

Architectural Decisions Document

1 Project Overview

My data set is a set of ECG data. The data includes both healthy ECG signals and ECG signals from patients who have one of 4 types of Arrhythmias. The goal of my project is to classify the heartbeat signals into the 5 possible categories.

1.1 Data Source

1.1.1 Technology Choice:

I chose to use the MIT-BIH data.

1.1.2 Justification:

I am using this data set because it is a highly reputable data set which as be use in countless studies of heartbeat arrhythmias.

1.2 Data Repository

1.2.1 Technology Choice:

<https://www.kaggle.com/datasets/taejoongyoon/mitbit-arrhythmia-database>

1.2.2 Justification:

I am using this version of the data set because it is in text format and simple to import into python and I can view the data directly. However, it hasn't been altered in any other way from the original source.

1.3 Discovery and Exploration

1.3.1 Technology Choice:

Each patient has two associated files one with the heartbeat signal and one with information about the heartbeat. In my initial data exploration, I will first use these two files to isolate different types of beats and plot them.

1.3.2 Justification:

This will give me a good idea of the features I will be extracting in the next step and if I am correctly extracting them. It will also help me build an intuition of what a normal heartbeat should look like, and what the arrhythmias look like.

1.4 Features

1.4.1 Choice:

Each feature will be a single heartbeat based on the location of the R-peak. It will start at the previous R-peak and go a little 33% into the next heartbeat.

1.4.2 Justification:

This choice seems to best capture the defining features of each heartbeat.

1.5 Actionable Insights

1.5.1 Show that ML can discover heart rate arrhythmias reliably.

- 1.5.2 I will demonstrate this using 3 metrics for scoring the predictions of the ML algorithm. First the accuracy, and F1 score for quick comparison of the different algorithms. Next, I will visualize the outcomes using a confusion matrix, to give deeper insight into the outcomes.