Crime Pattern Analysis and Prediction Using Machine Learning

Pradeep Nazareth
Department of Artificial Intelligence and Machine Learning
Alva's Institute of Engineering and Technology
Mangalore, India
pradeepn@aiet.org.in

Lathesh Kumar S R
Department of Artificial Intelligence and Machine Learning
Alva's Institute of Engineering and Technology
Mangalore, India
latheshkumar06@gmail.com

Sanket Patil
Department of Artificial Intelligence and Machine Learning
Alva's Institute of Engineering and Technology
Mangalore, India
sanketpatilsp360@gmail.com

Shivamani M Nayak
Department of Artificial Intelligence and Machine Learning
Alva's Institute of Engineering and Technology
Mangalore, India
shivamaninayak5757@gmail.com

Tejashwini Shailendhra Murdeshwar
Department of Artificial Intelligence and Machine Learning
Alva's Institute of Engineering and Technology
Mangalore, India
murdeshwartejashwini@gmail.com

***Abstract*—Urban crime rate changes that are influenced by tiny spatiotemporal factors are typically outside the realm of a usual police investigation. This paper reflects the authors' hierarchical machine learning workflow to understand crime trends through the "Crime in India" open dataset from 2001 to 2020. Our approach very detailed stages a preprocessing, moves the Synthetic Minority Over-sampling Technique (SMOTE) to solve the class imbalance problem, and has multilayer feature extraction to capture the spatial and temporal aspects so that the classifier's prediction quality can be brought up to standard again.The authors tried different algorithms to have a quantitative comparison of their performances and measured accuracy, precision, recall, and F1-score. The algorithms were K-Nearest Neighbours, Support Vector Machine, Decision Tree, Random Forest, and Naive Bayes. Random Forest was the most accurate classifier as it achieved the predictive accuracy of 89 % and being very stable, it varied almost evenly across all the evaluation metrics. In addition, the method leads to the generation of the interactive hotspot maps that show not only the locations with high crime rates but also repetitive temporal cycles.The research reveals that a quantitative crime prediction model facilitates the implementation of proactive policing as well as the delegation of the resources processes. The article states that data being properly prepared, training on balanced datasets, and model interpretability as the major elements for the development of reliable crime analytics systems for smart-city governance are acknowledged by them.**

***Index Terms*—Crime Prediction, Machine Learning, Spatiotemporal Analysis, Data Imbalance, Ensemble Learning, Smart Policing, ST-CrimeNet**

I. Introduction

The rapid urban expansion, socioeconomic inequalities, and changing demographic trends have caused modern cities to deal with various complex crimes of different kinds. Bigger cities, in general, have higher rates of crimes than smaller ones, and the different types of urban crimes make it difficult for law enforcement to use their usual methods to find patterns. While past crime data is rich in spatiotemporal details, traditional methods of analysis have difficulty identifying nonlinear relationships and hidden deeper structures in large crime datasets. Consequently, law enforcement agencies have to turn to more sophisticated and faster methods.

Advances in computational intelligence have led to a crime forecasting system of the future, locality patrol, behavioral pattern, and crime mapping algorithms. Research after research, scholars have illustrated the effectiveness of supervised and unsupervised methods for crime localization and prediction of future incidents [9], [10]. Yet, the deployment of these methods in practice is a minefield of obstacles which, for the most part, shake the trustworthiness of such models to a great extent. The extreme imbalance of classes between, e.g., ordinary crimes of theft and unusual criminal activities, which results in predictive models being very highly biased toward the majority classes and having a limited capacity for generalization, is one of the most obvious problems [1], [5].

Furthermore, the spatiotemporal factor of criminal events, which is indistinguishable from the phenomena, makes the problem even more difficult because the happening of crimes depends on the location but also the factors of the neighborhood, season, and time. To correctly model these relations one should have feature representations that do not only consider the categorical attributes but also go further to incorporate the environmental context [13], [20], [21]. Moreover, the question of the model's interpretability for humans has become a critical point, particularly since police, as the main users of the predicted results, are the ones that are most concerned with transparency and being accountable when making decisions [16].

