International Journal of Future Generation Communication and Networking
Vol. 13, No. 4, (2020), pp. 2574–2586

Empirical Data Analysis of Crime Rate using Enhanced Ensemble & Genetic Algorithms

K.Rajasekhar Rao
Associate Professor, Sridevi Women's Engineering College,
V.N.Pally, Hyderabad
Email: krsraohyd@gmail.com

Abstract

Crimes are the increasing danger to the mankind and India is facing one of the vast effects among them. There are numerous crimes that happens normal interim of time. Since Crimes are expanding, there is a need to illuminate the cases and processes in a lot quicker way. The Crime exercises have been expanded at a quicker rate and it is the obligation of crime analysis department to control and decrease the crimes exercises. There is a need of innovation through which the case explaining could be quicker. Procedure and techniques of crime analysis is the process of recognizing and sorting conclusions communicated in a piece of content, particularly with the end goal to decide if the item is positive, negative or unbiased along with comparing with Unsupervised and supervised processes of Machine learning to get better accuracy, precision and recall percentages along with graphs and charts to represent current state of crime activities in the locality mentioned in data. Paper explains "how the fetching data from the source", pre-processing, sorting is done to implement and classify. The whole process gives a brief of the advanced subjects of Machine Learning Algorithms that describes the subject of Crime analysis and comparison with ensemble of classification techniques and unsupervised algorithms. The supervised and unsupervised algorithms will be compared with each other on the basis of accuracy. This paper explains the use of Genetic algorithm for the selection of features that when used for a particular classifier will result in highest possible accuracy. The more the accuracy come across according the dataset given the more the result will be up to the mark. The comparison and the use of Genetic algorithm provides the best accuracy and list out the features that have higher co-relation with the crime type.

Keywords: *Naïve Bayes, SVM, Gradient boosting, Adaboost, Random Forest, Crime, Genetic, Ensemble*

1. Introduction

Crimes happen from little town, town to colossal urban territories additionally in enormous urban communities. India as we probably am aware is one of the urban areas loaded with savagery, debasements by legislative issues. Infringement are of different sorts – Burglary, murder, attack, trap, battery, bogus confinement, commandeering, homicide, robbery, youngster misuse, Child Trafficking, Molestations and so forth. The Crime practices have been extended at a speedier rate and it is the commitment of police division to control and diminish the violations works out. Since Crimes are extending there is a need to light up the cases and procedures in much snappier manner. Crimes conjecture additionally, criminal unmistakable evidences are the significant issues to the police office as there are huge proportions of Crimes event data that exist. So in this way it is expected to investigation the Crimes of each state and territories of India by months, years and time of present to forestall the progressing crimes. Machine Learning is the study of having PCs settle on choices without human mediation. As of late, Machine Learning has been applied in self-driving vehicles, discourse acknowledgment, web search, and an improved comprehension of the human genome. It has likewise made anticipating Crimes dependent on referenced information

ISSN: 2233-7857 IJFGCN

attainable. Arrangement is an administered forecast system which takes into account ostensible class marks. Grouping has been utilized in numerous areas including climate anticipating, medicinal consideration, funds and banking, country security, and business insight. Machine Learning based Crimes examination for the most part includes information assortment, order, design distinguishing proof, forecast, and representation. Customary information mining methods - affiliation examination, grouping and expectation, bunch investigation, and exception investigation - recognize designs in organized information while fresher strategies distinguish designs from both organized and unstructured information. The essential goal of this work is to make a forecast model that can precisely anticipate Crimes. This paper speaks about the field of law prerequisite has formed into an entirely flighty calling with unlimited zones of specialty and capacity. Crimes examination could be viewed as the most present development to the field. This venture executes an approach to check the Crimes exercises and discover the measure of continuous things and alongside that it predicts whether there will be any rough Crimes event will be there or not all that that it very well may be forestalled.

In this paper, we first try to describe the previous works done on the said topic with their merits and demerits. Then the paper moves forward to describe the proposed method and about the different algorithm used, software requirements and data set to achieve the said topic. Then we describe the methodology that we took up in solving the said task, we also provide the viewer with proper block diagram of the project for better understanding. The paper further provides the comparative analysis and its findings and then it finally concludes the aim of the project with providing the future aspects

1.1 Literature Review

a) [1]In the paper titled *"A proposed framework for analyzing crime data set using decision tree and simple k-means mining algorithms."* Al-Janabi, Kadhim B. Swadi tells how cluster accuracy can be improved to capture the local correlation structure by associating each cluster with the combination of the dimensions as independent weighting vector and subspace span which is embedded on it. Recent research works on crime analysis Includes-Adderley and Musgrove applied Self Organizing Map (SOM) to link the offenders of serious sexual attacks

