COMPSCIX 415.1 Homework 1

Sanatan Das

April 3, 2018

Code and Documents Git Repository

All the work can be found in the below Git repository location:

https://github.com/sanatanonline/compscix-415-1-assignments

Sanatan's Profile

- ***Summary :*** *My name is Sanatan Das. Currently I am working as a software developer at Wells Fargo, living in East Bay, Fremont, California. I have more than ten years of experience working in the software industry, mainly in Banking/Finance, Media and Retail domain. I have worked on different roles like developer, designer, architect etc. Currently I am a consultant for building web based banking solutions.*

- ***Education :*** *I have a Bachelors degre in Computer Science and Engineering and a Masters degree in Software Systems.*

- ***Objectives :*** *Almost two years back, I have started growing interests in Statisticical Learning and Data Science. Since then I have been taking online tutorials and watching videos on YouTube on these topics (Big Data, Hadoop, Mongo DB, Machine Learning and other processes and tools used in Data Science Methodologies). Almost a year back, I have joined UCBE Data Science certificate program and this is my third course. I am trying to change my career path to data science field as I have interests in statistical learning and data science. From this course, my expectation is to get through knowledge on Data Science Principles and best practices.*

Project Idea

As I am working on Banking and Finance domain, I have selected the below dataset (**National Student Loan Data System**). This is regarding national student loan data. My objective is to do a through analysis of the data (EDA) and create a model to predict the likelihood of a student loan approval, help Wells Fargo by predicting a warning message if the student can be a defaulter in future etc. More I will come to know once I start analysing it.

https://catalog.data.gov/dataset/national-student-loan-data-system