This study introduces a machine learning framework that systematically analyzes criminal patterns. The study utilizes a dataset called "Crime in India" which is openly available and covers the time period from 2001 to 2020. The authors of the paper initially direct their attention to exhaustive data preprocessing in their method. Furthermore, they apply the Synthetic Minority Over-sampling Technique (SMOTE) to reduce the class distribution skewness [5] and also merge the enhanced spatiotemporal feature engineering to increase the model discrimination capability. To obtain the best performance in criminal category classification and hotspot tendency identification, we have compared five famous classification algorithms such as K-Nearest Neighbours, Support Vector Machine, Decision Tree, Random Forest, and Naive Bayes. The core goal is to demonstrate the effect of carefully planned data preparation, balanced learning distributions, and strategically engineered features on classification performance and model interpretability.

This paper is outlined with a comprehensive survey of the previous studies, the description of the dataset, and the methodological framework employed. Afterward, it presents the results of the experiment and discusses them. The research finally ends with an emphasis on how data-driven methods might be employed as a powerful tool to deliver proactive public safety in smart-city settings.

II. Literature Review

The main component of computational criminology is machine learning usage in crime analytics. Many studies have been performed on labeling, hotspot identification, and trend forecasting. The first experiments have demonstrated that simple algorithms such as K-Nearest Neighbour, Naive Bayes, and Decision Trees can give quite good prediction results when time- and space-related variables are combined with crime data [9], [10]. Besides that, Support Vector Machines can be used to help in clarifying complex crime categories and differentiating safe areas from dangerous ones since they can detect non-linear relationships in the feature space. [10].

Ensemble learning methods, especially the use of Random Forest, is one of the most acknowledged techniques in this regard, which has been renowned for its strong predictive accuracy and interpretability, therefore, being a good instrument both for crime classification and for getting insight into the influential variables [2], [20]. By going beyond supervised methods, one can discover unsupervised methods like K-Means clustering which is a tool for finding spatial aggregations and crime hotspots. At the same time, association rule mining reveals the repeating patterns of linking the offense types, locations, and time periods. [13]. In addition to this, some people have also taken regression-based analyses as a tool, not only for forecasting crime trends over time, but also for explaining demographic and socioeconomic correlates of crime [21].

Along with these methodological enhancements, the field is still facing numerous obstacles. The most frequent issue is that imbalanced crime datasets are used without the taking of necessary corrective measures, thus the models that result are heavily biased towards the most frequent crime classes and cannot perform well on minority categories [1]. What is more, the spatial-temporal context in feature engineering is not being utilized quickly enough and, until now, geospatial clustering analyses have hardly been used, which restricts the capacity to model local criminal dynamics [13]. Conversely, the ethical problems like data privacy, algorithmic bias, and the social effects of predictive policing that have been recognized by very few studies only, are also present [16], [17].

To sum up, the literature to date on the use of machine learning in crime analysis has acknowledged the effectiveness of these methods, but it has also pinpointed various obstacles. These obstacles comprise the requirement of correct data preprocessing, the skilful handling of class imbalance, and the creation of models that can be easily understood and are ethical. The resolution of these problems is, therefore, the main reason for the existence of the staged analytical pipeline which is put forward in this paper.

III. Methodology

ThiS four-stage structured analytical workflow is sequential and highly integrated: it comprises data preparation, feature engineering, model implementation, and evaluation. These stages were designed as a single unit of work to result in models which are more stable and which can signal significant spatiotemporal changes of crime.

A. Data Description and Preprocessing

The study employs the open-source dataset "Crime in India" which is available on Kaggle and details crimes that have taken place across the whole country from 2001 to 2020 (Fig. 4). The data file features nearly 1.2 million rows that explain the crime categories, time, and location identifiers, along with some demographic variables. The initial preprocessing aimed at more uniform feature ranges also involved the operations of filling missing values, encoding categorical variables, and normalizing numerical attributes by Min–Max scaling. The process of outlier removal was done through the Interquartile Range (IQR) method to reduce their influence on the learning stage.A drastic imbalance between classes where common crimes were several times more than rare ones was among the most apparent problems of the dataset. To soften this blow, the Synthetic Minority Oversampling Technique (SMOTE) was implemented to generate artificial minority samples and thus improve classifier generalization. The cleaned and balanced data served as the point of departure for modeling.

B. Feature Engineering and Selection

Feature engineering in this study focused on extracting richer and more informative spatial and temporal attributes. During exploratory analysis, redundant or weakly correlated variables were removed, while key features such as hour, day, district, and crime type were retained because prior research has already shown their strong relevance to crime prediction. Additional engineered features, including temporal aggregates, time-of-day groupings, and zonal crime-density indicators, were also created to help uncover deeper behavioral patterns within the data.

