Crime Prediction using Machine Learning

T.Prathima, Madan Vijay Karnati, Vislavath Srinath
Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India

Abstract: For a developing country like India, it is not new that people hear of crimes happening quite often. With the rapid urbanization of cities, we have to constantly be aware of our surroundings. In order to avoid the unfortunate, we will try to observe crime rates by the KNN prediction method. It will predict, tentatively, the type of crime, when, where and at what time it may take place. Recognizing the patterns of a 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 is to use machine learning with the help of python to classify a criminal incident by type, depending on its occurrence at a given time and location. To be better prepared to respond to criminal activity, it is important to understand patterns in crime. In our project, we analyze crime data from the city of Indore, scraped from publicly available Kaggle website. At the outset, the task is to predict which category of crime is most likely to occur given a time and place in Indore. The use of AI/ML in predicting crimes or an individual's likelihood for committing a crime has promise but is still more of an unknown. The biggest challenge will probably be "proving" to politicians that it works. When a system is designed to stop something from happening, it is difficult to prove the negative. Companies that are directly involved in providing governments with AI tools to monitor areas or predict crime will likely benefit from a positive feedback loop. Improvements in crime prevention technology will likely spur increased total spending on this technology. We also attempt to make our classification task more meaningful by merging multiple classes into larger classes. Finally, we report and reflect on our results with different classifiers, and dwell on avenues for future work.

Key words: Naïve Bayes, K-NN, Random Forest, SVM, Crime Prediction

Introduction

With the rapid urbanization and development of big cities and towns, the graph of crimes is also on the increase. This phenomenal rise in offences and crime in cities is a matter of great concern and alarm to all of us. There are many types of crimes occurring daily in all parts of the world which are increasing day by day and it is a big task to police to stop the crimes. Different types of crimes can be robberies, murders, rapes, shop lifting, physical and emotional abuse, etc.. The frequent and repeated thefts, burglaries, robberies, murders, killings, rapes, shoplifting, pick pocketing, drug- abuse, illegal trafficking, smuggling, theft of vehicles etc., have made the common citizens to have sleepless nights and restless days and fearing what crime may happen tomorrow as crimes are increasing day by day. They feel very insecure and vulnerable in the presence of anti-social and evil elements. The criminals have been operating in an organized way and sometimes even have nationwide and international connections and links. Now a days many cyber crimes are taking place where the criminals are operating everything from different countries and the people who are getting cheated are from another countries and it is being a big task to police to catch those criminals. Much of the current work is focused in two major directions they are predicting surges and hotspots of crime and understanding patterns of criminal behavior that could help in solving criminal investigations. If we can predict what crime could happen based on previous crime patterns then it will be easy to control the crimes at those places. The main objective of our work is predicting crime before it takes place so that everyone can be on alert, predicting hotspots of crime so that more protection can be provided at those hotspots eventually the crime rate decreases at those hotspots, understanding crime pattern so that we can know what type of crimes can occur and classify crime based on location.

Machine learning

The term machine learning refers to a machine learning meaningful patterns from the data. In the past couple of decades it has become a common tool in almost any task that requires knowledge extraction from large data sets. We are surrounded by a machine learning based technology: search engines learn how to bring us the best results (while placing pro_table ads), anti-spam software learns to filter our email messages, and credit card transactions are secured by a software that learns how to detect frauds. Digital cameras learn to detect faces and intelligent personal assistance applications on smart-phones learn to recognize voice commands. Cars are equipped with accident prevention systems that are built using machine learning algorithms. Machine learning is also widely used in scientific applications such as bioinformatics, medicine, and astronomy. One common feature of all of these applications is that, in contrast to more traditional uses of computers, in these cases, due to the complexity of the patterns that need to be detected, a human programmer cannot provide an explicit, fine detailed specification of how such tasks should be executed. Taking example from intelligent beings, many of our skills are acquired or renewed through learning from our experience (rather than following explicit instructions given to us). Machine learning tools are concerned with endowing programs with the ability to learn and adapt [1]. We have five different classifications problems to solve, which we proceeded to attack with an assortment of classification algorithms. The following are the algorithms which we are using:

