

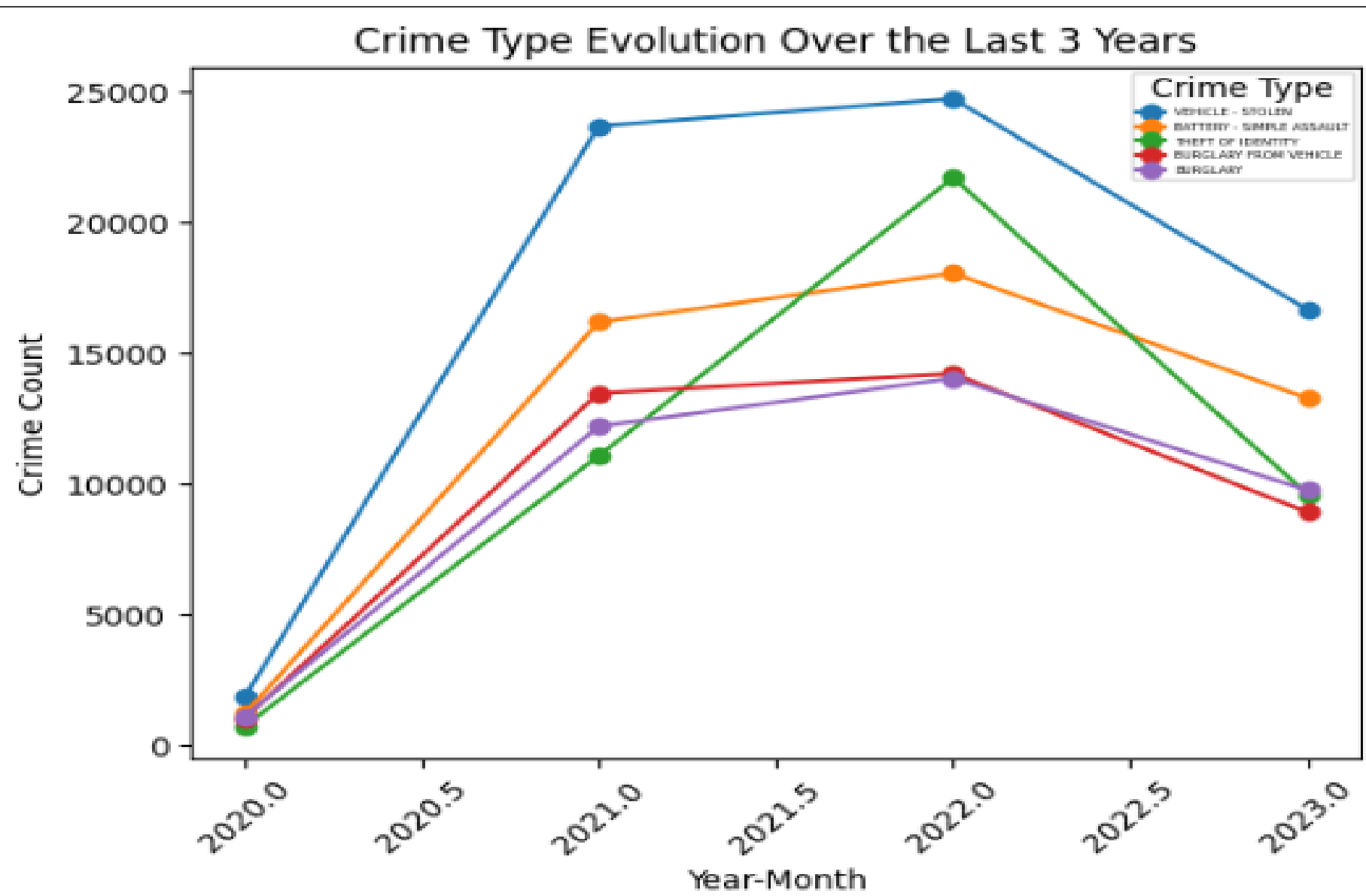
Analyzing Crime Trends in Los Angeles

San Diego State University
CS 577, Fall 2023



Problem Motivation

- Over the past 3 years crimes in LA have rapidly increased.
- Crimes are considered as difficult to predict because people could commit crime at random, however it would help reduce crime rate and future loss.



