TEST Engineering & Management

March - April 2020
ISSN: 0193-4120 Page No. 11280 - 11286

Big Data Analytics and Mining for Effective Visualization and Trends Forecasting of Crime Data

[1]Rohit S, [2]Pradeep S, [3]K Sree Kumar

[1,2,3]Department of Computer Science, SRM Institute of Science and Technology, Kattankulathur, India
[1]rsmack57@gmail.com, [2]pradeepcampion@gmail.com, [3]sreekumk@srmist.edu.in

Article Info
Volume 83
Page Number: 11280 - 11286
Publication Issue:
March - April 2020

Article History
***Article Received**: 24 July 2019*
***Revised**: 12 September 2019*
***Accepted**: 15 February 2020*
***Publication**: 15 April 2020*

Abstract

Big data analytics (BDA) is a deliberate methodology for investigating and distinguishing various examples, relations, and patterns inside a huge volume of data. Right now, apply BDA to criminal data where exploratory data investigation is directed for perception and patterns prediction. A few the state - of - the workmanship data mining and profound learning procedures are utilized to investigate and foresee. Following factual examination and perception, some intriguing realities and examples are found from criminal data over the United States of America. These promising results will profit for police divisions and law authorization associations to more readily comprehend wrongdoing issues and give experiences that will empower them to follow exercises, foresee the probability of occurrences, adequately convey assets and improve the dynamic procedure.

Keywords: *Big Data Analytics (BDA),Prediction, Crime Data, Prediction, Visualization, Data Mining,*

1. Introduction

As of late, Big Data Analytics (BDA) has become a rising methodology for examining data and removing data and their relations in a wide scope of utilization zones. Because of nonstop urbanization and developing populaces, urban areas assume significant focal jobs in our general public. Be that as it may, such advancements have additionally been joined by an expansion in vicious wrongdoings and mishaps. To handle such issues, sociologists, examiners, and security foundations have dedicated a lot of exertion towards mining potential examples and elements. Comparable to open approach in any case, there are numerous difficulties in managing a lot of accessible data.

Understanding the components that anticipate crime is important on the grounds that despite the fact that crime percentages have been by and large declining since the mid nineties, late years have begun to see lower paces of decay and even some upward variances after 2010 . In addition, reports of immediate and roundabout exploitation and exposures to wrongdoing stay extremely high. For example, more than two-fifths of kids and youth in an ongoing national study announced a physical attack in the earlier year. Understanding the local setting of wrongdoing is especially significant in light of the fact that exploitation and different types of wrongdoing exposures have numerous serious results. Past the high doctor's visit expenses and vicious passing, results incorporate conduct and psychological wellness issues, animosity, substance misuse, post-traumatic pressure issue, and suicide, lower scholastic accomplishment, and taking part in further crime.

Right now, study the issue of crime percentage surmising across networks. We select Chicago as the objective of study for the accompanying explanation. Chicago has more murders and non-careless homicide rates per 100,000 inhabitants than New York and Los Angeles as per the FBI wrongdoing insights for 2013 and has encountered no decrease in the previous decade contrasted with the other two enormous urban communities, which have been on a moderate declining incline.

Generally, specialists have utilized segment data dependent on Decennial Census (e.g., populace destitution level, financial burden, racial synthesis of populace) to appraise the crime percentage in a network. Nonetheless, such data just contains the social data of inhabitants in the

March - April 2020
ISSN: 0193-4120 Page No. 11280 - 11286

areas and misses data on day by day populace elements inside and between neighborhoods. In our analyses, negative binomial and geologically weighted negative binomial models that solitary utilize segment highlights and a capture bring about a general blunder of as much as 30 percent for crime percentage estimation in Chicago. Existing examinations likewise feature the significance of topographical impact in evaluating crime percentages, i.e., the wrongdoing in the close by networks can be engendered to the central network. Be that as it may, contingent upon the geographic size of examination, topographical impact doesn't contribute a lot in improving the wrongdoing deduction on segment highlight (with at most 0.4 percent relative improvement in our trials concentrated on bigger geographic units than enumeration tracts). This is presumably in light of the fact that the close by networks likewise share comparable socioeconomics, which restrains the extra advantage of land impact.

