



# Crime Data Analysis using Machine Learning

Dr. P. Sriram Chandra | N. Sowmya | M. SyamaSundari | A. Prahruyani | K. Nitesh

Department of Computer Science and Engineering, Godavari Institute of Engineering and Technology(A), JNTUK, Kakinada.

## To Cite this Article

Dr. P. Sriram Chandra, N. Sowmya, M. SyamaSundari, A. Prahruyani and K. Nitesh. Crime Data Analysis using Machine Learning. International Journal for Modern Trends in Science and Technology 2022, 8(S03), pp. 13-15. <https://doi.org/10.46501/IJMTST08S0304>

## Article Info

Received: 26 April 2022; Accepted: 24 May 2022; Published: 30 May 2022.

## ABSTRACT

*The evildoer cases in India are becoming quickly because of which assortment of examples forthcoming are furthermore stacking up. This constant extend in the criminal occasions is ending up intense to be marked and to be addressed. Perceiving the hoodlum practice examples of an area is fundamental to thwart it from working out. The wrongdoing fixing organizations can accomplish a higher work assuming they have a right thought about the example of evildoer activities that are happening in a particular region. This can be achieved via the utilization of work area concentrating via utilizing selective calculations to find the examples of the hooligan activities in an exceptional region. This paper utilizes wrongdoing insights set and predicts such violations in a particular area which helps in running up the arrangement of hoodlum cases and continue accordingly. This paper utilizes the measurements of past 18 years that is accumulated from in excess of a couple of relied upon sources. Information pre-handling is all around as essential as residual forecast, this paper utilized work choice, discarding invalid qualities and name encoding to simple and support the information. This query gives a climate accommodating processing gadget inclining life sized model for anticipating the ensuing criminal case.*

## 1. INTRODUCTION

As of now, the hoodlum cases that are forthcoming in India are quickly developing with the wide assortment of violations committed are expanding. To cure a case dependent absolutely on a particular records there should be a careful examination and assessment that will be performed inside. With the amount of wrongdoing measurements that is current in India by and by the assessment and choice creation of these hoodlum occurrences is excessively hard for the authorities. Distinguishing this an overwhelming issue this paper focuses on fostering a solution for the choice creation of wrongdoing that is perpetrated. AI is the part of science the spot PC frameworks decide aside from human mediation. In current occasions Machine Learning is being utilized in excess of a couple of spaces one of the instances of such cases is mechanized or self-driving

vehicles. By Machine Learning calculations there is a way the spot we can anticipate sure results dependent absolutely on our bits of feedbacks offered and outfit a response to fixing wrongdoing examples in India. The two successive sorts of forecast techniques are characterization and relapse. This wrongdoing insights forecast is a region the spot grouping is applied. Arrangement is an administered expectation approach and it has been utilized in a scope of areas like determining stock, restorative region, etc. The most significant goal of this paper is to ponder on consideration a few calculations which can be utilized to foresee and examine the wrongdoing realities and improve the exactness of these designs through data handling to procure higher outcomes. The reason is to train the expected life sized model to anticipate the records the utilization of the instructing measurements

set through approval of the check realities set [3]. The designs which are being utilized here are Logistic Regression, Decision Tree order, Random Forest arrangement.

## 2. REALATED WORK

McClendon, Lawrence, and NatarajanMeghanathan. "Utilizing PC getting to know calculations to investigate wrongdoing information." *Machine Learning and Applications: An International Journal (MLAIJ) 2.1*

