https://doi.org/10.48047/AFJBS.6.13.2024.6079-6110

AFJBS

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

Crime rate prediction in Afro-American Society using Machine Learning Technique

Sulekha Das[1], Avijit Kumar Chaudhuri[2] , Nobhonil Roy Choudhury[3], Debanjan Paul[4], Akash Modak[5]

[1]Assistant Professor, Computer Science and Engineering, Techno Engineering College Banipur,

[2]Associate Professor, Department of CSE, Brainware University, West Bengal, India

[3]Assistant Professor, Department of Computer Science & Engineering, Brainware University, Barasat

[4]UG-Computer Science and Engineering, Techno Engineering College Banipur,

[5]UG-Computer Science and Engineering, Techno Engineering College Banipur

Email: [1]shu7773sea@gmail.com, [2]c.avijit@gmail.com, [3]nobhonilnew@gmail.com [4]debanjanpaul607@gmail.com, [5]akashmodak266@gmail.com

Volume 6, Issue 13,2024
Received: 15 jun 2024
Accepted: 25 July 2024
Published: 15 August 2024
doi:10.48047/AFJBS.6.13.2024.6079-6110

ABSTRACT

This paper discusses the long-standing problem of household segregation within the black area and its correlation with crime rates. Despite the progress in the area of socioeconomic improvement, domestic isolation continues to be one of the major obstacles faced by African Americans. The study aims to predict crime rates accurately, and it investigates the use of data techniques, particularly those developed from the UCI repository[1]. African Americans experience a higher degree of racial profiling within the penal system than whites. Given the increasing crime magnitude all over the globe, analyzing crime data is an urgent necessity to reduce crime and speed up how cases are solved. The paper uses a recursive feature elimination technique for feature selection through a dataset containing around 2000 records and 128 fields. The Multiple Linear Regression(MLR) method is used for feature selection, followed by the analysis using the Logistic Regression Algorithm (LRA). The proposed model is applied to the "Communities and Crime Data Set" from UCI Machine Learning Datasets to achieve a remarkable accuracy of 99.92%. This implies the success of the approach towards forecasting or prediction of crime patterns and emphasizes the significance of data-driven models in dealing with critical issues within communities, like crime.

[1]https://archive.ics.uci.edu/dataset/183/communities+and+crime

Keywords- Crime, Multiple Linear Regression, Logistic Regression, Confusion Matrix, 10Fold Cross Validation

Introduction

The tremendous progress and success of machine learning technology that has been demonstrated in the fields of speech(Deng, Hinton, & Kingsbury, 2013) and face recognition(Taigman, Yang, Ranzato, & Wolf, 2014; Sun, Wang, & Tang, 2014), automatic driving(Hoermann, Bach, & Dietmayer, 2018), and medical analysis(Choy et al., 2018) has been awe-inspiring. Although it may be very effective, machine learning technology often functions in a way that nobody can understand or explain because the working process is not transparent and not explainable(Apicella, Isgrò, Prevete, & Tamburrini, 2020). This obscurity causes complex issues, most importantly in assessing and predicting crime.

Although the interpretability of machine learning in both social sciences and AI tasks has been widely acknowledged, machine learning-based prediction models in crime are still criticized for their opacity(Miller, 2019). This opacity casts doubt on the reliability of the outcomes the models are generating among practicing professionals(Alves, Ribeiro, & Rodrigues, 2018; Rummens, Hardyns, & Pauwels, 2017; Zhang et al., 2022).

Hence, the need to create understandable and transparent crime prediction models to address this impasse cannot be overemphasized. These models help practitioners know the logical background and factors taken into account in the prediction process, thereby generating confidence in the final results. More specifically, the spatial variations of variables inside these models should be taken into account, as they may have a significant impact on crime opportunities(Lan, Liu, & Eck, 2021). Therefore, this will not only increase the reliability and credibility of crime prediction models but will also pave the way for more targeted crime prevention and law enforcement efforts.

