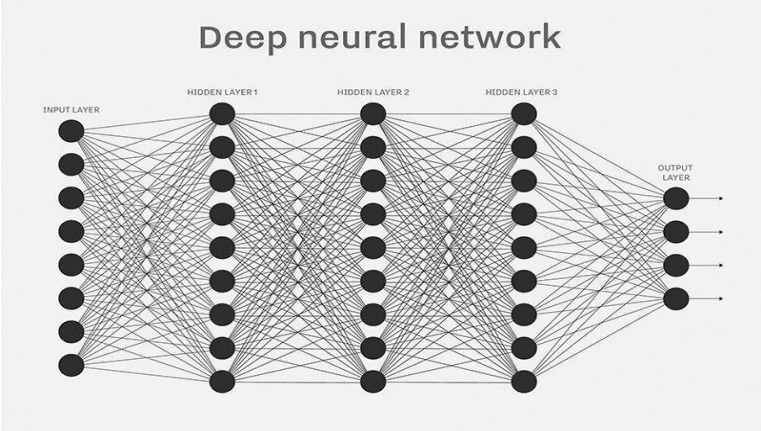
1. **INTRODUCTION**

### 1.1 Project Overview:

According to the World Health Organization (WHO), cardiovasculardiseases (CVDs) are the number one cause of death today. Over 17.7million people died from CVDs in the year 2017 all over theworld which isabout 31% of all deaths, and over 75% of these deaths occur in low andmiddle-income countries. Arrhythmia is a representative type of CVD thatrefers to any irregular change from the normal heart rhythms. There areseveral types of arrhythmia including atrial fibrillation, premature contraction, ventricular fibrillation, and tachycardia. Although a singlearrhythmia heartbeat may not have a serious impact on life, continuousarrhythmia beats can result in fatal circumstances. In this project, we buildan effective electrocardiogram (ECG) arrhythmia classification method using a convolution al neural network (CNN), in which we classify ECG intoseven categories, one being normal and the other six being different typesof arrhythmia using deep two-dimensional CNN with grayscale ECGimages. We are creating a web application where the user selects theimage which is to be classified. The image is fed into the model that istrained and the cited class will be displayed on the webpage

### 1.2 Purpose

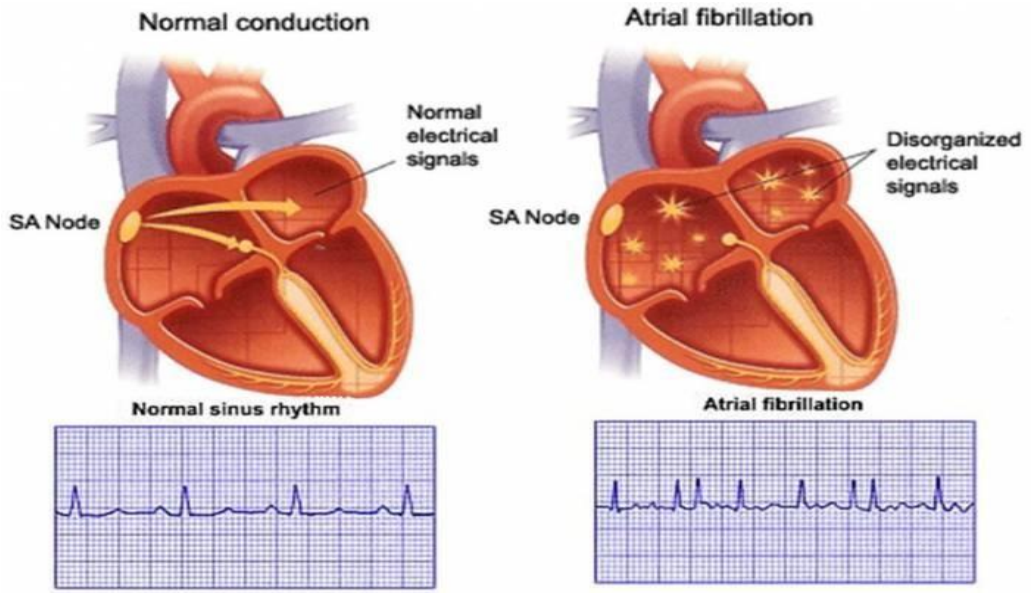
In the past few decades, Deep Learning has proved to be a compelling tool because of its ability to handle large amounts of data. The interest to use hidden layers has surpassed traditional techniques, especially in pattern recognition. One of the most popular deep neural networks is Convolutional Neural Networks.




