DSP PROJECT TEAM #1 – ECG (NORMAL & PVC)

MEMBERS:

- JANA HANI MOHAMED ZAKI EL-SHEIKH || **ID:** 2022170647
- TOKA KHALID MOSTAFA ABD-ELHAMEED || **ID:** 2022170106
- JANA ESSAM ABDELFATAH ERFAN || **ID:** 2022170112
- RAWAN MOHAMED TAHA AWAD ALLAH || **ID:** 2022170160

OVERVIEW:

Preciseness is key when it comes to processing ECG signals as one simple mistake can cost a person their life. Thus, when it comes to training a machine learning model to detect the delicate abnormalities that are sometimes found in a person's heart, being precise and accurate becomes all the more critical.

In this project, our team strives to:

- Deliver a machine learning model capable of processing ECG signals as well as differentiating between a normal heartbeat and a diseased one.
- Provide an accurate assessment of signals provided from ECG machine readings.

So, without further ado, we shall now move on to the first step in our project.

FIRST: Pre-Processing:

Our team first started by concatenating the two training files into one dataset. We also generated appropriate labels for the data to differentiate between Normal and PVC signals.

We then displayed one of the samples to get an idea of the signal shape. The normal signals were concatenated first and, thus, the plot revealed a normal ECG signal.

Moving on to the actual pre-processing process.

1. BANDPASS FILTER

We first started by passing our data through a bandpass filter. We chose to do it with this filter to eliminate noise outside the desired range of our data (the range being 0.5 to 40).

FILTER PARAMETERS:

- Signal
- Lower Cutoff: 0.5 Hz
- Upper Cutoff: 40 Hz
- Sampling Rate/Fs: 360 Hz
- Order of Filter: 5

The first four parameters were specified in the project document. The order of filter, however, was not and we had to look through other sources as well as experiment with the order to achieve our desired end. We wound up choosing a filter order of 5 as that was what gave us the best accuracy as well as the best predictive results.

Thus, our bandpass filter was complete.

2. NORMALIZATION

After removing the noise via the bandpass filter, we then move onto the normalization of our signal. This is an important step in pre-processing the data as it prevents numerical instabilities and makes it so that the data is on an equal scale.

NORMALIZATION PARAMETERS:

- MIN-MAX: -1 TO 1

Between the Standard Scalar and the min-max normalization methods, the min-max worked better for our model and provided us with a more accurate performance.

We chose the range -1 to 1 as it centers the data around 0 and makes it so the bipolar nature of the ECG Signal (its peaks and troughs) is preserved. It makes it more true to the signal itself.

SECOND: Feature Extraction

After pre-processing the data and ensuring it is well and truly free of noise we now move on to the next step and that is the Feature Extraction. Feature Extraction is a process in which we extract certain features from our data to reduce redundancy in the data we pass to our model. It helps the model converge faster and reduces computation power making for an efficient and accurate prediction.

The Feature Extraction method used was the Daubechie Family of Wavelets (db4). We decided on this iteration of the Daubechie Family of Wavelets due to research as multiple research projects have utilized this specific configuration.[i]

DAUBECHIE WAVELET PARAMETERS:

- Signal
- Wavelet: db4
- Level: 3

Our team decided to go about specifying the level through manually performing the branching method. The final output reached was 3 levels and that is ultimately what we chose to go with.

FINALLY: Model Building

We decided to work with the KNN model as, not only was it specified in the document, but also one of the better options for what we wanted to achieve.

Some of the observations we had when working with the KNN models were:

- The model was very consistent regardless of how many k neighbors were used. Be it 1 or 3 or 5 the accuracy remained consistent.
- The model provided an incredibly high accuracy after the data preprocessing and feature extraction were both complete.
- The model predicted with incredible efficiency, passing all the trials with accurate predictions.
- The final accuracy provided by the KNN model was 100%.

CONCLUSION

When processing ECG signals it is important for one to precise. When training a model to make predictions based on such signals it becomes critical to provide only accurate assumptions with every trial.

Our model utilizes data that has been passed through a bandpass filter through a passband range of 0.5 to 40 (eliminating all noise outside that range), normalized from -1 to 1 (making it perfectly symmetrical and preserving the shape of the signal), and had its features extracted using the Daubechie Family of Wavelets over 3 levels of decomposition.

All of this work ensured that model was capable of predicting with near faultless precision making it so any margin of error in this process was mitigated.

REFERENCES

---

[i] **Article Detailing ECG Processing Project**: https://www.researchgate.net/publication/258237553_FEATURE_EXTRACTION_OF_ECG_USING_DAUBECHIES_WAVELET_AND_CLASSIFICATION_BASED_ON_FUZZY_C-MEANS_CLUSTERING_TECHNIQUE