KNN (K-Nearest neighbors)

A powerful classification algorithm used in pattern recognition K nearest neighbors stores all available cases and classifies new cases based on a similarity measure (e.g. distance function). One of the top data mining algorithms used today. A non- parametric lazy learning algorithm (An Instance based Learning method).In pattern recognition the *k*-nearest neighbors algorithm (*k*-NN) is a non parametric method used for classification and regression. In both cases, the input consists of the *k* closest training examples in the feature space. The output depends on whether *k*-NN is used for classification or regression[2].

Both for classification and regression, a useful technique can be to assign weights to the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones. The neighbors are taken from a set of objects for which the class (for k-NN can be classification) or the object property value (for k- NN regression) is known. This thought of as the training set for the algorithm, though no explicit training step is required.

KNN: Classification Approach

i. An object (a new instance) is classified by a majority votes for its neighbor classes.
ii. The object is assigned to the most common class amongst its K nearest neighbors, measured by distance function

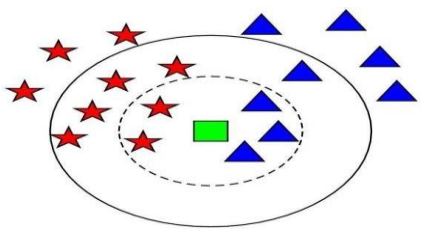

Fig 1. Principle diagram of KNN

KNN works by finding the distances between a query and all the examples in the data, selecting the specified number examples **(K)** closest to the query, then votes for the most frequent label (in the case of classification) as shown in fig.1 or averages the labels (in the case of regression). Distance function can be Euclidean, Camberra, Minkowsky, Chebychev, Manhattan, etc. A distance function provides distance between the elements of a set. If the distance is zero then elements are equivalent else they are different from each other

Decision Tree

As the name says all about it, it is a tree which helps us by assisting us in decision- making. Used for both classification and regression, it is a very basic and important predictive learning algorithm. A decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is one way to display an algorithm that only contains conditional control statements. A decision tree is a flowchart-like structure in which each internal node represents a "test" on an attribute (e.g. whether a coin flip comes up heads or tails), each branch represents the outcome of the test, and each leaf node represents a class label (decision taken after computing all attributes). The paths from root to leaf represent classification rules.

i. It is different from others because it works intuitively i.e., taking decisions one- by-one.
ii. Non-parametric: Fast and efficient. It consists of nodes which have parent-child relationships:

Decision tree considers the most important variable and splits dataset based on it. It is done to reach a stage where we have homogenous subsets that are giving predictions with utmost surety.

Random Forest

Random Forests is a very popular ensemble learning method which builds a number of classifiers on the training data and combines all their outputs to make the best predictions on the test data. Thus, the Random Forests algorithm is a variance minimizing algorithm that uses randomness when making split decision to help avoid overfitting on the training data. A random forests classifier is an ensemble classifier, which aggregates a family of classifiers h(x|θ1),h(x|θ2),..h(x|θk). Each member of the family, h(x|θ), is a classification tree and k is the number of trees chosen from a model random vector.

Also, each θk is a randomly chosen parameter vector. If D(x,y) denotes the training dataset,each classification tree in the ensemble is built using a different subset Dθk(x,y) ⊂ D(x,y) of the training dataset. Thus, h(x|θk) is the kth classification tree which uses a subset of features xθk ⊂ x to build a classification model. Each tree then works like regular decision trees: it partitions the data based on the value of a particular feature (which is selected randomly from the subset), until the data is fully partitioned, or the maximum allowed depth is reached. The final output y is obtained by aggregating the results thus, where I denotes the

$$y = \text{argmax}_{p \in \{h(x_1)..h(x_k)\}} \{ \sum_{j=1}^{\kappa} (I(h(x|\theta_j) = p)) \}$$

indicator function.