- Advantage: The pre-processed data were used to find out different crime and criminal trends and behaviors, and crimes and criminals were grouped into clusters according to their important attributes.
- Disadvantage: Complex and involve high level mathematics that requires a statistical program to analyze the data.

b) [2]In this paper titled *"Crime analysis and prediction using data mining." Sathyadevan, Shiju, and Surya Gangadharan* describes Crime analysis and prevention is a systematic approach for identifying and analyzing patterns and trends in crime. Their system can predict regions which have high probability for crime occurrence and can visualize

ISSN: 2233-7857 IJFGCN

crime prone areas. Using the concept of data mining we can extract previously unknown, useful information from an unstructured data. Instead of focusing on causes of crime occurrence like criminal background of offender, political enmity etc they are focusing mainly on crime factors of each day.

- Advantage**:** This setting is motivated by many cases in which there exist laws that disallow a decision that is partly based on discrimination. Naive application of machine learning techniques would result in huge fines for companies.
- Disadvantage**:** Cannot handle large and complex dataset accurately.

c) [3]In this paper titled *"Using machine learning algorithms to analyze crime data" Lawrence McClendon and Natarajan Meghanathan* implements the Linear Regression, Additive Regression, and Decision Stump algorithms using the same finite set of features, on the communities and crime un normalized dataset to conduct a comparative study between the violent crime patterns from this particular dataset and actual crime statistical data for the state of Mississippi that has been provided by neighborhoodscout.com

- Advantage**:** The scope of this project was to prove how effective and accurate machine learning algorithms can be at predicting violent crimes, there are other applications of data mining in the realm of law enforcement such as determining criminal "hot spots", creating criminal profiles, and learning crime trends.
- Disadvantage**:** Utilizing these applications of data mining can be a long and tedious process for law enforcement officials who have to sift through large volumes of data.

d) [4]In this paper titled *"Crime Analysis Through Machine Learning"* Suhong Kim , Param Joshi, Parminder Singh Kalsi, and Pooya Taheri created a prediction model that can accurately predict crime. In this research, two classification algorithms, K-Nearest Neighbour (KNN) and boosted decision tree, were implemented to analyse the VPD crime dataset compiled between 2003 and 2018 with more than 560,000 records.

- Advantage**:** Geographic Information System (GIS) has been used as a powerful analytical tool for crime mapping. It shows the locations of crime series with various geographic information on one map, which helps police officers to make decisions for operational and tactical purposes

e) [5]In this paper titled *"Naïve Bayes Approach for the Crime Prediction in Data Mining"* Mrinalini Jangr, Shaveta Kalsi Uses the Data mining approach to build a model for crime prediction. The process builds representative model in which observational data is implemented. The prediction model serves two purposes. Firstly, the output is predicted on the basis of the input variables and secondly, this model helps in understanding the relationship between the output variable and all input variables.

- Advantage**:** This analysis works together with patrolling officers and detectives to provide support to police officials and detectives to identify and understand particular and immediate crime issues.
- Disadvantage**:** Cannot handle large and complex dataset accurately.

2. Proposed Method

In this paper, multiple algorithm such as Naïve Bayes, KNN, Gradient boosting, Logistic

ISSN: 2233-7857 IJFGCN

regression, Random forest, Genetic Algorithm are used on same dataset. From the crime

ISSN: 2233-7857 IJFGCN

dataset, it can be observed that a classification of crime depends on various factors such as: Date, Category, Description of the crime, Day of the Week on which crime occurred, Police station near closest to the crime scene, Resolution of the crime and co-ordinates (i.e. longitude and latitude) of the crime. Different factors influence the classification of the crime differently. Some factors such as crime co-ordinates and time of the crime has high co- relation with the crime type, whereas other factors such as "Day of the week" has low co- relation with the crime. Hence, using those factors which has high correlation with crime type will improve the accuracy of any algorithm. Results varies for different classifier, hence genetic algorithm is used to figure out the features for a particular classifier with which it provides good accuracy for classification. With the help of dataset belonging to different cities, pictorial representation of crime is obtained for the better understanding of the crime type and which part of the city travails with what type of crime. Crime wise "Heat maps" are generated to pin-point the part of the city on the map and the extent of particular crime in that part.

Genetic algorithm treats various features as "chromosomes" and treats them as such. It is recursive in nature and includes the process of "mutation" and "cross-over". Initially, it will make various stochastic sets of "features", then find accuracy for a particular classifier with various sets. It maintains a vector of "sets of features" which provides best accuracies. After finding the accuracy, it does mutation and crossover of different sets of features. Finally, it provides with the set of features which will give the best accuracy for particular classifier.

