

ASSIST CRIME PREVENTION USING MACHINE LEARNING

Nair Swati S. *, Soniminde Saloni Ajit, Sruthi Sureshababu, Apurva Chandrakant, Sagar Kulkarni
(University of Mumbai – Pillai College of Engineering, New Panvel)

Abstract:

Crime rate is increasing significantly over the years. Crime prevention is the attempt to reduce and deter the crimes and the criminals. The government must go beyond law enforcement and criminal justice to tackle the risk factors that cause crime because it is more cost effective and leads to greater social benefits. The data driven method is used which is based on the broken windows theory, having an enormous impact on the working of the police department. The theory links disorder and incivility within a community to subsequent occurrences of serious crimes. Predictive policing is used by the law enforcement stakeholders for taking proactive measures against crimes. This will help the police departments to efficiently focus their resources on the potential crime hotspots. The model is built to predict the crime rate based on demographic and economic information of particular localities using decision trees, linear classification, regression and spatial analysis.

Keywords:

Crime, Broken Windows Theory, Decision Trees, Classification, Regression, Spatial Analysis

Submitted on: 15/10/2018

Revised on: 15/12/2018

Accepted on: 24/12/2018

***Corresponding Author Email:** swatisanair15e@student.mes.ac.in Phone: 77386246914

I. INTRODUCTION

Crimes are increasing day by day which means that there should be measures to avoid them. Crime prevention refers to recognizing that a crime risk exists and taking some corrective action to eliminate or reduce that risk. Using machine learning approach we will assist the local authorities in preventing crime and to take the necessary actions against crime.

There are numerous types of crimes taking place at different locations. Some areas have crimes occurring frequently whereas there are some places where occurrence of crime is negligible. Therefore potential crime hotspot areas require much more security than those areas where crime rate is comparatively less. For example, Crimes like chain snatching occur mostly at lonely places so that criminals could escape easily from that location. Detecting the crime hotspot areas helps the police officials to decide what kind of security strength will be required for that particular place.

The system is based on the broken windows theory. Broken windows theory is an academic theory proposed by James Q. Wilson and George Kelling in 1982 that used broken windows as a metaphor for disorder within neighbourhoods. Their theory links disorder and incivility within a community to

subsequent occurrences of serious crime [12].

II. METHODOLOGY

A. Preprocessing

Pre-processing is the process of cleaning and preparing the text for classification.

Algorithm for Pre-processing module:

1. Accept the data set in .csv (comma separated value) format.
2. Remove corrupt data.
3. Impute missing data.

The communities-crime-full.csv dataset is used. The dataset consists of the crime records of the communities within the United States. The dataset is for per-capita crime rates around the country. Our task is to build models to predict the crime rate based on demographic and economic information about the particular locality.

The data is given in the file “communities-crime-full.csv”. It includes data fields with missing values (indicated by “?”), which have to be removed.

Table 1:- Dataset before cleaning

	A	B	C	D	E	F
1	state	county	communit	communit	fold	populatio
2	8 ?	?	Lakewood	1	0.19	
3	53 ?	?	Tukwilacit	1	0	
4	24 ?	?	Aberdeen	1	0	
5	34	5	81440 Willingbo	1	0.04	
6	42	95	6096 Bethleher	1	0.01	
7	6 ?	?	SouthPass	1	0.02	
8	44	7	41500 Lincolntov	1	0.01	
9	6 ?	?	Selmacity	1	0.01	
10	21 ?	?	Henderso	1	0.03	

Table 2: Dataset after cleaning

	A	B	C	D	E	F
1	state	communit	fold	populatio	househol	racepctbl
2	1 Alabaster	7	0.01	0.61	0.21	
3	1 Alexander	10	0.01	0.41	0.55	
4	1 Anniston	3	0.03	0.34	0.86	
5	1 Athenscity	8	0.01	0.38	0.35	
6	1 Auburncit	1	0.04	0.37	0.32	
7	1 Bessemer	6	0.04	0.44	1	
8	1 Birmingham	2	0.41	0.37	1	
9	1 Cullmanci	1	0.01	0.3	0	
10	1 Daphnecit	7	0	0.39	0.31	

B. Processing

1. Decision tree

The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features. We will use the clean dataset to predict whether the crime rate in a locality is greater than 0.1 per capita or not. A new field "highCrime" is created which is true if the crime rate per capita (ViolentCrimesPerPop) is greater than 0.1, and false otherwise.

