Crime Insight: A Dynamic Crime Prediction and Analysis

Rakshitha G, MCA Student, PES Institute of Technology and Management, Shivamogga, Karnataka, India

Mrs. Tejaswini A, Assistant Professor, Dept. of MCA, PES Institute of Technology and Management, Shivamogga

ABSTRACT

Crime Insight is the particular circumstances of crime has turned into a critical problem that to be dealt with in an efficient manner, it makes a significant effect on people and national welfare. The prevalence of criminal behaviors spoils the peace of a country. Millions of crimes are being committed every day; these repeated crimes have made the life of general people restless. So, preventing the occurrence of the crime is a critical job. Therefore, these raises a paramount necessity of thoroughly analyzing crime patterns to effectively combat such predictive crimes. The role of crime analysis based on data sources from varied datasets, an open-source platform, to enable the prediction of the latest criminal events.

This research engages the application of Random Forest Algorithm here presents a strong framework for differentiating among different crime patterns and behaviors. In summary, this work strives to impart useful insights into modern crime dynamics during the formulation of proactive measures directed towards reducing criminal activities.

KEYWORDS- *Machine Learning, Crime, Prediction, Crime Patterns, Random Forest, Decision Tree Algorithm.*

I. INTRODUCTION

Crime is a serious issue in most urban centers and towns. It impacts the security and health of individuals. With all the crime statistics around, it is sometimes difficult for police and officials to keep up with patterns or prevent crimes from occurring. Crime Insight is an intelligent initiative that employs machine learning and data analysis to analyze historical crime data and forecast where and when crimes may occur in the future. This enables the police to better plan, act early, and make people safer. The system is also displayed in an easy-to-see and understand manner, making it both official-friendly and accessible to the public. With the increasing availability of digital crime records and advances in data science, there is a growing interest in using ML techniques to analyze. Crime prediction involves studying past forecast future incidents based on factors like location, time, day, and social conditions. Accurate crime forecasting informed decisions, allocate resources efficiently, and take preventive measures in high-risk areas. Traditional crime analysis methods are often limited in handling large and complex datasets. However, ML algorithms such as SVM, Decision Tree, Random Forest, Logistic Regression provide powerful tools to uncover hidden trends and accuracy of crime. This project aims crime enhance public safety.

Through frequency of crime was number of factors like criminals' intelligence, security of the place etc. the work has adopted the steps which utilized in data analysis, where the crucial steps like data gathering, data categorization, pattern finding, forecasting and visualization. Predictive analysis, aided by sophisticated machine learning algorithm, allows future instances of crime past trends and background factors. Predictive models may pick up high-risk locations and be deployed to forestall criminal activity. But crime analysis provides actionable recommendations by suggesting what can actually be done in light of the patterns identified and outcomes projected.

II. LITERATURE SURVEY

Different machine learning algorithms like Random Forest, Decision Tree, and many more to predict and analyze crimes. While certain methods have their concentration on crime detection. There have been a host of studies carried out aiming to tackle the increasing problem of crime, each suggesting various ways through which more accurate predictions can be made. The accuracy and success of these models are greatly based on the source of data used, i.e., whether historical crime data or demographic data, and the features or attributes chosen for prediction. Selecting the appropriate data and pertinent attributes is crucial when creating efficient and credible crime forecasting systems.

[1] "An Exploration of Crime Prediction Using Data Mining on Open Data" Ginger Saltos and Minhaela Cocea (2017).

Crime information is upgrade by gathering and data analytics gave rise to research methods that attempt to learn from crime records to gain insights into criminal behavior and, ultimately, prevent crimes. Few of these methods focus on crime forecasting models, although most of them use association rule mining and clustering algorithms. Here we explore forecasting models for the number of anti-social behaviors by Lower Layer Super Output Areas code, a UK police administrative area system. Three algorithms from three separate groups of methods—decision trees, regression, and instance-based learning—are applied.

[2] "Survey paper on crime prediction using ensemble approach" Ayisheshim Almaw, Kalyani kadam (2018).