3. Experiment Analysis

3.1 Data

The data of various cities for the experiment is obtained from various state monitored websites where they were uploaded by the respected authority of the state. For few states whose data wasn't available on their respected website, their crime dataset can be obtained from "kaggle". The format of obtained dataset is more or less similar. Similarity of the dataset is necessary to obtain a uniform outcome. Mandatory column of the dataset are: "Date of occurrence of crime", "category of crime", "Description of crime", "Dayof theWeek", "Police station", "Resolution off the crime", "Longitude and Latitude of the place". The project includes the classification of the crime into groups that are present in "Category of crime". Various groups of classification are: "Warrants", "Theft", "Larceny", "Vandalism", "Robbery", "Warrants", "Robbery", "Sex offenses", "Missing person", "Narcotic", "Assault" etc. There are 37 such classes among which various crimes are classified using all the above mentioned features.

5/11/2015 20:30	VEHICLE T	STOLEN AU	Monday	PARK	NONE	1100 Block of HA	-122.441	37.77079
5/11/2015 20:19	WARRANT	WARRANT	Monday	TENDERLOIN	ARREST, B	100 Block of TUF	-122.412	37.78305
5/11/2015 20:19	SUSPICIOU	INVESTIGA	Monday	TENDERLOIN	ARREST, B	100 Block of TUF	-122.412	37.78305
5/11/2015 20:15	BURGLARY	BURGLARY	Monday	BAYVIEW	NONE	700 Block of KAN	-122.403	37.76025
5/11/2015 20:15	ASSAULT	BATTERY,	Monday	TARAVAL	ARREST, B	1500 Block of 24	-122.482	37.75865
5/11/2015 20:10	VEHICLE T	STOLEN AU	Monday	TARAVAL	NONE	ULLOA ST / 46TH	-122.47	37.74142
5/11/2015 20:05	LARCENY/	GRAND TH	Monday	SOUTHERN	NONE	1ST ST / STEVENS	-122.399	37.79057
5/11/2015 20:00	LARCENY/	PETTY THE	Monday	PARK	NONE	2600 Block of GE	-122.447	37.78225
5/11/2015 20:00	VEHICLE T	STOLEN AU	Monday	INGLESIDE	NONE	300 Block of COL	-122.421	37.74091
5/11/2015 20:00	LARCENY/	PETTY THE	Monday	PARK	NONE	100 Block of CRE	-122.459	37.75533
5/11/2015 20:00	VANDALIS	MALICIOU	Monday	TARAVAL	NONE	300 Block of ASH	-122.462	37.7251
5/11/2015 20:00	BURGLARY	BURGLARY	Monday	PARK	NONE	1600 Block of GR	-122.444	37.77524
5/11/2015 20:00	LARCENY/	GRAND TH	Monday	TENDERLOIN	NONE	OFARRELL ST / M	-122.41	37.7862

Fig. 1. Dataset For San Francisco

ISSN: 2233-7857 IJFGCN

Dates	Category	Descript	DayOfWee	PdDistrict	Resolution	Address	X	Y
5/13/2015 23:53	WARRANTS	WARRANT	Wednesda	Sholinganallur	ARREST, B	Plot 155, 292/132, Ground Fl	80.22684	12.90145
5/13/2015 23:53	OTHER OFFENSES	TRAFFIC V	Wednesda	MRC Nagar	ARREST, B	Somerset Greenway, Lords A	80.2707	13.01716
5/13/2015 23:33	OTHER OFFENSES	TRAFFIC V	Wednesda	T. Nagar	ARREST, B	40, 2nd Floor, Bazullah Road,	80.23702	13.04681
5/13/2015 23:30	LARCENY/THEFT	GRAND TH	Wednesda	Velachery	NONE	Shop 17-18, Rajalakshmi Nag	80.21822	12.98606
5/13/2015 23:30	LARCENY/THEFT	GRAND TH	Wednesda	Alwarpet	NONE	40, Maharaja Surya Rao Roac	80.25384	13.03941
5/13/2015 23:30	LARCENY/THEFT	GRAND TH	Wednesda	Ispahani Centre, N	NONE	Bearing Door 123 - 124, Ishpa	80.24971	13.05505
5/13/2015 23:30	VEHICLE THEFT	STOLEN AU	Wednesda	Navallur	NONE	1A-1B, Rajiv Gandhi Salai, Old	80.22635	12.85154
5/13/2015 23:30	VEHICLE THEFT	STOLEN AU	Wednesda	Nungambakkam	NONE	18/24, Yafa Towers, Khader N	80.24971	13.06076
5/13/2015 23:00	LARCENY/THEFT	GRAND TH	Wednesda	Porur	NONE	Survey 38/1, Mount Poonam	80.17038	13.0289
5/13/2015 23:00	LARCENY/THEFT	GRAND TH	Wednesda	Anna Nagar East	NONE	189, AI Block, Shanthi Colony	80.21198	13.08209
5/13/2015 22:58	LARCENY/THEFT	PETTY THE	Wednesda	T. Nagar	NONE	102/82, GN Chetty Road, T. N	80.23667	13.04314
5/13/2015 22:30	OTHER OFFENSES	MISCELLAI	Wednesda	Alwarpet	NONE	6,Seshadri Road, Venus Color	80.25584	13.0397
5/13/2015 22:30	VANDALISM	MALICIOU	Wednesda	Semmancheri	NONE	3/62, 2nd Floor, Tech Pump T	80.22732	12.87509