For records containing location coordinates, local areas were derived from K-Means clustering regions, depicting the geographical zones of crime hotpots based on the concentration of the crime in a particular area. This is consistent with the methods employed in the studies referenced in [10], [13]. The new variables that have been created not only gave more background information but also improved the models' ability to uncover hidden spatiotemporal patterns.

A. Machine Learning Algorithms

To select the best one for use in multiclass crime classification five different supervised learning algorithms were experimented with. To detect local patterns of similarity K-Nearest Neighbours was used along with the Euclidean distance metric:

$$d(x, x_i) = \sqrt{\sum_{j=1}^{n} (x_j - x_{ij})^2} \tag{1}$$

Though KNN is a good method for neighborhood-based reasoning, it is susceptible to noise and high dimensionality [10]. Support Vector Machines with a radial basis function (RBF) kernel were utilized to create nonlinear class boundaries in the crime data, in compliance with the methods referred to in the previous comparative studies [9].

Decision Trees were used to facilitate the understanding of the model's decisions since they can show the rule-based classification paths, however, they are inherently vulnerable to overfitting if not properly regularized [12]. Random Forest, operating as an ensemble of decision trees,

$$\hat{y} = \mathrm{mode}\{h_1(x),\, h_2(x),\, \ldots,\, h_T(x)\} \tag{2}$$

could bring about increased stability and predictive power as a result of the ensemble technique, hence, it was in agreement with the findings of similar crime prediction works [2]. Naive Bayes served as a quick probabilistic classifier that is based on Bayes' theorem. It can be represented mathematically as:

$$P(C_k \mid x) = \frac{P(C_k) \prod_{i=1}^{n} P(x_i \mid C_k)}{P(x)} \tag{3}$$

$$\hat{C} = \arg\max_{C_k} \; P(C_k) \prod_{i=1}^{n} P(x_i \mid C_k) \tag{4}$$

The model was capable of producing good results for categorical variables.

B. Experimental Setup and Evaluation

All experiments were conducted in Python 3.11 with Jupyter Notebook and the standard data science libraries such as Scikit-learn [4], pandas, NumPy, matplotlib, seaborn, plotly, and Folium. The data for the research was divided into a 70–30 train–test split, and 5-fold cross-validation was implemented to allow for unbiased performance estimation. The tuning of hyperparameters was done by GridSearchCV.

The evaluation primarily relied on accuracy, precision, recall, and F1-score, which together help assess both the overall performance of the model and its ability to correctly identify minority classes. In addition, confusion matrices, classification reports, and interactive hotspot maps were generated to visually present the model's behavior and to evaluate how practical the system is for real-world crime analysis applications.

IV. Results and Discussion

The research systematically evaluated five machine learning model K-Nearest Neighbour (KNN), Support Vector Machine (SVM), Decision Tree, Random Forest, and Naive Bayes that were experimentally deployed on the preprocessed crime dataset.

The experiment uses a 70:30 train-test split with 5-fold cross-validation to obtain the performance results. Accuracy, precision, recall, and F1-score were four different measures used to evaluate the models. The measures used enable global prediction capability to be combined with discrimination at the class level which was especially important since the crimes were distributed in an imbalanced manner from the very beginning.

The Random Forest classifier was the best with a maximum accuracy of 89%, thus reflecting a wide range of generalization of different crime categories. The Decision Tree classifier had an accuracy of 85%, thus demonstrating its capability of deriving transparent rule-based structures from the spatiotemporal attributes. The SVM classifier yielded 81% accuracy where non-linear feature boundaries could be modeled effectively by its RBF kernel, thus being in line with earlier research on crime prediction that is characterized by high-dimensional and non-linear feature interactions [9]. KNN was able to reach 78% accuracy and thereby perform reasonably well in local neighborhood structures while showing reduced effectiveness in inter-class boundary areas. Naive Bayes was at 74% accuracy the worst performer of the five but still remained favorable in terms of computational efficiency because of its probabilistic formulation.

Random Forest's winning move is in line with the ensemble learning, where the combination of decorrelated trees does not only lessen variance but also increase the predictive power of the resulting model in classification tasks with noisy or mixed-type features [2].

