Spartan Security: Crime Forecasting using Time Series

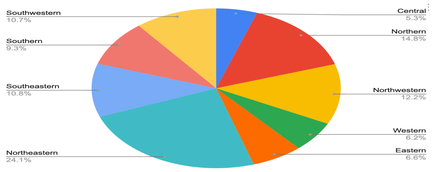
Computer Engineering Department

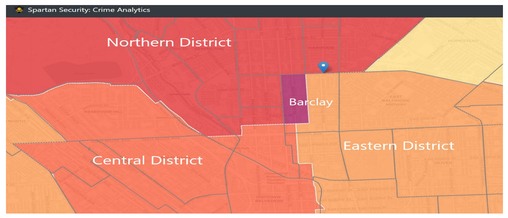
Project Advisor: Mahima Agumbe Suresh

Bairagi, Nikita(MS Software Engineering)
Madappa, Tejas (MS Computer Engineering)
Sidhabathuni, Divya
(MS Software Engineering)
Swaroop Mishra, Swayam
(MS Computer Engineering)

Introduction

The purpose of this project is to research and develop a machine learning model to forecast the next crime incident and help provide more awareness and safety for those areas at risk. In this project we aim to use the data points such as crime location, time and nature of crime from victim based crime data and combine these data points with the census data.

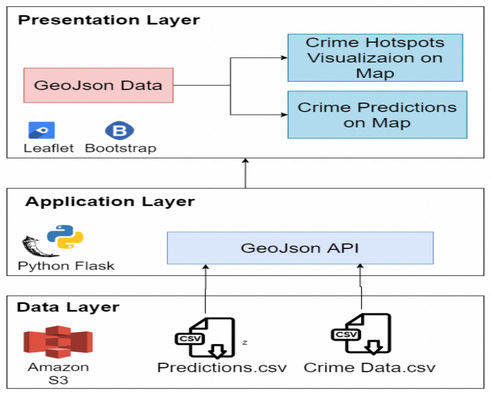




- Baseline research paper divided the city based on clustering. We divided the city based on police districts and have 9 sub-units unlike 20 that is mentioned in the baseline paper.
- The authors considered only crimes with highest priority. We on the other hand considered every type of crime i.e. crime with highest priority as well as crime as low priority crime.

Methodology

Web Architecture

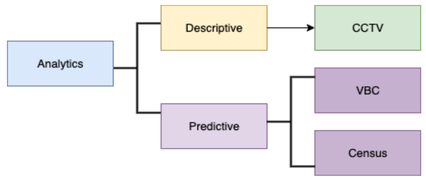


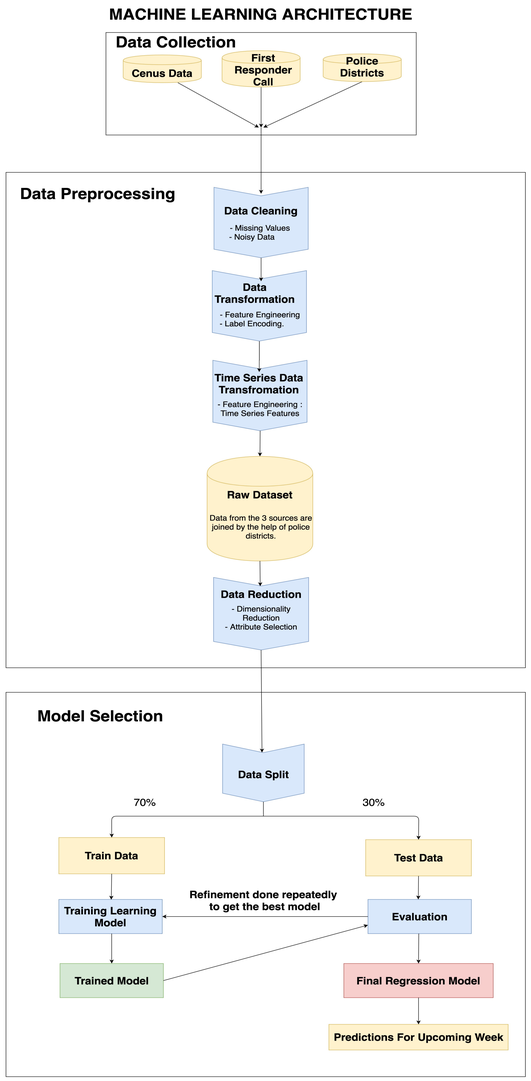
Methodology

Machine Learning Architecture

Our model was trained for 284 weeks or 4 years of historical data and tested on 52 weeks or 1 year of data.

