Capstone Project Proposal

Introduction

The Chicago Data Portal publishes a dataset reflecting reported incidents of crime that occurred in the City of Chicago from 2001 to present, minus the most recent seven days.

In this project, through machine learning, I would like analyze the crime dataset, classify the data and build a predict model for different types and trends of crimes in city blocks.

The Chicago and Police department would be the clients for this project. They can use my analysis report and model to identify the crime trend in each block, and therefore mobilize resources more efficiently and effectively to reduce crime rate. Meanwhile, they can use my work to understand the impact of community on crime rate and identify a way to improve the community environment.

Data Source

The data is available for download from the Chicago Data Portal as follows:

https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2

Approach

I will perform the following major steps to finish this project:

- Examine the dataset
- Apply Data Wrangling to the data
- Analyze the data statistics
- Apply Machine Learning techniques
- Extract conclusions

Deliverables

By completion of this project, I will have the following deliverables:

- Python codes for data analysis and machine learning
- A document describe the approach to be adopted
- A slide deck for the whole project and conclusion