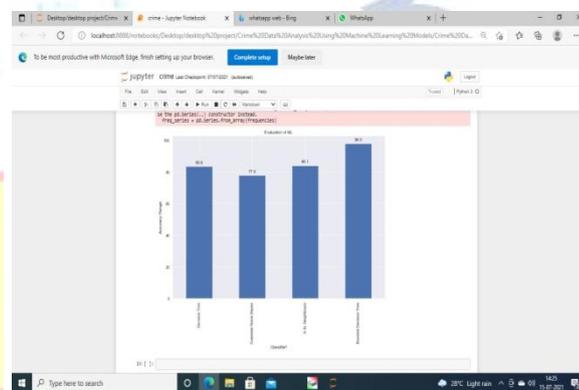
Information mining and work area examining have arise as a basic period of wrongdoing identification and anticipation. In this examination, we use WEKA, an open stockpile records mining programming, to conduct a near learn about between the vicious wrongdoing designs from the Communities and Crime Unnormalized Dataset outfitted through the University of California-Irvine storehouse and veritable wrongdoing measurable measurements for the nation of Mississippi that has been outfitted through neighborhoodscout.com. We applied the Linear Regression, Additive Regression, and Decision Stump calculations the utilization of the equivalent limited set of elements, on the Communities and Crime Dataset. Generally, the straight relapse calculation completed the quality among the three picked calculations. The extent of this task is to show the way that great and right the PC acquiring information on calculations utilized in data mining assessment can be at anticipating brutal wrongdoing designs.

## 3. PROPOSED WORK

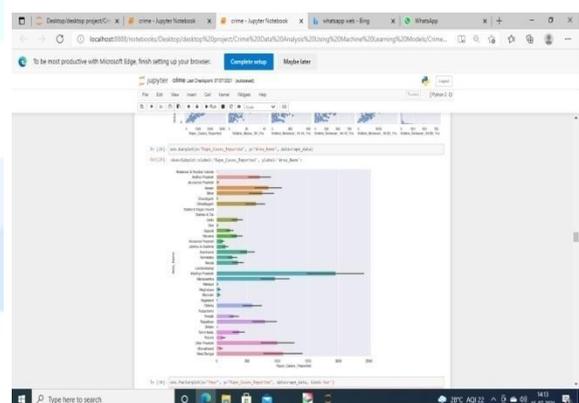
This device is based on the query work that is accomplished by browsing through various types of documentation. Essentially, each of the wrongdoings is anticipated based on the area and the kind of infractions that occur in these areas.[5] Linear Regression, Decision Tree, and Random Forest are used in this paper to predict wrongdoings since they tend to deliver sure exactness when compared to previous works. The data used in this article was obtained from data.world.com. The set of measurements comprises distinct types of wrongdoings committed in India in relation to the realm and a year separately. This paper accepts many types of wrongdoings as input and provides a space where infractions are dedicated as a result. Data cleansing, work determination, removing incorrect characteristics, and

records scaling by normalising and normalising are all part of the realities pre-handling process. Following reality pre-handling, the data is freed of incorrect qualities that could impair the precision of the life-sized model, and trademark decision is used to select only the necessary components that would not affect the model's exactness.[6] After pre-processing the data, the chosen designs, such as Logistic Regression, Decision Tree, and Random Forest, are trained by splitting the records into educated and investigated data. As a result of the need for a specific expense grouping style, it is used here. The forecasting of insights is done in Python.

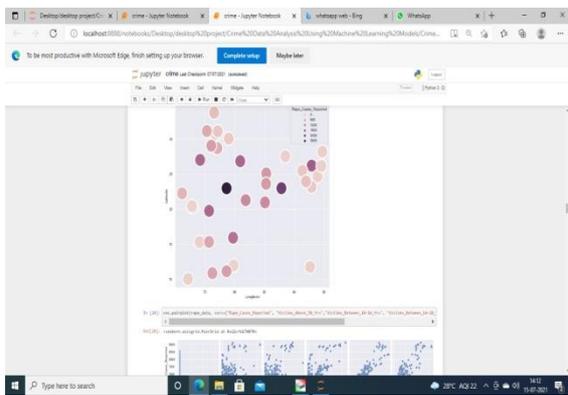
## 4. RESULTS



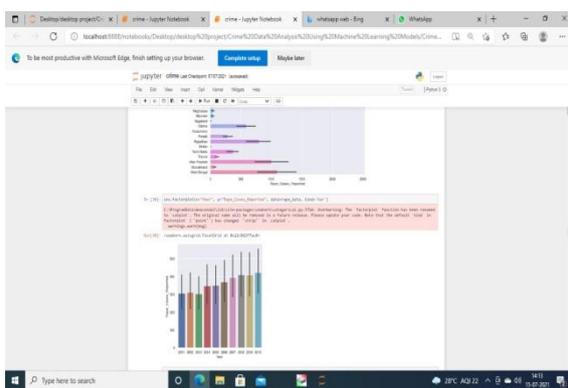
From the above the Boosted DT are giving better accuracy for prediction