Crime has ever been a priority concern for persons, society, and governments. This study discusses different ensemble learning techniques widely applied crime reports. This research emphasizes distinct methods used for criminal datasets identifying patterns and making future crimes predictions. Crime prediction seeks to identify and assist in preventing crimes through past data to anticipate when and where predict crime occur—depending on year, time, day, season, location. Yet, with the speed and uncertainty with which crimes are increasing nowadays, it is still a major problem to attain high precision in predicting crimes. This makes the construction of better prediction models an essential research area.

[3] "Systematic Literature Review of Crime Prediction" Falade Adesola and Ambrose Azeta (2019).

Varying methods have to be employed for the different methods of information that report illicit behavior in crime datasets. In their crime feedback forecast initiatives, Falade et al. (2019) demonstrate how various machine learning methods have been applied to multiple datasets such as police reports, social media, the news, and criminal records. publishers underscore various tasks and problems that associated with every crime method in dataset, including the society are seldom organized and reality of FIRs are reliable but unstructured.

[4] "Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques" Wajiha Safat, Sohail Asghar, Saira Andleeb Gillani (IEEE-2021).

Crimes and offenses have to be governed since they directly threaten public safety and justice. Precise forecasting of crimes and future crime patterns through computational techniques can significantly enhance urban security and law enforcement operations. Still, early crime prediction is usually hindered by the inability of human beings to handle immense and complex data efficiently. With increasing availability of crime records from the past, there exist numerous opportunities—and challenges—of detecting crime rates, crime types, and crime hotspots through machine learning. Even though research continues, there remains prediction models that can direct police patrols and prevent crimes in a timely and cost-effective manner.

III. PROPOSED METHODOLOGY

The methodology of the Crime Insight system proposed here consists of a number of major steps for making accurate predictions, criminal activity through machine learnings useful. To do this, crime data like historical crime records, locations, dates, and demographic data is gathered from dependable sources. This data is next preprocessed by cleaning, eliminating duplicates, dealing with missing values, and formatting it to be used in machine learning. Once preprocessed, divided datasets into test and training sets. Feature selection is done to select specific context and what to measure relevant in the prediction of crime. Different classical ML algorithms like Logistic Regression, SVM, Random Forest, Decision Tree are utilized to model training using measures such as Accuracy, Precision, Recall, and F-beta Score in order to identify the top performer. Lastly, the chosen model is combined with a comfortable user interface that enables law officials to enter information and get predictions or crime trend analysis so they can assist in making decisions preventing crimes.

Proposed Model Diagram

Advantages of existing system were discussed many sections. Researchers constructed hybrid domain systems, which seek to inherit strengths and avoid weaknesses, by fusing both methods for enhancing domain outcomes. In general terms, system recommender's hybrid combines multiple methods of recommendation with the aim of maximizing synergy between them. Our activity will mainly be the integration of CF and CBF approaches, although there are numerous practical suggestion methods which may be integrated.