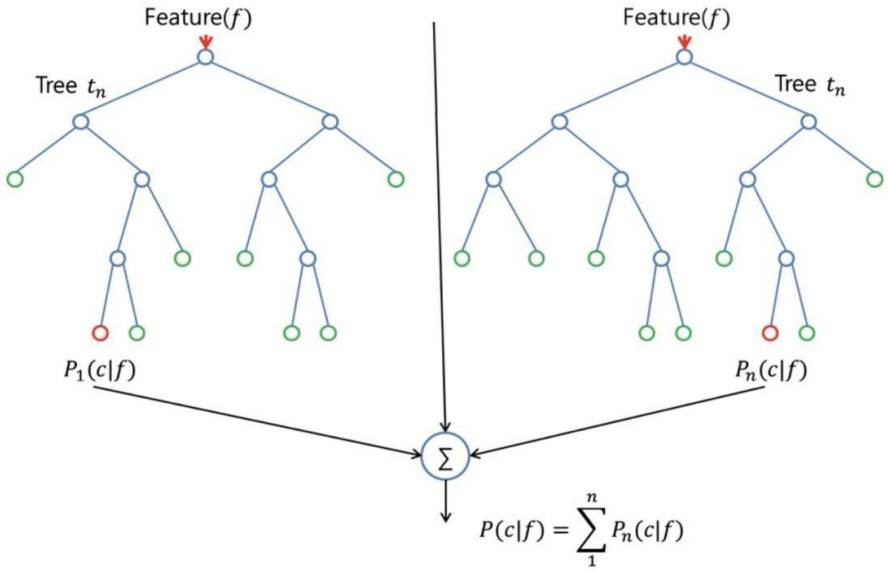

Fig 2: Random Forest Example

Naive Bayes

Gaussian Naive Bayes is a supervised classifier that uses naive assumption that there is no dependency between two features. This classifier is implemented by applying Bayesian Theorem. According to the theorem, class y and a dependent feature vector consisting of x0 , x1 ,. , xn , has the following relationship:

This probability model, along with a decision rule construct Naive Bayes Classifier. There are different types of Naive Bayes classification algorithms based on data distribution. In our project, Gaussian Naive Bayes is used, where data is assumed to be distributed according to a Gaussian distribution.

Support Vector Machine (SVM)

A support vector machine (SVM) is machine learning algorithm that analyzes data for classification and regression analysis. SVM is a supervised learning method that looks at data and sorts it into one of two categories. An SVM outputs a map of the sorted data with the margins between the two as far apart as possible. We used svm algorithm in our project to determine the accuracy of this model .

Inputs to algorithms

The inputs to our algorithms are time (hour, day, month, year), place (latitude and longitude), class of crime. The date which includes day, month, year and time which includes hour and minute and the longitude and latitude of the place where crime is happened, these are the inputs given to the above said algorithms. The class of crime includes :

i. Act 379-Robbery
ii. Act 13-Gambling
iii. Act 279-Accident
iv. Act 323-Violence

v. Act 302-Murder
vi. Act 363-Kidnapping

The output is the class of crime that is likely to have occurred.

