Capstone Project Proposal

Introduction

Dialysis is a treatment that does some of the things done by healthy kidneys. It's a treatment that takes over your kidney functions if those organs stop doing their job. There are several dialysis facilities registered with Medicare in the country where patients often visit. These facilities contain separate health data of patients which include mortality ratios (deaths), hospitalizations, blood transfusions, incidents of hypercalcemia (too much calcium in the blood), percentage of waste removed during hemodialysis in adults and children, percentage of waste removed in adults during peritoneal dialysis, percentage of AV fistulas, percentage of catheters in use over 90 days and others.

The goal of this capstone project is to predict a pattern of mortality rate to the separate health statistics listed on the above paragraph. Health institutions will primarily benefit from this project's analysis.

Data

A list of all dialysis facilities registered with Medicare will be utilized for this project. The data was collected from 2012 to 2015 and will be downloaded from the data.medicar.gov [*https://data.medicare.gov/Dialysis-Facility-Compare/Dialysis-Facility-Compare-Listing-by-Facility/23ew-n7w9/data*].

Method/Approach

I will treat this project as a unsupervised learning problem. The following libraries will be used for data loading, wrangling, cleaning, feature selection, array data structure, matrix manipulation, data visualization, classifier models, model evaluation etc.

- pandas
- numpy
- matplotlib
- scikit-learn

Deliverables

- Code along with document
- Power point slides followed by presentation