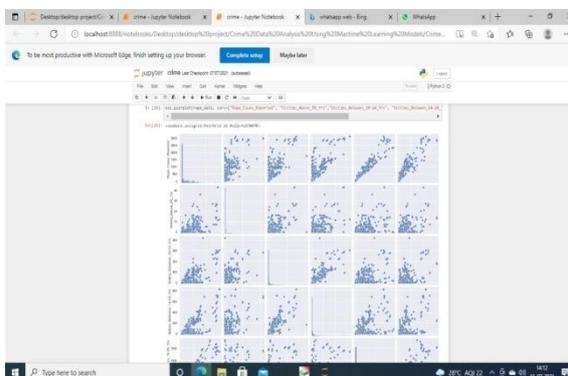
Crime rate in each state



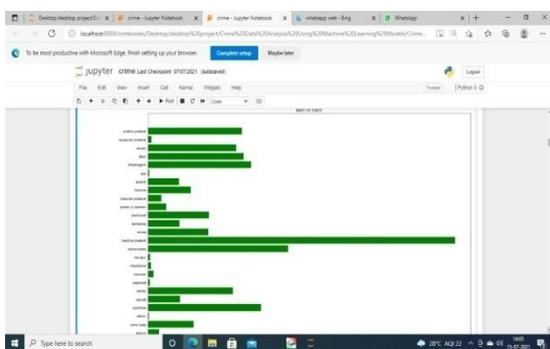
Crime prediction along latitude and longitude



Crime rate from 2001 - 2010



Statistical representation of crime vs age



Crime predicting across state vs state

## 5. CONCLUSION

Clearly, the most important indications of evildoer actions in a region include signs that will be used with the assistance of PC learning vendors to order a hooligan endeavour provided a location and date. When the training specialist encountered unequal classes in the dataset, it was able to solve the problem by oversampling and under-examining the data. This paper gives a wrongdoing records expectation by accepting such violations as enter and giving are in which these violations are committed as a result of the utilisation of Colab wallet with python as a centre language and python award worked in libraries, for example, Pandas and Numpy, through which the work will be completed faster and Scikit gives all of the methodology of how to utilise unmistakable libraries presenting through the python. Expectations have obvious consequences for unique calculations, and the precision of the Random Forest Classifier was found to be enticing, with an accuracy of above 95%.

### Conflict of interest statement

Authors declare that they do not have any conflict of interest.

### REFERENCES

- [1] McClendon, Lawrence, and Natarajan Meghanathan. "Utilizing AI calculations to analyze wrongdoing information." *Machine Learning and Applications: An International Journal (MLAIJ)* 2.1(2015): 1-12.
- [2] Alkesh Bharati, Dr Sarvanaguru RA. K, "Crime Prediction and Analysis Using Machine Learning" in *International Research Journal of Engineering and Technology (IRJET)*, Volume: 05 Issue: 09 | September 2018
- [3] McClendon, L., & Meghanathan, N. (2015). Utilizing AI calculations to break down crimedata. *AI and Applications: An International Journal (MLAIJ)*, 2(1), 1-12.
- [4] Chen, Hsinchun, et al. "Wrongdoing information mining: an overall structure and some examples." *PC 37.4* (2004): 50-56.
- [5] Parvathi, D. S. L., Leelavathi, N., Ravikumar, J. M. S. V., & Sujatha, B. (2020, July). Emotion Analysis Using Deep Learning. In *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 593-598). IEEE.
- [6] Kumar, J. R., Sujatha, B., & Leelavathi, N. (2021, February). Automatic Vehicle Number Plate Recognition System Using Machine Learning. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1074, No. 1, p. 012012). IOP Publishing."