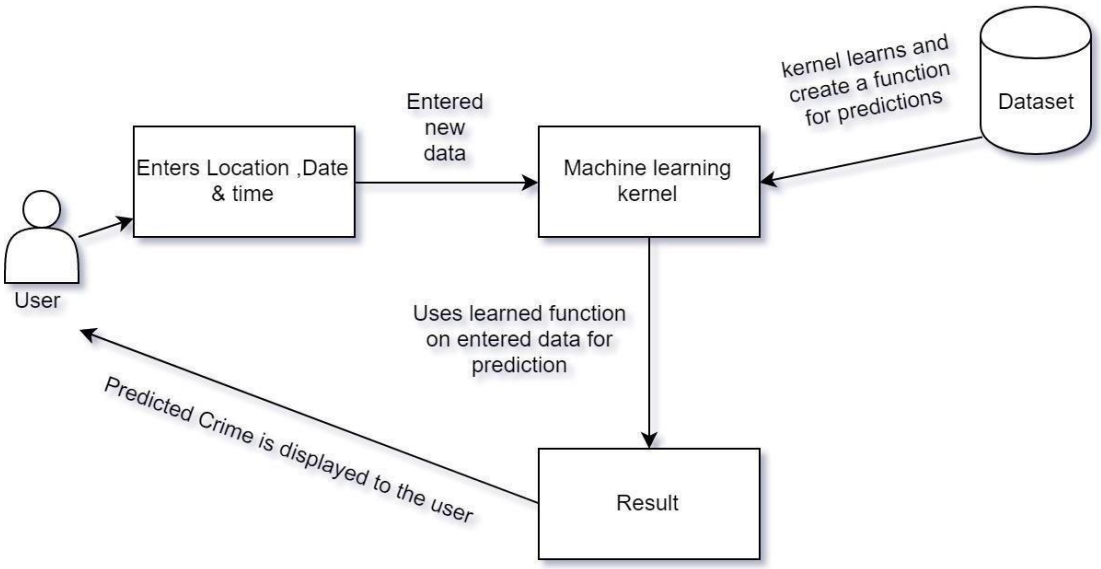

Fig3:System Architecture

The paper is organized as follows: section 2 presents existing work, section 3 describes methodology, section 4 presents analysis of the results is presented, section 5 deals with conclusion and future scope of the work.

2. Literature Survey

There is no system which uses machine learning models at present, the police can predict only at given important situations like any big festival at that robbery may take place , only predicting with experience which is of very less accuracy. There is currently no existing system to predict crimes that can take place in future . So, it is difficult to law enforcement officers to prevent to growth of crimes that are increasing day by day. Officers only get to know to know about the crimes if they have any informers, so it is difficult to predict crimes normally. So this machine learning models came into existence which can predict crimes above 90% also. This is the existing system, so this project is very much helpful in predicting the crimes and help law enforcement officers and safeguarding the public. Various researchers have addressed the problems regarding crime control and have proposed different crime-prediction algorithms. The accuracy of prediction depends on the attributes selected and the dataset used as a reference.

A comparison between two classification algorithms, Decision Tree and Naïve Bayesian, was performed using WEKA, an open-source data mining software, and 10-fold cross-validation. The socio-economic, law-enforcement, and crime datasets for this study were compiled from the 1990 US Census, the 1990 US LEMAS survey, and the 1995 FBI UCR, respectively.

The road accident patterns in Ethiopia was studied considering various circumstantial factors like the driver, weather, car, and road conditions. Three different classification algorithms,

KNN, Naïve Bayesian, and Decision tree were used on a dataset of 18,288 accidents. The prediction accuracy for all three algorithms was between 79% to 81%.

In [3] the authors presented that Criminal activity is gradually rising in India and has a significant and negative social impact. The recent spurt in the nation has put everyone wondering as to what will happen in the future. Cases of murder, abduction, rape, and fatal accidents have skyrocketed. The need of the hour is to make people of the nation realize the issue. Machine learning advancements and deep learning algorithms can find new patterns in various data sets and reveal new information. Crime prediction and identifying criminals are the one of the top priority problems to the police department because there is a tremendous amount of data related to crime that exists. There is a need for technology through which the case-solving could be faster . The idea behind this project is that crimes can be easily predicted once we are able to sort through a huge amount of data to find patterns that are useful to configuring what is required. The recent developments in machine learning makes this task possible. One will give date, time, location (longitude, latitude) as input and the output will be generated which will give us information about which crime is likely to happen in that area. It basically gives us the hotspots of crime. The data is taken considering the time and type of crime that happened in the past. KNN algorithm then uses its approach which assumes that similar things exist in close proximity and classifies new cases based on similarity measures. The next step was to decide which algorithm to use. K Nearest Neighbor Classifier is a supervised machine learning algorithm useful for classification problems. It works by finding the distances between a query and all the examples in the data, selecting the specified examples that are closest to the query, and then votes for the most frequent label. It is not parametric which implies that it does not make any supposition on the primary data distribution. To put it in simple words, the model structure is decided by the data. It's pretty useful because in reality, most of the data does not follow the typical theoretical norms made. Hence, we decided to use K-Nearest-Neighbor Algorithm. They reported an accuracy of around 89%

