Final Project Proposal - Heart Disease Presence

TEAM H - Chatzistefanidis Hlias - Adam Argyris - Koutsoni Eleni

November 2019

Problem and its significance to modern Medical Applications. There is a vast amount of medical data, much of it goes unanalysed or under analysed because when medical practices purchase software to house data, the purpose is mostly placed in the sector of record keeping. Clinics, pharmacies and the like store data in order to serve patients on an individual level via identification verification and medical records. Future analysis wasn't at the forefront of anyone's mind, but it is now entering the picture.
Machine Learning is used across many spheres around the world. The healthcare industry is no exception[1]. Machine Learning can play an essential role in

1. Identifying Diseases and Diagnosis[2][3]
2. Drug Discovery and Manufacturing[4]
3. Personalized Medicine
4. Machine Learning-based Behavioral Modification
5. Smart Health Records[5]
6. Better Radiotherapy
7. Outbreak Prediction

Such information, if predicted well in advance, can provide important insights to doctors who can then adapt their diagnosis and treatment per patient basis.

The data. The data we decided to choose are **Heart Disease UCI**. This database was created by

1. Hungarian Institute of Cardiology. Budapest: Andras Janosi, M.D.
2. University Hospital, Zurich, Switzerland: William Steinbrunn, M.D.
3. University Hospital, Basel, Switzerland: Matthias Pfisterer, M.D.
4. V.A. Medical Center, Long Beach and Cleveland Clinic Foundation: Robert Detrano, M.D., Ph.D.

One file has been "processed", the one containing the Cleveland database. That's the one we will be using.

The model. We will be using many methods that we learnt throughout the course of PSE in order to fit our model and make predictions. We will be using Regression Methods, PCA, Classifiers and Neural Networks.

The goal. Our goal in this project is to use different Machine Learning algorithms to decide which are the best in order to predict the presence/absence of a heart disease in a patient.

References

[1] Cleophas TJ.
Machine learning in medicine. New York: Springer; 2013.

[2] Kralj, K. and Kuka, M.
"Using machine learning to analyze attributes in the diagnosis of coronary artery disease". In Proceedings of Intelligent Data Analysis in Medicine and Pharmacology-IDAMAP98, Brighton, UK, 1998.

[3] "Strausberg, J. and Person, M.
"A process model of diagnostic reasoning in medicine". International Journal of Medical Informatics, 54, 9–23, 1999.

[4] R. Medina Marrero, et al.
QuBiLs-MAS method in early drug discovery and rational drug identification of antifungal agents SAR QSAR Environ. Res., 26 (2015), pp. 943-958

[5] Kaelber DC Jha AK Johnston D et al.
A research agenda for personal health records (PHRs). J Am Med Inform Assoc2008;15:729–36.