

# Crime Prediction Using Machine Learning Algorithms

**Shrinath Dalave<sup>1</sup>, Prajakta Lawande<sup>2</sup>, Prachi Panasare<sup>3</sup>, Prof. Deepali Sale<sup>4</sup>**

Students, Department of Computer Engineering<sup>1,2,3</sup>

Navsahyadri Education Society's Group of Institute, Pune, Maharashtra, India

shrinathdalave999@gmail.com<sup>1</sup>, awandeprajakta12@gmail.com<sup>2</sup>, pansareprachi2156@gmail.com<sup>3</sup>

**Abstract:** *The most serious security challenges we face in these turbulent times are terrorist attacks and the transmission of disease. length and breadth are measured in hundredths of a centimetre. On a daily basis, we see the most minor offences committed by ordinary citizens. Details of breaches and recurring cases of items should be applied to files to ensure that they are up to date. When it is known that a crime has been committed, people believe that disciplinary action will be taken, even if there is no means of knowing which one. The study of criminology helps to broaden our understanding of who is likely to become a suspect. In the midst of his attempts to identify and deter alleged criminals from reoffending the legal system, he is incorporating both computer science and deep learning. Anyone interested in learning more about the workings of the Chicago Police Force should visit "The Chicago Police Department Site." The Crime Timeline will keep track of all criminal activity as well as the time and date of any incident that occurs. The data collection and modelling have been completed; all that remains is on-line modelling and compilation. To address this question, we must first determine if the case history of K-grooming and other related methods will help with criminal prediction. The invention is typically used as a testing tool, but it can also be used in conjunction with other technologies. Based on internal or external metrics, an algorithm can estimate how easily law enforcement authorities may be able to track, anticipate, and cope with, or preempt, risks, such as the ratio of those sentenced to those arrested, with a life sentence to those awaiting the risk of life imprisonment.*

**Keywords:** Stock Prediction, Data Analysis, Natural Language Processing Machine

## I. INTRODUCTION

To understand this, it should be noted that criminals can be harmful to our wellbeing Nobody is actually imprisoned for just for ordinary crimes but a wide variety of misdemeanours occur on a daily basis. While nearly all galaxies are rotating at over the course of a million times a second, this galaxy can advance approximately 70 million miles a year. In small towns, crime is not just as common as in rural counties as it is in large cities. However numerous their names may be, all these actions may be, attempted murder, armed robbery, larceny, false imprisonment, sexual assault, false conviction, and battery are all identical. Criminal activity has increased, it is the responsibility of the police to act quickly and stop the problem from spreading. since all past criminal offences contain specific information and some indications of future criminal behaviour, the task of policing the city is difficult to predict, the cops' abilities to forecast criminal activity are highly limited [T] To minimise the number of incidents, there are solutions that need to be employed to increase the speed of determination. Using an existing research approach, such as identifying places where crimes are most likely to occur, is a good way to better predict where potential ones will occur The definition of crime in this research uses different terminology throughout; terms related to the places it is incorporated into various geographic divisions are mentioned as well. the likelihood of committing a certain offence can be predicted based on knowing the circumstances in which it is more specific crimes have occurred before lots of crimes have happened in a certain locations also aids in the identification of hotspots, as crime prevention is sought in locations where crimes have been committed most frequently According to recent criminal rate figures, that must have happened with the same methodology as before. The use of a strong and highly functional cyber forensics platform is imperative for allowing us to detect trends in the database in a timely manner that can be viewed immediately so bugs can be addressed quickly. Concomplementing this phrase: Extortion, robbery, financial terrorism, and vandalism may also be examples of other kinds of criminal activity that are considered to be new for the industry, which

may be described as unnatural. One important function of new technologies is the consumers get out of using it is to open doors to all the types of crooks, and customers getting a return on their investment out of the addition of new technology. Risks to financial, which include the possibility of financial losses due to cybercrime, an inability to transact business, and data leaks are covered under three distinct categories: 1) financial loss due to cybercrime; 2) incapacity to manage business due to technological factors; 3) possible data leakage.

