NOTE: If you have never used the AstroSkin before please start by reading the Appendix

**Gathering Data**

**Introduction:**

The point of this internship was to apply machine learning to ECG analysis. To start with that the first step was to accumulate data. Knowing that the AstroSkin can record data, it was the first tool used.

Recording healthy data was easy, since anyone who wore the shirt was healthy. The main issue was getting unhealthy data; luckily we have a dummy in the lab (named Bernie) that can generate up to 31 different conditions. While working with the dummy it was noticed that the signal is relatively clean and non-variant. So we need to get more data or our system will only work on simulated data. That’s why I had to go to Physionet.org to collect data.

**Working with Physionet:**

Initially I started by searching for the data that contains Lead I, Lead II, and Lead III to match the AstroSkin data. After that I narrowed the search to three conditions as a start, the conditions are:

* Myocardial Infarction
* Atrial Fibrillation
* AV Block

These conditions were selected since they are important, and can be found easily from both Physionet and from the dummy.

Two Physionet datasets were used:

* Ptbdb (<https://physionet.org/physiobank/database/ptbdb/>)
* Incartdb (<https://physionet.org/physiobank/database/incartdb/>)

To download these datasets, wget(https://www.physionet.org/physiotools/utilities/wget/) was used. Which download the whole dataset in .dat format. After that the wfdb library was used to convert .dat to mat format. This was followed by a transformation of the .dat folders to .csv files.

To convert .dat files to .mat files the Matlab on the laptop was used (Since it has wfdb installed on it) and the script called ConvertMatToCsv.m was used, this loops through a file and converts all .dat files to .mat files. After that Python was used to convert the .mat files to csv, by the use of mat\_to\_csv.py which loops through a folder and opens all folders that are in it, then loops through every file in the folder and converts .mat files to csv. This makes them ready to be processed.

**Processing:**

The data has to be split into heart beats. And knowing that we have three ECG leads, we need to store all three. So we designed a 3D matrix containing three ECG leads and different conditions. For every lead 204 samples (204 was selected since it is enough to contain all PQRST waves) were taken with the R interval being at the middle. Along with this matrix we need another matrix that identifies the condition we are currently working with; this matrix is a 2D matrix in the form:

[0001]

[0010]

[0100]

[1000]

Every row in the above matrix specifies a condition. The script that does that is called astro\_save\_all.py.

Before running astro\_save\_all.py it is important to set the *big\_file\_name* to a folder that contains other folder/folders. For example big file name might be called diseases and it contains folders named Healthy, MI, AV, etc. And every smaller folder should ONLY contain three ECG signals named: ECG\_I.csv, ECG\_II.csv, and ECG\_III.csv. After setting *big\_file\_name* count the number of conditions we have in the above example we have three conditions (Healthy, MI, AV). So we set *num\_of\_diseases = 3.*

If *num\_of\_diseases* = 3 the 2D matrix will contain arrays like this:

[001]

[010]

[100]

Along with the astro\_save\_all.py, another script was used to save Physionet heart beats (physio\_save\_all.py), but the problem with this script is that it needs to be modified for every Physionet dataset, since every folder is different.

**APPENDIX:**

**Saving Data with the AstroSkin:**

Start by pairing the module to your phone or Ipod. Then log in using your account, by doing that the module will know to which account that data will be saved. This has to be done only once unless another device was connected to the module.

Now to record the data, wear the shirt and plug in the module. And to record data from Bernie use alligator clips to connect the sensors in the shirt to the Dummy’s four metallic parts located near the chest and the belly.

**Getting the Saved Data:**

Now after we have used the AstroSkin to save data it is time to get that data. Start by connecting the AstroSkin module to a computer and use the HxServices App to upload the data online. Just by plugging in the module and opening the app, the data starts uploading. After the Upload is complete, the data can be accessed online. Log in to <https://my.hexoskin.com/en/login> and go to records and select the record that you need and on the right top corner you will find. Click on it and press download BINARY data, this will download a zip file containing .wav files.

**Converting wav to csv:**

After you download export it from the zip file. Then go to Matlab and run a script called convert.m giving it the path of the downloaded folder ex: convert(‘Desktop/Record1’). This script was downloaded from the hexoskin website. The job of this script is to convert all the .wav files to .csv.