2. Cross Validation

Cross-validation is a statistical method which has a single parameter called k that refers to the number of groups that a given data sample is to be split into. Algorithm for Cross Validation is:

1. Shuffle the dataset randomly.
2. Split the dataset into k groups
3. For each unique group:
 - a. Take the group as a hold out or test dataset
 - b. Take the remaining groups as training data.
 - c. Fit a model on the training set and evaluate it on the test set.
 - d. Retain the evaluation score and discard the model
 - e. Summarize the skill of the model using the sample of model evaluation scores

We will apply cross-validation (cross_val_score) to do 10-fold cross-validation to estimate the out-of-training accuracy of decision tree learning. We will find out what are the 10-fold cross-validation accuracy, precision and recall.

3. Classification

In machine learning, classification is the problem of identifying to which set of categories a new

observation belongs, on the basis of a training set of data containing observations whose category membership is known.

Linear SVM

A Support Vector Machine (SVM) is a discriminative classifier formally defined by a separating hyperplane. The LinearSVC is used to make a linear Support Vector Machine model learn to predict highCrime.

- i. The 10-fold cross-validation accuracy, precision, and recall for this method is found.
- ii. The 10 most predictive features are identified.
- iii. The results are the compared with results from decision trees.

Gaussian Naive Bayes

Bayes' Theorem provides a way that we can calculate the probability of a hypothesis given our prior knowledge.

The GaussianNB is used to make a Naive Bayes classifier learn to predict highCrime.

- i. The 10-fold cross-validation accuracy, precision, and recall for this method is found.
- ii. The 10 most predictive features are identified.
- iii. The results are the compared with results from decision trees.

4. Regression

Regression is used to predict continuous values. We perform regression analysis to understand which among the independent variables are related to the dependent variable. [11] Regression will be used for predicting the crime rate per capita (ViolentCrimesPerPop). The following errors are calculated:

1. RMSE(Root Mean Square Error)
2. MAE(Mean Absolute Error)
3. R²(R Square Error)

Ridge Regression

Ridge Regression is a technique for analyzing multiple regression data that suffer from multicollinearity.

SVM Regression

SVM constructs a hyperplane in multidimensional space to separate different classes. SVM generates optimal hyperplane in an iterative manner, which is used to minimize an error.

Random Forest Regression

It is a meta estimator that fits a number of classifying decision trees on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting.

XGBoost Regression

XGBoost stands for eXtreme Gradient Boosting. It is an implementation of gradient boosted decision trees designed for speed and performance.

KNN Regression

K nearest neighbors is a simple algorithm that stores all available cases and predict the numerical target based on a similarity measure.

Lasso Regression

Lasso (Least Absolute Shrinkage and Selection Operator) penalizes the absolute size of the regression coefficients.

Decision Tree Regression

Decision tree builds regression or classification models in the form of a tree structure with decision nodes. It breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed.

Algorithm for predicting crime:

1. Taking input dataset which is .csv file (In our example we have US based dataset).
2. Perform cleaning and pre-processing. Save the cleaned file and use this for further analysis.
3. Based on various conditions, apply appropriate decision tree and infer the results.
4. Split the data into train and test by using cross validation.
5. Apply various Classification and Regression models. Analyze them using evaluation metrics and select one which gives best results.
6. Perform spatial analysis using GeoPanda.
7. Based on the results obtained we can identify the area of high crime and assist police.

5. Feature Extraction

Determining a subset of the initial features is called feature selection. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data.

C. Spatial Analysis

Spatial analysis is a type of geographical analysis which seeks to explain patterns of human behavior and its spatial expression in terms of mathematics and geometry, that is, locational analysis.

GeoPandas is the geospatial implementation of the big data oriented Python package called Pandas. GeoPandas enables the use of the Pandas data types for spatial operations on geometric types. The potential crime hotspots are plotted on the map which gives better visualization of results.

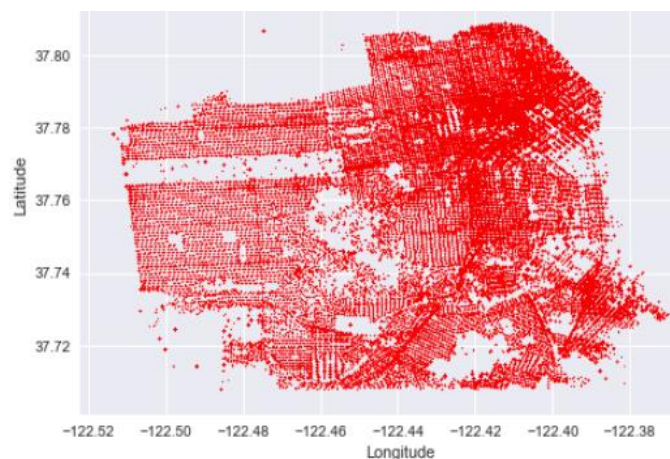


Fig. 1:- Plot of Crime Hotspots

III. EXPERIMENTATION

System architecture

The system architecture shows the overall flow of the System. There are 3 modules:

1. Preprocessing
2. Processing
3. Analyzing

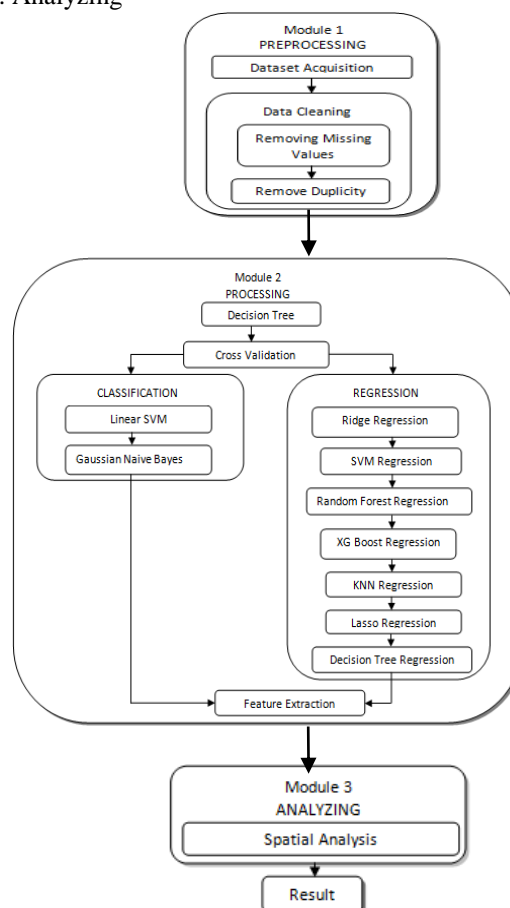


Fig. 2:- System architecture

IV. RESULTS AND DISCUSSION

These systems are typically measured using accuracy, precision and recall.

Table 3: Prediction Outcomes

Condition Positive (P)	The number of real positive cases in the data
Condition Negative (N)	The number of real negative cases in the data
True Positive (TP)	Equivalent to hit
True Negative (TN)	Equivalent to correct rejection
False Positive (FP)	Equivalent to Type I error
False Negative (FN)	Equivalent to miss, Type II error

Precision: A measure of exactness, determines the fraction of relevant items retrieved out of all items retrieved. Precision (P) It is given in Equation 2.

$$P = \frac{TP}{TP + FP} \quad \dots(2)$$

Accuracy: Accuracy is the proximity of measurement results to the true value; precision, the repeatability, or reproducibility of the measurement. Accuracy (A) is given in Equation 3.

$$A = \frac{TP + TN}{P + N} \quad \dots(3)$$

Recall: a measure of completeness, determines the fraction of relevant items retrieved out of all relevant items. Recall (R) is given in Equation 4.

$$R = \frac{TP}{TP + FN} \quad \dots(4)$$

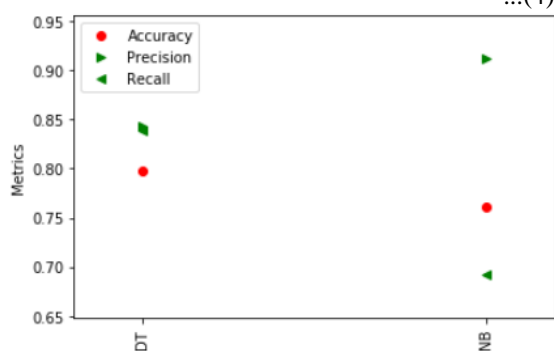


Fig. 3:- Plot of Accuracy, Precision and Recall

APPLICATIONS

Technical applications

Assist police department: The Crime Prevention System will assist police department in maintaining law and order, as the model will give a pictographic view of crime hotspots based on the data set provided of that region.

Crime Reports for newspapers: This system can be used by news reporters or journalists to give a brief analysis about crime occurrences at a particular place stating about the type of crime and its frequency.

Predicting crimes from news feeds: Crime patterns can be analyzed and crimes can be predicted from news feeds. The news feeds for a particular time span can be collected like for 20 years and this news feeds corpus can be used to predict future events.