In [4,5,6] the authors reported that crime, in all its facets, is a well-known social problem affecting the quality of life and the economic development of a society. Studies have shown that crime tends to be associated with slower economic growth at both the national level and the local level, such as cities and metropolitan areas. Crime-related information has always attracted the attention of criminal law and sociology scholars. Dating back to the beginning of the 20th century, studies have focused on the behavioral evolution of criminals and its relations with specific characteristics of the neighborhoods in which they grew up, lived, and acted. The study of the impact on behavioral development of factors like exposure to specific peer networks, neighborhood characteristics (e.g., presence/absence of recreational/educational facilities) and poverty indexes, has provided a wealth of knowledge from both individual and collective standpoints. Existing works in the fields of criminology, sociology, psychology and economics tend to mainly explore relationships between criminal activity and socio-economic variables such as education, ethnicity, income level, and unemployment. Several studies in criminology and sociology have provided evidence of significant concentrations of crime at micro levels of geography, regardless of the specific unit of analysis defined. It is important to note that such clustering of crime in small geographic areas (e.g. streets), commonly referred to as hotspots, does not necessarily align with trends that are occurring at a larger geographic level, such as communities. Research has

shown, for example, that in what are generally seen as good parts of town there are often streets with strong crime concentrations, and in what are often defined as bad neighborhoods, many places are relatively free of crime. The problem of crime hotspot forecasting as a binary classification task. For each Smart steps cell, we predict whether that particular cell will be a crime hotspot or not in the next month. In this section we provide details of the experimental setup that we followed. As the Smart Steps cell IDs, the borough profiles and the crime event locations are not spatially linked in the provided datasets, we first geo-referenced each crime event by identifying the Smart steps cell which is the closest to the location of the crime. We carried out a similar process for the borough profile dataset. As a result, each crime event and the borough profile information were linked to one of the Smart steps cells. Fivefold validation technique Logistic regression Svm, Neural networks, Decision tree. Random forest. In this random forest has more accuracy 70%. As phone data is personal chances of misusing it are more.

In [7,8] the authors reported that crime is one of the major issues is continuing to grow in intensity and complexity . In the recent years, crime is one of the social problems influencing the nature of life and economic development in a community. Crime can be divided into a few types such as crime against properties (theft, burglary, and robbery) and crime of aggression (homicides, assaults and rape). The availability of information technologies has enabled law enforcement to collect detailed information of crime data . With the increasing numbers of crimes nowadays, crime analysis is needed which comprises measure and procedure that intend to reduce the risk of crime . Crime analysis can be done through both quantitative and qualitative methods . Qualitative approaches in predicting crime such as scenario writing or environmental scanning are valuable in identifying the future of criminal activity . Meanwhile, quantitative method is used to predict the crime rates in future specifically. Moreover, crime analysis is a practical approach to analyze and identify the pattern of crimes . The study by stated that crime analysis is part of crime prevention which has the tasks of discovering and detection of crimes and their relation with criminals. However, a considerable challenge faced by law enforcements is to analyze the increasing numbers of crime data accurately and efficiently . The ability to analyze amount of crime data without using computational support will put strain on human because human mechanism is incapable of comprehending with millions of data. Since the high dimensional of crime data, traditional crime analysis methods cannot be applied for identifying the future crimes. The powerful system in predicting crime is required when the more data and the complex queries are being processed . Therefore, developing a systematic crime analysis tool to identify the crime patterns effectively and promptly for future crime patterns detection is needed. This includes the capability of the system to analyze or evaluate a large amount of data and information from variety sources. Recently, support vector machine (SVM) is a nonlinear model, and it has been performed well for time series prediction, classification tasks and regression. The performance of the SVM is compared with the performance of both recurrent neural network model (RNN) and autoregressive moving average (ARMA) in time series prediction . ARMA model is a linear behavior but it is easy and fast in use while RNN can be in principle model linear, but they are usually difficult to train and does not lead to one unique or global solution due to differences in their initial weight set. Besides that, SVM has been performed well in prediction of time series because they can model nonlinear relations in a stable and efficient way . Furthermore, the SVM is educated as a convex optimization problem resulting in a global solution which in many cases output unique