Fig. 2. Dataset For Chennai

Figure 1 and Figure 2 depicts the crime dataset for city of San Francisco [13] and Chennai [15] respectively. This experiment includes graphical representation of the crime of various cities such as “New York”, “Chicago”, “Vancouver”, etc. Crime dataset for other cities can also be found on “Kaggle” or official pages of respected cities.

4. Methodology

4.1 Modules

a) Data Collection (Cloud based, local or live web sites)

b) Data Pre-processing (Tokenization)

c) Processed Data warehousing (Storage)

d) Algorithm extraction:

- Supervised Learning -
  a) Naive Bayesian prediction Algorithms
  b) Random Forest prediction Algorithms
  c) Support Vector Machine prediction
  d) Logistic Regression

- Unsupervised Learning -
  a) K Nearest Neighbour

A special case of bio-inspired algorithm, “Genetic Algorithm” is also use to optimize the existing machine learning algorithm.

e) Features of algorithm extraction
f) Compare the result based on accuracy
g) Analyze the performance
h) Graph generates (Bar graphs according to data provided)
i) Heat Map Generates

ISSN: 2233-7857 IJFGCN

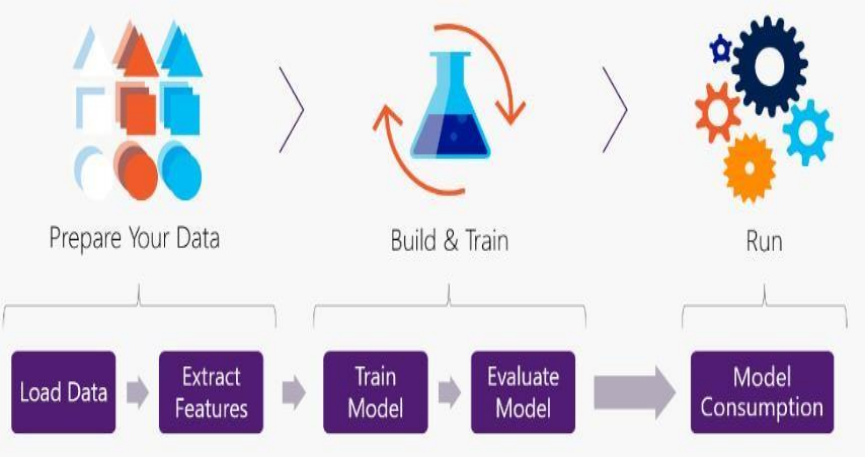


Fig 3: Flow Diagram of implementation process by ensemble of algorithms

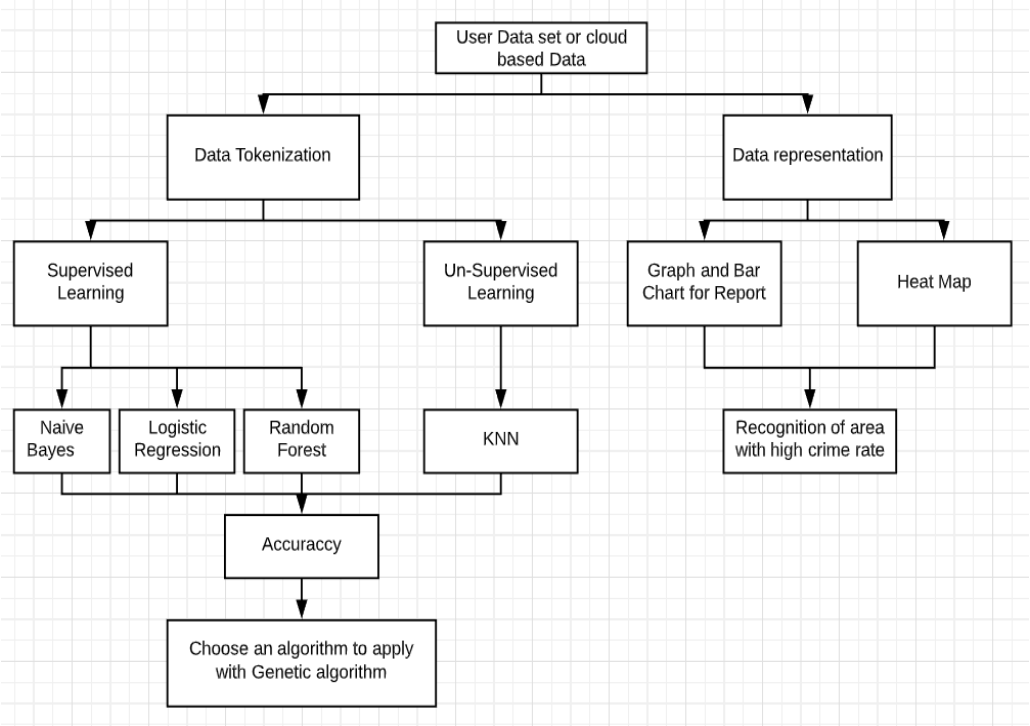


Fig.4 Flow diagram of the process mentioned in the paper

In figure 4, all blocks are attached and connected with each other as in all the functions of steps and in coding for the implementation is done in a similar manner. Different modules are implemented separately on Data set. The data set is represented in the form of bar graph and heat map to recognize the area with high crime rates.

4.2 Short Brief Of Each Module,

ISSN: 2233-7857 IJFGCN

a) Dataset type and resources: This part defines that it can permit the utilization of table labels and names of state where the crime Occurring, client reference and URLs. These require unexpected preparing in comparison to different data in all set. The

amount of vast dataset or base can be gathered of local datasets from any political or local government organization for records or any cloud based database such as tinyDB from thinkable platform or live fetching environment using python libraries (e.g. - tweepy) for the provided terms and data.

