Machine Learning based modeling and interpretation of data from wearable sensor

Patel Devarshi Chandrakant | 18CS10040

Supervised by : Professor Sudip Misra
Mentored by : Mrs. Saswati Pal

Introduction

- The sensors used for IoT-based healthcare monitoring involve 3-lead ECG sensors, pulse oximetry sensors and many more but the analysis and information extracted from these sensors are not at par to match the clinical purposes
- The project aims at using state-of-art algorithms and techniques to analyse, predict, and reconstruct the sensor data to match the clinical standards
- For example, predicting 12-lead ECG signal from 3-lead ECG sensors to predict any abnormalities

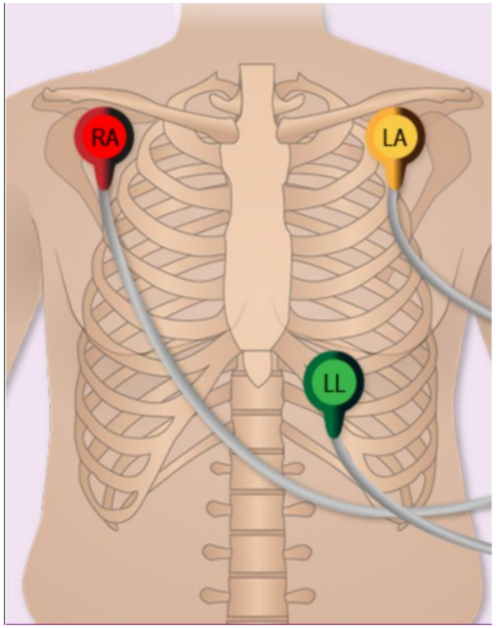
What did I learn this semester?

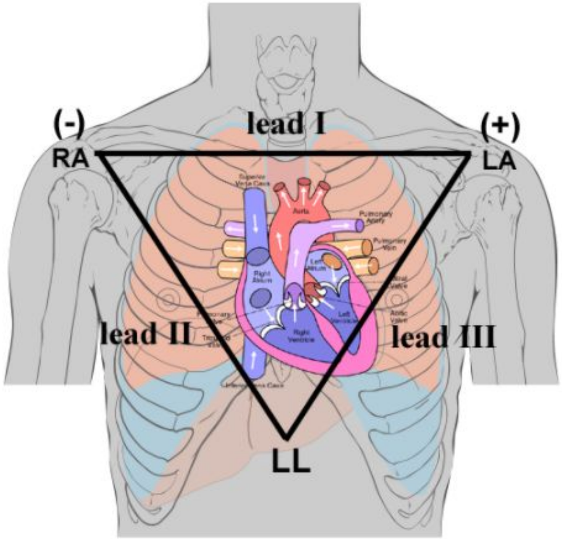
- Got acquainted with the 3-lead ECG and 12-lead ECG medical practices
- Learned Neural Networks and Deep Learning for prediction purposes
- Worked on prediction on 12-lead ECG signals from 3-lead ECG signals

3-Lead ECG

- 3-lead ECG uses limb leads namely RA(right arm), LA(left arm) and LL(left leg). 3-lead ECG signals are generated from this bipolar leads

  Signal 1 : RA(-) ------> LA(+)
  Signal 2 : RA(-) ------> LL(+)
  Signal 3 : LA(-) ------->LL(+)





Credits : https://researchguides.library.vanderbilt.edu/c.php?g=156859&p=1125029

12-lead ECG

Bipolar limb leads (frontal plane):

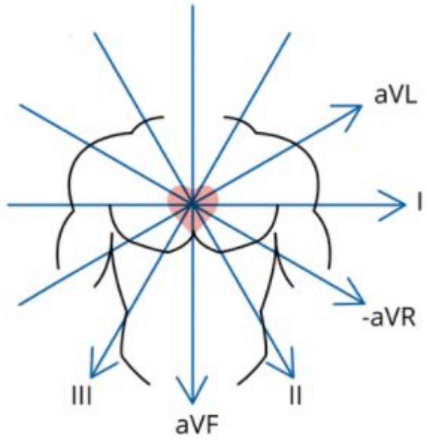
- Lead I: RA (-) to LA (+) (Right Left, or lateral)
- Lead II: RA (-) to LL (+) (Superior Inferior)
- Lead III: LA (-) to LL (+) (Superior Inferior)