## **II. PROBLEM STATEMENT**

The models for the crime prediction methods will be used to search through data found in the police archives for data on specific types of crimes as well as the records on different details of their occurrences, with variables that influence the probability of occurrence studied to get better.

## **III. LITERATURE SURVEY**

Alkesh Bharati, Dr Sarvanaguru RA, "Crime Analysis and Prediction Using Fuzzy C-Means Algorithm"[1], The data was/is presented as Crime research is a tool used to define criminal activities and study them. If the research conducted so far can be seen to be more specifically useful, it is mostly because it indicates which criminal types are useful in controlling crime, then, mostly they would it be places where violent crimes are reduced. It is an excellent method for measuring the crime rate because of each region can be broken down by procedure and the data is collected for any of each process to be examined. Through the rapid increase in information technology, crime analysts will be able to continue to enhance the investigations and help them interpret the evidence. on the sample clustering and preprocessing to get unstructured evidence, and then look for crimes inside it Thus, persons formerly investigated and then arrested or identified as having committed the same criminal behaviour may often be looked at at for patterns such as criminal history, or incident reports, rather than only offences themselves. This is simply intended to direct law enforcement resources to where crimes can occur, without attention to identifying who is responsible. Bayesian classifiers were used as the current scheme was in use in place In the current methodology, the fuzzy C-Means algorithm will be used to group the crime data for all items that are apprehensible, apprehension of, physical assault, larceny-theft, and crime of women, as well as all criminal offences such as kidnapping, in the dataset.

Shubham Agarwal, Lavish Yadav, "Crime Prediction based on Statistical Models"[2], The data was/is presented as In societies that have less criminal activity, the increase in the number of various offences is still a matter of concern. [Through] developments in technology, freely accessible records, and services,] these people manage to go undetected by society and continue their illegal deeds, even in the act of involving far more people. As a result, crime is rising in countries with a steep increase in incidence in [either/many] developed and [or/in] and [under] developing Because of the preceding year's criminal events in Indian states, we provide two models —Working Average Geometric Progression and WAGP and Seasonal Augmented EP WAG with known past criminal occurrences in order to anticipate the crime activity that is likely to occur in following years. These recent examples are used to help researchers understand how crime has changed in Indian states between the years 2001 and 2013. It was discovered that information from police reports covering the years 2001 to 2011 was a valuable in forecasting data that followed. The calculated expected crime values were compared to crime data for the same years, as well as for 2014 and 2015. To estimate our prediction to be within 85% of the correct value obtained from real data, the difference between the real and expected values of each needs to be doubled.

Chitra Lekha, "Data Mining Techniques in Detecting and Predicting Cyber Crimes in Banking Sector"[3] appearing as, also represented as Data mining is used in many industries such as client segmentation and efficiency, credit scoring, predicting payment failures, preventing fraud, and predicting which customers will default, but are also in the fields of advertisement and marketing Data Mining in banking applications is presented in this paper as a basic concept that serves as a foundation for understanding a range of diverse cyber criminals. It offers an inclusive look at the most powerful and current approaches used data mining practises for the purposes of criminal data analysis. the aim of stealing personal data is to be able to identify behavioural habits in order to save you from doing what's likely to happen, and avoid the illegal things you have already done This paper employs novel data mining techniques like K-Means, which uses current events as a context signal to guide the challenge, and influences, and a novel data mining algorithm known as J48 Prediction Tree, to investigate sets of data about cyber crime and cyber threats. Cluster expansion is being used in unsupervised learning even though it is affected by the association (influenced) Using the J48 algorithm, the classifier finds the initial centres of things



that may be indicative of cyber criminals and populates the list of candidate results with the predicted instances of those centres. In light of the above, a general intelligence outcome over all criminal justice and a prediction tree will definitely have expanded, and improved, and a significant general learning result for banking sector is a K-Means and J48 Incorporation. As a result, we need our law enforcement agencies to be properly equipped to face and avoid and deter online criminal activities.