Classical crime prediction techniques are based on historical crime data using the supposition that crime incidents are repeatable in both space and time. These models, for example, the near-repeat prediction model and the density estimation model(Chainey & Ratcliffe, 2013; Kalinic & Krisp, 2018), employ spatiotemporal information from previous crimes to forecast future criminal activities. They are good for creating crime risk on a large scale when detailed environmental and behavioural data is unavailable. On the other hand, their performance gets weaker if they are used to more fine-grained spatial and temporal scales.

Over the last few decades, researchers have started taking environmental factors into account in their crime models. Some of the key methodologies used are risk terrain modeling (RTM)(Caplan & Kennedy, 2010), Bayesian models, and discrete choice models(Hu, Zhu, Duan, & Guo, 2018; Law, Quick, & Chan, 2014). Among others, precisely RTM is in the focus. RTM is focused on establishing the environmental features that correlate with crime and assessing how the presence of these features simplifies crime opportunities. It starts with spatial feature selection and critical incident-related environmental characteristics assessment. Then, a statistical model is used to differentiate areas with more statistical certainty of crime; these areas are referred to as "risk terrains."(Kennedy, Caplan, & Piza, 2011; Wheeler & Steenbeek, 2021). RTM adequately pinpoints the impacts of individual factors on the environment but fails to take into account the possible interactions among such factors. Besides, this prediction may be wrong because there has been little mention of timing influences.

The term ‘violent crime’ is used to describe the types of offenses where a person uses physical force or threatens to use force to violate the rights of other people. Besides physical harm, psychological trauma, and sometimes even death, these horrendous acts also bring about profound social and emotional consequences that go on for extended periods. The scope of crimes that involve violence covers a wide range of offenses, including assault, robbery, homicide, sexual assault, and domestic abuse.

Communicating and dealing reasonably with violent crime is crucial to maintaining the security and health of the people and the community. Coordination among law enforcement agencies, legislators, and social groups is essential for the prevention of violent crimes and their intervention. By cooperating, different tools are used, which include crime prevention programs, community engagement initiatives, legislative measures, and law enforcement operations.

It is important to keep in mind that the study of violent crime requires a multidimensional analysis of several factors. These elements embrace the origin of the violence, the effect of socioeconomic conditions, the availability of weapons, mental health problems, substance abuse, and other determinants mediating the act. Through analyzing these fundamental structural issues carefully, societies can keep striving to deter the occurrence of violent crime and nurture more peaceful and secure living spaces for their people.

One of the most recent achievements in machine learning is its use as a tool for crime analysis as well as crime prevention. Using machine learning, an indication of data allows a machine to solve issues without the rules-based routine. Machine learning algorithms can extract patterns and predict future occurrences by training them on relevant datasets, thereby providing input to decision-making processes (Zhang et al., 2022).

Multiple regression analysis, an approach in machine learning, helps to analyze the relationship between variables. Using multiple regression analysis, it is possible to understand the interrelationship between one dependent variable and various independent variables in the context of violent crime. The ultimate goal of multiple regression analysis is to find independent variables that are known to efficiently predict the value of the dependent variable.

This paper will focus on the implementation of machine learning methods, especially regression analysis, to explain violent crimes and possible solutions. By using data-driven approaches, we also hope to promote the ongoing endeavor of reducing the high levels of violent crimes and help make communities safer for everyone.

Relevant Literature

Nasridinov & Park(2014) discussed and scrutinized decision trees, neural networks, SVM, k-NN, and Naïve Bayes algorithms based on real South Korean crime data. Their aim was to determine which of the two algorithms has a higher accuracy in crime prediction.

Dubey & Chaturvedi(2014) explained theoretical compression techniques that can be used for future crime prediction. They compared SVM and BayesNet, considering the accuracy, and drawbacks of each method.