We apply different relapse models to 5 years of wrongdoing data in Chicago. The most every now and again utilized model is straight relapse; be that as it may, on the grounds that wrongdoing check can't be negative, we additionally utilize negative binomial relapse. We exhibit that negative binomial model for the most part performs superior to the straight relapse. Furthermore, including POI and taxi stream highlights decreases the relative mistake by in any event 5 percent in our investigations. This shows the new urban data give extra data about the networks which are not secured by conventional highlights

As an expansion to the gathering adaptation, we explore models that consolidate geographic heterogeneity; that is, we don't anticipate that similar highlights should have a similar connection to wrongdoing in each area since wrongdoing episodes in various locales might be related with various socialeconomic factors. Truth be told, there are a few neighborhoods where negative binomial model gives poor prediction. This reveals to us that a worldwide model, for example, the negative binomial model, which expect a consistent relationship amongst wrongdoing and watched highlights, would not yield precise assessments. Hence, we further propose to utilize a graphically weighted relapse way to deal with catch the non-stationary property of wrongdoing. The instinct behind this procedure is to prepare numerous neighborhood models rather than one worldwide model to anticipate the wrongdoing. The geologically weighted relapse is a valuable structure on the best way to pick tests and weight them for neighborhood model preparing. We therefore apply topographically weighted relapse in blend with negative binomial model, and the trials show further enhancements over the worldwide model.

2. State of the Art (Literature Survey)

2.1 Title: Dynamic Network Model for Smart City Data-Loss Resilience Case Study: City-to-City Network for Crime Analytics/2017

Authors: Olivera Kotevska ; A. Gilad Kusne; Daniel V. Samarov ; Ahmed Lbath ; Abdella Battou

Summary: Crime data analytics might be utilized to enhance the appropriation of police, clinical, and crisis administrations. Be that as it may, as shrewd city administrations become subject to data, they additionally become vulnerable to disturbances in data streams, for example, data misfortune because of sign quality decrease or because of influence misfortune during data assortment. This paper presents a powerful system model for improving assistance versatility to data misfortune.

2.2 Title: Non-Stationary Model for Crime Rate Inference Using Modern Urban Data/2017

Author: Hongjian Wang;Huaxiu Yao; Daniel Kifer; Corina Graif; Zhenhui Li

Summary: Enormous scope Point-Of-Interest data and taxi stream data in the city of Chicago, IL in the USA. We watch fundamentally improved execution in crime rate surmising contrasted with utilizing conventional highlights. The connections amongst crime and different watched highlights are not steady over the entire city. So as to address this geospatial non-stationary property, we further utilize the geologically weighted relapse on negative binomial model (GWNBR). Investigations have indicated that GWNBR outflanks the negative binomial model.

2.3 Title: Crime Analysis using Open Source Information/2019

Author: Sarwat Nizamani Nasrullah Memon, Azhar Ali Shah, Sehrish Nizamani, Saad Nizamani, Imdad Ali Ismaili

Summary: The analysis depends on notable clustering and affiliation methods. The outcomes show that the proposed technique for wrongdoing analysis is proficient and gives a wide image of the violations of a region to investigator absent a lot of exertion. The analysis is assessed utilizing manual methodology, which uncovers that the outcomes created by the proposed approach are similar to the manual analysis, while a lot of time is spared.