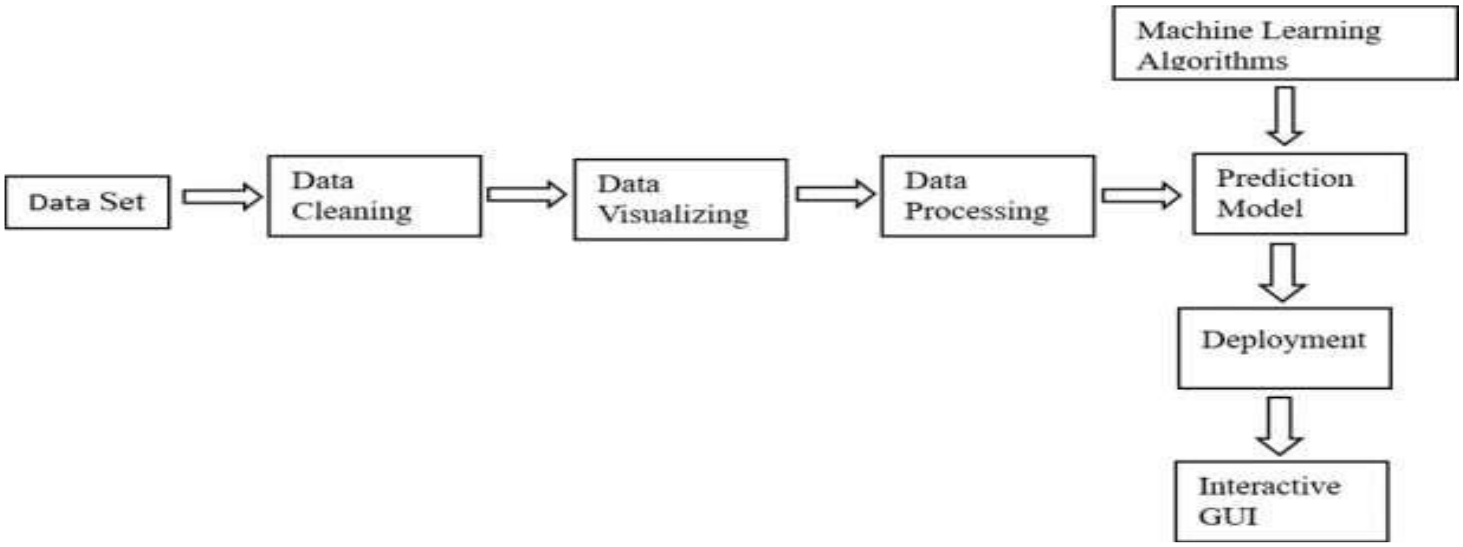


FIGURE 1.1: Proposed Model Diagram for Crime Prediction

Block Diagram of ML Model

A block diagram is an easy and efficient method to illustrate system operates basic graphical components. It graphically illustrates the path of information within the system by indicating the input information entering into the system, the different processes or operations carried out on such information, and the resulting output generated. Block diagram are essential tools in system modeling because it makes it easy to comprehend structure and behavior of the system. The elements displayed in a block diagram typically comprise the processes of the system, how information moves from one part to another. Through the application of a block diagram, intricate systems are decomposed into easier components, which are simpler to analyze, design, and describe.

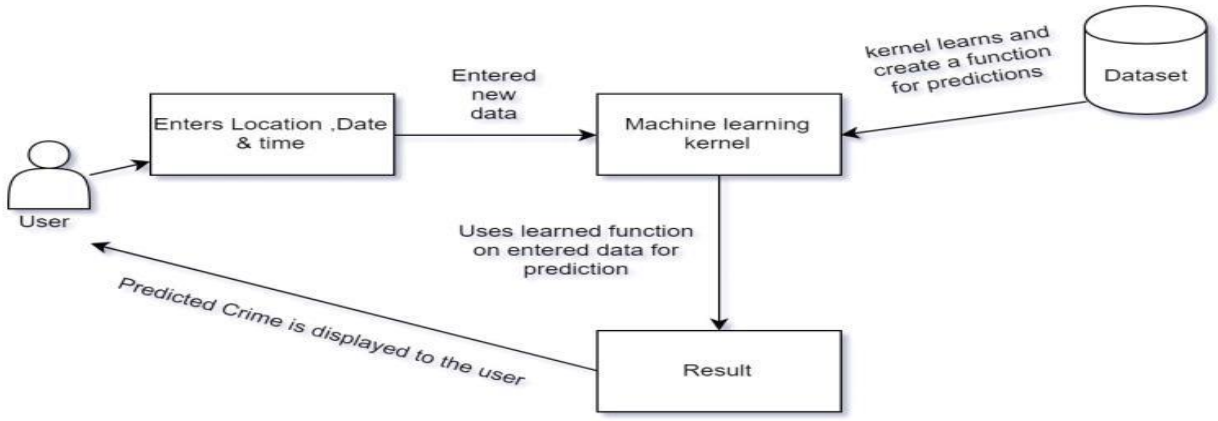


FIGURE 1.2: Block Diagram for Crime Prediction

Crime prediction process & datasets

The crime database is gathered directly from field work. Based on this data, crucial information such as the year, age, name of crime type, location of crime, date of offence, name of crime group, and name of unit are selected. This selected information is applied as input features to the system.