Social Applications

Combat drug addiction and other related crime: This system will help to identify the predominant drug and other related crime hotspots and then the government can set up rehabilitation centres and camps. NGOs can also conduct awareness drives.

Urban planning: Once the crime hotspots are identified the government can take measures to redevelop those areas by implementing urban planning so as to improve the social neighbourhood of a person by which there is no or minimal indulgence in criminal or illegal activities. Bad urban planning can lead to an increase in crime rate.

Analyzing crime through social media: The tweets and social media posts can be analyzed for a certain timespan. From this corpus certain deductions can be made about the crime patterns and criminal instincts. By further enhancements on the model using Natural Language Processing, the crimes can be prevented from happening by assessment of social media posts.

V. CONCLUSIONS

The system uses different Machine Learning approaches to assist in crime prevention by predicting whether a particular area is a potential crime hotspot or not. The community crimes dataset of the US is used this purpose. As the dataset collected consists of missing values, it has to be cleaned and pre-processed. Decision trees can then be used to make decision about a high crime area. The classification models are applied to the system and the topmost features can be predicted. Different regression models are applied aiming for the least error. The model with the least error will be the winning model. Accuracy, Precision and Recall are considered for evaluation of the system. Geospatial analysis can then be done to plot the potential crime hotspots across the longitudinal and latitudinal positions over a map. This plot will assist the police department in deciding which area requires greater attention and hence larger security forces could be deployed at that crime hotspot.

REFERENCES

- [1] Ayisheshim Almaw, Kalyani Kadam (2018), "Survey Paper on Crime Prediction using Ensemble Approach", As appeared in International Journal of Pure and Applied

Mathematics", Pune, India, Vol. 118 No. 8, ISSN: 1311-8080 (printed version), ISSN: 1314-3395 (on-line version).

[2] Ying-Lung Lin, Liang-Chih Yu, Tenge-Yang Chen (2017), "Using Machine Learning to Assist Crime Prevention", Taiwan, INSPEC Accession Number: 17375465.

[3] N.D. Waduge, Dr. L. Ranathunga (2017), "Machine Learning Approaches To Detect Crime Patterns", Sri Lanka.

[4] Hyeon-Woo Kang, Hang-Bong Kang (2017), "Prediction of crime occurrence from multimodal data using deep learning", Plos One, Bucheon, Gyonggi-Do, Korea.

[5] Lawrence McClendon, NatarajanMeghanathan (2015), "Using Machine Learning Algorithms To Analyze Crime Data", Machine Learning and Applications: An International Journal (MLAIJ), USA.

[6] Harsha Perera, Shanika Udeshini, Malith Munasinghe (2014), "Criminal short listing and crime forecasting based on modus operandi", 14th International Conference on Advances in ICT for Emerging Regions (ICTer) ,Colombo, SriLanka INSPEC Accession Number: 15058519.

[7] Devendra Kumar Tayal, Arti Jain, Surbhi Arora et.al., "Crime detection and criminal identification using data mining" (2014), Springer-Verlag, London, ISSN: 0951-5666.

[8] Shiju Sathyadevan, Devan M.S, Surya Gangadharan S, "Crime Analysis and Prediction Using Data Mining" (2014), First International Conference on Networks & Soft Computing, Guntur, India.



[9] Andrey Bogomolov, Bruno Lepri, Jacopo Staiano et.al, "Once Upon a Crime: Towards Crime Prediction from Demographics and Mobile Data" (2014), Proceedings of the 16th International Conference on Multimodal Interaction, Istanbul, Turkey, pp. 427-434.





[10] Lenin Mookiah, William Eberle and Ambareen Sira, "Survey of Crime Analysis and Prediction" (2014), Proceedings of the Twenty-Eighth International Florida Artificial Intelligence Research Society Conference, Cookeville, Tennessee.

[11] <http://scikit-learn.org> , last accessed on 28th October, 2018.

[12] <https://www.britannica.com> , last accessed on 29th October, 2018.

Author Biographical Statement

	Nair Swati Sasindrakumar B.E. Computer Engineering Student Pillai College of Engineering
	Soniminde Saloni Ajit B.E. Computer Engineering Student Pillai College of Engineering

	
	Sruthi Sureshababu B.E. Computer Engineering Student Pillai College of Engineering
	Apurva Chandrakant Tamhankar B.E. Computer Engineering Student Pillai College of Engineering
	Prof. Sagar Kulkarni Professor Pillai College of Engineering