Further feature importance investigation through the model revealed that both temporal variables such as hour-of-day and day-of-week, and spatial variables like district-level crime frequency were the key predictors. Similar findings were reported in the literature of the spatiotemporal crime prediction, where the structuring through temporal cycles and geospatial context was claimed to largely determine the occurrence of crime patterns [10], [13].

The Decision Tree, which was less performant than the Random Forest, nevertheless obtained a great amount of value from the interpretability it offered. The decision rules derived from the tree were very straightforward in terms of criminological systems and were the concentration of repeated offenses in certain temporal intervals and geographic clusters, among other things, a finding that corroborates the environmental criminology frameworks [8]. The SVM kept up its performance level but was very resource-consuming during the model fitting stage. KNN depended heavily on the scaling of features and local neighborhood boundaries, thus it was less trustworthy in areas where classes were heavily overlapped. Naive Bayes was only a baseline for comparison and showed the expected performance limitations due to the model's strong independence assumptions [16].

Random Forest was essentially the main factor that brought the highest and most accurate prediction performance with a slight typical resistance to overfitting, and it also produced easy feature explanations that correspond to the real-world behavior.

The findings of this study constitute a perfect demonstration of the ways in which reliable and useful crime prediction systems can be built by means of proper data preprocessing, the correct use of spatial and temporal features, and ensemble-based modeling. Such systems, which can be very lightly used by the police and incorporated into modern urban safety apps, have a great potential.

V. Applications and Use Cases

The crime analysis system which was the major focus of this research is, in fact, the key operational device that is the basis of the solution development, which, eventually, gets applied in the real world, e.g., become very useful to police departments, public safety authorities, and urban planners. Essentially, the system is, therefore, doing a deep dive into the data to bring out these stark truths that are at the core of the whole process of evidence-based policing and strategic decision-making by the integration of spatiotemporal feature analysis with predictive modeling.

Most notably, the system's practical application in spotting crime hotspots is what comes first to mind. By combining geographical characteristics with temporal indicators the system is one of a kind in being able to realize not only the static but also the dynamic maps that show the areas of the highest and most persistent crime incidence with great accuracy (see Figure 3). Such representations of the hotspot have been validated through scientific methods in the previous research as a means of facilitating the efficient implementation of targeted surveillance and resource allocation whereby the police agencies can deploy personnel with great effectiveness [8], [10]. Apart from that, the spatial results that have been obtained can be utilized in an efficient manner to enhance patrolling activities, install surveillance devices, or simply give the green light for the intervention of localities with high crime rates yet have been left out.

The framework, in addition to the identification of the hotspots, also goes on to produce prediction models that describe the likely crime types at the given time. By an in-depth analysis of the previous occurrences, the system makes a forecast of the probability of certain crimes happening at particular hours or on specific days (see Fig. 4). The temporal factors of the predictions relate to the lifestyle patterns that have been confirmed and thoroughly investigated in the criminological research and at the same time, they can be very useful in the implementation of risk prevention strategies [13]. As an example, the anticipations of an upsurge in theft during the wee hours or a rise in the number of assault cases during the weekends enormously assist in the taking of preventive measures well ahead of time.

Along with the temporal predictions and identification of hotspots, the organization also creates decision-making reports that assist in the bigger picture of the decision-making process. These characteristics include the capabilities of finding out the neighborhoods that are most in need, the effectiveness of the anti-crime measures that have been put in place, and also providing guidance for community outreach programs. Neighborhood crime trend evaluations based on local data result in the implementation of less opaque and more accountable policing measures. Besides these facts, they point to a fresh investigation that necessitates providing the justice sector with user-friendly analytical tools [16]. If the management team members are heavily dependent on forecasts during a specific time, they will, eventually, find a way to come up with long-term plans for crime prevention and public safety ensuring.

Overall, the system presented serves as a strong example of how machine learning can be effectively combined with well-designed spatiotemporal modeling to generate meaningful intelligence. Such insights are highly valuable to law enforcement, supporting operational deployment, strategic decision-making, and community-focused policing efforts.

VI. Challenges and Limitations

The artificial intelligence framework mapped out in this article is essentially a commendable tool for the study of criminal patterns, nevertheless, it acknowledges some procedural and functional constraints. These constraints point to the problems which are not only crucial for the comprehension of the findings but also for the viability of the system's outdoor experimental use situations..