Year	Age	Crime Areas	Crime Type
2016	14	Huvinahalli	POCSO
2016	16	Muradi	Theft
2016	6	Bagalkot	Murder
2024	11	Badami	POCSO
2016	13	Girisagar	Theft
2017	5	Hunagund	POCSO
2018	16	Banahatti	Murder
2019	16	Arakeri	Theft
2016	18	Gangur	Theft
2017	35	Aminagad	Rape
2018	22	Hunagund	Crimes Related to Women
2017	48	Arakeri	Vehicle Accidents
2018	18	Hubli	Cheating
2016	30	Bagalkot	Forgery

TABLE 1.1: A Simple Dataset employed in Crime Predicting

Machine learning-based crime prediction entails a number of significant steps. The initial step is data accumulation, where pertinent data like crime records, demographics, and even climatic patterns is accumulated. Data preparation is the subsequent step, involving cleaning, categorizing, and information in structure to render it model training-perfect. Following preprocessing, model is trained and tested against the testing set. Feature engineering is a critical step in which relevant variables (features) are chosen or derived dataset. These features enable patterns and make good predictions. These methods are important to building an effective and trustworthy crime prediction system.

Offence from date	Crime group name	Unit name
2016-02-15 10:00:00	POCSO	Amengad PS
2016-03-28 11:45:00	Election	Amengad PS
2016-11-10 20:00:00	Motor Vehicle Accidents	Amengad PS
2015-01-20 08:00:00	Theft	Amengad PS
2015-12-15 21:30:00	Attempt to Murder	Amengad PS
2014-02-11 10:00:00	CrPC	Amengad PS
2016-07-25 09:12:00	Riots	Amengad PS
2017-10-08 29:00:00	Burglary-Night	Amengad PS
2017-05-12 04:45:00	Motor Vehicle Accidents	Amengad PS
2017-02-28 24:37:00	POCSO	Amengad PS
2017-10-26 05:30:00	Theft	Amengad PS
2018-04-15 02:00:00	Murder	Amengad PS
2018-08-08 10:37:00	CrPC	Amengad PS
2017-06-18 27:30:00	Election	Amengad PS

TABLE 1.2: Crime Prediction of Datasets

Crime prediction based on ML Models

The classic ML algorithm has been reliable tools for crime prediction based on past data and pattern recognition. Such classic algorithms- Decision Trees, SVM, are frequently analyses the crime patterns and the prediction of criminal activity. Such classic approaches are less demanding in terms of data requirements and easier to interpret and understand compared to ML models. For example, prediction of logistic regression in many types of crime happening, given characteristics like location, day of the week, and demographic characteristics in population. Beyond prediction, outlier analysis and anomaly detection, which can assist in identifying suspicious patterns in the data. Identification of such anomalies can point towards suspicious patterns, allowing for prompt preventive measures by the authorities. Overall, classic machine learning offers a real-world and interpretable model of crime prediction and analysis.

• Machine learning regression techniques for crime prediction

Regression techniques are utilized in predicting several criminal detection situations. Researchers primarily examined typical crimes such as loss of property, motorbike robberies, and crime in cities. There could be many reasons for the increase in motorbike robberies. For example, population density and growth, transportation modes, do they use a bike, etc. These situations render it hard for the police to monitor and carry out regular screenings as they require precise prediction of robbery likelihood.

• Machine learning classification techniques for crime prediction

Traditional regression techniques more reliable for crime prediction, but they can efficiently measure the efficiency of the variables. ML models in predicting crimes has been established by authors. They may be more efficient at ascertaining what factors are essential in forecasting crime. We also researched classification techniques, like analyzing criminal reports, to predict different criminal events.