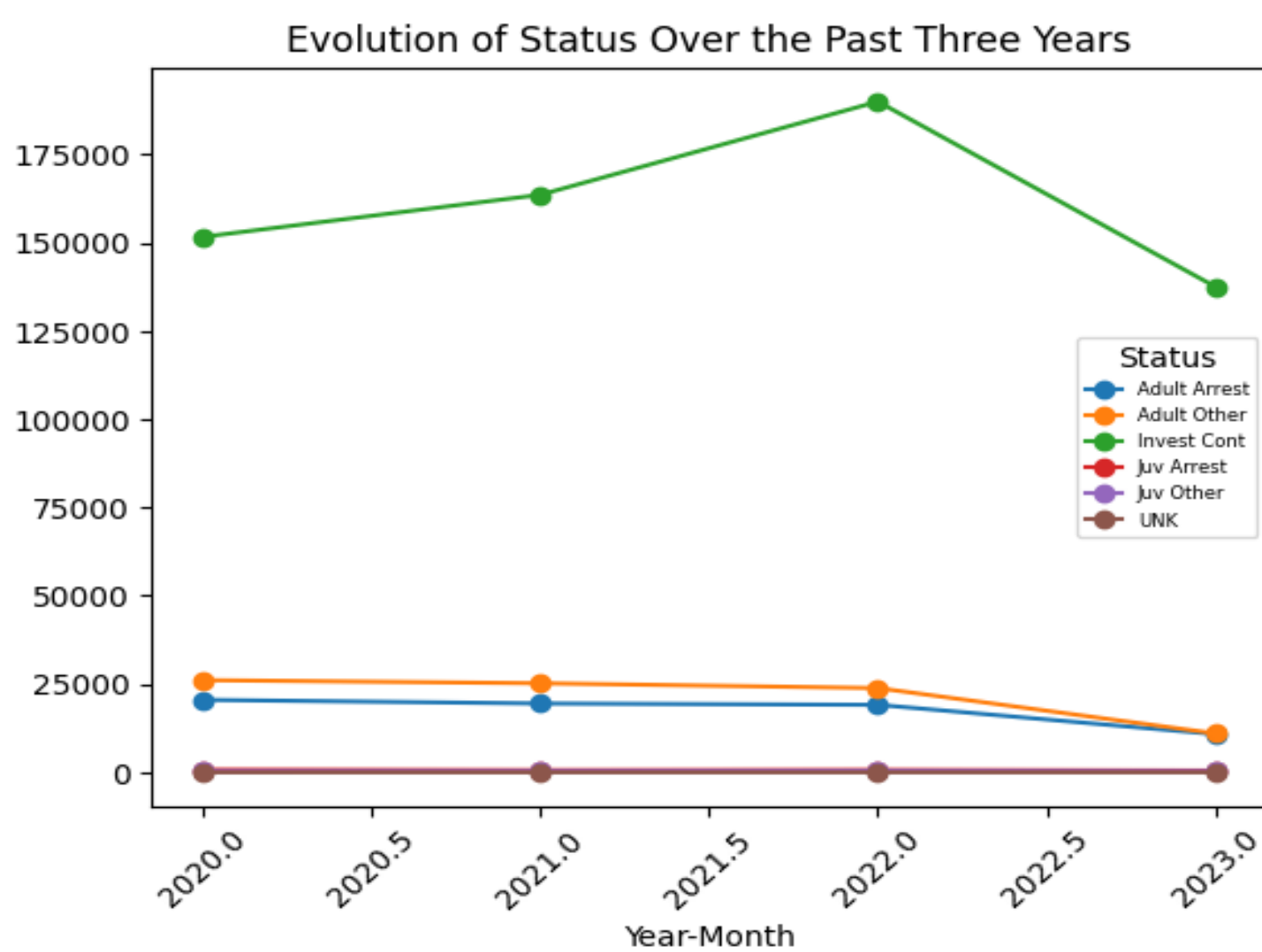
- Our motive is to find out insightful data extraction, analysis, and the discovery of patterns, trends, and possible correlations within the crime data.
- For the purpose of forecasting and predicting the top 5 criminal activities in the city, we are utilizing multiple machine learning models.
- Our ultimate goal in undertaking this project is to make Los Angeles a more secure and safe place.