Nafiz Mahmud, Khalid Ibn Zinna, "CRIMECAST: A Crime Prediction and Strategy Direction Service"[4], The data was/is presented as The wide variety of studies on criminology is valuable in providing us with a new information on criminal psychology. Criminals don't live in uncertain territory; they wait before they have an easy target to commit offences, in which case there are clustered areas like hotspots of people or strangers. It is possible to simulate a crime forecasting model using evidence that can be checked in the fact of past crimes, as long as it has been publicly available, there is enough time enough to verify. This paper intends to show how the CCRIMBA's artificial Neural Network has been broadened to include the CRIMAST, a crime prediction and threat management service which assists law enforcement in training and testing criminologists to work with Neural models. the CriMA employs spatial techniques, which concentrate on legitimate crime patterns of crime and generate defence strategies, designates areas vulnerable to criminal action, and then broadcasts security warnings. Our simulation with a large dataset shows that CRIEC can be much more effective than other models in terms of predicting crime.

Mary Shermila A, "Crime Data Analysis and Prediction of Perpetrator Identity using Machine Learning Approach"[5], The data was/is presented as Prevention is one of the most prominent and important tasks we have in the realm of civilization. In addition to being a means of identifying and researching the usual patterns and developments of violence, it is also a systemic, scientific approach. the aim of this model is to make systems more effective at detecting and apprehending criminals This statistical model can be employed at the crime scenes to discover crime dynamics and to forecast the description of the criminal most probable perpetrator to be present based on inferences drawn from the site. This is a long, involved process of both physically expanding and philosophically advancing. Predictions on who is likely to commit crimes and how serious such crimes are The phase includes determining the number of open offences, which gauges the importance of different variables, such as the year, month, the weapons used, and the social class or demographics of the perpetrators. the prediction process is able to deduce how old, whether the suspect is male, female, and/how many years they've known the victim There are several theories from the investigation based on the information gathered on this area. The method uses multinomial regression, k-neigh regression, and neural networks for classifiers like Multilearate Regression, kNeighbors, and KNeighbors for static entity definition. The machine learning algorithm was developed and thoroughly tested using the San Francisco homicide dataset (1981-2014) and then deployed using Python.

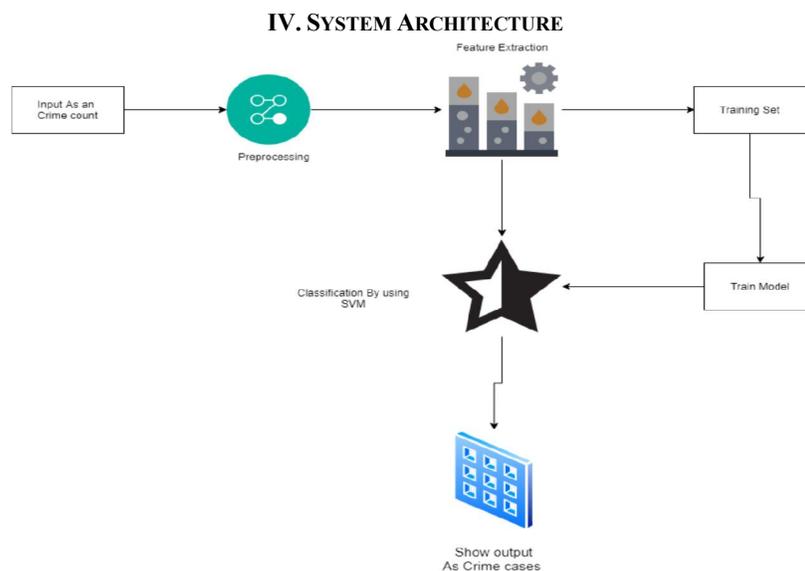


Figure: System Architectures

#### 4.1 Modules

An improvement on image data unwilling to distort, but necessary for further processing is achieved by means of geometric transformations, including rotation, scaling, and translation. Likewise, the method of feature selection is feature extraction, which breaks the raw data into usable classes and thereafter divides it into a series of features is a source of data reduction. When you want to work on it, you will find it more agreeable. Also, one of the primary characteristics of these massive data sets is that it has a large numbers of variables. Through supervised learning machines (i.e. classification or regression) can both classify and predict the future, it by getting information from past training examples, they are more useful for regression and less useful for prediction. Classification issues account for the vast majority of the instances, however. We depict each object of the data with n-dimensional space (which could be described as a set of features with each having a value for each of those features), and the features are identified with various coordinates. And once the hyper-plane has distinguished the two classes, we use the classifier to locate the other point(s) that sits(s) in between the