IV. Mathematical Formulas

To evaluate how well our classification models work in predicting crime, we use four main metrics: Accuracy, Precision, Recall, and the F-beta Score. These help us understand the model can work from different angles.

- Precision (q): Precision tells us the predicted positive cases were how actually correct.

$$\text{Precision (q)} = \frac{tv}{tv+fv}$$

Where, tv = True Positives (correctly predicted crimes), fv = False Positives (wrongly predicted as crimes)

- Recall (rc): Recall measures actual positive cases (real crimes) were how many models can correctly identify.

$$\text{Recall (rc)} = \frac{tv}{tv+fnv}$$

Where, fnv = False Negatives (missed crimes)

- Accuracy: Accuracy tells how many total predictions (both crime and non-crime) were correct.

$$\text{Accuracy} = \frac{tv+tnv}{tv+fv+fnv+tnv}$$

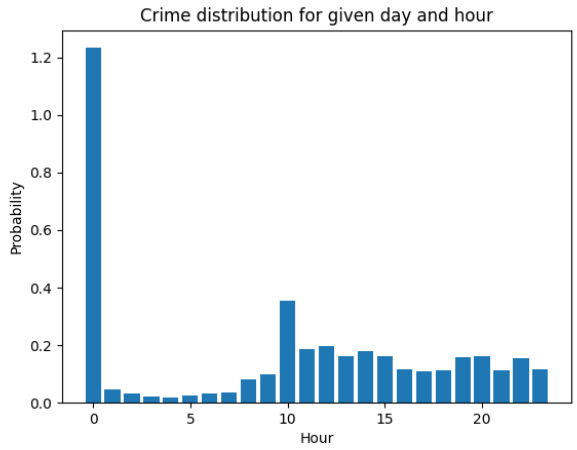
Where, tnv = True Negatives (correctly predicted non-crimes)

- F-beta Score: F-beta Score gives a better overall picture by combining precision and recall. their more reliable than accuracy, respectively the data is imbalanced.

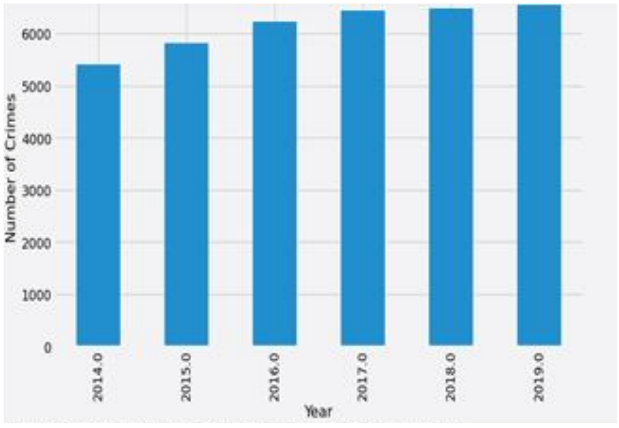
$$\text{F-beta Score} = \frac{2*(rc*q)}{rc+q}$$

This score helps measure how well the model balances between correctly identifying crimes (recall) and avoiding false alarms (precision).

V. Graphs



GRAPH 1.1: Distribution of Crime Analysis



GRAPH 1.2: Distribution of Number of Crimes in Year

GRAPH 1.3: Distribution of Research publication in crime prediction trends in year

From the graph drawn on historical records, it is obvious that peak crime. The information drawn here is taken from historical records. The graph varies with time when we incorporate more data into the records. Graph 1.1 and 1.2 illustrates crime in locations such as airport, temples, bus station, railway stations, bank, casino, jewellery shops, bar, ATM, highways etc. On x axis the year and hour during which crime occurred played an important role in graph plotted whereas in y axis crime probability is plotted. Graph 1.3 indicates the trend in research publication on crime prediction in distribution of crime between 2018 and 2022 (i.e., the past five years). It indicated increasing trends between 2018 and 2021 and decreasing trends in 2022. The years 2020 and 2021 are the peaks of COVID-19, which might current researchers have taken advantage in investigate more dynamic crime research.