Limitation of Past Work

- Feature Selection**
If some factors are excluded, this model won't be able to completely capture the intricacy of crime patterns.
- Temporal Considerations**
The model's performance may be sensitive to changes in crime patterns and the factors that influence crime over time.
- Confidentiality and Ethical Considerations**
Confidentiality must be maintained when working with features pertaining to suspects and victims. Make sure that the relevant legal requirements, privacy concerns, and ethical considerations are taken into account.

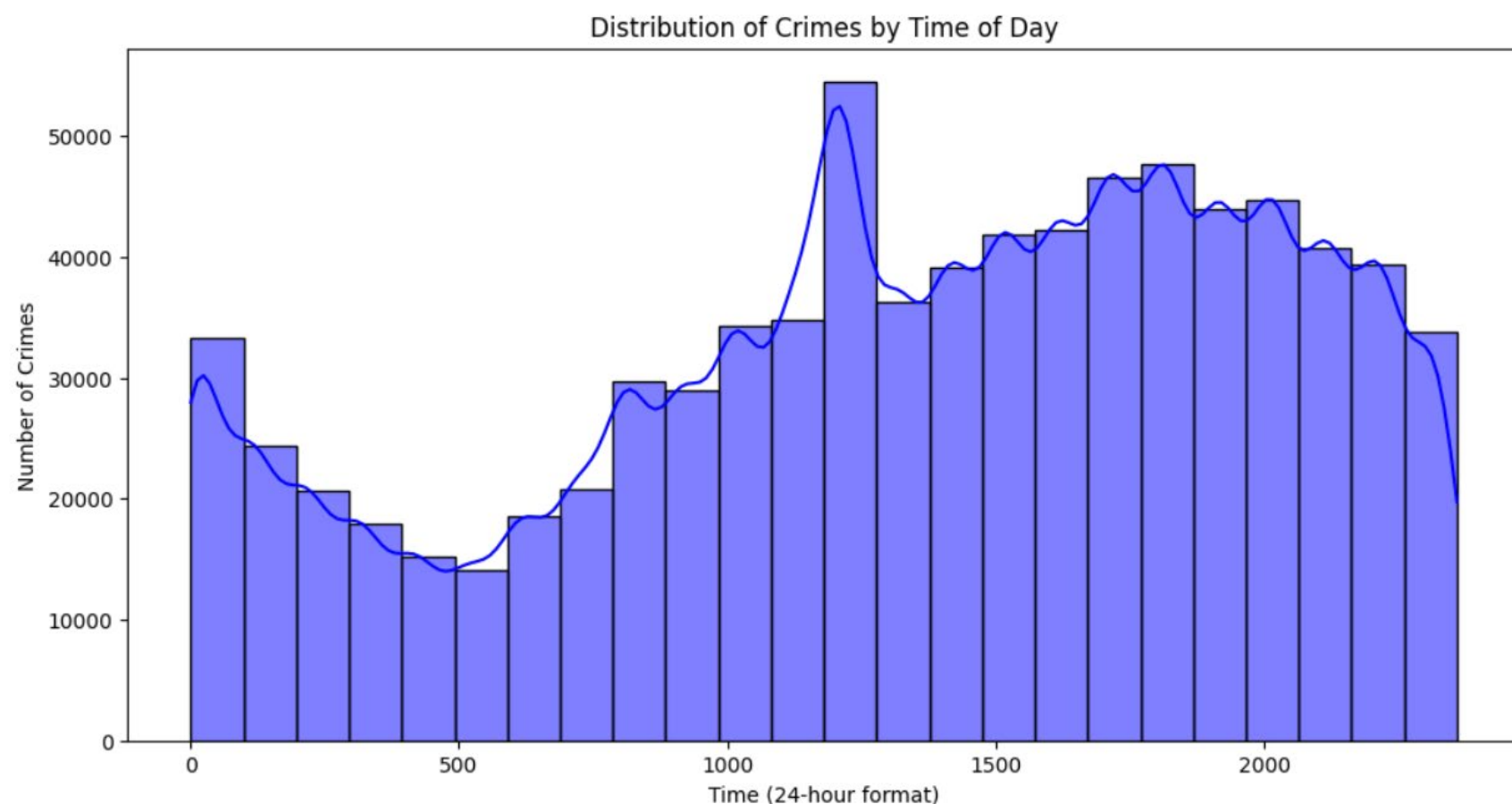
Understanding the Data

- This dataset has criminal activities data from 2020-till present having approx. 1million records.
- The data consists of around 28 columns having Crime description, Age, Sex, Location, etc.
- Our goal is to use information from past crimes to help create a model that will predict the category or type of crime that might occur given a set of features.