2.4 Title: Hotspot Analysis of Crimes Using GIS; A Case Study of District Abbottabad/2019

Author: Mehran Khan,Rabi Azhar

Summary: Crimes hotspot maps are proficient strategy for mapping the high force of wrongdoing in a territory and the utilization of GIS methods in spatiotemporal investigation demonstrates as a powerful apparatus to appreciate the understood connection among occasions. This apparatus works by taking a gander at each component inside the setting of neighboring highlights. An element with a high worth appears to be conspicuous yet may not be a factually noteworthy problem area. Hotspot analysis give an approach to show the relationship and whether high or low qualities group spatially.

3. Implementation

The following algorithms have been used to analyze and visualize the data that has been obtained and the results were analyzed.

3.1) K Means

K-means clustering is one of the least complex and well known unaided AI calculations. Regularly, solo algorithms make deductions from datasets utilizing just info vectors without alluding to known, or named, outcomes. A bunch alludes to an assortment of information focuses totaled together in light of specific likenesses.

You'll characterize an objective number k, which alludes to the quantity of centroids you need in the dataset. A centroid is the nonexistent or genuine area speaking to the focal point of the bunch. Each datum point is assigned to every one of the groups through decreasing the in-bunch whole of squares. At the end,K-means recognizes k number of centroids, and afterward allots each datum point to the closest group, while keeping the centroids as little as would be prudent. The 'means' in the K-means alludes to averaging of the information; that is, finding the centroid.
To process the learning information, the K-means calculation in information mining begins with a first gathering of haphazardly chosen centroids, which are utilized as the starting focuses for each group, and afterward performs iterative (monotonous) counts to advance the places of the centroids It ends making and upgrading clusters when either:
1)The centroids have balanced out — there is no adjustment in their qualities in light of the fact that the clustering has been effective.
2)The characterized number of cycles has been accomplished.
K-means clustering is a widely utilized strategy for information cluster investigation. Besides, it conveys preparing results quickly. Nonetheless, its presentation is normally not as serious as those of the other advanced clustering strategies since slight varieties in the information could prompt high change.

$$J = \sum_{i=1}^{m} \sum_{k=1}^{K} w_{ik} \| x^i - \mu_k \|^2 \tag{1}$$

where wik=1 for data point xi if it belongs to cluster *k*; otherwise, wik=0. Also, μk is the centroid of xi's cluster.
3.2) Random Forest:
The essential idea driving random forest is a straightforward yet ground-breaking one — the shrewdness of groups. In data science terms, the explanation that the random forest model works so well is:
An enormous number of generally uncorrelated models (trees) working as a board of trustees will beat any of the individual constituent models.
The low relationship between models is the key. Much the same as how ventures with low connections (like stocks and securities) meet up to shape a portfolio that is more prominent than the entirety of its parts, uncorrelated models can create gathering expectations that are more precise than any of the individual forecasts. The purpose behind this great impact is that the trees shield each other from their individual mistakes (as long as they don't continually all fail a similar way). While a few trees might not be right, numerous different trees will be correct, so as a gathering the trees can move in the right heading. So the essentials for random forest to perform well are:

1)There should be some real sign in our highlights with the goal that models assembled utilizing those highlights show improvement over random speculating.
2)The expectations (and accordingly the blunders) made by the individual trees need to have low connections with one another.

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (fi - yi)^2$$

Where *N* is the number of data points,
fi is the value returned by the model and
yi is the actual value for data point *i*.

3.3) Linear Regression

Linear regression endeavors to demonstrate the connection between two factors by fitting a linear condition to watched information. One variable is viewed as an illustrative variable, and the other is viewed as a reliant variable. For instance, a modeler should relate the loads of people to their statures utilizing a linear regression model.
Before endeavoring to fit a linear model to watched information, a modeler should initially decide if there is a connection between the factors of intrigue. This doesn't really suggest that one variable causes the other (for instance, higher SAT scores don't cause higher school grades), however that there is some critical relationship between the two factors. A scatterplot can be a useful instrument in deciding the quality of the connection between two factors. On the off chance that there has all the earmarks of being no relationship between the proposed informative and subordinate factors (i.e., the scatterplot doesn't demonstrate any expanding or diminishing patterns), at that point fitting a linear regression model to the information presumably won't give a helpful model. A significant numerical proportion of relationship between two factors is the connection coefficient, which is an incentive between - 1 and 1 demonstrating the quality of the relationship of the watched information for the two factors.
Linear regression is an approach to demonstrate the connection between two factors. You may likewise perceive the condition as the incline formula. The condition has the structure Y= a + bX, where Y is the reliant variable (that is the variable that goes on the Y axis), X is the free factor (for example it is plotted on the X axis), b is the slope of the line and an is the y-intercept. The initial phase in finding a linear regression condition

