Springboard - Capstone Project proposal

By Sujata Pathak

Dated : 9/8/2017

* Problem Description –
  + The idea of this problem is from Kaggle competition
  + This competition ran from 2011-2013.
  + Problem Statement - Predict number of hospitalization days for each patient within the next year
  + The goal is to create an "early warning system" for managed care providers and provide a better way to identify which patients need care immediately to improve their health—and save on high costs that would be required during a hospital visit.
  + predict how many days each person will spend in the hospital in the one year after that three-years of data.
* Client and why to care about this problem –
  + Every managed care health organization’s main goal is to help improve overall health of patients by improving quality of care for their member population
  + The client will be managed health care organization who will be able to use this model by identifying high risk patients and by reducing unnecessary hospitalizations
  + The changes in intervention by identifying high risk patients will prove preventive and ultimately reduce costs for managed care organization as well as patients
* Data description and availability
  + Three years of historical medical data from anonymized real-life people
  + This data is available for public online
  + Specific data includes claim data for 3 years, Patients days in hospital file, Patients demographic information
* Approach to solving this problem

Work on different algorithms such as

* K-Nearest Neighbours
* Logistic Regression
* Support Vector Regression
* Random Forests
* Gradient Boosting Machines
* Neural Networks
* Deliverables
  + Documentation including code and algorithm used
  + Presentation slides