Capstone Project 1 Proposal

Chicago Crime

Abstract

The main objective of my project is to anticipate if a particular neighbor in a city, at a specific point in the day will be a crime hotspot or not, with an acceptable rate of accuracy. The dataset I am using comes from the City of Chicago. It reflects reported incidents of crime from 2001 to present, minus the most seven recent days. The Data is extracted from the Chicago Police Department system. This reports all types of crimes for the exception of murders where data exists for each victim. Here is the URL of the data set

https://www.kaggle.com/chicago/chicago-crime?select=crime. This is a project that hits home for me as I lived in Chicago for 3 years and I saw how bad the crime can get there over time. Chicago has a very large homicide rate and if I had to guess it is probably larger than most other major American cities. I will use several graphs to analyze the dataset, and apply predictive analytics to solve classification and prediction problems in our dataset. My model evaluation will be the following: A multi classification model to determine what is the exact crime we are looking for. I will also develop a binary classification model to predict if the crime was serious or not. In order to increase reliability, we need to compare the accuracy rate among various machine learning classifiers. Some of the variables in the dataset include whether an arrest was made, longitude, latitude, date crime occured, police district involved etc. It may be helpful to use some census data to get a better idea what the socio economic status is of each neighborhood

https://data.cityofchicago.org/Health-Human-Services/hardship-index/792q-4jtu/data .

This study will provide helpful insights into the patterns of crime in Chicago.