Ahmedabad University

CSE523 - Machine Learning

Section: 1

Heart Disease Prediction - Report

Submitted to Prof. Mehul Raval

Group name: Bias_Variane_Dilema

Enrollment no	Name
AU1940155	Shubham Patel
AU1940110	Shrey Patel
AU1940207	Keyur Nagar

❖ Things we have done:

- We have tried to find important features using PCA (Principal component analysis).
- We explored the weka tool for applying various Machine learning classifiers and dimensionality reduction algorithms.
- Using PCA, we didn’t get satisfactory results(i.e. could not find a particular feature using lower dimensional space).
- So, We explored the CFS algorithm for finding the ranking of the feature. Basically, CFS is a simple filter algorithm that ranks feature subsets according to a correlation-based merit evaluation function. The bias of the evaluation function is toward subsets that contain features that are highly correlated with the class and uncorrelated with each other.
- Using the CFS algorithm, we come up with the following features which are important for our dataset.

  Age, ap_hi, ap_lo, cholesterol, gluc, active and BMI

❖ Things to be done:

- We tried to find why accuracy is low in the case of KNN and logistic regression classifier. We also incorporate another classifiers if needed.

- In the end, we will try to deploy our model on the front-end website where the user can enter their values for the 13 parameters, and the model will detect the disease and show relevant results. This will ensure our model is widely used.