2. **LITERATURE SURVEY**

## 2.1 Existing problem

Cardiovascular diseases (CVDs) are the number one cause of death today.Over 17.7 million people died from CVDs in the year 2017 all over the worldwhich is about 31% of all deaths, and over 75% of these deaths occur inlow and middle-income countries. Arrhythmia is a representative type ofCVD that refers to any irregular change from the normal heart rhythms.There are several types of arrhythmia including atrial fibrillation, premature
contraction, ventricular fibrillation, and tachycardia.




## 2.2 References

| Topic | Authors | Published Year | Abstract |
|---|---|---|---|
| Classification of Arrhythmia by Using Deep Learning with 2-D ECG Spectral Image Representation | Amin Ullah,Syed Muhammad Anwar,Muhammad Billal,Raja Mehmood | 2020 | In this study they have proposed a two-dimensional(2D) convolutional neural network (CNN) model for the classification of ECG signals into the eight classes. The eight classes are the normal beat, premature ventricular contraction beat, paced beat, right bundle branch block beat, left bundle branch block beat, atrial premature contraction beat, ventricular flutter wave beat and ventricular escape beat. The proposed CNN classifier was implemented in Python with the open-source library Tensor Flow which was developed by Google for deep learning. This |

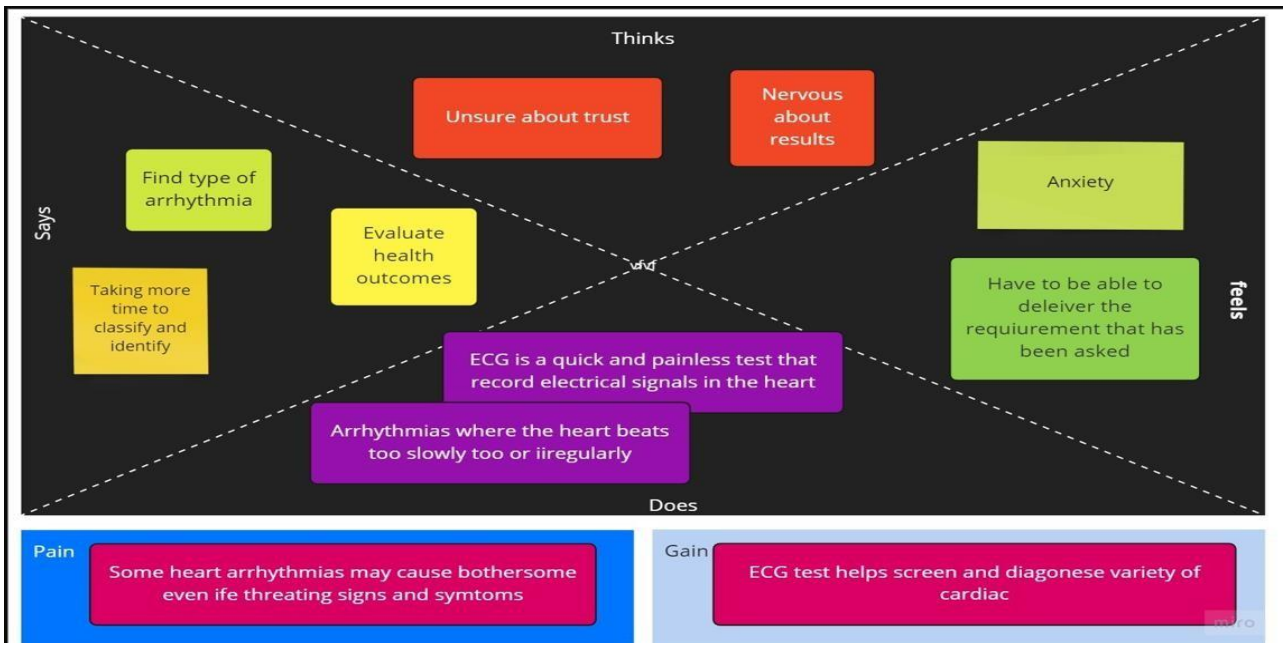
| | | | |
|---|---|---|---|
| | | | research paper uses only a single-lead ECG signal. |
| Cardiac Arrhythmia Detection from 2D ECG Images by Using Deep Learning Technique | E. Izci, M. A. Ozdemir, M. Degirmenci and A. Akan | 2019 | In this research paper, they proposed a deep learning approach using convolutional neural network for classifying the five different types of arrythmia disease. They segmented the heartbeats out of the ECG radio signals and all the images are converted into grayscale images. They implemented novel preprocessing, extraction and selection. They achieved an accuracy of 97.52% |
| Cyberbullying Detection Using Machine Learning | Aaminah Ali, Adeel M.Syed | 2020 | In this study they used numeric data for training, So the text was first converted into the numerical form using a label encoder. After that, the dataset divided into 80% training set, and 20% test and then classification algorithms were applied. The algorithms were used in classification are SVM, naïve Bayes, Random Forest and then an ensemble approach. The ensemble approach was a hybrid model, In this approach, a soft voting criterion was used. Which predicts the class label utilizing the maximum sum of the predicted probabilities. |
| Social Media Cyberbullying Detection using Machine Learning | John Hani , Mohamed Nashaat , Mostafa Ahmed | 2019 | In this research paper, they have proposed an approach to detect cyberbullying using machine learning techniques. And they evaluated their model on two classifiers SVM and neural network and they used TFIDF and sentiment analysis algorithms for features extraction. The dataset has been taken from Kaggle. Thus, a larger cyberbullying data is needed to improve the performance. Hence, deep learning techniques will be suitable in the larger data as they are proven to outperform machine learning approaches over larger size data. |