b) Data tokenization: Data Tokenization is the process used to replace sensitive data with unique identification symbols that retain all the essential information about the data without compromising its security. It will give in numeric state whether the data is in one of the crime class

c) Machine Learning Algorithm: Data gathered from data warehouse will go through processing of ensemble algorithms and coded in python language and provide outputs of individuals. This will show how efficient can one platform is. In end the terminal of console shows the accuracy of supervised and unsupervised categorized program to the developer and compare of all algorithms.

d) Data Diagrams: Client will see the process through chart and graphs about the state of the crimes amount in individual areas. Take the visualization in particular axis and precautions can be taken. In this way predicted areas can be revealed and in future the crimes can be avoided.

e) Heat map**:** Implementation time and processing will return a graph and chart of affected areas along with that it will provide the heat map prediction of effected with the accuracy of individual algorithms and areas according to months and years of data gathered to prevent crime happenings later on.

f) Comparison: The supervised and unsupervised algorithms will be compared with each other for checking in efficiency. The more the accuracy come across according the dataset given the more the result will be up to the mark. The comparison gives an output of one outcome so clients don't have to go with individual checking and come up with one answer which is more time convenient and less efficient.

5. Comparative Analysis

5.1. Pre-Processing

For comparative analysis, it is very important to convert all the alphabetical data into numerical data so that analysis will be uniform for each and every algorithm used. For the conversion of data into numerical value, "Label encoder" has been used which assigns a specific number to an alphabetical character

5.2. Feature Wise Accuracy

ISSN: 2233-7857 IJFGCN

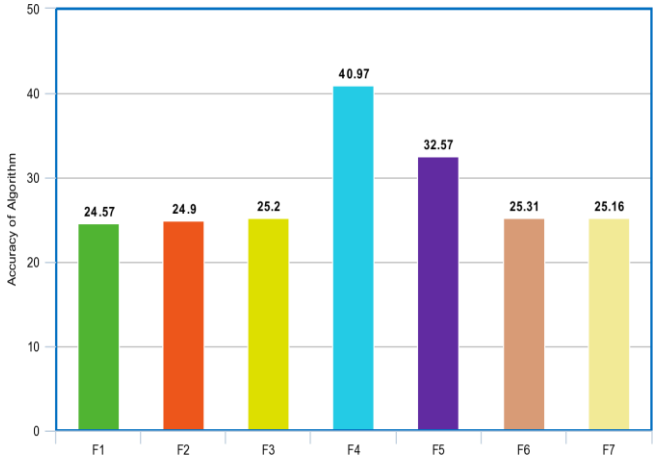


Fig. 5 Feature wise accuracy

Figure 5 depicts the feature wise accuracy for Naïve Bayes algorithm. In the above figure various features are as follows: - f1:"Hour", f2:"Day of the week", f3:"Police station", f4:"Crime description", f5:"Resolution of the crime", f6:"Address of the crime", f7:"Co-ordinates of the crime"

