Crime Prediction Using Neural Network

Toukir Ahmed Sojib	17.01.04.095
Rizeya Rabbi Reyad	16.01.04.082
Ishtiaque Ahmed Nuhash	16.01.04.128
Nakib Uddin Ahmed	16.01.04.137

Project Report

Course ID: CSE 4238

Course Name: Soft Computing Lab

Semester: Fall 2020

Department of Computer Science and Engineering

Ahsanullah University of Science and Technology

Dhaka, Bangladesh

October 2021

ABSTRACT

Recognizing the patterns of criminal activity of a place is paramount in order to prevent it. Law enforcement agencies can work effectively and respond faster if they have better knowledge about crime patterns in different geological points of a city. The aim of this project is to use deep learning technique to classify a criminal incident. The experimentation is conducted on a dataset containing San Francisco's crime records from 2003 - 2015. For this multiclass problem we built a neural network to classify crimes into different categories.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Criminal activities are present in every region of the world affecting quality of life and socio-economical development. As such, it is a major concern of many governments who are using different advanced technology to tackle such issues. Crime Analysis, a sub branch of criminology, studies the behavioral pattern of criminal activities and tries to identify the indicators of such events. Deep learning agents work with data and employ different techniques to find patterns in data making it very useful for predictive analysis. Law enforcement agencies use different patrolling strategies based on the information they get to keep an area secure. A deep learning agent can learn and analyze the pattern of occurrence of a crime based on the reports of previous criminal activities and can find hotspots based on time, type or any other factor. This technique is known as classification and it allows to predict nominal class labels. Classification has been used on many different domains such as financial market, business intelligence, healthcare, weather forecasting etc. In this project, a dataset from San-Francisco Open Data [1] is used which contains the reported criminal activities in the neighborhoods of the city San Francisco for a duration of 12 years. We used a neural network model to classify crimes into different categories.

Chapter 2

Literature Reviews

As combating criminal activity has always been a priority for governments around the world, many researches has been done to effectively find countermeasures and indicators of crime prior to happening. Criminologists have been pursuing to identify hotspots that need major attention from law enforcement agencies.

Analyzing the usage of mobile network infrastructure and demographic information of people living in different areas of London, a group of researchers were able to predict if particular areas of London would become a criminal hotspot [2]. They have implied that anonymized data collected by mobile networks contain indicators for predicting crime levels.

Combining two datasets - 1990 US LEMAS and crime data 1995 FBI UCR and applying classification techniques like Decision Tree and Naive Bayesian algorithm, 83.95accuracy have been achieved when asked to predict a crime category for different states of
USA [3]. However, this paper does not disclose if there were any imbalanced classes of crime category. The same databases were also explored by Somayeh et al [4] who employed a number of machine learning algorithms, where k-Nearest Neighbor algorithm performed better than other algorithms by having an accuracy of 89.50feature to improve the feature selection.

Wang et al [5] proposed the Series Finder, a machine learning agent that tried to find patterns in crime committed by same offender or groups of offenders. Clustering has also been used to study patterns of criminal behavior and geographic criminal history.

Remond and Baveja [6] have worked on the data noise problem and studied how some police reports or cases are idiosyncratic and do not contain good indicative matrices. Their proposed system called Case-Based Reasoning (CBR) filtered out these cases, and using this system, they were able to predict better compared to not having any filters on the data.

Chapter 3

Data Collection & Processing

The experiment is conducted on a specific dataset. The dataset is provided by SF Opendata from SFPD Crime Incident Reporting System [1]. It provides information on crime incidents that occurred in San Francisco for the period of 1/1/2003 to 5/13/2015. The dataset is a csv file containing 878049 rows. The attributes are given below.

Datetime	A timestamp when the given crime occurred.
Category	Type of crimes. This is the target label for the data. There are 39 types ofcrime listed in the data
Crime Description	A detailed description of a specific type of crime
Day	Day of week
pDistrict	Name of police department district. There are total 10 Police Districts in thedata
Resolution	How the crime was solved. 17 types of resolution
Address	The approximate street address for the incident
X	It signifies the latitude of the location of the crime.
Y	It signifies the longitude of the location of the crime.