### 2.3 Problem Statement Definition

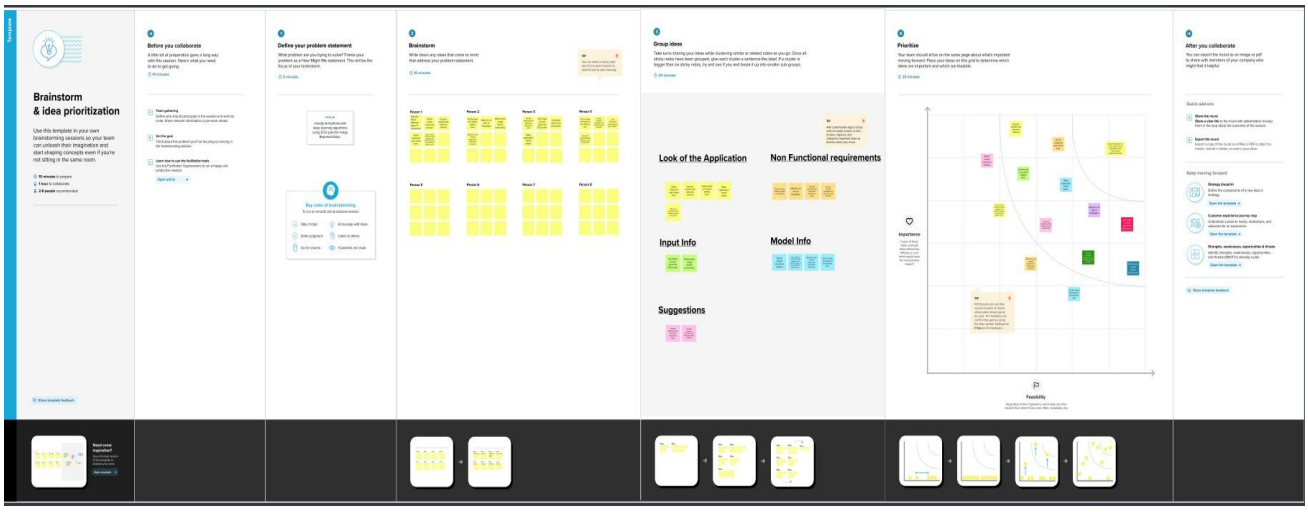
The presented problem wants us to identify and classify the types of Arrhythmias provided in the dataset using spectral images of the Electrocardiogram that are employed for prediction of cardiovascular diseases