Classifier	F1	F2	F3	F4	F5	F6	F7
KNN	12.93	14.62	22.28	59.91	12.56	22.57	22.13
NBB	25.22	24.9	25.2	22.05	29.09	25.33	25.86
NBG	24.57	21.66	18.66	40.97	32.57	25.31	25.16
LR	24.82	24.9	24.9	32.17	32.6	25.33	26.22
AB	24.78	25.15	24.92	36.02	36.63	25.18	25.43
RF	24.99	27.37	25.85	59.92	37.15	29.38	30.27
ETC	25.07	24.96	25.16	57.35	37.65	26.46	27.95

Fig. 6 Feature wise accuracy

Fig.6 provides information about the feature wise accuracy for other algorithm. In the table, various mentioned algorithm are: - KNN: K nearest neighbour, NBB: Bernoulli Naïve Bayes, NBG: Gaussian Naïve Bayes, LR: Logistic Regression, AB: AdaBoost, RF: Random Forest, ETC: Extra Tree Classifier. It can be observed that "Crime description" and "Resolution of crime" are among the features having good accuracy whereas features such as "Day of week" and "Hour" are having least accuracy. Features with good correlation with the crime is used to do the classification. This helps in improving the accuracy of almost each and every algorithm to the highest extent.

5.3. Finding Accuracy

"Sklearn" python library is used to build various machine learning models. Dataset is tested simultaneously against multiple model and then accuracy and implementation time of the algorithm is compared.

ISSN: 2233-7857 IJFGCN

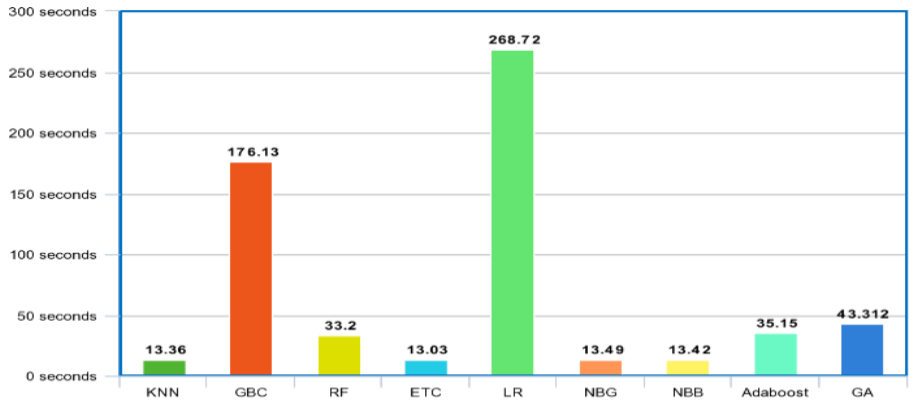


Fig 7. Execution time

Figure 7 depicts the execution time for each algorithm so that particular algorithm can be chosen on various factors such as efficiency, execution time, resources required, etc.

Algorithm	KNN	Gradient Boosting Classifier	Decision Tree		Logistic Regression
Classifier	-	-	Random Forest	Extra Tree	-
Accuracy	54.25	92.45	88.38	59.87	37.9189

Fig. 8 Accuracy Table

Naïve Bayes		Ada-boosting	Genetic Algorithm
Gaussian	Bernoulli		-
49.64	29.057	34.35	97.21

Fig. 9 Accuracy Table

Fig. 8 and Fig. 9 depicts the accuracy of the various algorithm that has been observed in the experiment. It can be seen that many algorithms are tested with various classifier present for a particular algorithm. Genetic algorithm is used to optimize the Naïve Bayes – Gaussian Classifier. Genetic algorithm helps is selection of various features that will improve the accuracy of the classifier.

6. Result Analysis

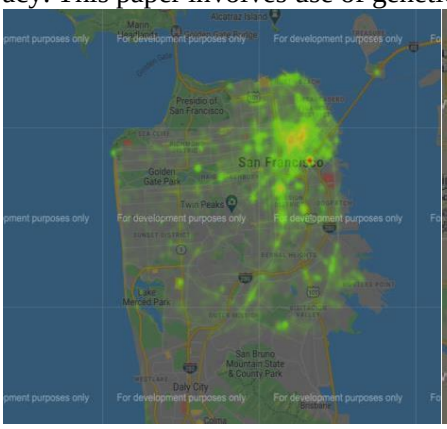
Analyzing crime data using various machine learning algorithm helps to understand that which algorithm is efficient for the classification for different crimes. If classification of crime is done properly, then it will make it possible to visualize which area is inflicted by which type of crime and what policy change is required in a particular area to fight a specific crime.