Table 3.1: Attributes of the crime dataset

From the list of attributes in 3.1, the features and label can be determined. The target label that needs to be predicted is the Category of a crime incident. The attributes: Crime Description and Resolution are also related to the target label. Hence, all other attributes apart from these three attributes are used as features.

Category	Frequency
LARCENY/THEFT	174900
OTHER OFFENSES	126182
NON-CRIMINAL	92304
ASSAULT	76876
DRUG/NARCOTIC	53971
VEHICLE THEFT	53781
VANDALISM	44725
WARRANTS	42214
BURGLARY	36755
SUSPICIOUS OCC	31414
MISSING PERSON	25989
ROBBERY	23000
FRAUD	16679
FORGERY/COUNTERFEITING	10609
SECONDARY CODES	9985

Table 3.2: Frequency of top 15 crimes

There are 39 types of crime in the San Francisco Crime Dataset. These types are considered as classes and having 39 classes makes it a multi class problem. Frequency of top 15 crimes are given3.2

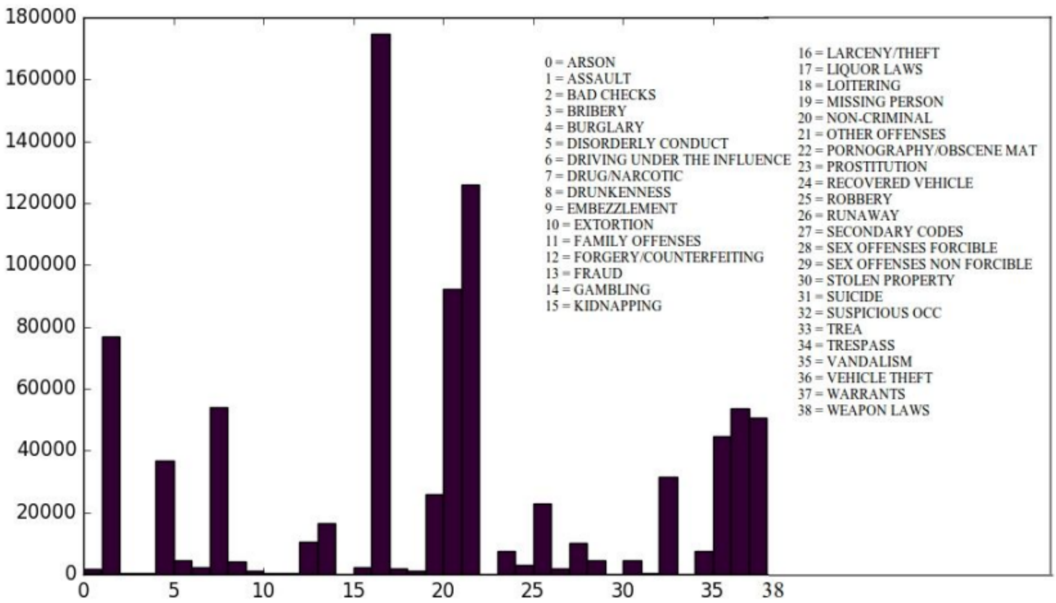

Figure 3.1: : Frequency of crime categories

Most crimes occur during afternoon-evening. From midnight to morning, reports of criminal activities are low. There is an upsurge of criminal activities at 6 PM and 8 PM. Criminal activities are drastically reported around 9 AM and it continues to show a gradual increase throughout the day peaking at 6 PM, after which it starts to decrease.