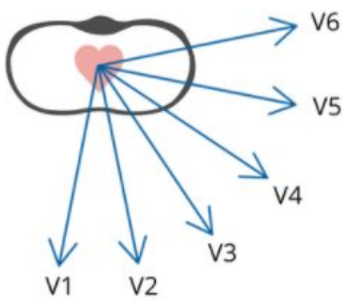
Augmented unipolar limb leads (frontal plane):

- Lead aVR: RA (+) to [LA & LL] (-) (Rightward)
- Lead aVL: LA (+) to [RA & LL] (-) (Leftward)
- Lead aVF: LL (+) to [RA & LA] (-) (Inferior)

Unipolar (+) chest leads (horizontal plane):

- Leads V1, V2, V3: (Posterior Anterior)
- Leads V4, V5, V6:(Right Left, or lateral)





Credits : https://ecg.utah.edu/lesson/1 , CardioSecur

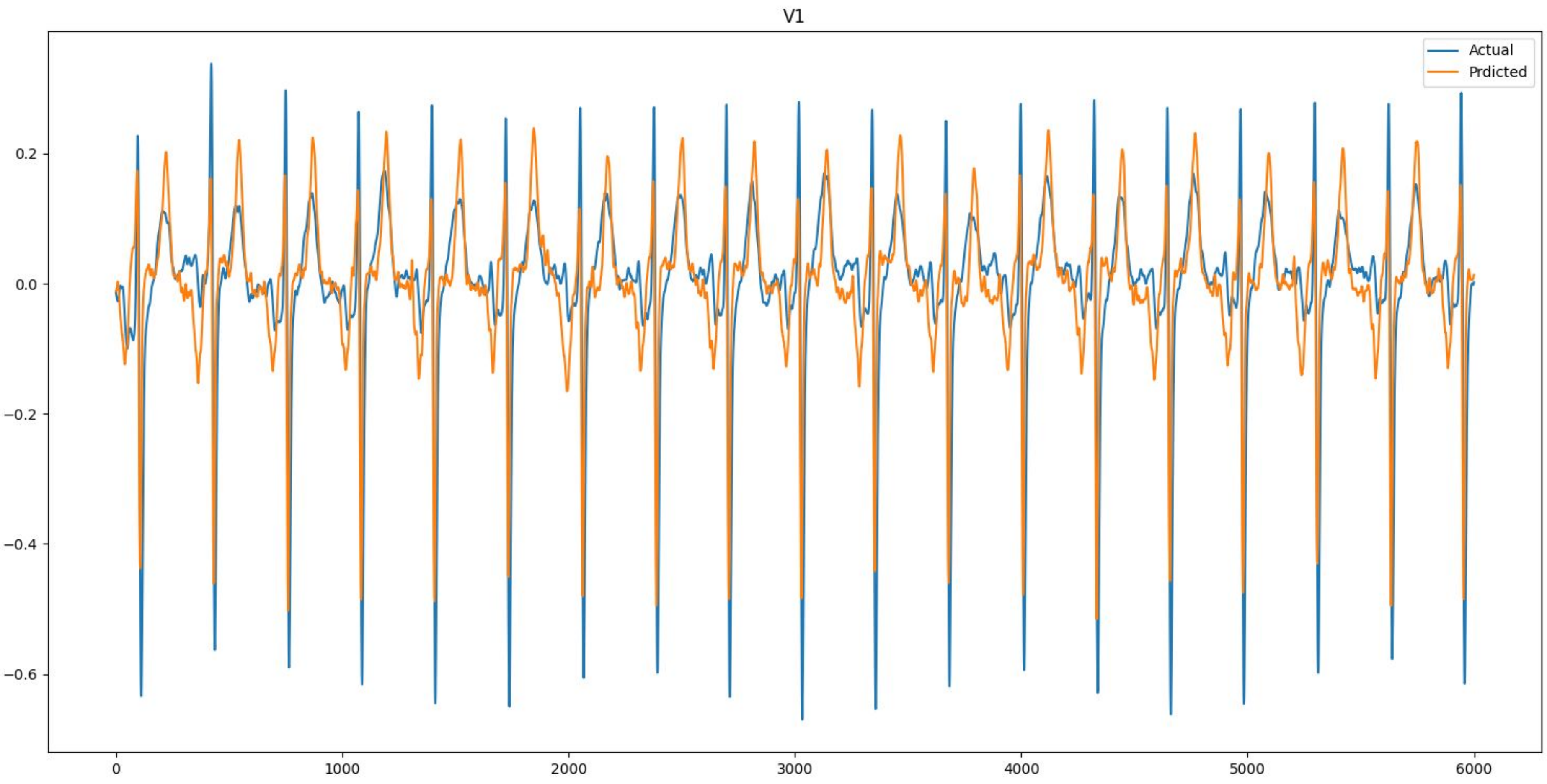
3-lead ECG to 12-lead ECG signal

- Since the bipolar limb leads are common in both the practises they it can be concluded that the output in both the methods will be the same. So leads I, II, III can be accurately predicted with minimal manual error while recording