#### V. CONCLUSION

Since computers are now being used for both on- and off-site to help in data analysis, it is possible to recognise patterns using artificial intelligence. the bulk of the research in this project focuses on is focused on detecting crimes that have already occurred It was processed using the machine learning technique of data cleaning and normalisation. According to the theory, the prediction, this type of crime would have an accuracy of 7.88%. Expanding on the original definition, it can be defined as "to assist in making better use of the dataset". When you look at the shapes of diagrams, you're often examining their properties, not when you determine what sort of information you're searching for. We designed multiple images to go along with this concept It was essential in obtaining datasets in Chicago to show patterns and trends in criminal behaviour in order to discover different factors that could further improve the likelihood of avoiding or combat crime. The new uses of AI and big data analysis to clarify and illustrate dynamic relationships and relationships that are based on big data have gained prominence in recent years the vast majority of the research done in this project has to this has been dedicated to estimating crimes that have already occurred It was through the use of the machine learning techniques that we developed a model using the data that has been through the whole data cleaning and data analysis.

#### REFERENCES

- [1]. McClendon, Lawrence, and Natarajan Meghanathan. "Using machine learning algorithms to analyze crime data." *Machine Learning and Applications: An International Journal (MLAIJ)* 2.1 (2015).
- [2]. Kiani, Rasoul, Siamak Mahdavi, and Amin Keshavarzi. "Analysis and prediction of crimes by clustering and classification." *Analysis* 4.8 (2015).
- [3]. Heartfield, Ryan, George Loukas, and Diane Gan. "You are probably not the weakest link: Towards practical prediction of susceptibility to semantic social engineering attacks." *IEEE Access* 4 (2016): 6910-6928.
- [4]. Sivaranjani, S., S. Sivakumari, and M. Aasha. "Crime prediction and forecasting in TamilNadu using clustering approaches." *Emerging Technological Trends (ICETT), International Conference on. IEEE, 2016.*
- [5]. Kansara, Chirag. "Crime mitigation at Twitter using Big Data analytics and risk modelling." *Recent Advances and Innovations in Engineering(ICRAIE), 2016 International Conference on. IEEE, 2016.*
- [6]. [6] Kim, Suhong, Param Joshi, Parminder Singh Kalsi, and Pooya Taheri. "Crime Analysis through Machine Learning." *In 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON),pp. 415-420. IEEE, 2018.*
- [7]. Shah, Riya Rahul. "Crime Prediction Using Machine Learning." (2003).
- [8]. Lin, Ying-Lung, Tenge-YangChen and Liang-ChihYu. "Using machine learning to assist crime prevention." *In 2017 6th II AI International Congress on Advanced Applied Informatics (IIAI-AAI), pp. 1029-1030. IEEE, 2017.*
- [9]. M. V. Barnadas, Machine learning applied to crime prediction, Thesis, University at Politecnica de Catalunya, Barcelona, Spain, Sep. 2016. ‘
- [10]. Williams, Matthew L., Pete Burnap, and Luke Sloan. "Crime sensing with big data: The affordances and limitations of using open-source communications to estimate crime patterns." *The British Journal of Criminology* 57, no. 2 (2017): 320-340. 11 Agarwal, Shubham, Lavish Yadav, and Manish K. Thakur. "Crime Prediction Based

on Statistical Models.”In2018Eleventh International Conference on Contemporary Computing (IC3), pp. 1-3.  
IEEE, 2018. College Short Form Name, Department of Computer Engineering