solutions. The performance of prediction models can be evaluated using a variety of different prediction methods such as support vector machine, multivariate time series and artificial neural network. Crime analyses are required to reveal the complexities in the crime dataset.

4. Methodology

Below is the block diagram of the process adopted in this work.

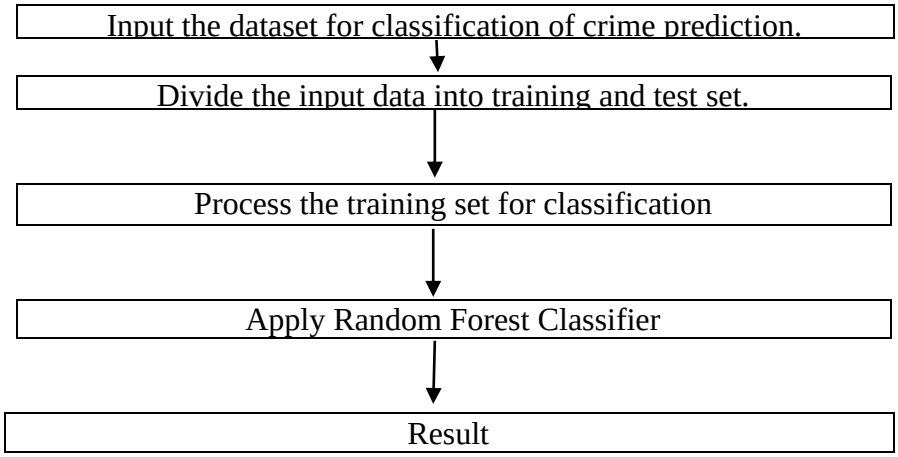

Fig 4. Proposed system

The above block diagram gives a brief flow of the working of the project. The steps include 1.input the dataset for classification of crime prediction the preprocessing of dataset.2. we divide the dataset for training and for testing set. Here 20% of the values are used in the training and the remaining 80% are used in the testing of the model. 3. Process the training set for classification we train and test on different models and calculated accuracy for the models. 4.we deploy the project into flask while deploying we used the random forest classifier for prediction which gives good accuracy and predict the outcome.5.finally we will get required output of the project results of predicted crime.

Module Description Loading Dataset

First the dataset is loaded which is a csv file that has 2091 rows that consists of date , timestamp , latitude longitude and different types of crimes that took place on given date and timestamp .

Preprocessing:

Before implementing machine learning algorithms on our data, we went through a series of preprocessing steps with our classification task in mind. These included:

- Dropping features such police station, station number, Complainant name & address, Accused name & address.
- Dropping features such as Resolution, Description and Address: The resolution and description of a crime are only known once the crime has occurred, and have limited significance in a practical, real-world scenario where one is trying to predict what

kind of crime has occurred, and so, these were omitted. The address was dropped because we had information about the latitude and longitude, and, in that context, the address did not add much marginal value.

- The timestamp contained the year, date and time of occurrence of each crime. This was decomposed into five features: Year (2018), Month (1-12), Date (1- 31), Hour (0-23) and Minute (0-59).

	timestamp	act379	act13	act279	act323	act363	act302	latitude	longitude
0	28-02-2018 21:00	1	0	0	0	0	0	22.737260	75.875987
1	28-02-2018 21:15	1	0	0	0	0	0	22.720992	75.876083
2	28-02-2018 10:15	0	0	1	0	0	0	22.736676	75.883168
3	28-02-2018 10:15	0	0	1	0	0	0	22.746527	75.887139
4	28-02-2018 10:30	0	0	1	0	0	0	22.769531	75.888772

Fig 5. Dataset

Theoretical Foundations/Algorithms Decision Tree

Decision Tree algorithm belongs to the family of supervised learning algorithms. Unlike other supervised learning algorithms, the decision tree algorithm can be used for solving regression and classification problems too. The goal of using a Decision Tree is to create a training model that can use to predict the class or value of the target variable by learning simple decision rules inferred from prior data(training data).