VI. EXPERIMENTAL RESULT

While testing the Crime Insight system, we employed a dataset of historical crime records with attributes such as location, time, crime type, and other. These data had been cleaned and pre-processed prior to application in ML models training and testing. Various algorithms like SVM, Decision Tree, Random Forest were implemented to compare their accuracy. Among all the models that were experimented, Random Forest performed best with the highest accuracy and improved performance in predicting crime type and location. The system was successful in identifying high-risk zones and data trends accurately. Accuracy scores, confusion matrices, and classification reports were employed to assess the models. These results indicate that the system can provide accurate predictions and assist decision-making for crime prevention.

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**Training Time (s)**
Decision Tree	83.2%	81.5%	80.7%	81.1%	1.5
Random Forest	89.6%	88.3%	87.5%	87.9%	2.3
Logistic Regression	78.4%	75.9%	76.3%	76.1%	1.2
Support Vector Machine	85.7%	84.2%	83.9%	84.0%	3.1
K-Nearest Neighbours	80.5%	79.0%	78.1%	78.5%	0.9

TABLE 2.1: Performance Comparison of Crime Using ML Techniques

VII. CONCLUSION

As technology has evolved, crime has evolved and become more sophisticated, giving law enforcers difficult challenges to deal with. With the emphasis on learning sequences and movements in crime incidents, the use of ML to predict crime by researchers has increased in the last few years. This report scrutinizes more than 150 articles to review the various machine learning approaches employed in crime prediction. Crime Insight is a powerful tool that is meant to assist in forecasting and analyzing crime through data analysis. The tool assists police officers and the public to learn about crime patterns and take preventive action. Through examination of historic crime data, it can give insightful information on where and when crimes are most likely to occur.

The system enhances public safety through enhanced planning and quick response. With the incorporation of future upgrades such as real-time information, mobile functionality, and sophisticated algorithms, Crime Insight used in future that enhanced. Overall, intelligent move in creating safer and more secure communities through the leveraging of technology. Thus, our findings contribute to understanding the implications of various machine learning methodologies. The ongoing research community will also benefit from our suggested datasets and areas for future work as they pursue their research into the area of crime prediction.

VIII. FUTURE ENHANCEMENTS

In the future, Crime Insight can be enhanced in various ways that we can use in intelligent. It can be linked with live crime statistics from the police and news websites to provide real-time information and improved forecasts. Advanced algorithm is used in machine learning that employed by the system to improve accuracy. A map function can be included that displays crime-ridden areas in the form of heatmaps. A mobile application can be designed so that users and police can easily access it at any time. Individuals can also be permitted to report crimes directly via the system. Social media can be scanned for early indications of crime or issues. The system can speak in various languages and even include a voice assistant to assist users. Warnings can be sent to individuals in risky areas. Providing additional security, feedback mechanisms, and coupling it with CCTV or sensors can aid in monitoring in real-time and swift action. These improvements will make Crime Insight more effective and beneficial for the public as well as law enforcement.

Through performance of these models to real circumstances, law enforcement agencies, legislators, and city planners are able to derive pragmatic insights towards optimizing resources, focused intervention, and anticipatory crime prevention. Agencies are able to design evidence-driven policies for enhancing public security and reducing criminality by mapping crime hotspots, understanding latent socioeconomic determinants leading to crime, and forecasting future crime patterns.

REFERENCES

[1] S. Kim, P. Joshi, P. S. Kalsi, and P. Taheri, ‘‘Crime analysis through machine learning,’’ in *Proc. IEEE 9th Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON)*, Nov. 2018.

[2] “An Exploration of Crime Prediction Using Data Mining on Open Data” Ginger Saltos and Minhaela Cocea (2018).

[3] “Survey paper on crime prediction using ensemble approach” Ayisheshim Almaw, Kalyani kadam (2018).

[4] L. Elluri, V. Mandalapu, and N. Roy, ‘‘Developing machine learning based predictive models for smart policing,’’ in *Proc. IEEE Int. Conf. Smart Compute. (SMARTCOMP)*, Jun. 2019.