March - April 2020
ISSN: 0193-4120 Page No. 11280 - 11286

is to decide whether there is a connection between the two factors. This is regularly an informed decision for the analyst. You'll additionally require a rundown of your information in x-y format.

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

3.4) K-NN

k-nearest neighbors algorithm (k-NN) is a non-parametric strategy utilized for characterization and regression,in pattern recognition. In the two cases, the info comprises of the k nearest preparing models in the component space. The yield relies upon whether k-NN is utilized for classification or regression. In k-NN classification, the output is a class enrollment. An item is characterized by a majority vote of its neighbors, with the article being doled out to the class generally normal among its k closest neighbors (k is a positive number, commonly little). In the event that k = 1, at that point the article is basically allotted to the class of that solitary closest neighbor.
In k-NN regression, the yield is the property estimation for the article. This worth is the normal of the estimations of k nearest neighbors. k-NN is a kind of case based learning, or sluggish realizing, where the capacity is just approximated locally and all calculation is conceded until work assessment.
Both for characterization and regression, a valuable method can be to dole out loads to the commitments of the neighbors, so that the closer neighbors contribute more to the normal than the more far off ones. For instance, a typical weighting plan comprises in giving each neighbor a load of 1/d, where d is the separation to the neighbor.
The neighbors are taken from a lot of items for which the class (for k-NN arrangement) or the article property estimation (for k-NN regression) is known. This can be thought of as the preparation set for the algorithm, however no express preparing step is required.
A characteristic of the k-NN algorithm is that it is sensitive to the nearby structure of the data.

$$\mathrm{d}(\mathbf{p},\mathbf{q}) = \mathrm{d}(\mathbf{q},\mathbf{p}) = \sqrt{(q_1-p_1)^2 + (q_2-p_2)^2 + \cdots + (q_n-p_n)^2} = \sqrt{\sum_{i=1}^{n}(q_i-p_i)^2}.$$

3.5) Decision Tree:
Decision Tree calculation has a place with the group of directed learning calculations. In contrast to other regulated learning calculations, the decision tree calculation can be utilized for taking care of relapse and characterization issues as well.

The objective of utilizing a Decision Tree is to make a preparation model that can use to anticipate the class or estimation of the objective variable by taking in basic decision rules surmised from earlier data(training data).
In Decision Trees, for predicting a class name for a record we start from the base of the tree. We look at the estimations of the root characteristic with the record's property. Based on examination, we follow the branch comparing to that worth and hop to the following node.The decision of making key parts intensely influences a tree's precision. The decision criteria are diverse for order and regression trees.
Decision trees utilize various calculations to choose to part a node into at least two sub-nodes. The production of sub-hubs expands the homogeneity of resultant sub-nodes. At the end of the day, we can say that the virtue of the hub increments as for the objective variable. The decision tree parts the hubs on every single accessible variable and afterward chooses the split which brings about most homogeneous sub- nodes.

