Predicting Crime in Boston

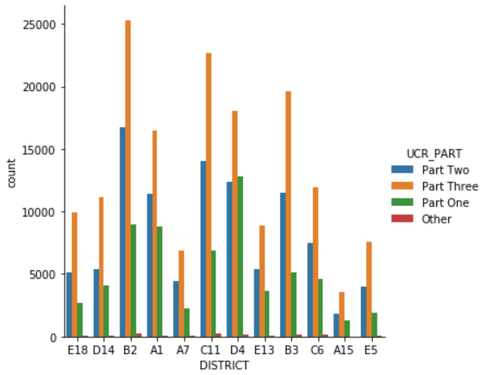
Julie Osborne, Chris McCooey, Mary Wishart

Abstract

Boston is the largest city in New England. Being a major metropolitan area means that crimes occur around the city every day. Our project takes a look into the ”Crimes in Boston” dataset. This is a dataset of police reports from the last From the information in these data, we used machine learning techniques to see if given a specific crime if we could predict where that crime is most likely to happen within the various neighborhoods in Boston.

1 Introduction

- Purpose: Try to predict where various crimes are likely to happen within the city of Boston
- Use Multi-class logistic regression to predict where a specific crime is likely to happen
- Create a map of Boston which visualises our results



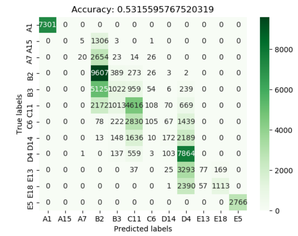
A graph showing the police districts along with the amount of crime organized by UCR part number.

2 Setup to our model

- Each location of city is organized by BPD districts
- 12 districts in total
- Individual crimes are given an offense code number. These are then put into offense groups
- We predicted crimes based off of their offence group, not the individual offense code number
- 67 offense groups in total

3 Multi-class logistic regression

We created a multi-class logistic regression algorithm to try to compute where a crime was most likely to happen. This model had an accuracy of 0.53. Thus it correctly predicted what district a crime was most likely to happen around half the time.



A Confusion Matrix representing our mutli-class logistic model

3.1 Visualizations

Citations

Jain, A. (2018, February). Crimes in Boston, Version 3. Retrieved November, 2019 from https://www.kaggle.com/ankkur13/boston-crime-data.

Boston Police. “The Boston Police Department’s Virtual Community.” Bpdnews.com, 7 Dec. 2019, https://bpdnews.com/.