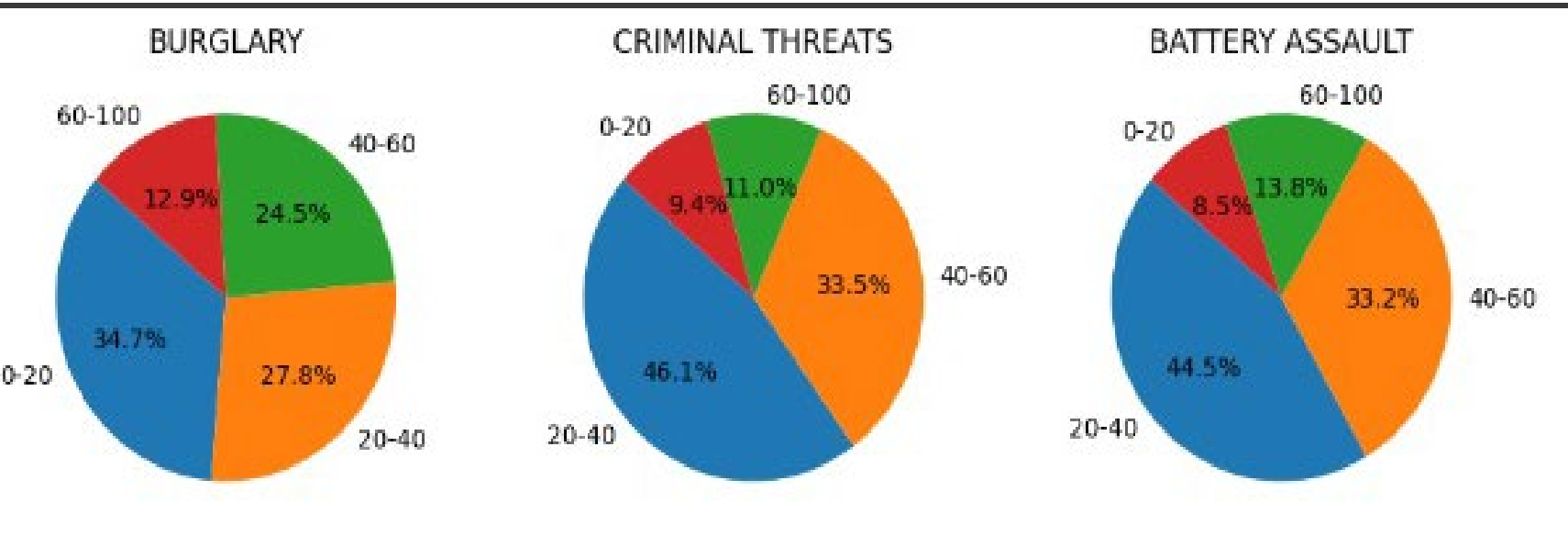


Approach and Data Analysis

- Our approach to deal with this large dataset involved multiple steps like: Handling missing values, convert into categorical data(hot encoding), Feature selection and model training.
- We did data analysis and found out insightful finding regarding the crime in LA.