- So we have lead signals I, II, III as input from 3-lead ECG
- Augmented unipolar limb leads aVR, aVL and aVF are linearly related with leads I, II, III as it lies in the same plane(frontal plane).
  Upon derivation,
  aVR = -½ * (I+II)
  aVL = ½ * (I-III)
  aVF = ½ * (II + III)
  Thus this signals vectors can also be predicted with minimal error induced while recording lead signals I, II, III
- However lead signals in horizontal plane (namely V1 to V6) are not linearly related with frontal plane vectors(I, II, III, aVR, aVL, aVF)

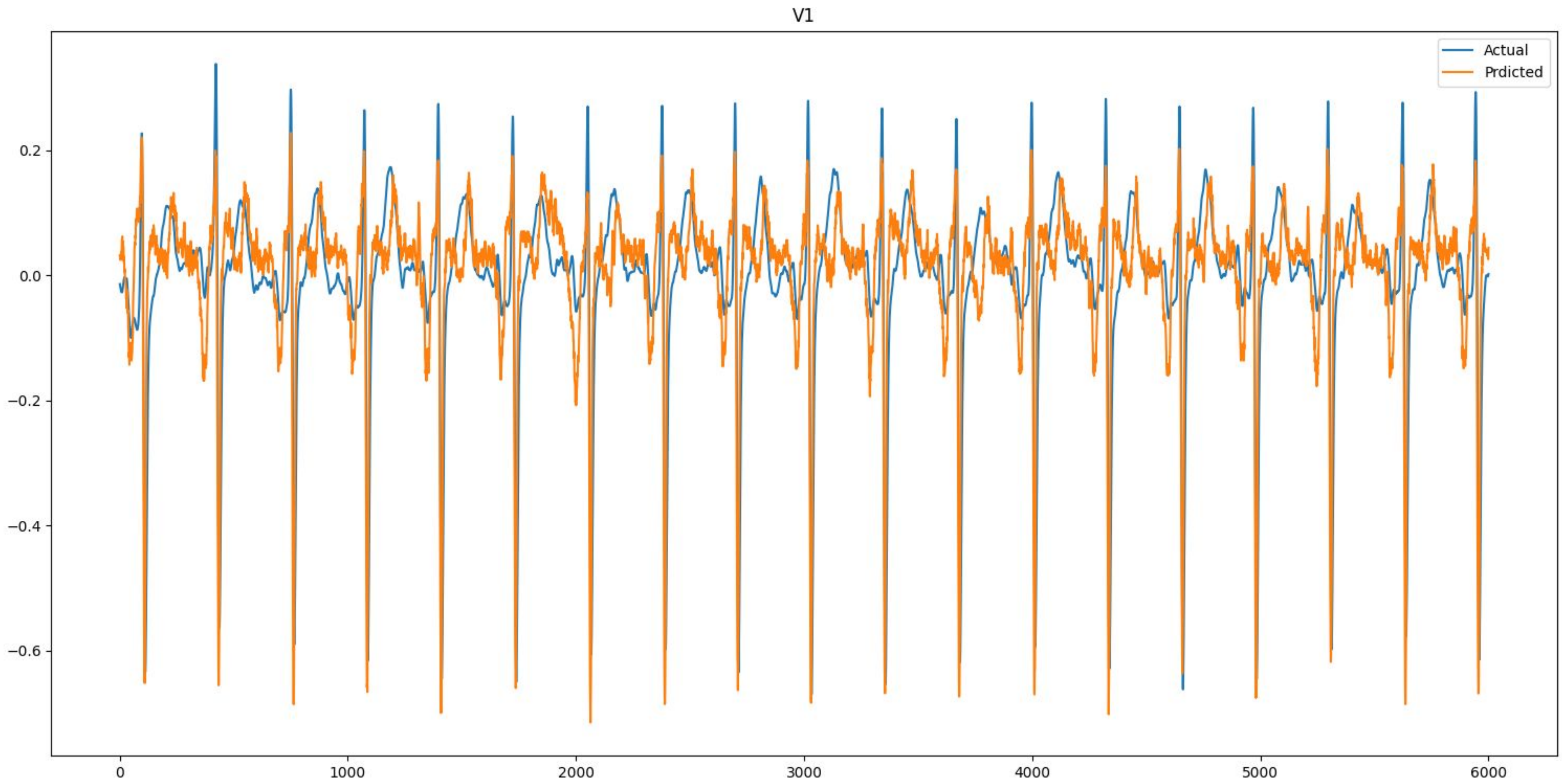
3-lead ECG to 12-lead ECG signal (contd.)