In another research, Sathyadevan et al.(2014) used the Naïve Bayes algorithm, Decision Tree, and Apriori Algorithm along with Mongo DB and Graph DB to predict and analyze crime data. The Naïve Bayes algorithm was able to forecast the probability of crimes taking place in certain districts on given days with a 90% accuracy rate. The Apriori Algorithm followed crime patterns within specific districts, which were repeated crimes. High crime rate districts were represented using GraphDB, and data was managed in MongoDB.

Cesario et al. (2016), in their study, used two different sets of data from different sources to predict crime level categories. They used the Naïve Bayes classifier and neural networks to analyze the crime level at certain points and draw comparisons between the two methods. Their research showed that the Naïve Bayes classifier is superior to the neural network in terms of accuracy(90.22)%.

Joshi & Srivastava, 2014 used classification techniques such as BF Tree, J48, Decision Stump, and CART with the bagging ensemble method. These techniques were integrated with the datasets, including UCI Anneal, Credit, Iris, Wine, Zoo, Vowel, and Dermatology, to optimize the precision of individual decision tree-based classifiers. After conducting a series of experiments and comparisons between different datasets, they concluded that the bagging ensemble model with decision tree classifiers had the highest performance compared with other configurations.

Hassan & Abdel-Qader(2015) proposed a computer simulation tool to analyze the efficiency of a majority voting integrator. Their theoretical approach dealt with analyzing how the dependence between classifiers influences the error rate of ensemble models. According to their observations, the performance of the ensemble method deteriorates when the output of the constituent classifiers is dependent. This research thus gives an understanding of the features of ensemble learning and why it is relevant to consider independence among classifiers.

Almaw & Kadam(2018) analyzed the performance of Naïve Bayes, Decision Tree, K-Nearest Neighbor, and Multilayer Perceptron classifiers. They concentrated on evaluating the advantages of multiple classifier systems in building ensemble-based decision-making frameworks for future use. In their study, six significant parameters were considered to determine the best ensemble learning algorithm.

In their research, Zhang et al.(2016) proposed an improved strategy for identifying high-crime-intensity areas with five heat levels. To reduce features they used linear discriminant Analysis (LDA) for feature reduction and predicted the heavy crime intensity area by using the K-nearest neighbor algorithm (KNN).

Retnowardhani & Triana (2016) proposed the Crime Prevention Decision Support System (CreP-DSS); CreP-DSS is a web-based system developed in PHP. The CreP-DSS is intended to help law enforcement agencies manage and support decision-making tasks regarding crime prevention. The anticipated crime patterns are segmented into intervals through the system about the level of offense, hence improving decision-making for the prevention of crime.

In their study, Gupta et al. (2016) analyzed six classification algorithms, namely, Naïve Bayes, OneR, Decision Table, J48, JRip, and BayesNet, for classifying crime and accident records of Denver, USA. Their aim was to find out which of the algorithms to use when it comes to prediction accuracy and time when forecasting seasonal trends in crime. The comparative assessment of the tested methods revealed that JRip provided the best results, with a relatively high accuracy of 73.71%. This study provides important information on the enhancement of crime prediction methods by employing machine learning approaches.

Rojarath et al. (2016) investigated ensemble models using Naïve Bayes, Decision Trees, Multilayer Perceptron, and K-Nearest Neighbor classifiers. They used some datasets from UCI, such as Harberman, Urban, Chronic Kidney, Mammographic, Phoneme, and Pima. They proposed the M-ensemble learning framework that included the 3-ensemble and 4-ensemble models, where the number of base models was odd and even, respectively. The objective was to increase classifier accuracy using a majority vote. Surprisingly, the 3-ensemble models outperformed the other models by achieving an accuracy of 83 percent. Only 13% of the clients were satisfied with the performance of AI in providing unknown data in the forecast. The study shows that an ensemble can increase the predictive accuracy of a given model.