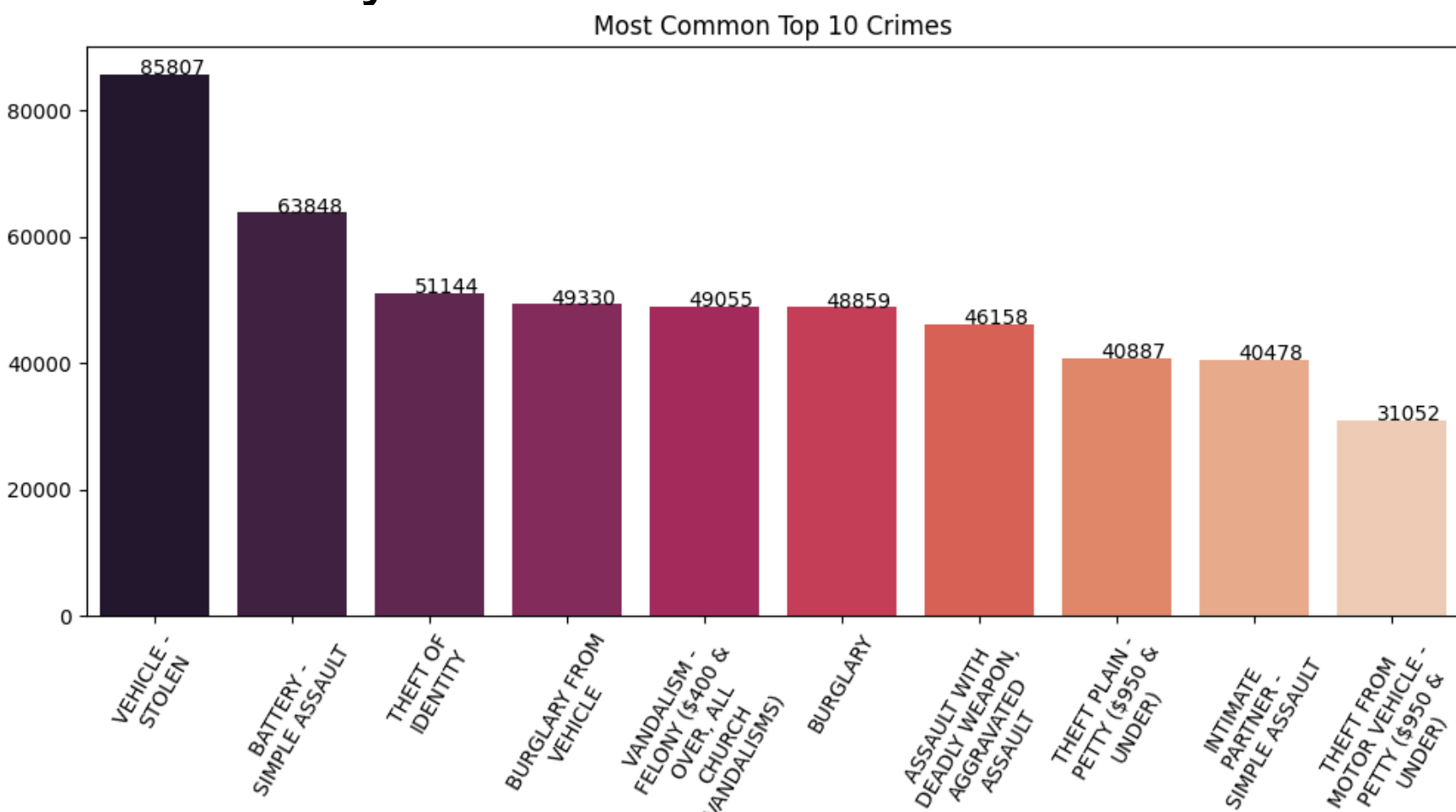


- Most of the crimes are taking place in daytime.
- Various type of crime occurrence in specific age group.