In Decision Trees, for predicting a class label for a record we start from the root of the tree. We compare the values of the root attribute with the record's attribute. On the basis of comparison, we follow the branch corresponding to that value and jump to the next node.

Random Forest: Radom Forests is a very popular ensemble learning method which builds several classifiers on the training data and combines all their outputs to make the best predictions on the test data. Thus, the Random Forests algorithm is a variance minimizing algorithm that uses randomness when making split decision to help avoid overfitting on the training data.

Support Vector Machine: A support vector machine (SVM) is machine learning algorithm that analyzes data for classification and regression analysis. SVM is a supervised learning method that looks at data and sorts it into one of two categories. An SVM outputs a map of the sorted data with the margins between the two as far apart as possible. We used svm algorithm in our project to determine the accuracy of this model[6] .

Naive Bayes: Gaussian Naive Bayes is a supervised classifier that uses naive assumption that there is no dependency between two features. This classifier is implemented by applying Bayesian Theorem. According to the theorem, class y and a dependent feature vector

consisting of x0 , x1 ,. , xn , has the following relationship:

$$P(y|x0,x1,...\ldots\ldots xn)=P(x0,x1\ldots\ldots xn|y)/P(x0,x1,\ldots\ldots xn)$$

Bayes theorem provides a way of calculating the posterior probability, P(c|x), from P(y), P(x), and P(x|y). Naive Bayes classifier assume that the effect of the value of a predictor (x) on a given class (y) is independent of the values of other predictors. This assumption is called class conditional independence[9].

KNN (K-Nearest neighbors): A powerful classification algorithm used in pattern recognition K nearest neighbors stores all available cases and classifies new cases based on a similarity measure (e.g., distance function). One of the top data mining algorithms used today. A non- parametric lazy learning algorithm (An Instance based Learning method).In pattern recognition the *k*-nearest neighbors algorithm (*k*-NN) is a non-parametric method used for classification and regression . In both cases, the input consists of the *k* closest training examples in the feature space. The output depends on whether *k*-NN is used for classification or regression.

Dataset Description: The dataset is available on Kaggle. It consists of data like latitude longitude, different types of crimes, timestamp which includes date and time of the crime. The different types of crimes that are present in the dataset are Robbery, Accident, Murder, kidnapping, violence, and gambling.

Testing process: Testing the model is the step wherein our model is given the test dataset which it has not been trained on during the training process. Here the model gives us its prediction of the crime in the dataset, and we determine whether the output given by the model is the one which is expected or not. If not, we determine how much it is deviating from the actual outcome and calculate the efficiency/accuracy of the model based on the outcomes given by our model. The trained model is tested on the test data, and we make use of several functions for determining the loss of the model on test data which can thereby be used in order to determine the accuracy of the model on the test data.

5. Results and Discussions

The dataset has been tested on different permutations and each of the models has tested and given different performance measures on the same dataset. The following table provides an overview about various accuracies that have been recorded for these models.