Verma and Mehta (2017) examined ensemble learning techniques comprising AdaBoost with decision stump bagging as well as stacking methods using J48, Naïve Bayes, and random forest classifiers. They used datasets from UCI such as breast cancer, dermatology, hepatitis, liver disorders, and Splice datasets. They used bagging, boosting, and stacked generalized ensembles for the comparison, with equal emphasis on the overall results instead of the individual ones. Bagging proved to be optimal for accuracy, whereas boosting and stacking were optimal for root mean square error. This research proves how ensemble methods work well in different dataset conditions.

To assess the performances of multiple machine learning models, Zhang et al. (2020) gathered the public property crime statistics of a city in southeastern China from 2015 to 2018. They identified that the LSTM, model achieved a precision of 64% on crime data only while KNN, random forest, support vector machine, Naive Bayes, and convolutional neural networks yielded lower precision. Furthermore, better accuracy in predicting crime locations was achieved by using points of interest (POIs) and urban road network density as additional covariates in the LSTM model than in the baseline LSTM model, which only used historical crime data for prediction. It also demonstrates how environmental components associated with criminological theories need to be incorporated into future attempts at crime prediction, as well as differences in efficiency between various machine learning approaches.

Bandekar & Vijayalakshmi, (2020) investigated the model that forecasts the types of crime incidents to be committed based on certain geographical precedents. Machine learning was used to develop a sample model with a cleaned and transformed training data set. Descriptive statistics and data visualization techniques were used to determine the characteristics of the dataset and the factors that are relevant for analysis. It established risks and developed measures of prediction to improve the safety of society. Furthermore, to meet the objectives of the research, various clustering algorithms, optimization techniques, and statistical tests were used.

Research Methodology

In this paper, the dataset is obtained from the UCI repository[1] website, concerning crimes against African Americans. It comprises a set of variables to evaluate algorithms that either choose or learn weights for attributes. Only attributes with plausible connections to crime (N=122) were selected, along with the attribute to be predicted: Violent crimes per capita. It includes demographic variables, which are the percentage of the urban population and median family income, and policing variables, including the number of police officers per capita and the proportion of officers belonging to the drug unit. The Per Capita Violent Crimes variable was calculated using population data and the sum of crimes considered violent in the United States: murder, rape, robbery, and assault. Some of the states had different numbers of rapes or had missing data which caused gaps and distortions on the per capita violent crime in those regions. Therefore, these cities, mainly from the Midwestern USA, were not included in the dataset.

The data original values were normalized to the decimal scale 0.00 – 1.00 using an unsupervised, equal interval binning method. This normalization preserved the distribution and skewness of the attributes. For instance, an attribute like 'mean people per household' is normalized to a range of 0 to 1. The population attribute has a mean value of 0.06, reflecting that most communities are small. Table 1 provides a summary of studies on the Crime dataset and Table 2 presents a summary of selected features using multilinear regression.

Table 1: Crime dataset description

Attributes	Definition
community	not predictive and many missing values (numeric)
communityname: community name	not predictive (numeric) --
population	population for community: (numeric - decimal)
householdsize	mean people per household (numeric - decimal)
racepctblack	It is African American percentage of population (numeric - decimal)
racePctWhite	It is Caucasian percentage of population (numeric - decimal)
racePctAsian	It is the percentage of population of Asian heritage (numeric - decimal)
racePctHisp	It is the percentage of population of Hispanic heritage (numeric – decimal)
agePct12t21	It is the percentage of population 12-21 in age (numeric - decimal)
agePct12t29	It is the percentage of population 12-29 in age (numeric - decimal)
agePct16t24	It is the percentage of population 16-24 in age