## 3. IDEATION & PROPOSED SOLUTION

### 3.1 Empathy Map Canvas




### 3.2 Ideation & Brainstorming



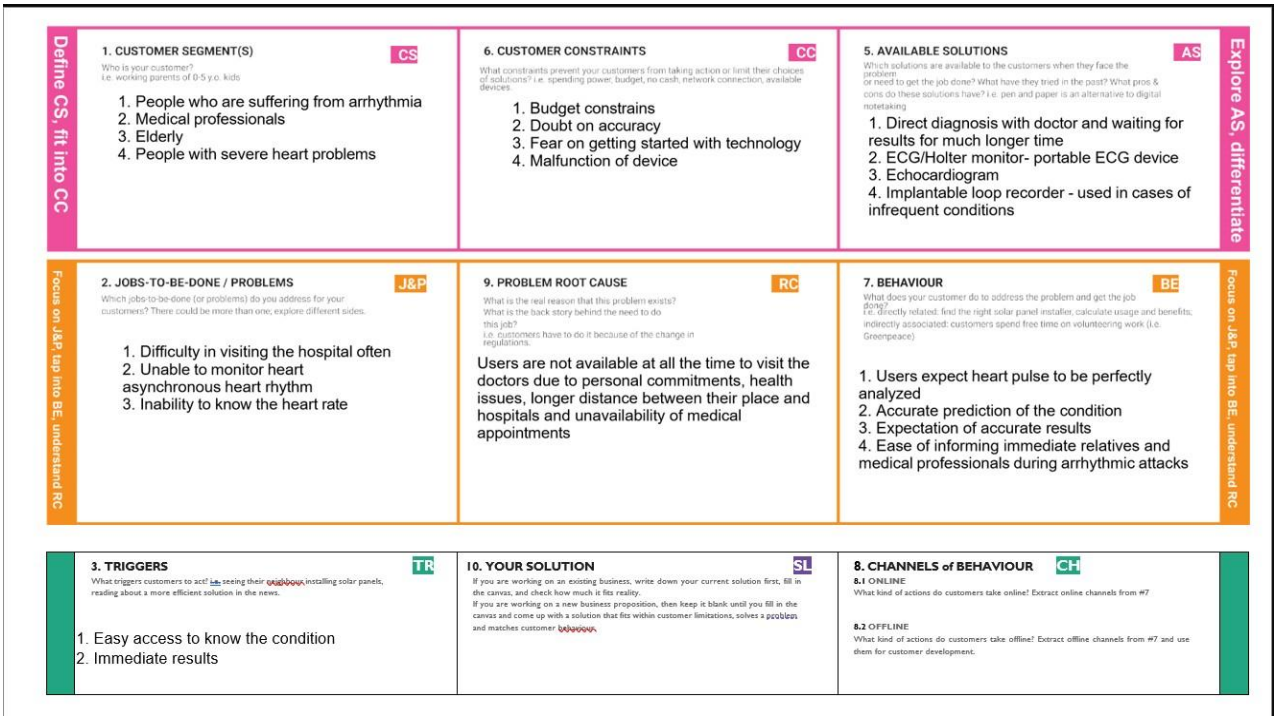
### 3.3 Proposed Solution

| S.No. | Parameter | Description |
|---|---|---|
| 1. | Problem Statement (Problem to be solved) | The presented problem wants us to identify and classify the types of Arrhythmias provided in the dataset using spectral images of the Electrocardiogram that are employed for prediction of cardiovascular diseases. |
| 2. | Idea / Solution description | The given dataset for this problem initially undergoes various data pre-processing steps to identify various forms of noise in the dataset and denoise them to make the data suitable for training a deep learning model. We will employ 2-Dimensional Convolutional Neural Network Model to carry out this classification. |
| 3. | Novelty / Uniqueness | • Deploys the model to a mobile application by assigning all heavy pre-processing to the cloud through an API.<br>• Usage of cloud-based ML-training services.<br>• Proposed solution will be deployed in the cloud making it easily accessible over the internet to people across the globe.<br>• Provide well detailed instructions or point the user to well-equipped hospitals to get good treatment. |