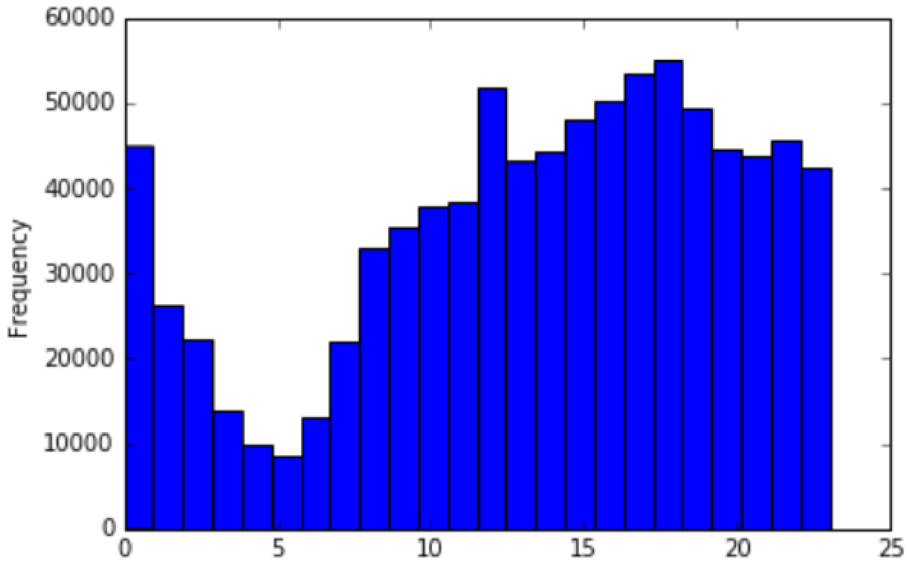

Figure 3.2: : Criminal activities occurring in different hour of day (in 24 hours format)

'Latitude' (X), 'Longitude' (Y) have more than 30,000 unique entries of the total 878049 entries, while 'Address' has total 23228 unique Entries. One problem with the location features is, there are 26,533 entries for a specific address. Which is the location of San Francisco Police Officer's Association. This default address gives data a low variance problem.

Among the ten police districts, criminal activities in the southern district is higher than any other district.

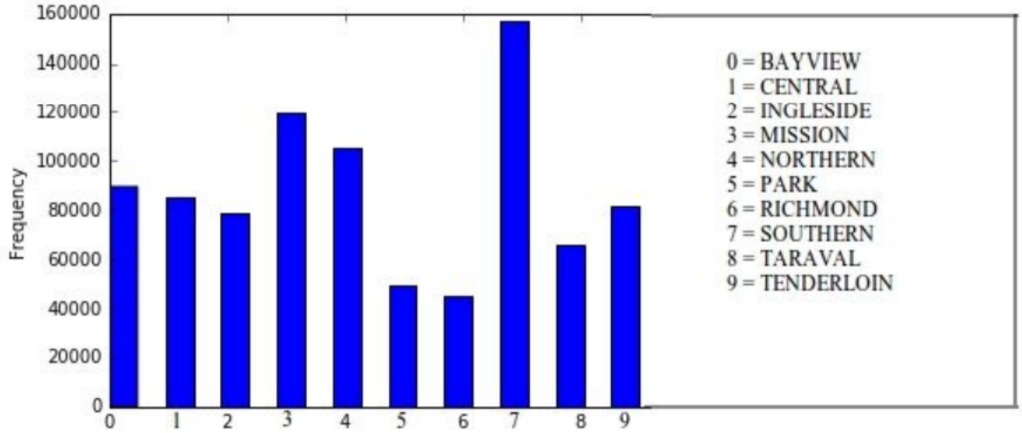

Figure 3.3: : Crimes occurring in different police district

From datetime stamp,four main features are extracted. - year, month, day, hour.Frequency of crime over time are given below.

