

Final Project Proposal

Katelynn Cao (kwcao@ucdavis.edu), Kelvin Chao (kkwchao@ucdavis.edu), Kai Shen (kzshen@ucdavis.edu)

Instructor: Krishna Balasubramanian

Datasets:

The dataset obtained from Kaggle, was extracted from the Chicago Police Department's CLEAR (Citizen Law Enforcement Analysis and Reporting) system. It contains more than 6,000,000 records along with 22 features which reflect on the reported incidents of crime that occurred in the City of Chicago from 2001 to 2017.

Project Topic:

- I. Topic Intro [Chicago Crime Data] The city of Chicago has a reputation as one of America's most dangerous cities. While its crime rate has declined in 2018, the city continues to experience a higher homicides rate than the average U.S. cities. By analyzing the city's historical crime data from the past ten years and applying the method of K-mean clustering, we are hoping to sectionalize these crime records in order to identify the patterns of crimes across time and space, and construct a predictive model that can help the city in better allocate police forces and to prevent crimes from happening.
- II. Goal
 - A. The goal for this project is to find when, what, and where most crimes are committed in Chicago. Through the use of k-means clustering, we will highlight the general crime evolution over the timespan of 2007 to 2017. Following that, we are hoping to find the correlation between the types of crimes committed and the month of the year. And to finalize our goal, we would like to find the relationship between arrest rates and police districts from the datasets.
 - B. Our second goal is to predict what crime will create a model that will predict what specific crime will occur at a general location and time. Furthermore, it will predict whether or not an arrest will be made.
- III. Statistics Model and Algorithms
 - A. Data Preparation and Serialization
 - B. Basic Visualization of Characteristics and Trends
 - C. K-means Clustering
 - D. Linear Regression vs. Random Forest Regression vs. Gradient Tree Boosting
- IV. Reference
 - A. Crime dataset: <https://www.kaggle.com/currie32/crimes-in-chicago>
 - B. IUCR: <https://data.cityofchicago.org/Public-Safety/Chicago-Police-Department-Illinois-Uniform-Crime-R/c7ck-438e/data>