| | | |
|---|---|---|
| 4. | Social Impact / Customer Satisfaction | • Promotes Simplicity.<br>• Promotes Self-Diagnosis.<br>• Requires minimal effort and time.<br>• Proposed solution abides by privacy laws and no private information of user is stored.<br>• Delivers highly accurate results(classification of arrhythmia) in a short span of time. |
| 5. | Business Model (Revenue Model) | • Our business model primarily covers the expense we incur by deploying the service in cloud platforms.<br>• Primary consumers of our proposed service are hospitals who seek immediate consultation or use our service as a reference.<br>• Our service can be used by anybody who has access to internet services.<br>• Most of the competing products do not offer a Revenue is generated from Corporate Editions which has a monthly subscription, whereas the Community Edition is free for individuals.<br>• Users who would like to not travel to hospitals to get an ECG or get a selfdiagnosis can rent/buy an ECG Machine through our service which will be the secondary source of income. |
| 6. | Scalability of the Solution | • Increasing the dataset used for model training will in turn increase the application's scalability.<br>• Making the model more reliable will lead to hospitals using this application, this also eliminates human error.<br>• More powerful Cloud Instances for Concurrent use of the application.<br>• Periodically expanding the dataset and updating the model to increase scalability and reliability. |

| | | |
|---|---|---|
| | | • Cloud services guarantee high availability so there is very little probability that the service will face serious down times. |

### 3.4 Problem Solution fit




## 4. REQUIREMENT ANALYSIS

### 4.1 Functional requirement

| FR No. | Functional Requirement (Epic) | Sub Requirement (Story / Sub-Task) |
|---|---|---|
| | | |

| | | |
|---|---|---|
| FR-1 | User registration | Registration through website through email for verification. |
| FR-2 | User Confirmation | Verification of the OTP on the website. |
| FR-3 | User data | Uploading the image of ECG used for arrhythmia diagnosis. |
| FR-4 | User diagnosis results (Single and multiple input) | Results of the arrhythmia prediction and classification displayed in the website. |
| FR-5 | User diagnosis results (multiple input) | Results of the arrhythmia prediction and classification sent to registered email. |
| | | |

Following are the functional requirements of the proposed solution.