1) Inherent dataset limitations: The analytical framework is the principal reason locally that the bias description of crime reporting data skewness is that most of the crimes committed locally such as thefts are more than a hundred times than a serious but rare crime like homicide. The differences in distributions lower the predictive accuracy of the framework for the least frequent crime categories, which is recognized as an issue in the field of imbalanced learning scenarios [1], [6]. In addition to this, the changes in the wayseparated by distance and

thus the model's capacity of handling the heterogeneous geographic contexts is decreasing.

Computational and infrastructural requirements:

The existing version is only capable of batch processing and is limited to data from the past. A person intending to improve the system for real-time predictions would need to connect it with live data streams. To give an example, these could be the records of the fire brigade, police reporting systems, and other information streams that are continuously updating. Such a step requires a considerable amount of computing infrastrs of reporting time and regions make different parts of the dataset that are ucture, data pipelines that are in good working order and have low latency, and the system being appropriately set up for scalability.

The lack of adequately skilled employees remains a problem that is heavily affecting agencies that are poorly funded or have weak digital infrastructures. In these environments, the scarcity of technical staff may readily result in the overburdening of the already existing workforce as each of these factors exacerbates the situation.

Socio-ethical implications:

Using predictive analytics to define police strategies raises significant ethical questions that primarily concern privacy, algorithm fairness, and the risk of perpetuating the past. Often, crime data sets reflect the social and institutional inequalities that exist, and if there is no thorough oversight, machine learning models may unintentionally exacerbate these inequalities [17]. Moreover, the decision to employ predictive tools in law enforcement should entail that there are unequivocal and well-organized policies addressing transparency, interpretability, and the community's role as a control body.

These measures are not only necessary for retaining the community's trust but also for ensuring that these technologies are used correctly [16].

latency, and the system needs to be properly designed to be scalable. Each of these factors may overwhelm very few technical personnel- agencies that are technically under-resourced or have weak digital infrastructures.

Socio-ethical implications:

The use of prediction analytics by the police for decision-making raises a large number of questions regarding the morality of the situation. The primary concerns are questions relating to privacy, the fairness of the algorithm, and the risk of bias. Simply put, crime data constitute the vehicle through which social or institutional inequalities that exist are passed on; thus, machine learning models may become instruments that, even though unintentionally, are the causes of the further development of these inequalities if they are not closely supervised [17]. Besides that, the use of prediction tools for some jobs should be accompanied by strict rules regarding disclosure, interpretability, and the provision of public control mechanisms so that community trust is not lost, and the implementation is performed in a

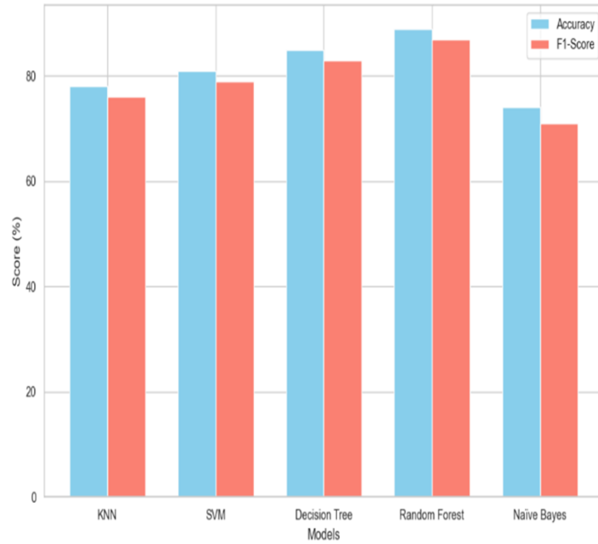

Fig. 1: Model Comparison: Accuracy vs F1-Score

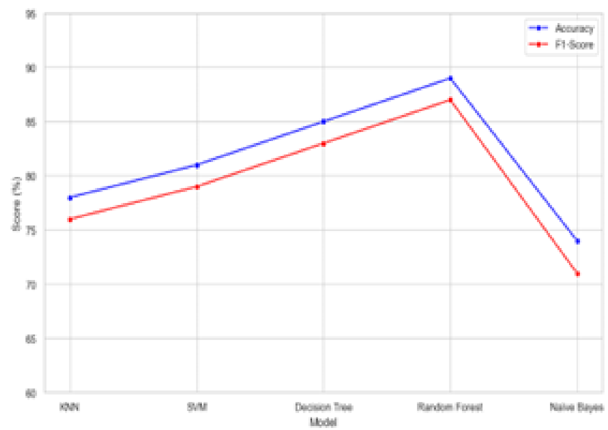

Fig. 2: Model Performance Line Plot

rightful way [16].

