# Project Report - Machine Learning based modeling and interpretation of data from wearable sensor

Patel Devarshi Chandrakant

18CS10040

Mentored by:

Mrs. Saswati Pal

Supervised by:

Prof. Sudip Misra

## 1 Introduction

The sensors used widely for IoT-based healthcare monitoring involve 3-led ECG sensors, pulse oximetry sensors and many more. Hover, the analysis and information extracted from these sensors are not at par to match the clinical purposes. The existing low cost IoT devices involve 3-lead ECG sensors but to predict any abnormalities there exists the need to predict the 12-lead ECG signal. The proposed thesis will investigate the methods to predict 12-lead ECG signals from 3-lead ECG signals.

## 2 Literature Survey

The positioning and interpretation of 12-lead signals show that there are 6 signals in frontal plane, 3 bipolar limb leads(I, II, III) and 3 augmented uni-polar leads(aVR,aVL,aVF), and 6 signals in horizontal plane (V1, V2, V3, V4, V5 and V6). (Yongbo Liang and Chen, 2020) show that how SVM classifier and a combination ofCNN and BiLSTM can be used to classify cardiovascular diseases with great accuracy and specificity.

## 3 Work Done

We aim to develop a machine learning based modeling approach to analyse, predict, and reconstruct the sensor data to match the clinical standards. The 3-lead ECG monitoring have only three bipolar limb leads leads I, II, III which are common in 12-lead ECG monitoring and other 9 lead output needs to be predicted. Other augmented uni-polar limb leads aVR, aVL, aVF can be predicted using linear combination of limb leads I, II, III as the signal vector lies in the same plane. Upon derivation and confirming the same with the actual dataset (Feifei Liu and Kwee, 2018) we have deduced that aVR = $\frac{1}{2}$ * (I+II), aVL = $\frac{1}{2}$ * (I-III), aVF = $\frac{1}{2}$ * (II + III). For deriving the other six uni-polar chest leads(V1 to V6) I have used linear regression approach where the input parameters are lead signals I, II, III, d(I)/dt, d(II)/dt, d(III)/dt. Approach used is
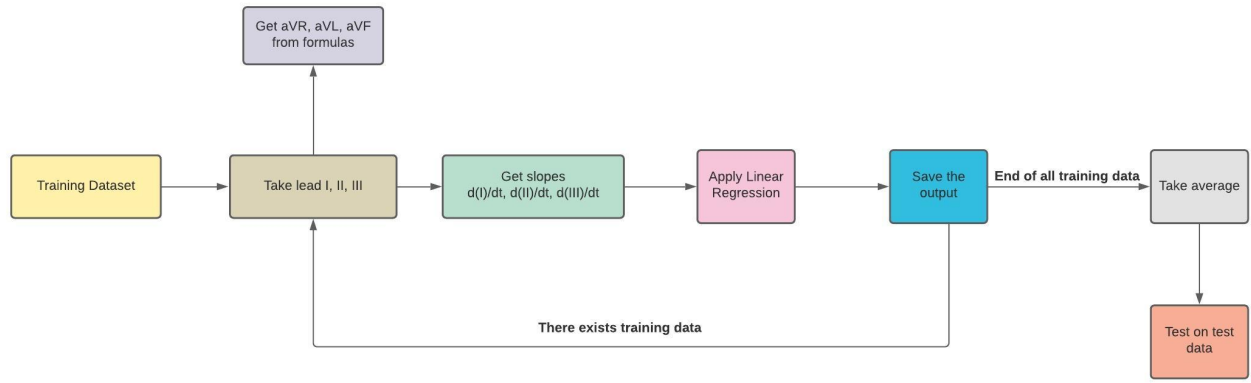
For each training set the fitting curve should be accurate enough so that the hypothesis parameter resulted can be used later for test cases. In one training case the resulting lead V1 and V2 fitting curves were as below
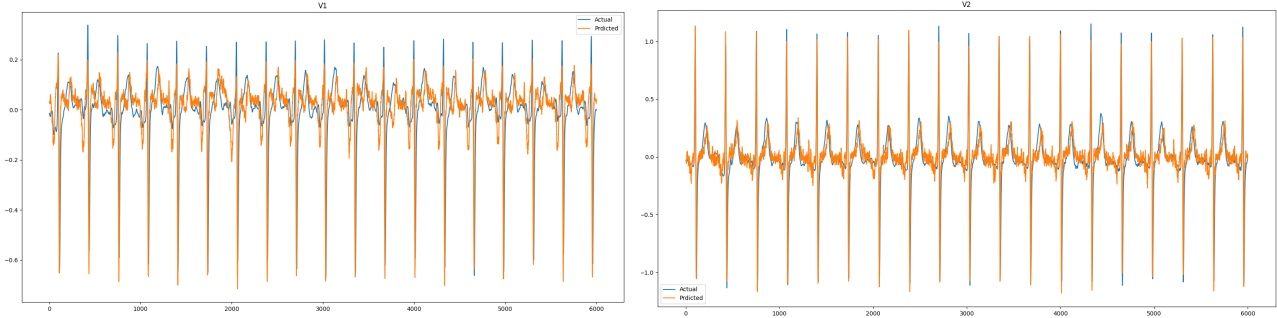


Once we get the corresponding hypothesis parameters for each training instance it can be averaged together maybe by the weight inversely proportional by their errors in fitting to get an overall hypothesis parameter which can be tested on testing data. Also we should try to predict two out of 6 signal vectors in horizontal plane and then deduce other four from those two and compare the corresponding result with linear regression fitted approach.

# 4 Problems to be solved

Linear regression was applied on a single training set to get hypothesis parameters but there is no guarantee that the weighted average may lead to acceptable model so as to apply in real world scenarios. **The first problem** is how should training instance be treated so as to handle the issue of over-fitting or under-fitting? **The second problem** is that the fitting curve got for each instance is time varied which needs to be processed so as to reduce error and remove excessive fluctuations. At last, I need consider other prediction techniques like neural networks and SVM classifiers to see their accuracy rate and proceed further.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?

# References

Q. T. Z. Z. M. E. Yongbo Liang, Shimin Yin and Z. Chen, 2020.

L. Z. X. Z. X. W. X. X. Y. L. C. M. S. W. Z. H. J. L. Feifei Liu, Chengyu Liu and E. N. Y. Kwee, *Journal of Medical Imaging and Health Informatics Vol. 8, 1368–1373*, 2018.