$$Gain(S,A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|}.Entropy(S_v)$$

3.6) R Squared

R-squared is a factual measure that speaks to the extent of the fluctuation for a needy variable that is clarified by a free factor or factors in a relapse model; it is also represented as R2. Though connection clarifies the quality of the connection between a free and ward variable, R-squared discloses to what degree the change of one variable clarifies the fluctuation of the subsequent variable. Thus, in the event that the R2 of a model is 0.50, at that point around half of the watched variety can be clarified by the model's sources of info. In contributing, R-squared is commonly deciphered as the level of a store or security's developments that can be clarified by developments in a benchmark file.
R-Squared formula is:

$$R^2 = 1 - \frac{\text{Explained Variation}}{\text{Total Variation}}$$

R-squared qualities extend from 0 to 1 and are usually expressed as rates from 0% to 100%. A R-squared of 100% implies that all the dependent variable are totally clarified by the changes found in the independent variable.
It doesn't tell you whether your chosen a good model or a bad model, nor will it tell you whether the data and results are biased. A R-square score isn't necessarily good or bad, as it doesn't show the reliability of the model, nor whether you've chosen the right regression. There are chances to get a low R-squared for a good model, or a high R-square for a poorly fitted model.

TEST Engineering & Management
January-February 2020
ISSN: 0193-4120Page No. 564- 568

3.7) Mean Squared Error

The Mean Squared Error (MSE) or Mean Squared Deviation (MSD) of an estimator measures the normal of error squares for example the normal squared contrast between the assessed value and true value. It is a risk function, comparing to the expected result of the squared error loss. It is consistently non – negative and values near zero are better. The MSE is the second moment of error (about the origin) and in this manner joins both the variance of the estimator and its bias.

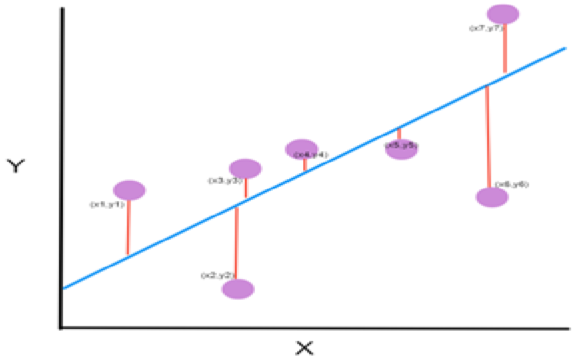


Figure 5: MSE graph

The purple dots are the points on the graph. Each point has an x-coordinate and a y-coordinate. The blue line is our prediction line. this is often a line that passes through all the points and fits them within the best way. This line contains the anticipated points. The red line between each purple point and therefore the prediction line are the errors. Each error is that the distance from the purpose to its predicted point.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2$$

4. Results Discussion

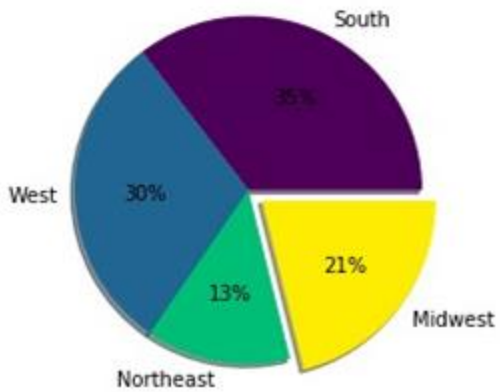


Figure 1: Crime rate Share by region

Fig(1) shows a very interesting result. We can infer that states in the lowest and highest end of the economic spectrum like Texas,New Mexico,Virginia and Florida,Chicago,California,New York,Washington D.C. have the highest share of crime in the country. This can be inferred as an imbalance in the availability of jobs and the amount of population living within these states. The southern part of the country has the highest share of crime. This is possibly due to the higher immigrant population that is present within these states and the inability of locals to accept the immigrant population due to patriotic feelings that are strong in states like Texas[5]. The western part of the country however has a severe job opportunity imbalance that is being improved day by day. The northeastern part of the country has several small states with a severe population imbalance causing people to resort to criminal activity to run their day to day life. The mid-west however has lots of farming and barren areas with low income population being the maximum amount to occupy this area.