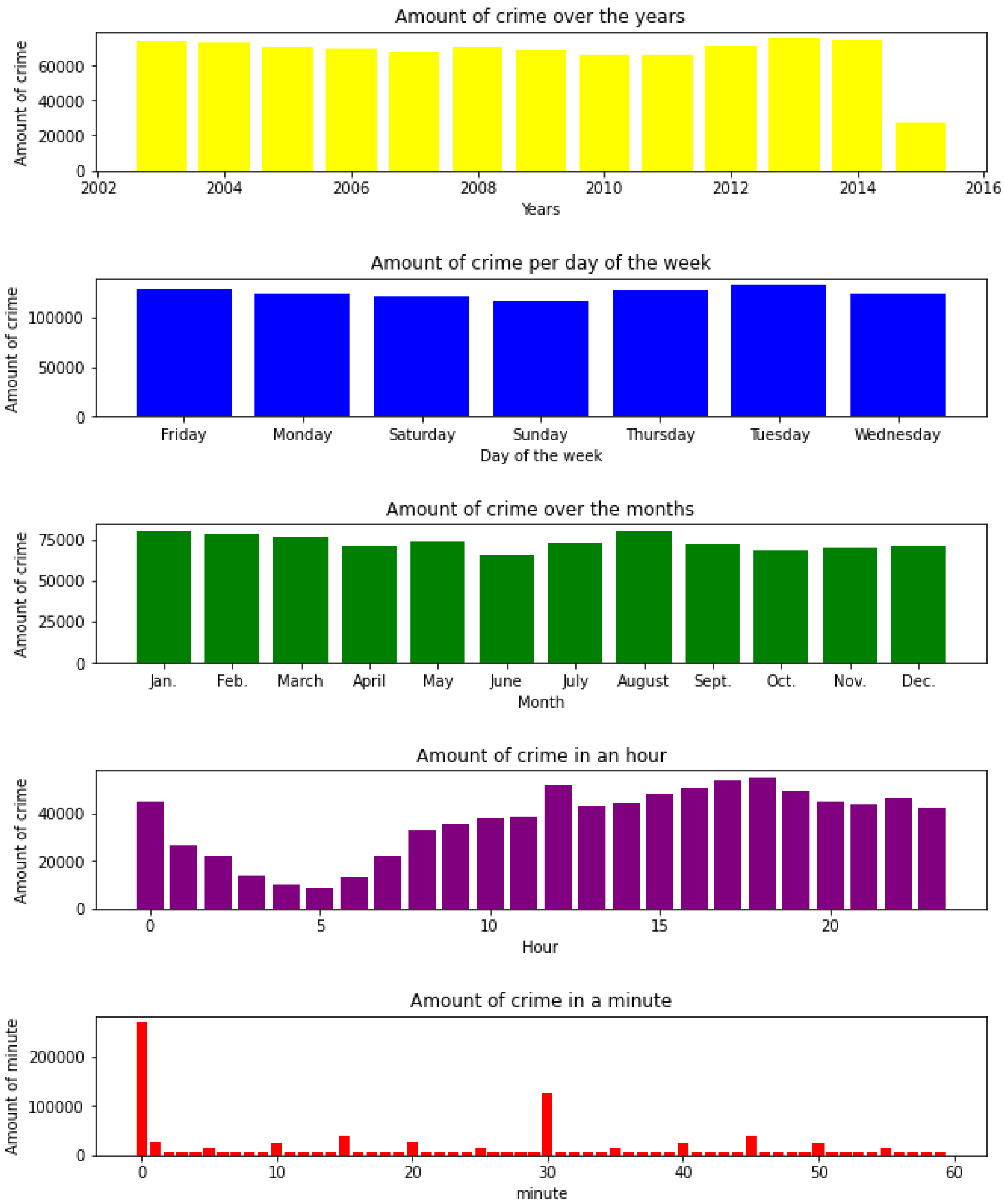

Figure 3.4: : Frequency of crime over time

Chapter 4

Methodology

Attributes with string data type are “DayOfWeek”, “Category”, “Addresss”,“PdDistrict” columns. Scikit-learn has a preprocessing package "LabelEncoder" that converts string data into numeric data. This package gives an integer value to each unique item after sorting items in ascending alphabetical order. Datetime attribute is also a string data type, however this is converted into a datetime object and four different attributes are obtained from it:“Minute”, “Hour”,”Day”,”Month” and “Year”.
To avoid overfitting and getting more realistic accuracy, the dataset is divided into two portion: testing dataset and training dataset. Training dataset contains all features along with the target label. Testing dataset only contains the features from which neural network model predicts the target label. In order to feed the data to our model, we must first make y categorical and standardisetrain and test data.Then we created and tuned a neural network and measured the results over epochs and plotted the reults in graphs.

Chapter 5

Experiments and Results

For our multiclass problem we built a sequential neural network model and tuned it through different dense layer, batch normalization and dropout. After running the model results are plotted into graphs.
Results 5.1of some of the epochs are given below.

Epoch	Loss	Accuracy	Val_loss	Val_accuracy
1	2.7085	0.2063	2.5538	0.2315
10	2.4606	0.2561	2.4825	0.2489
20	2.4340	0.2635	2.4484	0.2592
33	2.4262	0.2644	2.4423	0.2604
40	2.4149	0.2699	2.4503	0.2557
59	2.4113	0.2715	2.4547	0.2544
74	2.4004	0.2731	2.4328	0.2607
86	2.3962	0.2743	2.4396	0.2581
100	2.3962	0.2752	2.4220	0.2674

Table 5.1: Results over random epochs

From the above table we can see that Accuracy increases over epochs.

plotted graphs are given below.