	(numeric - decimal)
agePct65up	It is the percentage of population 65 and over in age (numeric - decimal)
numbUrban	People living in areas classified as urban (numeric - decimal)
pctUrban	percentage of people living in the areas classified as urban (numeric - decimal)
medIncome	median value of household income (numeric - decimal)
pctWWage	percentage of resident with wage or salary income in 1989 (numeric - decimal)
pctWFarmSelf	percentage of residents with farm or self-employment income in 1989 (numeric - decimal)
pctWInvInc	percentage of residents with investment / rent income in 1989 (numeric - decimal)
pctWSocSec	percentage of residents with social security income in 1989 (numeric - decimal)
pctWPubAsst	percentage of residents with public assistance income in 1989 (numeric - decimal)
pctWRetire	percentage of homey with retirement income in 1989 (numeric - decimal)
medFamInc	median family income (differs from household income for non-family residents) (numeric - decimal)
perCapInc	Average income (numeric - decimal)
whitePerCap	Average income for Caucasians (numeric - decimal)
blackPerCap	Average income for African Americans (numeric - decimal)
indianPerCap	Average income for native Americans (numeric - decimal)
AsianPerCap	Average income for people with Asian heritage

	(numeric - decimal)
OtherPerCap	Average income for people with 'other' heritage (numeric - decimal)
HispPerCap	Average income for people with hispanic heritage (numeric - decimal)
NumUnderPov	No. of people under the poverty level (numeric - decimal)
PctPopUnderPov	people percentage under the poverty level (numeric - decimal)
PctLess9thGrade	people percentage 25 and over with less than a 9th grade education (numeric - decimal)
PctNotHSGrad	People percentage 25 and over that are not high school graduates (numeric - decimal)
PctBSorMore	People percentage 25 and over with a bachelor's degree or higher education (numeric - decimal)
PctUnemployed	percentage of people 16 and over, in the labor force, and unemployed (numeric - decimal)
PctEmploy	People percentage 16 and over who are employed (numeric - decimal)
PctEmplManu	People percentage 16 and over who are employed in manufacturing (numeric - decimal)
PctEmplProfServ	People percentage 16 and over who are employed in professional services (numeric - decimal)
PctOccupManu	People percentage 16 and over who are employed in manufacturing (numeric - decimal)
PctOccupMgmtProf	People percentage 16 and over who are employed in management or professional occupations (numeric - decimal)
MalePctDivorce	Males percentage who are divorced (numeric

	- decimal)
MalePctNevMarr	Males percentage who have never married (numeric - decimal)
PersPerFam	Average number of people per family (numeric - decimal)
PctFam2Par	This are the percentage of families (with kids) headed by two parents (numeric - decimal)
PctKids2Par	percentage of kids in family housing with two parents (numeric - decimal)
PctYoungKids2Par	percent of kids 4 and under in two parent households (numeric - decimal)
PctTeen2Par	percent of kids age 12-17 in two parent households (numeric - decimal)
PctWorkMomYoungKids	percentage of moms of kids 6 and under in labour force (numeric - decimal)
PctWorkMom	percentage of moms of kids under 18 in labour force (numeric - decimal)
NumIlleg	kids born to never married (numeric - decimal)
PctIlleg	kids born to never married (numeric - decimal)
NumImmig	Sum of number of people known to be foreign born (numeric - decimal)
PctImmigRecent	immigrants percentage who immigrated within last 3 years (numeric - decimal)
PctImmigRec5	immigrants percentage who immigrated within last 5 years (numeric - decimal)
PctImmigRec8	immigrants percentage who immigrated within last 8 years (numeric - decimal)
PctImmigRec10	immigrants percentage who immigrated within last 10 years (numeric - decimal)
PctRecentImmig	Percentage of population_ who have immigrated within the last 3 years (numeric - decimal)

