Public Safety through GIS : Analysis and Predictions of Criminal Activities

Khanjan Shah | Dr Amrit Pal | School of Computer Science and Engineering

Motivation / Introduction

The idea of this poster came to the picture while observing problems and challenges faced by **G**eographical **I**nformation **S**ystem(GIS) in context of localized crime predictions. Public safety is very much important and due to increasing number and usage of technologies, day by day people are more dependent on smart devices. Use of GIS technology along with analytical and visualization approaches can help in facing the social challenge of public safety. The problems identified in this context are as follows:

- Accurate prediction of geographical areas where crimes can happen.
- Development of an alarming system, when any crime happens in a particular area.
- A quick response system from the police using GIS.
- A GIS based record of crimes and visualization of the same.

To overcome these challenges, models and techniques are need to be developed.

SCOPE of Project

This project proposes prediction and visualization techniques for spatial mapping of the crimes using GIS:

- The project is very helpful for government organizations.
- Government can use this project to track crime hot spots.
- This project can be used to set alarm on crime places so people can be made aware.

METHODOLOGY

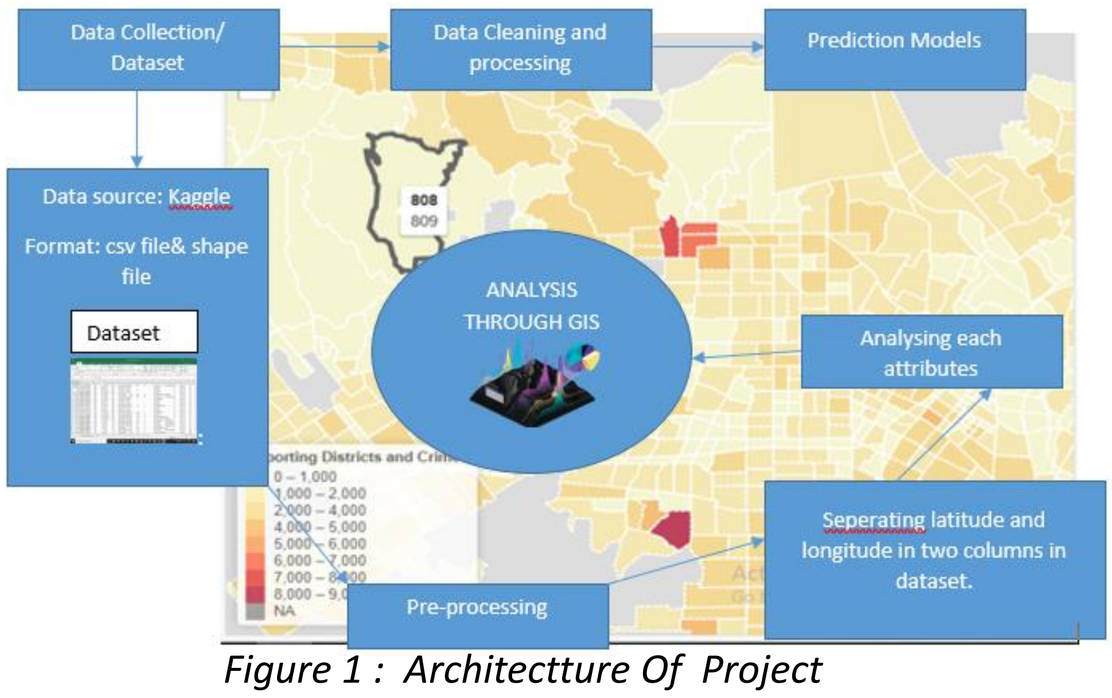

Figure 1 : Architectture Of Project

Step 1: To collect datasets

Step 2: Pre-processing done on the dataset (like : cleaning dataset by deleting NaN Values, Eliminate columns that are not required for analysis. It takes my 80 % of time. The most effective step in data science.

Step 3: Analysis should be done for types of crimes in particular cities and total number of crimes happened in particular city.

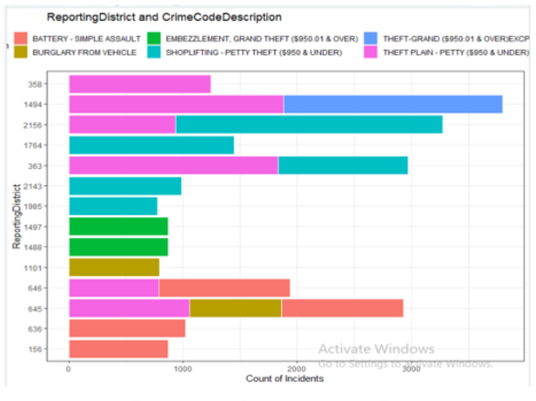

Figure 2: Analysis of type of crime in cities

Figure 3: Crime Counts In Particular city

Step 4: Prediction on dataset to predict crime rates in particular area and which crime happens in particular area. (using TPOT tool, random forest, decision tree and neural networks algorithms and models.

$$\text{Base Rate} = \#incedents\left(\frac{\#incedents}{Relevant\ Population\ Size}\right) \times Base$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

RESULTS

Pre-processing and Analysis Results

Figure 4 : Analysis Of Data On Map

Important Features	Inputs	Results			
METHOD_CODE	0.395665	Precision	Recall	F1-Score	Support
AGE	0.243054	0.53	0.80	0.64	22202
PSA	0.120549	0.54	0.58	0.56	17549
HOUR	0.082480	0.43	0.06	0.11	4222
LONGITUDE	0.045136	0.65	0.48	0.56	3634
LATITUDE	0.044909	0.62	0.29	0.39	5164
WARD	0.020519	0.33	0.02	0.03	4406
YEAR	0.017460	1.00	0.02	0.03	193
SHIFT_MIDNIGHT	0.013668	0.00	0.00	0.00	467
SHIFT_EVENING	0.010839	0.00	0.00	0.00	39
DAY_OF_YEAR	0.003062				
DAY	0.002004		Avg/ Total		
WEEK	0.000645				
MONTH	0.000010	0.52	0.54	0.50	57876
QUARTER	0.000000				

Table 1 : Inputs and Confusion Matrix Results

CONCLUSION

From my project on public safety through GIS, I have made analysis of crimes using R-studio tool on crime dataset and also did predictions through machine learning approaches and found random forest algorithm to be more accurate. Now I decide to increase the accuracy of my model using deep neural network.

CONTACT DETAILS

Student mail id: khanjanjayraj.shah2020@vitstudent.ac.in
Guide mail id: amrit.pal@vit.ac.in