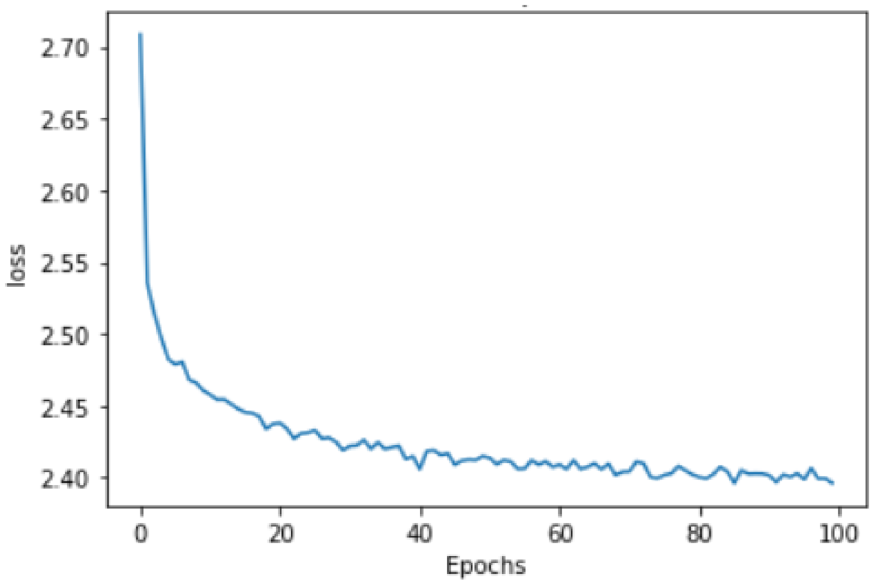

Figure 5.1: : Loss over epochs

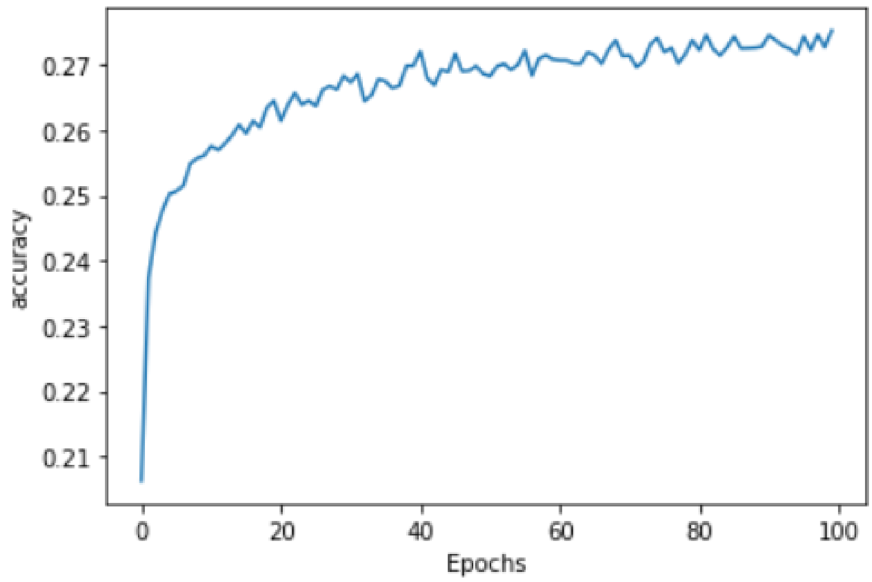

Figure 5.2: : Accuracy over epochs

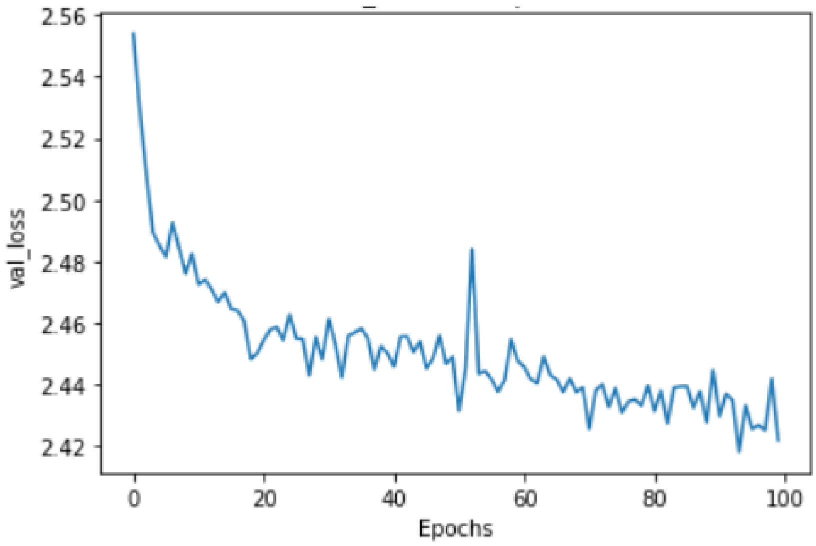

Figure 5.3: : Val_loss over epochs

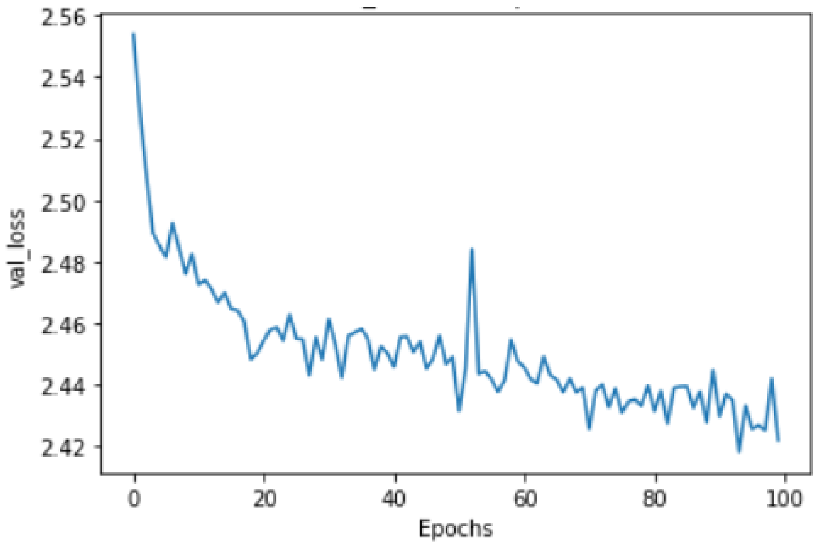

Figure 5.4: : Val_accuracy over epochs

We applied different machine learning algorithms in our pattern lab project and gained around 22 percent accuracy on average. In this project we applied neural network model and gained around 27 percent accuracy on the same dataset.

Chapter 6

Future Work and Conclusion

Neural network gives around five percent more accuracy than machine learning algorithms. By preprocessing the data in other better ways and tuning the neural network more accurately may increase accuracy.

References

[1] Kaggle, *San Francisco Crime Classification*. PhD thesis, https://www.kaggle.com/c/sf-crime/overview, 2017.