VBC – Victim Based Crime





Analysis and Results

Forecasting without census data using expanding window outperformed by 87.5% for all crime types by giving us an average MSE of 273.097 and average MAE of 12.608 for all the districts.

Evaluation Metrics		
Model Name	Our MSE	Research Paper's MSE
MLPR	736.438	1718.571
SVR	736.438	961.326
RFR	2253.103	897.149

TABLE I
PERFORMANCE COMPARISON OF BASELINE IMPLEMENTATION BETWEEN OUR APPROACH AND RESEARCH PAPER.

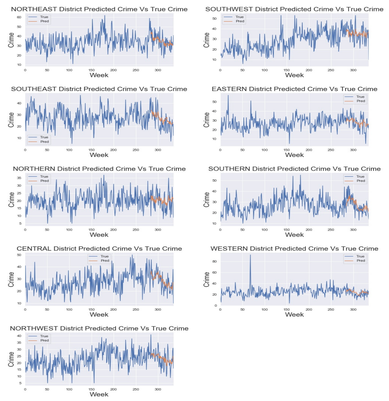
The below table represents the performance of Regressor (RFR), Multi-Layer Perceptron Regressor (MLPR) and Support Vector Regressor (SVR) on 1 year of time series data. In which RFR outperformed all other models.

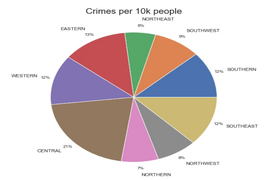
Evaluation Metrics		
Model Name	MSE Value	MAE
MLPR	22448716335.04	73574.90
SVR	1285.96	29.14
RFR	824.18	22.84

TABLE II
PERFORMANCE OF DIFFERENT REGRESSION MODEL ON 1 YEAR OF TIME SERIES DATA USING EXPANDING WINDOW FORECASTING.

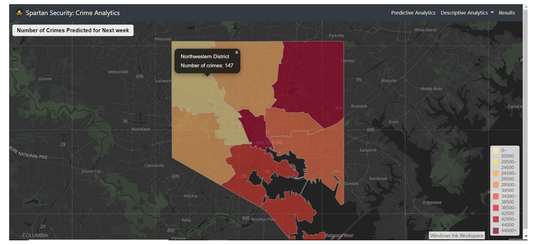
The graphs display results of True crime vs Predicted crime for Random Forest below for each of the police districts.

1) Forecasting without census data: Our approach of expanding window forecasting from forecasting frames outperformed the our baseline by 87.5% for all crime types





The predictions for the week are displayed as on overlay on the Police districts.



Summary/Conclusions

- Combination of Census data and police district information.
- Integration of CCTV dataset into web application.
- Prediction of crime with and without census data.
- Division of area based using intersection of overlapping areas.

Key References

[1] J. Borges, D. Ziehr, M. Beigl, N. Cacho, A. Martins, A. Araujo, L.Bezerra, S. Geisler, "Time-Series Features for Predictive Policing," 2018 IEEE International Smart Cities Conference (ISC2), Kansas City, MO, USA, 2018, pp. 1-8.

[2] Perry, Walter L., Brian Mcinnis, Carter C. Price, Susan C. Smith, and John S. Hollywood. Predictive Policing. RAND Corporation, 2013. Web.

[3] A. Malik, R. Maciejewski, S. Towers, S. McCullough, and D. S. Ebert, "Proactive spatiotemporal resource allocation and predictive visual analytics for community policing and law enforcement," Visualization and Computer Graphics, IEEE Transactions on, vol. 20, no. 12, pp. 1863–1872, 2014.

[4] Districts, Baltimore Police Department, Accessed on:June 14,2020. [Online]. Available:https://www.baltimorepolice.org/districts/find-my-district

Acknowledgements

The authors are deeply indebted to Professor Mahima Agumbe Suresh for her invaluable comments and assistance in the preparation of this study.
We would also like to thank Professor Dan Harkey for his continued support and guidance. We would like to thank Professor Kaikai Liu for his inputs on Industry Standards.