[5] A. Meijer and M. Wessels, ‘‘Predictive policing: Review of benefits and drawbacks,’’ *Int. J. Public Admin.*, vol. 42, no. 12, pp. 1031–1039, Sep. 2019.

[6] “Systematic Literature Review of Crime Prediction” Falade Adesola and Ambrose Azeta (2019).

[7] S. A. Chun, V. A. Paturu, S. Yuan, R. Pathak, V. Atluri, and N. R. Adam, ‘‘Crime prediction model using deep neural networks,’’ in *Proc. 20th Annu. Int. Conf. Digit. Government Res.*, Jun. 2019.

[8] S. Hossain, A. Abtahee, I. Kashem, M. M. Hoque, and I. H. Sarker, ‘‘Crime prediction using spatio-temporal data,’’ in *Computing Science, Communication and Security*. Gujarat, India: Springer, 2020.

[9] O. Kounadi, A. Ristea, A. Araujo, and M. Leitner, ‘‘A systematic review on spatial crime forecasting,’’ *Crime Sci.*, vol. 9, pp. 1–22, Dec. 2020.

[10] P. Tamilarasi and R. U. Rani, ‘‘Diagnosis of crime rate against women using K-fold cross validation through machine learning,’’ in *Proc. 4th Int. Conf. Compute. Methodologies Commun. (ICCMC)*, Mar. 2020.

[11] A. R. C. da Silva, I. C. D. P. Junior, T. L. C. da Silva, J. A. F. de Macedo, and W. C. P. Silva, ‘‘Prediction of crime location in a Brazilian city using regression techniques,’’ in *Proc. IEEE 32nd Int. Conf. Tools Artif. Intell. (ICTAI)*, Nov. 2020.

[12] N. Shah, N. Bhagat, and M. Shah, ‘‘Crime forecasting: A machine learning and computer vision approach to crime prediction and prevention,’’ *Vis. Compute. Ind., Biomed., Art*, vol. 4, no. 1, pp. 1–14, Apr. 2021.

[13] S. S. Kshatri, D. Singh, B. Narain, S. Bhatia, M. T. Quasim, and G. R. Sinha, ‘‘An empirical analysis of machine learning algorithms for crime prediction using stacked generalization: An ensemble approach,’’ *IEEE Access*, vol. 9, pp. 67488–67500, 2021.

[14] C. Janiesch, P. Zschech, and K. Heinrich, ‘‘Machine learning and deep learning,’’ *Electron. Mark.*, vol. 31, no. 3, pp. 685–695, Apr. 2021.

[15] W. Safat, S. Asghar, and S. A. Gillani, ‘‘Empirical analysis for crime prediction and forecasting using machine learning and deep learning techniques,’’ *IEEE Access*, vol. 9, pp. 70080–70094, 2021.

[16] “Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques” Wajiha Safat, Sohail Asghar, Saira Andleeb Gillani (IEEE-2021).

[17] D. M. Raza and D. B. Victor, ‘‘Data mining and region prediction based on crime using random forest,’’ in *Proc. Int. Conf. Artif. Intell. Smart Syst. (ICAIS)*, Mar. 2021.

[18] M. Saraiva, I. Matijosaitiene, S. Mishra, and A. Amante, ‘‘Crime prediction and monitoring in Porto, Portugal, using machine learning, spatial and text analytics,’’ *ISPRS Int. J. Geo-Inf.*, vol. 11, no. 7, p. 400, Jul. 2022.

[19] M. L. Trinhammer, A. C. H. Merrild, J. F. Lotz, and G. Makransky, ‘‘Predicting crime during or after psychiatric care: Evaluating machine learning for risk assessment using the Danish patient registries,’’ *J. Psychiatric Res.*, vol. 152, pp. 194–200, Aug. 2022.

[20] A. K. Das and P. Das, ‘‘Graph based ensemble classification for crime report prediction,’’ *Appl. Soft Compute.*, vol. 125, Aug. 2022.