Machine Learning and the Ideal Algorithms for Determining Illness Across a Population

Data set

http://archive.ics.uci.edu/ml/machine-learning-databases/parkinsons/

Project Idea

It's almost common knowledge that computers are in most cases, faster, smarter and more efficient than humans in today's technologically filled world. The ideology that computers have already started excelling in jobs usually done by humans may scare many people, but in reality it may be more beneficial to us than ever expected. In regards to determine illnesses by statistical analysis computers are on the forefront of cracking down on early detection for a multitude of diseases.

The difference in time it may take some doctors to recognize symptoms versus a computer could be the difference between life and death. That is why for our project we would like to take a more indepth look at machine learning algorithms and how they classify certain diseases. The idea behind this project is to determine which algorithm or classification method is best to ensure no false positives or negatives, and also seeing how the outcome varies across a wide variety of diseases.

Software

The software needed will be a multitude of machine learning algorithms and classification methods to run through a dataset of a given illness. The software we plan on creating will help determine which algorithm used most successfully classified the illness and also ways to improve its results. All of this will be built on a text based user interface for ease of use and time constraints.

Papers To Read

Using Rule-Based Machine Learning for Candidate Disease Gene Prioritization and Sample Classification of Cancer Gene Expression Data

Classification of Diabetes Using SVM's

Parkinson Disease Gait Classification based on Machine Learning Approach

Teammates

Matthew Miller, Aashaka Desai, John Maloy

Milestone

At our halfway point we would like to have a skeleton structure of our UI and also each ML algorithm we plan to use and test our data with.