It can be easily observed that the random forest classifier is the best for classification of a crime dataset followed by Extra Tree classifier, Gradient Boosting algorithm. Genetic algorithm is based on the concept of mutation and hence is used to figure out features that helps in improving the accuracy of any particular classifier. The problem of genetic algorithm is that it is iterative in nature, therefore classifiers which have high execution time or are themselves

ISSN: 2233-7857 IJFGCN

iterative in nature such “Logistic Regression”, “Gradient Boosting Classifier”, “Random Forest Classifier” should not be used with Genetic Algorithm. Using

genetic algorithm with classifiers such as KNN or Naïve Bayes results in efficiently increase of accuracy. This paper involves use of genetic algorithm on KNN.



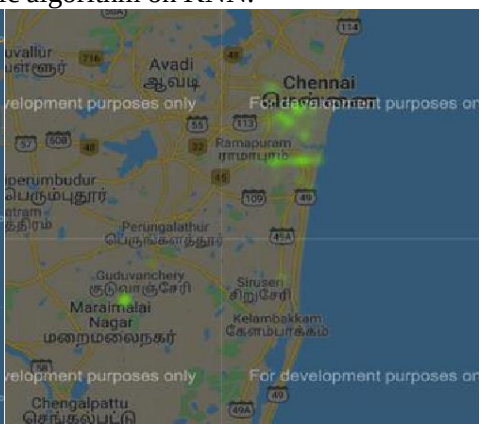


Fig.10 San Francisco **Fig.11 Chennai**

Figure 10 shows the crime density of the San Francisco and figure 11 shows the same for Chennai. The part of the city which can be seen in Red is highly inflicted by the crime and the part of the city which doesn’t have any color is crime free. Using the dataset of other cities, their crime density map can also be generated. Crime wise heat map can also be generated to understand which area is infested by what crime.

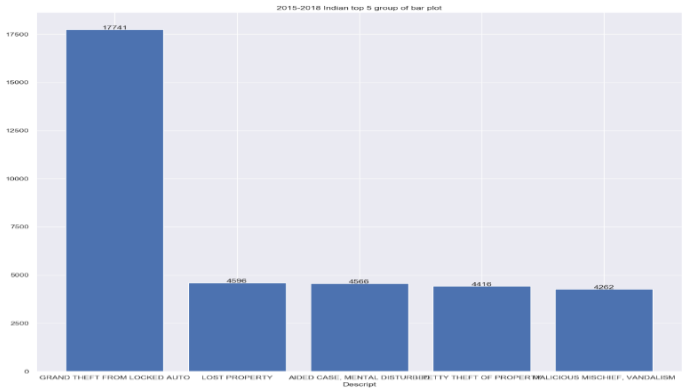
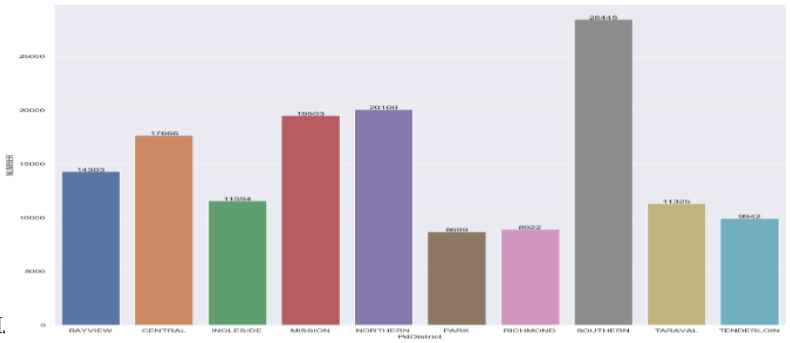


Fig. 12 Crime Bar graph

Figure 12 shows the top 5 crime type that are very common in Chennai. The graph is obtained after processing the crime dataset of India belonging to the year 2018-2018. Understanding what crime is more prominent in an area can help the respective police station to make policies according to the crime type.



ISSN: 2233-7857 I

Fig 13. Police station wise crime

From the dataset, figure 13 has been obtained which depicts the crime according to the area monitored by particular police department. This helps in understanding the requirement of different resources by different police department. It also tells about the efficiency of various police departments. If number of crime is increasing or the number reported against a particular crime is high in an area, then it can be concluded that the respected police station is not working efficiently or rules and regulations for that area is needed to be reformed

7. Conclusion and Future Aspects

From this paper, it can be concluded that because of high number of classes in which the crime has been classified (37 classes), accuracy of various machine learning algorithm is low. There are multiple ML algorithm such as "Gradient Boosting" and "Random forest" which are able to classify crime with high accuracy. If types of crime is grouped and new classes are formed for classification, the accuracy of other algorithm such as Naïve Bayes and logistic regression can be improved. The use of "Genetic algorithm" with various classifier can lead to improve in the accuracy of the classifier to great extent. From this paper, it can be observed for KNN. With Genetic algorithm, accuracy of the classifier boosted to
97.21 from 49.91.The crime analysis part is also holding the prediction process therefore it provides results of individual algorithms and output will be shown. In a brief this particular project is scalable enough where we can add up ne new features along with implementation on Cloud Platform such as AWS (Amazon Web Service) Microsoft Azure etc, in future if needed. The particular application can also be added in few IoT related projects such as Image processing by using a drone or vehicle and analyze the environment thorough surveillance camera.