PctRecImmig5	percentage of population_ who have immigrated within the last 5 years (numeric - decimal)
PctRecImmig8	Percentage of population_ who have immigrated within the last 8 years (numeric - decimal)
PctRecImmig10	Percentage of population_ who have immigrated within the last 10 years (numeric - decimal)
PctSpeakEnglOnly	Percentage of people who speak only English (numeric - decimal)
PctNotSpeakEnglWell	Percentage of people who do not speak English well (numeric - decimal)
PctLargHouseFam	Percentage of family households that are large (6 or more) (numeric - decimal)
PctLargHouseOccup	Percentage of all occupied households that are large (6 or more people) (numeric - decimal)
PersPerOccupHous	Average persons per household (numeric - decimal)
PersPerOwnOccHous	Average persons per owner occupied household (numeric - decimal)
PersPerRentOccHous	Average persons per rental household (numeric - decimal)
PctPersOwnOccup	Percentage of people in owner occupied households (numeric - decimal)
PctPersDenseHous	Percentage of persons in dense housing (more than 1 person per room) (numeric - decimal)
PctHousLess3BR	Percentage of housing units with less than 3 bedrooms (numeric - decimal)
MedNumBR	median bedrooms number (numeric - decimal)
HousVacant	vacant households number (numeric - decimal)

PctHousOccup	Percentage of housing occupied (numeric - decimal)
PctHousOwnOcc	Percentage of households owner occupied (numeric - decimal)
PctVacantBoarded	Percentage of vacant housing that is boarded up (numeric - decimal)
PctVacMore6Mos	Percentage of vacant housing that has been vacant more than 6 months (numeric - decimal)
MedYrHousBuilt	median year housing units built (numeric - decimal)
PctHousNoPhone	Percentage of occupied housing units without phone (in 1990, this was rare!) (numeric - decimal)
PctWOFullPlumb	Percentage of housing without complete plumbing facilities (numeric - decimal)
OwnOccLowQuart	owner occupier housing - lower quartile value (numeric - decimal)
OwnOccMedVal	owner occupied housing - median value (numeric - decimal)
OwnOccHiQuart	owner occupier housing - upper quartile value (numeric - decimal)
RentLowQ	rental housing - lower quartile rent (numeric - decimal)
RentMedian	rental housing - median rent (Census variable H32B from file STF1A) (numeric - decimal)
RentHighQ	rental housing - upper quartile rent (numeric - decimal)
MedRent	median gross rent (Census variable H43A from file STF3A - includes utilities) (numeric - decimal)
MedRentPctHousInc	median gross rent as a percent of household income (numeric - decimal)
MedOwnCostPctInc	median owners cost as a percent of household

	income - for owners with a mortgage (numeric - decimal)
MedOwnCostPctIncNoMtg	median owners cost as a percent of household income - for owners without a mortgage (numeric - decimal)
NumInShelters	number of people in homeless shelters (numeric - decimal)
NumStreet	homeless people number counted in the street (numeric - decimal)
PctForeignBorn	Percentage of people foreign born (numeric - decimal)
PctBornSameState	Percentage of people born in the same state as currently living (numeric - decimal)
PctSameHouse85	Percentage of people living in the same house as in 1985 (5 years before) (numeric - decimal)
PctSameCity85	Percentage of people living in the same city as in 1985 (5 years before) (numeric - decimal)
PctSameState85	percentage of people living in the same state as in 1985 (5 years before) (numeric - decimal)
LemasSwornFT	number of sworn full time police officers (numeric - decimal)
LemasSwFTPerPop	sworn full time police officers per 100K population (numeric - decimal)
LemasSwFTFieldOps	number of sworn full time police officers in field operations (on the street as opposed to administrative etc) (numeric - decimal)
LemasSwFTFieldPerPop:	sworn full time police officers in field operations (on the street as opposed to administrative etc) per 100K population (numeric - decimal)
LemasTotalReq	total requests for police (numeric - decimal)
LemasTotReqPerPop	Total number of requests for police per 100K popuation (numeric - decimal)