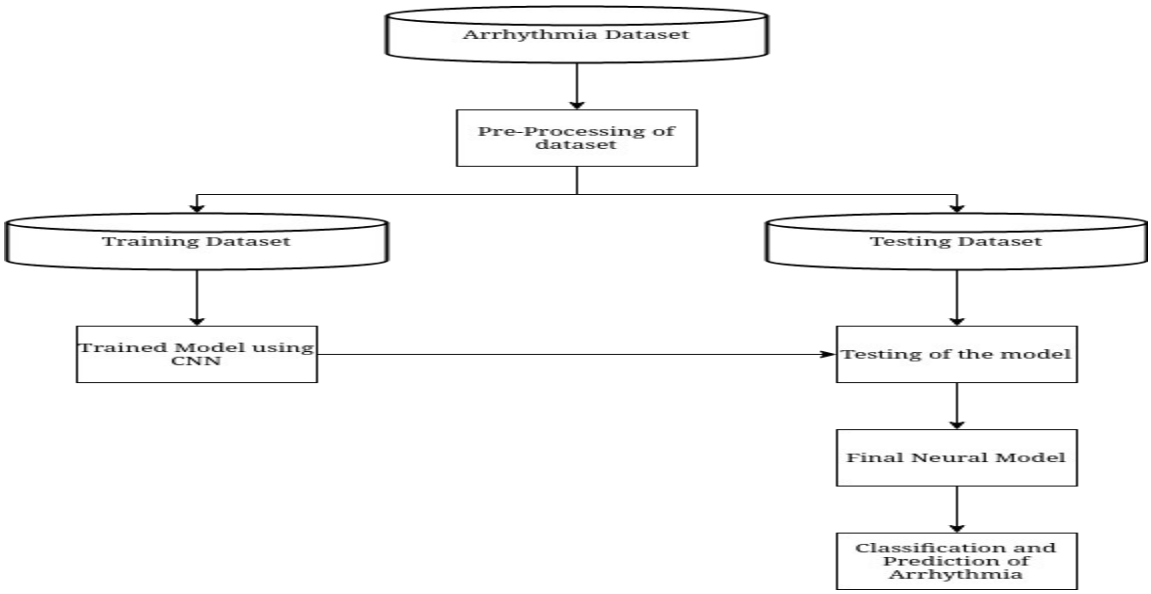
### 4.2 Non-Functional requirements

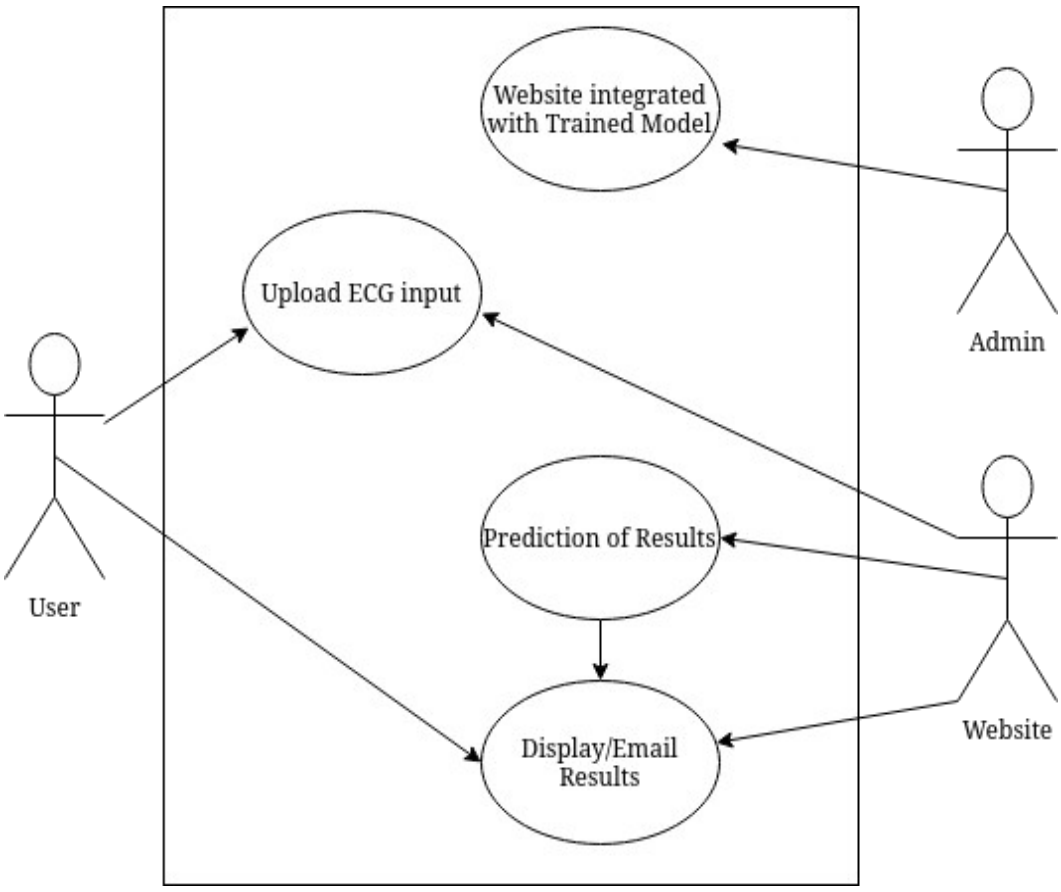
Following are the non-functional requirements of the proposed solution.

| FR No. | Non-Functional Requirement | Description |
|---|---|---|
| NFR-1 | **Usability** | Usability or ease of use refers to the many factors that constitute what users often describe as user-friendliness. |
| NFR-2 | **Security** | It assures all data inside the system or its part will be protected against malware attacks or unauthorized access. |
| NFR-3 | **Reliability** | It is the probability and percentage of the software performing without failure for a specific number of users or amount of time. |
| NFR-4 | **Performance** | It defines how fast a software system or a particular piece of it responds to certain users actions under a certain workload. |
| NFR-5 | **Availability** | It describes how likely the system is accessible to a user at a given point in time. |
| NFR-6 | **Scalability** | It is the ability to handle an increase in workload without performance degradation, or its ability to quickly enlarge. |