8. References

[1] Al-Janabi, Kadhim B. Swadi. "A proposed framework for analyzing crime data set using decision tree and simple k-means mining algorithms." Journal of Kufa for Mathematics and Computer 1.3 (2011): 8- 24.

[2] Sathyadevan, Shiju, and Surya Gangadharan. "Crime analysis and prediction using data mining." 2014 First International Conference on Networks & Soft Computing (ICNSC2014). IEEE, 2014.

[3] Lawrence McClendon and Natarajan Meghanathan, "Using machine learning algorithms to analyze crime data" Jackson State University, 1400 Lynch St, Jackson, MS, USA, Machine Learning and Applications: An International Journal (MLAIJ) Vol.2, No.1, March 2015

[4] Suhong Kim , Param Joshi, Parminder Singh Kalsi, and Pooya Taheri, *"*Crime Analysis Through Machine Learning", Fraser International College, Simon Fraser University British Columbia, Canada 2018

[5] Mrinalini Jangr, Shaveta Kalsi, "Naïve Bayes Approach for the Crime Prediction in Data Mining",
International Journal of Computer Applications (0975 – 8887) Volume 178 – No. 14, May 2019

[6] T. Almanie, R. Mirza, and E. Lor, "Crime Prediction Based on Crime Types and Using Spatial andTemporal Criminal Hotspots," International Journal of Data Mining & Knowledge ManagementProcess, vol. 5, pp. 01-19, 2015.

[7] Rasoul Kianin, Siamak Mahdavi, Amin Keshavarzi, "Analysis and prediction of crimes by clustering and classification", International Journal of Advanced Research in Artificial Intelligence (IJARAI) – volume 4 No. 8 2015.

[8] Almanie, Rsha Mirza and Elizabeth Lor, "Crime Prediction Based on Crime Types and

ISSN: 2233-7857 IJFGCN

Using Spatial and Temporal Criminal Hotspots", International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.5, No.4, July 2015.

[9] Prajakta Yerpude1 and Vaishnavi Gudur, "Predictive Modelling Of Crime Dataset Using Data Mining", International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.7, No.4, July 2017.

[10] S. V. Nath, "<Crime Pattern Detection Using Data Mining.pdf>."International Journal of Database Management Systems (IJDMS) Vol.9, No.4, August 2017.

[11] Keyvanpour, Mohammad Reza, Mostafa Javideh, and Mohammad Reza Ebrahimi. "Detecting and investigating crime by means of data mining: a general crime matching framework." Procedia Computer Science 3 (2011): 872-880.

[12] S. Fong, R. Wong, A.V. Vasilakos, Accelerated PSO swarm search feature selection for data stream mining big data, IEEE Trans. Serv. Comput. 9 (1) (2016) 33–45.

[13] https://www.kaggle.com/c/sf-crime

[14] https://data.vancouver.ca/datacatalogue/index.htm

[15] http://ncrb.gov.in/StatPublications/CII/CII2018/cii2018.html

[16] Zelenkov, Y., Fedorova, E., & Chekrizov, D. (2017). Two-step classification method based on genetic algorithm for bankruptcy forecasting. Expert Systems with Applications, 88, 393–401., 2017.

[17] Li, G., Jin, Y., Akram, M. W., Chen, X., & Ji, J. (2018). Application of bio-inspired algorithms in maximum power point tracking for PV systems under partial shading conditions – A review. Renewable and Sustainable Energy Reviews, 81, 840–873., 2017.

[18] Djenouri, Y., & Comuzzi, M. (2017). Combining Apriori heuristic and bio-inspired algorithms for solving the frequent itemsets mining problem. Information Sciences, 420, 1–15., 2017.

[19] Kar, A. K. (2016). Bio inspired computing – A review of algorithms and scope of applications. Expert Systems with Applications, 59, 20–32., 2016.

[20] Martín, D., Alcalá-Fdez, J., Rosete, A., & Herrera, F. (2016). NICGAR: A Niching Genetic Algorithm to mine a diverse set of interesting quantitative association rules. Information Sciences, 355-356, 208–228., 2016.

ISSN: 2233-7857 IJFGCN