VII. Future Scope

This research has unlocked a number of possible avenues for its subsequent evolution. The idea, perhaps, of the most significant impact, of going beyond the past analysis to actually forecasting the future in real-time by the integration of live data streams from the sources such as emergency dispatch systems, digital police records, and networked surveillance infrastructure, is one of the aspects that strongly influences the further development of this research. The infusion of a continuous data stream would not only raise the current analysis tool to a system of operations that can immediates warning, rapid response coordination support as well as prediction adjustment according to new crime patterns but also elevate the latter to a system of operations. To pull this off, basically they need not only thoroughly tested robust streaming pipelines but also stable computational architectures that can scale up.

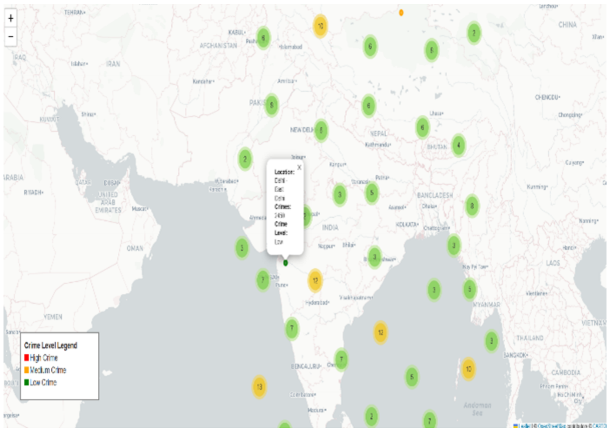

Fig. 3: Geospatial Crime Distribution Map

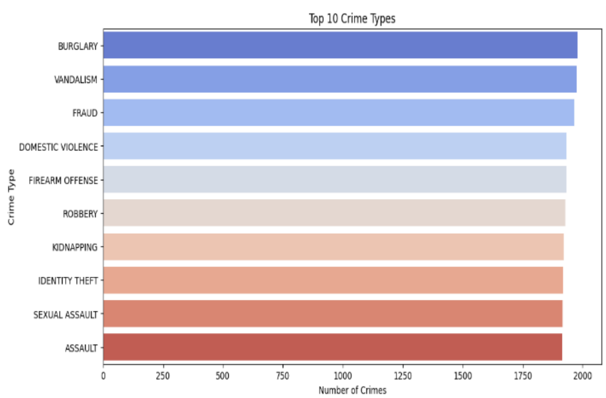

Fig. 4: Top 10 Crimes in India

This has a lot of potential for the development of a complex neural model that can explain the intricately intertwined spatiotemporal dependencies of urban crime patterns. For example, Long Short-Term Memory networks (LSTM) can be used for temporal sequence modeling, whereas Convolutional Neural Networks (CNN) can be utilized for spatial feature extraction. The breakthroughs in these fields have made them very promising for forecasting-related domains [21]. Hence, these inventions can serve as a means not only to reveal the long-term temporal trends but also the localized geographic variations more deeply.

Besides that, if we move our attention to the application side, subsequent research could be aimed at figuring out the ways of constructing deployment platforms that interpret the analytical outputs for the practical use of decision-support. For example, police officers can be tremendously facilitated by web or mobile-version interfaces, which integrate interactive visualization dashboards, a wide range of alerts, and secure reporting channels that not only can be accessed easily but also operated in a simple and user-friendly manner.

Similarly, these platforms facilitate the engagement of residents in public safety programs and, thus, becoming the venue of a more democratic crime prevention approach.

In a nutshell, these bold concepts of connecting temporally delayed systems with advanced deep learning techniques and user-centric platform creation not only indicate the enhancement of technical skills but also the possible socio-economic effects that data-driven crime analysis frameworks can bring about.

VIII. Conclusion

This paper is about one of the machine learning methods that can be used to analyze and predict crime patterns by processing a vast amount of historical crime data. The paper has shown through their comparative evaluation the use of various machine learning approaches such as K-Nearest Neighbour, Support Vector Machine, Decision Tree, Random Forest, and Naive Bayes not only for crime classification but also for finding significant spatiotemporal patterns. The main point of the methodological framework was the data preprocessing, the class imbalance problem addressed by SMOTE, and the intentional feature engineering for getting the most relevant temporal and spatial features.

