CIS490 Machine Learning Project Proposal

Group 8

Prediction and Knowledge Discovery of Heart Diseases using Machine Learning Techniques

Kevin D'Cruz(Leader)|Chetan Kumar|Abhishek Manoj Kumar|

Apoorva ShivaShankar|Madhuri Gawali

The scope of machine learning in the healthcare domain and prediction of diseases has been broadening and paving its way significantly as the years go by. This made us select datasets based on such a domain. The three datasets we selected are as follows:

1. Heart Disease Data Set: In this dataset we will focus on Cleveland database.
   Attributes: Age, sex, chest pain type, resting blood pressure, fasting blood sugar etc.
   Goal: Presence of Heart Disease in Patients
   https://archive.ics.uci.edu/ml/datasets/Heart+Disease

2. Statlog (Heart) Data Set
   Attributes: Age, Serum Cholesterol, Exercise Induced angina, Sleeping Blood Force
   Goal: Presence of Heart Disease in Patients
   http://archive.ics.uci.edu/ml/datasets/statlog+(heart)

3. SPECT Heart Data Set (Single Proton Emission Computed Tomography)
   Attributes: Overall Diagnosis, Clinical Patient Records, Image data
   Goal: Knowledge Discovery, by semi-automation of the Diagnostic Process
   http://archive.ics.uci.edu/ml/datasets/SPECT+Heart

All the three datasets are from the Machine Learning Repository of UCI. On these datasets, we plan on applying Logistic Regression (Common), Decision trees (Advanced) and Neural Network (Optional) as machine learning models. Depending on the results obtained, we may also train other models to check as to which would be more efficient.

References:

[1] Heart Disease Prediction Using Machine learning and Data Mining Technique: http://csjournals.com/IJCSC/PDF7-1/18.%20Tejpal.pdf

[2] Intelligent and Effective Heart Disease Prediction System using Weighted Associative Classifiers: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.302.6636&rep=rep1&type=pdf

[3] Classification and prediction of heart disease risk using data mining techniques of Support Vector Machine and Artificial Neural Network: http://ieeexplore.ieee.org/document/7724835/

[4] Diagnosis of heart disease patients using fuzzy classification technique: http://ieeexplore.ieee.org/document/7066746/

[5] Knowledge discovery approach to automated cardiac SPECT diagnosis: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.17.3773&rep=rep1&type=pdf