[2] A. Bogomolov, B. Lepri, J. Staiano, N. Oliver, F. Pianesi, and A. Pentland, "Once upon a crime: Towards crime prediction from demographics and mobile data," in *Proceedings of the 16th International Conference on Multimodal Interaction*, ICMI '14, (New York, NY, USA), p. 427–434, Association for Computing Machinery, 2014.

[3] R. Iqbal, M. A. A. Murad, A. Mustapha, P. H. S. Panahy, and N. Khanahmadliravi, "An experimental study of classification algorithms for crime prediction," *Indian Journal of Science and Technology*, vol. 6, no. 3, pp. 4219–4225, 2013.

[4] A. Nasridinov and Y.-H. Park, "A study on performance evaluation of machine learning algorithms for crime dataset," *Adv. Sci. Technol. Lett.-(Networking Commun. 2014)*, vol. 66, pp. 90–92, 2014.

[5] X. Wang, M. S. Gerber, and D. E. Brown, "Automatic crime prediction using events extracted from twitter posts," in *International conference on social computing, behavioral-cultural modeling, and prediction*, pp. 231–238, Springer, 2012.

[6] M. Redmond and A. Baveja, "A data-driven software tool for enabling cooperative information sharing among police departments," *European Journal of Operational Research*, vol. 141, no. 3, pp. 660–678, 2002.