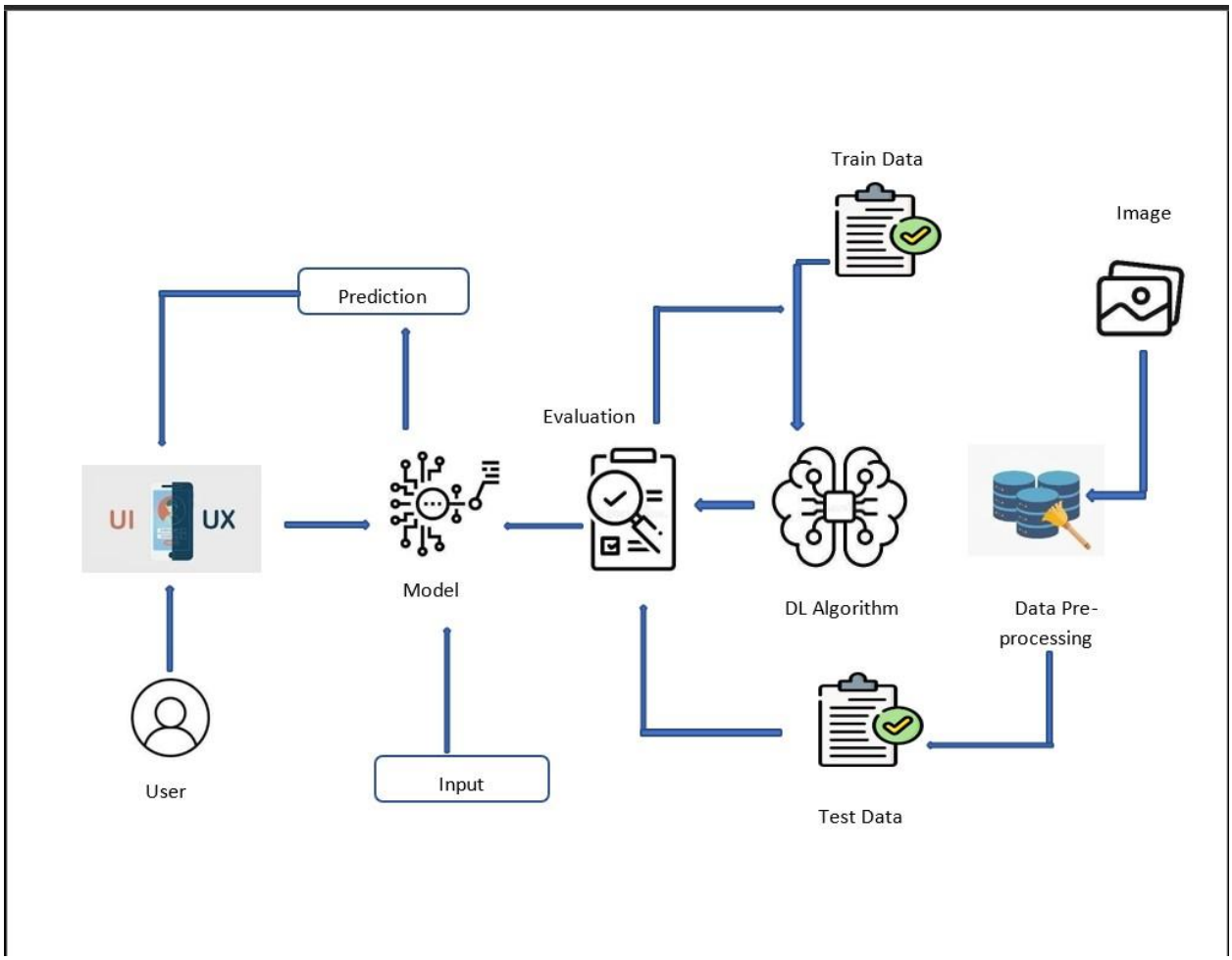
## 5. PROJECT DESIGN

### 5.1 Data Flow Diagrams



### 5.2 Solution & Technical Architecture




### 5.3 User Stories

| User Type | Functional Requirement (Epic) | User Story Number | User Story/Task | Acceptance Criteria | Priority | Release |
|---|---|---|---|---|---|---|
| Customer (Web User) | Registration | USN-1 | As a user I can register an account on the website | I can access classification page | High | Sprint-1 |
| | Login | USN-2 | As a user I can I login into the website using credentials | can access classification page | High | Sprint-1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Arrhythmia Classification page | USN-3 | User uploads the data | Input for the model | High | Sprint-1 |
| | Info page | USN-4 | Info page displays information about Arrhythmia | User can learn about Arrhythmia | Medium | Sprint-2 |