In Fig(2) we see that Larceny is highest in the state of Illinois which is known as “The Prairie state”,meaning it has a large cover of wetlands and farmlands where dairy farming and other similar lines of work are high. Larceny is the theft of personal property and as the value of the stolen items increase it is known as Grand Larceny. The most stolen property in the state of Illinois are livestock and items related to the same. There is also a sizable amount of elite living in this state which can be inferred from the great amounts of skyscrapers that can be found in major cities. This indicates a low job count but a high income rate hence causing more crime as the population is also around 1.2 million in a state with low occupiable area[7]. The second highest is observed to be in the state of Ohio which is a good state for business and hence a sizable amount of people with larger incomes occupy this state. Although the unemployment rate is among the lower spectrum, the crimes can be justified from an emotional standpoint as there are a sizable amount of identity thefts which happen in the state which leads to larceny. On the other side of the coin we have North and South Dakota with the lowest Larceny rate as they are the most sparsely populated states in the United States of America.

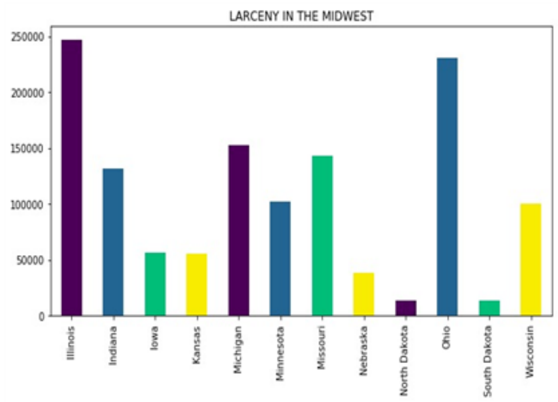


Figure 2: Larceny numbers in the Midwest region

Table 1: Number of Murders per 100k residents

STATE	MURDERS	POPULATION	MURDER RATE
Maryland	356	1300377	27.38

TEST Engineering & Management
March - April 2020
ISSN: 0193-4120 Page No. 11280 - 11286

District of Columbia	138	681170	20.26
Louisiana	368	1909852	19.27
Mississippi	114	859499	13.26
Missouri	452	3750564	12.05
Delaware	29	262570	11.04
Tennessee	394	3832161	10.28
Indiana	346	3400015	10.18
Georgia	361	3726113	9.69
Alabama	231	2499649	9.24

In Table(1) we tabulate the number of murders per 100k residents for the whole country and display the top 10 murder rates. Maryland leads the chart with 27.38% murder rate which is due to the mid level population which is disproportionate when counting the number of murders that have happened over the state. As of 2020 Maryland is considered to be one of the safest states to live in, this is due to the action of the law enforcement to crack down on crime and taking an effort to make the state a safer place[8]

.

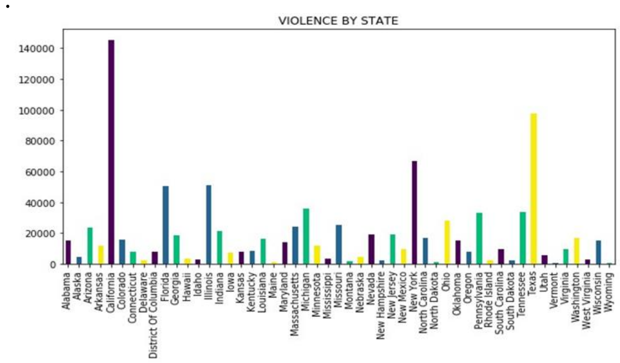


Figure 3: Violence by State

The second highest is the small District of Columbia which has a low population rate but the number of murders is a pretty high count comparatively. The highest amount of murders can be seen in the state of Missouri which has a high population when compared to the number of crimes which causes the rate of murder to drop.