- I have used linear regression to predict the other 6 signals from lead signals I, II, III
- For this I have used China 12-Lead ECG Challenge Database | Kaggle
- To verify the aVR, aVL, aVF with the linear relation I have applied it on the above dataset and the average absolute squared error for a particular test case are as below :
  aVR : 3.86×10^{-7}
  aVL : 1.75×10^{-7}
  aVF : 3.73×10^{-7}
- I used linear regression on input input values I, II, III to see the curve fit on the actual values and the result is shown in the next slide.

Linear Regression on I, II, III leads to see fit on V1

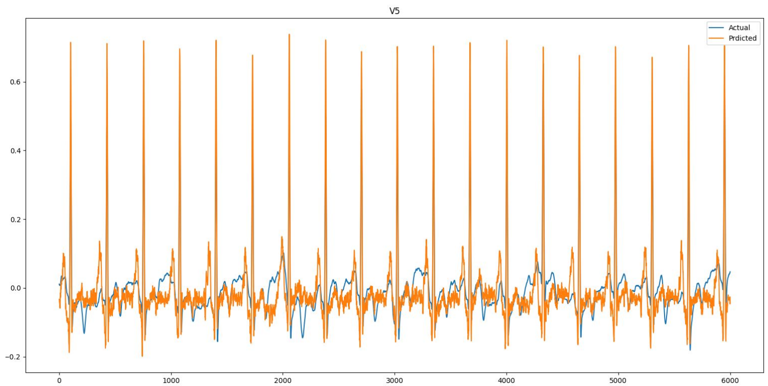


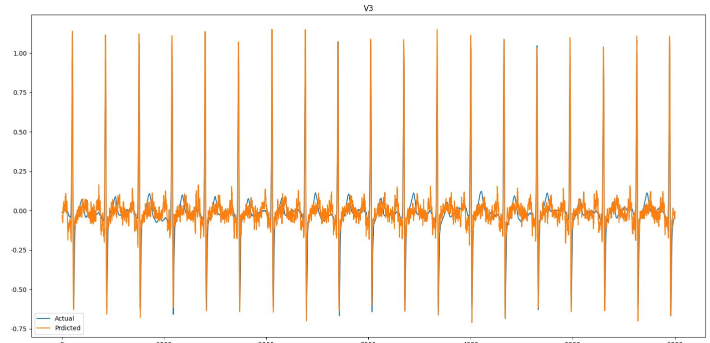
Linear Regression on I, II, III,d(I)/dt,d(II)/dt, d(III)/dt leads to see fit on V1

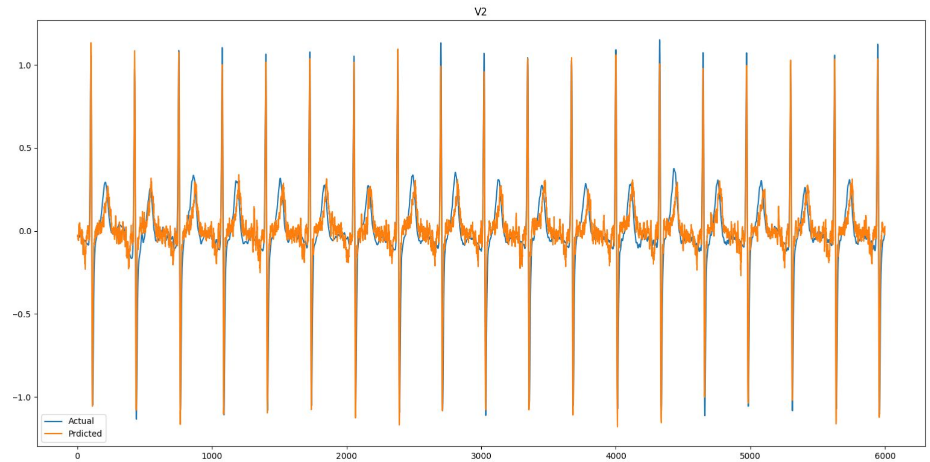


3-lead ECG to 12-lead ECG signal (contd.)