## 6. PLANNING & SCHEDULING

### 6.1 Reports from JIRA


SEP
OCT
NOV
DEC
IBM-1 Download The Dataset
IBM-2 Import The Image Data Generator Library
IBM-3 Configure Image Data Generator class
IBM-4 Apply the Image Data Generator functionalit...
IBM-5 Import Libraries
IBM-6 Initialize the Model
IBM-7 Adding CNN layer
IBM-8 Adding Dense Layer
IBM-9 Configure The Learning Process
IBM-10 Train the Model
IBM-11 Save the Model
IBM-12 Test the model
IBM-13 Create Html files
IBM-14 Build Flask Python code
IBM-15 Run the App
IBM-16 Register IBM Cloud
IBM-17 Train the model on IBM

## 7. CODING & SOLUTIONING (Explain the features added in the project along with code)

### 7.1 Application Building

We created a website using HTML, CSS which helps the user to classify the type of arrhythmia by inputting their ECG image into the website. Flask was also used to render our HTML files and also to integrate our model to the website which helps in classification of the disease.

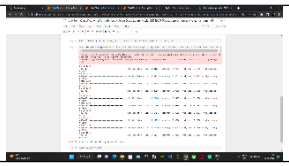
### 7.2 Model Building

Model Building is one of the most important parts of the application which helps to classify the arrhythmia disease. We imported various packages and libraries such as TensorFlow, Pandas, NumPy, etc. for data manipulation and pre-processing. Dataset was taken from Kaggle and we trained and tested using ImageDataGenerator.

We used the Sequential CNN (Convolution Neural Network) model as our classifying model which had 2 hidden layers. Model was made to run a total of 10 epochs for higher accuracy and we attained an accuracy of 92% which is a very high accuracy for disease prediction.

## 8. RESULTS

### 8.1 Performance Metrics

| S.No. | Parameter | Values | Screenshot |
|---|---|---|---|
| 1. | Model Summary | Model: "sequential"<br>__<br>Layer (type) Output Shape<br>#<br>==========<br>==<br>conv2d (Conv2D) (None, 62, 62, 32)<br>max_pooling2d (MaxPooling2D) (None, 31, 31, 32)<br>conv2d 1 (Conv2D) (None, 29, 29, 32)<br>max_pooling2d_1 (MaxPooling2D)(None, 14, 14, 32)<br>flatten (Flatten) (None, 6272)<br>dense (Dense) (None, 32)<br>dense_1 (Dense) (None, 6)<br>==========<br>==<br>Total params: 211,078<br>Trainable params: 211,078<br>Non-trainable params: 0 |  |
| 2. | Accuracy | Training Accuracy - 0.9655<br>Validation Accuracy -0.9136 |  |

## 9. ADVANTAGES & DISADVANTAGES

**ADVANTAGES:**

- Very fast and time-efficient
- Easy to access and communicate
- Cost-efficient

**DISADVANTAGES:**

- Low chance of error in classification
- No emotion in conveying the result
- Requires good internet access for accessing the website which may not be available in an urgent situation

## 10. CONCLUSION

Application was built using various web technologies such as HTML, CSS, Flask, etc and model was also build using Machine Learning Techniques such as Data Preprocessing and model building using TensorFlow and other packages to classify arrhythmia accurately and time-efficiently.

## 11. FUTURE SCOPE

This application can help other classification of diseases much more easier as it is nearly equivalent to this application and can further be researched to improve accuracy and other metrics.

## 12. APPENDIX

### Source Code

Github Link: https://github.com/IBM-EPBL/IBM-Project-4249-1658725786