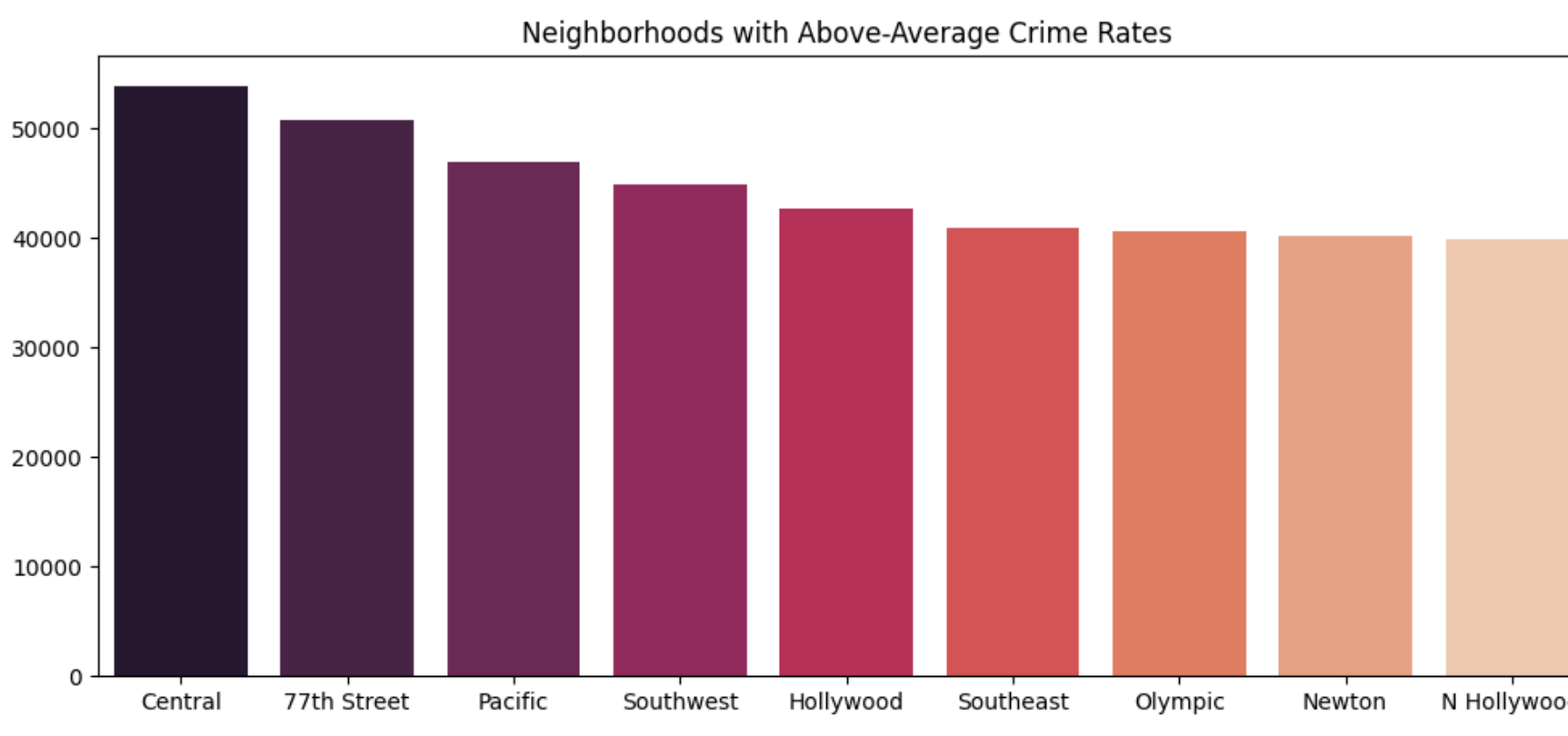


Exploratory Data Analysis

- Among all reported crimes, Vehicle Stealing is the most prevalent, followed by simple assault and Identity theft.



- Central LA, 77th Street and Pacific are neighborhoods that reported the highest incidents.