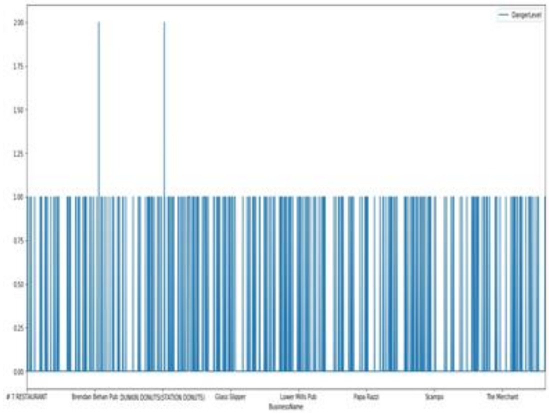
In Fig (3) a state by state analysis and visualization of crime is obtained. From the plot we can infer that the state of California with a population of around 39.5 million has the highest rate of crime with an average household income of around $75,000 which is the 7th highest among all the states. This indicates a lack of job opportunities and an imbalance in the share of wealth among the rich and poor. The main source of income foe the state comes from Los Angeles which is the hub for theater and arts as it is the base of Hollywood. The high crime rate can also be linked to human emotions such as jealousy and contempt for successful personalities who live there in high amounts. The second highest crime rate can be observed in Texas which is a farming state in general and also has its fair share of patriotic people and also immigrants who hold most of the jobs hence causing civil unrest and crime. New York, Florida, Illinois are states which have a high inflow of tourists and hence we can pin the crime that happens as petty crime such as theft and battery, but in great amounts. The states with the lowest crime rates are Vermont, Maine, North Dakota which are mountainsides that act as getaway destinations for locals and has a very low population as well.

In Table (2) we predict the danger levels of the shops and restaurants in the city of Boston based on the rate and frequency of crime in the areas around them and the location of the shop. Nearby police booths were also taken into consideration from the zip codes and as we can see the city of Boston is a relatively safe city with a low rate of crime. Few places have a danger level of 1 which means they are more prone to crime and must take precautions accordingly.

The areas that are indicated as danger level 1 in the above table are all observed to be crowded areas with a lot of shops and common places where the population is higher and the houses as a result are more expensive. This leads to the elite class of people taking residence in such areas and hence we can relate that to a higher instance of crime rate in areas like Newbury street and Boyleston street. The cause in the end comes down to feelings of resentment of the success of the people living in such areas by the less fortunate.

	BusinessName	Zip	DangerLevel
0	# 7 RESTAURANT	2132.0	0
1	20TH CENTRY BOWLING LANES	2136.0	0
2	21 ST. AMENDMENT	2108.0	0
3	224 BOSTON STREET	2125.0	0
4	29 Newbury Street	2116.0	1
5	412 Broadway	2127.0	0
6	49 Social	2111.0	0
7	5 Napkin Burger	2199.0	0
8	68 Chinese Fast Food	2108.0	0
9	7 Pond Coffee Bar	2130.0	0
10	73 Cafe	2108.0	0
11	75 CHESTNUT	2108.0	0
12	99 RESTAURANT AND PUB	2129.0	1
13	A & N PIZZA	2131.0	0
14	A @ Time	2134.0	0
15	A K's Take Out	2127.0	0
16	A K's Take Out & Delivery	2120.0	0
17	ABE & LOUIE'S	2116.0	1
18	AFC SUSHI @SIMMONS COLLEGE	2115.0	0
19	AFC Sushi @ Berklee Sch. of Music	2115.0	0
20	AFC Sushi @ Shaw's #539	2132.0	0
21	AFC Sushi @ Shaw's #586	2108.0	0
22	AFC Sushi @ Suffolk University	2114.0	0
23	AFC Sushi @ Walgreens #15390	2108.0	0
24	AL CAPONE	2110.0	0
25	AL DENTE RISTORANTE	2113.0	0