The experiments demonstrate that ensemble-based methods, especially Random Forest, can often achieve the highest predictive accuracy and overall stability. In addition, interpretable models such as Decision Trees, which are slightly less accurate, can still provide feasible and easily understandable supportable decision paths that law enforcement units are able to utilize. The results obtained show the mutual benefits of the systematic spatiotemporal feature extraction and the powerful learning algorithms in the enablement of data-driven public safety intervention strategies.

Research regarding model performance is only a very small part of the paper, which makes up the bigger argument for data-driven crime analytics versus traditional evaluation methods. Traditional methods mostly depend on fragmentary or anecdotal insights, while machine learning systems provide scalable, empirical analyses that can be used for strategic resource allocation, identifying areas for interventions, and drafting targeted crime prevention policies. The next generation of innovations with real-time data integration, sophisticated modeling techniques, and better dataset quality may not simply enhance these systems but also facilitate the emergence of more responsive, community-oriented public safety initiatives.

References

[1] N. V. Chawla *et al.*, "SMOTE: Synthetic Minority Over-Sampling Technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.

[2] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[3] G. O. Mohler *et al.*, "Self-Exciting Point Process Modeling of Crime," *J. Amer. Statist. Assoc.*, vol. 106, no. 493, pp. 100–108, 2011.

[4] F. Pedregosa *et al.*, "Scikit-Learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

[5] S. Barua *et al.*, "MWMOTE: Majority Weighted Minority Oversampling Technique for Imbalanced Data Set Learning," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 2, pp. 405–425, 2014.

[6] H. He and E. A. Garcia, "Learning from Imbalanced Data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, 2009.

[7] G. Haixiang *et al.*, "Learning from Class-Imbalanced Data: Review of Methods and Applications," *Inf. Sci.*, vol. 412–413, pp. 241–255, 2017.

[8] G. O. Mohler *et al.*, "Evaluation and Analysis of Hotspot Policing and Crime Forecasting Methods," *Security J.*, vol. 27, no. 4, pp. 431–450, 2014.

[9] J. Liu *et al.*, "Deep Learning for Crime Prediction: A Comparative Study," in *Proc. IEEE Int. Conf. Big Data*, 2020.

[10] L. Wang *et al.*, "Predicting Crime Using Spatial Features," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 2018.

[11] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 2016, pp. 785–794.

[12] J. H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, 2001.

[13] Q. Huang *et al.*, "Spatio-Temporal Crime Prediction with GIS and Deep Learning," *ISPRS Int. J. Geo-Inf.*, vol. 10, no. 9, p. 602, 2021.

[14] W. Liu *et al.*, "Temporal Pattern-Aware Crime Prediction with Dynamic Graph Convolutional Networks," in *Proc. AAAI Conf. Artif. Intell.*, 2022.

[15] S. Kim *et al.*, "Temporal Attention Networks for Crime Prediction," in *Proc. AAAI Conf. Artif. Intell.*, 2023.

[16] C. Rudin, "Why We Should Use Interpretable Models for High-Stakes Decisions," *Nature Mach. Intell.*, vol. 1, pp. 206–215, 2019.

[17] R. Berk *et al.*, "Fairness in Criminal Justice Risk Assessments: The State of the Art," *Sociol. Methods Res.*, vol. 50, no. 1, pp. 3–44, 2021.

[18] H. Guo *et al.*, "Dynamic Sampling Strategies for Multi-Class Imbalanced Data Streams," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 12, pp. 5678–5692, 2022.

[19] Y. Li *et al.*, "Adaptive Sampling for Imbalanced Data Learning," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, 2021.

[20] X. Chen *et al.*, "Ensemble Learning for Crime Prediction," *ACM Trans. Intell. Syst. Technol.*, vol. 11, no. 4, pp. 1–25, 2020.

[21] Y. Wang *et al.*, "Deep Learning for Spatio-Temporal Data Mining: A Survey," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 8, pp. 3681–3700, 2022.

[22] X. Chen *et al.*, "Crime Prediction Using Twitter Sentiment and Weather Data," in *Proc. AAAI Workshops*, 2018.

[23] National Crime Records Bureau, *Crime in India 2020*. Ministry of Home Affairs, Government of India, 2021. Available: https://ncrb.gov.in/