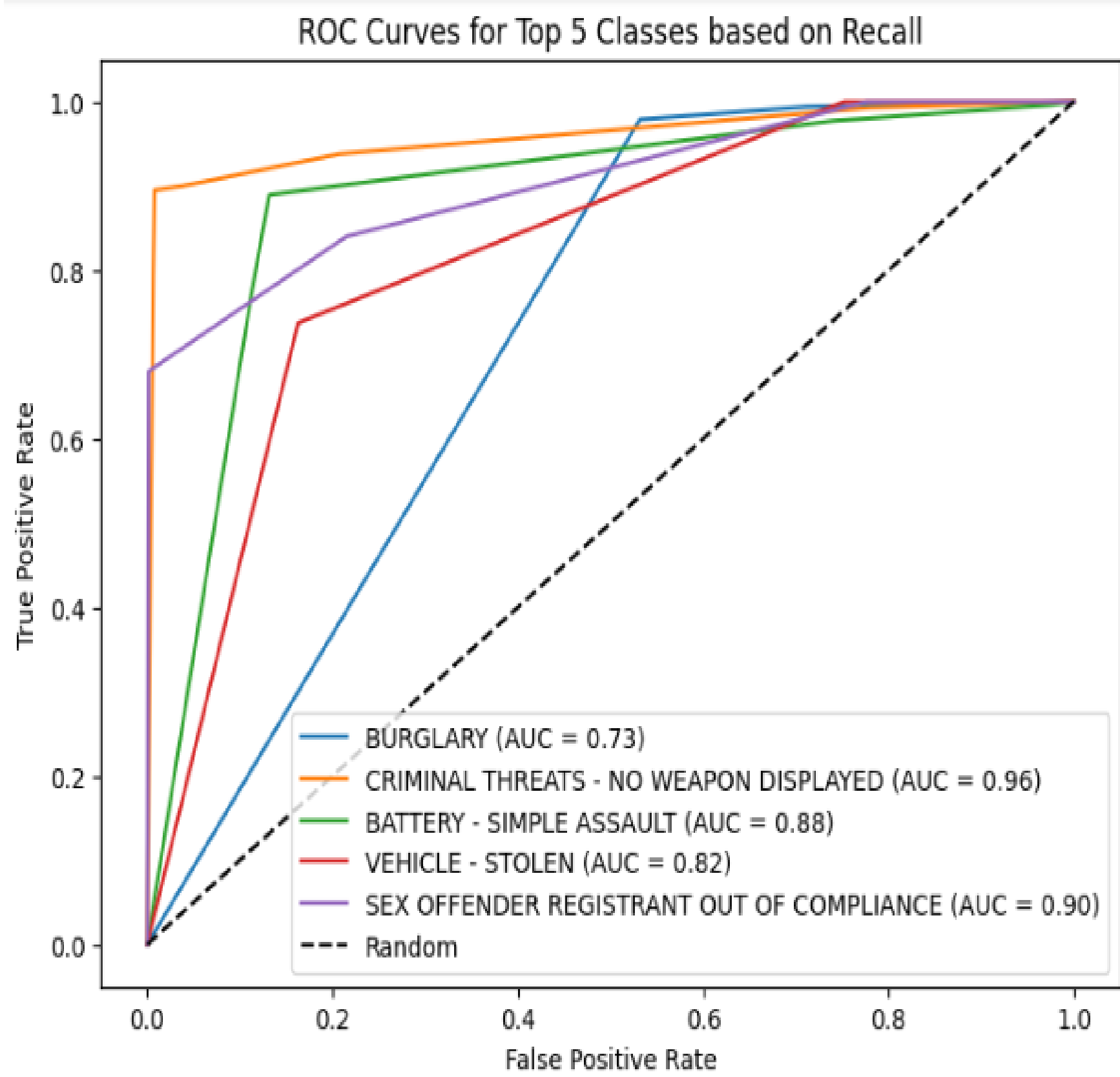
Experimental Results

- To predict and categorize the type of crime, we employ predictive models that makes use of LA Crime data (Crm Cd Desc) using several features, such as the incident's date and time, location, victim's age, victim sex, and description of crime.

Features	Precision	Recall	F1-score
BURGLARY	0.113241	0.979183	0.203005
CRIMINAL THREATS - NO WEAPON DISPLAYED	0.706215	0.894988	0.789474
BATTERY - SIMPLE ASSAULT	0.388238	0.888695	0.540397
VEHICLE - STOLEN	0.305009	0.737408	0.431528
SEX OFFENDER REGISTRANT OUT OF COMPLIANCE	0.404762	0.680000	0.507463

- The model tries to forecast what kind of crime is most likely to occur given details about the date, time, location, age of the victim, description of the weapon, and premises.
- We used diverse set of models for crime predictions including Random Forest, Neural Network, KNN, etc.

Experiment Results



Classifier	Accuracy
Random Forest	65.13%
Decision Tree	58.34%
K-Nearest Neighbors	57.18%
Neural Network	74.56%

Conclusion - Future Work

- These models can help us predict the type of crimes that are bound to happen in LA.
- Key features influencing crime prediction are time, location, victim's age, weapon, etc.
- Males are more likely to be victim of Battery Assault, criminal Threats and vehicle stolen.
- The most occurrence of crimes is observed to happen in the age group of 20-40.
- Future Work :- Accuracy of these models can be improved by employing a denser and more focused dataset.

References

- [1] Crime Analytics and Prediction in Los Angeles by Hye Lim Park, UCSD.
- [2] A Data-Driven Exploration of Crime Trends in Los Angeles, CSULA.