

ABSTRACT

Analyzing Crime Statistics using Machine Learning techniques for Smart City Applications

Modern cities and communities are not as effective as they could be with their services. Although a vast amount of data is collected, it isn't being utilized to distribute resources effectively. A smart community is capable of managing its resources and services solely based on data collected by sensors throughout the community. Montgomery County, MD is one of the richest counties in the United States. It is not currently a smart community. However, data collected within the community can help make Montgomery County smarter. This paper presents an analysis of possible predictors of crime based on Montgomery County's recorded crime incident database. The crime data was expected to have daily, weekly, monthly, and seasonal temporal trends, high positive correlations with spatial quantities such as population, average house prices, and urbanity. Additionally, it was hypothesized that a relationship between crime in multiple zip codes could be used to predict future crime frequencies in those zip codes. Using evidence from various data analysis methods and machine learning techniques, it was found that the crime data in Montgomery County has a weekly and daily seasonality is correlated to the level of urbanity, and that crime in certain zip codes can help predict crime in others. This project opens an avenue for further research regarding the crime patterns. In the future, these patterns can be exploited to create a safer county.