- Upon examining the two approaches the one where slopes where considered gives closer overall shape despite the numerical error is sometimes slightly higher
- But the output in this case is very time varied and requires smoothening process
- Similar fitting process is applied for other leads as well



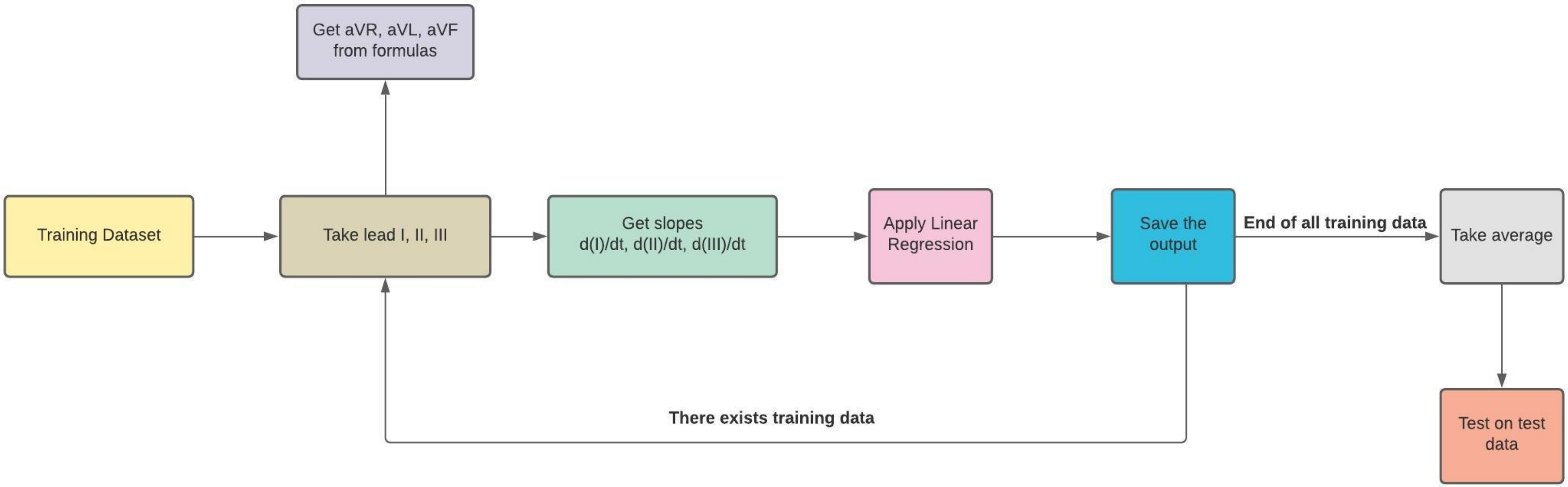




Challenges

- Linear regression was applied on a single training set to view the fitting of output graph but after training multiple test cases the average output does not guarantee the well fitting of the training data set.
- Absolute average squared error is very high (of the order of >180)
- The fitting graph is time varied with quickly changing values which needs to be smoothened for proper diagnosis.

Proposed Idea and Challenges



- Give weight to the linear regression output of each training set, inversely proportional to their error.
- Explore Neural Network and other approaches to get better results.
- In case of neural network approach the data will be cropped this might lead to data loss or might need to extrapolate this might give erroneous output. How to tackle such a situation?
- To explore the possibility to get two out of six leads in horizontal plane with high accuracy and then try to predict other 4 as they lie in the same plane and compare this approach with directly predicting those 4.

For Further Diagnosis

- According to paper,
  Deep Learning Algorithm Classifies Heartbeat Events Based on Electrocardiogram Signals
  Yongbo Liang, Shimin Yin, Qunfeng Tang, Zhenyu Zheng, Mohamed Elgendi and Zhencheng Chen (2020)
  https://doi.org/10.3389/fphys.2020.569050
  CVD(s) can be classified and predicted using CNN blocks and BiLSTM

Thank You