Table 1: Results obtained after experimentation

Model	Accuracy
Decision Tree	98.1%
KNN	97.7%
Random Forest	99.1%

6. Conclusion

The initial problem of classifying 6 different crime categories was a challenging multi-class classification problem, and there was not enough predictability in our initial data-set to obtain very high accuracy on it. We found that a more meaningful approach was to collapse the crime categories into fewer, larger groups, in order to find structure in the data. We got high

accuracy and precision on Prediction. However, the Violent/Non-violent crime classification did not yield remarkable results with the same classifiers – this was a significantly harder classification problem. Thus, collapsing crime categories is not an obvious task and requires careful choice and consideration. Possible avenues through which to extend this work include time-series modeling of the data to understand temporal correlations in it, which can then be used to predict surges in different categories of crime. It would also be interesting to explore relationships between surges in different categories of crimes – for example, it could be the case that two or more classes of crimes surge and sink together, which would be an interesting relationship to uncover. Other areas to work on include implementing a more accurate multi-class classifier, and exploring better ways to visualize our result.

Future Work

The goal of any society shouldn't be to just catch criminals but to prevent crimes from happening in the first place. Predicting Future Crime Spots: By using historical data and observing where recent crimes took place we can predict where future crimes will likely happen. For example a rash of burglaries in one area could correlate with more burglaries in surrounding areas in the near future. System highlights possible hotspots on a map the police should consider patrolling more heavily. Predicting Who Will Commit a Crime: Using Face Recognition to predict if a individual will commit a crime before it happens. The system will detect if there are any suspicious changes in their behavior or unusual movements. For example if an individual seems to be walking back and forth in a certain area over and over indicating they might be a pickpocket or casing the area for a future crime. It will also track individual over time. Pretrial Release and Parole: After being charged with a crime, most individuals are released until they actually stand trial. In the past deciding who should be released pretrial or what an individual's bail should be set at is mainly now done by judges using their best judgment. In just a few minutes, judges had to attempt to determine if someone is a flight risk, a serious danger to society, or at risk to harm a witness if released. It is an imperfect system open to bias. The media organization's analysis indicated the system might indirectly contain a strong racial bias. They found, "That black defendants who did not recidivate over a two-year period were nearly twice as likely to be misclassified as higher risk compared to their white counterparts (45 percent vs. 23 percent)." The report raises the question of whether better AI/ML can eventually produce more accurate predictions or if it would reinforce existing problems. Any system will be based off of real-world data, but if the real world data is generated by biased police officers, it can make the AI/ML biased.

REFERENCES

[1] Wang, Tong and Rudin, Cynthia and Wagner, Daniel and Sevieri, Rich. 2013. pages 515-530, Machine Learning and Knowledge Discovery in Databases

[2] Jiawei Han, Micheline Kamber and Jian Pei, "Data Mining: Concepts and Techniques", The Morgan Kaufmann Series in Data Management Systems, Morgan Kaufmann Publishers, 3rd Ed, ISBN 978-0123814791, (2011)

[3] Akash Kumar, Aniket Verma, Gandhali Shinde, Yash Sukhdev, Nidhi lal , "A KNN under sampling approach for data balancing" 2020, IEEE

[4] Bogomolov, Andrey and Lepri, Bruno and Staiano, Jacopo and Oliver, Nuria and Pianesi,Fabio and Pentland, Alex.2014. Once upon a crime: Towards crime prediction from demographics and mobile data, Proceedings of the 16th International Conference

on Multimodal Interaction.
[5] Nurul Hazwani Mohd Shamsuddin, Nor Azizah Ali, Razana Alwee, "An overview on crime prediction methods" 2017, IEEE
[6] Yu, Chung-Hsien and Ward, Max W and Morabito, Melissa and Ding, Wei.2011. Crime forecasting using data mining techniques, pages 779-786, IEEE 11th International Conference on Data Mining Workshops (ICDMW)
[7] Kianmehr, Keivan and Alhajj, Reda. 2008. Effectiveness of support vector machine for crime hot-spots prediction, pages 433-458, Applied Artificial Intelligence, volume 22, number 5.
[8] Toole, Jameson L and Eagle, Nathan and Plotkin, Joshua B. 2011 (TIST), "Spatiotemporal correlations in criminal offense records",volume 2, number 4, pages 38, ACM Transactions on Intelligent Systems and Technology
[9] Friedman, Jerome H. "Stochastic gradient boosting." Computational Statistics and Data Analysis 38.4 (2002): 367-378.