Table 2: Danger Level Prediction



January-February 2020
ISSN: 0193-4120Page No. 564- 568

Figure 4: Danger level prediction graph

In Fig(4) the plot indicates the danger levels of the different businesses in Boston and as we can see no danger level is less than 100 which is the minimum level. This is due to the fact that no business or area is completely immune from crime at any given instance and the business should be prepared. We observe only 2 businesses with a high danger level of 200 which is due to the fact that one is a pub where there are a high instance of previously recorded crimes like drunk and disorderly due to the fact that people do not know what they are doing when intoxicated[10]. The other is a gas station where there are a lot of burglary attempts which is common, hence we can see that even the owners themselves mostly carry gun licenses to protect themselves in case of any activity before the police arrive on the scene.

5. Conclusion

In this paper a series of state-of-the-art big data analytics and visualization techniques were utilized to analyze crime big data from US cities, which allowed us to identify patterns and obtain trends. By exploring the simple algorithms such as K-means, k-NN and Linear regression we find that they are useful in analyzing the data and finding areas of frequent crime that can be predicted easily and stopped in its tracks. We also found the optimal time period for the training sample to be 3 years, in order to achieve the best prediction of trends in terms of Decision tree and Random forest algorithms. Optimal parameters for the other models were also determined. Additional results explained earlier will provide new insights into crime trends and will assist both police departments and law enforcement agencies in their decision making. We also analyze the city of Chicago in detail and understand why the crime is in such high numbers so as to make it the crime capital of the world. Boston Business danger levels were also predicted and analyzed by relating it to previous crime records and area details.

References

[1] BJS, "“Rates of violent victimizations, 1993–2015” and “rates of property victimizations, 1993–2015”," Generated using the NCVS Victimization Analysis Tool, 2016. [Online]. Available: http:// www.bjs.gov, bureau of Justice Statistics, 17-Nov-2016.

[2] J. Truman and L. Langton, “Criminal victimization, 2014, ” Bureau of Justice Statistics (BJS), U.S. Department of Justice, Tech. Rep. NCJ 248973, 2015.

[3] D. Finkelhor,H. A. Turner, A. Shattuck, and S. L. Hamby, “Violence, crime, and abuse exposure in a national sample of children and youth: an update,” JAM Pediatrics, vol. 167, no. 7, pp. 614–621, 2013.

[4] C. Graif, “Toward a geographically extended perspective of neighborhood effects on children’s victimization,” Amer. Soc. Criminology Annu. Meet., Nov. 2015.

[5] C. Graif and R. J. Sampson, “Spatial heterogeneity in the effects of immigration and diversity on neighborhood homicide rates,” Homicide Stud., vol. 13, pp. 242–260, 2009.

[6] L. Anselin, “Under the hood: issues in the specification and interpretation of spatial regression models,” Agricultural Economics, vol. 27, no. 3, pp. 247–267, 2002.

[7] Y. Zheng, L. Capra, O. Wolfson, and H. Yang, “Urban computing: Concepts, methodologies, and applications,” ACM Trans. Intell. Syst. Technol. Special Section Urban Comput., vol. 5, no. 3, 2014, Art. no. 38.

[8] D. W. Osgood, “Poisson-based regression analysis of aggregate crime rates,” J.Quantitative Criminology, vol. 16, no. 1, pp. 21–43, 2000.

[9] T. Wang, C. Rudin, D. Wagner, and R. Sevieri, “Learning to detect patterns of crime,” in Machine Learning and Knowledge Discovery in Databases. Berlin, Germany: Springer, 2013.

[10] R. J. Sampson, S. W. Raudenbush, and F. Earls, “Neighborhoods and violent crime: A multilevel study of collective efficacy,” Sci., vol. 277, no. 5328, pp. 918–924, 1997.

[11] X. Wang, M. S. Gerber, and D. E. Brown, “Automatic crime prediction using events extracted from twitter posts,” in Social Comput., Behavioral-Cultural Modeling and Prediction. Berlin, Germany: Springer, 2012, pp. 231–238.