PolicReqPerOffic	total requests for police per police officer (numeric - decimal)
PolicPerPop	Officers of police per 100K population (numeric - decimal)
RacialMatchCommPol	A measurement of the racial match between the community and the police force High values indicate proportions in community and police force are similar (numeric - decimal)
PctPolicWhite	police percentage that are caucasian (numeric - decimal)
PctPolicBlack	Police percentage that are african american (numeric - decimal)
PctPolicHisp	Police percentage that are hispanic (numeric - decimal)
PctPolicAsian	population of police that are asian (numeric - decimal)
PctPolicMinor	Police percentage that are minority of any kind (numeric - decimal)
OfficAssgnDrugUnits	Officers No. assigned to special drug units (numeric - decimal)
NumKindsDrugsSeiz	Total number of different kinds of drugs seized (numeric - decimal)
PolicAveOTWorked	police mean overtime worked (numeric - decimal)
LandArea	Area of land in square miles (numeric - decimal)
PopDens	Density population in persons per square mile (numeric - decimal)
PctUsePubTrans	People percentage using public transit for commuting (numeric - decimal)
PolicCars	police cars (numeric - decimal)
PolicOperBudg	police operating financial plan (numeric - decimal)

LemasPctPolicOnPatr	Sworn percentage full time police officers on patrol (numeric - decimal)
LemasGangUnitDeploy	Team unit deployed (numeric - decimal - but really ordinal - 0 means NO, 1 means YES, 0.5 means Part Time)
LemasPctOfficDrugUn	Officers percentage assigned to drug units (numeric - decimal)
PolicBudgPerPop	police operating financial plan per population (numeric - decimal)
ViolentCrimesPerPop	Totality of violent crimes per 100K popuation (numeric - decimal)-Class variable

Table 2. Crime dataset with 8 Features and 1 Target Variable

Attributes	**Definition**	**Mean**	**Standard deviation**
ViolentCrimesPerPop (Dependent)	Total number of violent crimes per 100K population.	0.214	0.206
pctWInvInc	Percentage of households with investment or rent income in 1989.	0.494	0.181
MalePctDivorce	Percentage of males who are divorced.	0.460	0.184
FemalePctDiv	Percentage of females who are divorced.	0.487	0.175
TotalPctDiv	Percentage of population who are divorced.	0.494	0.185

PctHousOwnOcc	Percent of household's owner occupied.	0.546	0.183
PctVacMore6Mos	Percent of vacant housing that has been vacant more than 6 months.	0.427	0.206
MedOwnCostPctInc	Median owners cost as a percentage of household income - for owners with a mortgage.	0.445	0.185
PctBornSameState	Percent of people born in the same state as currently living.	0.609	0.206

3. Methodology

Recursive Feature Elimination (RFE)

Recursive Feature Elimination (RFE) is a way of feature selection that distinguishes more informative features from the datasets. It recursively applies a model that is usually a machine learning algorithm and removes the features with the lowest importance until the best subset is achieved.

We choose MLR as the main model in RFE to find out the more influential predictors among the set of independent variables related to the explanation of the dependent variable. MLR is one of the statistical procedures used for modeling the relationships between several independent variables and a dependent variable. Here's a detailed elaboration of the steps involved in recursive feature elimination using MLR:

Model Selection: First of all, the MLR model is selected as the main model. This model presents the basis for the assessment of feature significance.

Full Model Fitting: The MLR model is fitted using all the features that are present in the dataset. We do the first step by preparing the baseline performance, which will consume the entire feature set.

Feature Ranking: Then the model is fitted, and the contribution of every feature is analyzed. This can be done by looking at the coefficients present for each independent variable in the regression model. Variables with smaller coefficients or insignificant levels of influence would fall into this category.

Feature Elimination: The least significant feature, according to the feature ranking procedure, is deleted from the dataset. The next step in this process is to enhance the performance of the model by concentrating on the most important features(Chaudhuri et al., 2021).

Iterative Process: Steps 2-4 are continued repeatedly until a predetermined stopping criterion is satisfied. This could be the number of features needed to maintain an important enough level of performance or the stage where further removal of features has no more positive effect on model performance, as the case may be.

Performance Evaluation: Each iteration is evaluated using a validation set or through cross-validation as part of the gradual process. Therefore, such consideration guarantees that the chosen features do not degrade the model's performance on data that it has not seen before.

Final Model Refinement: After screening for the most suitable features, the MLR model is trained using the chosen subset of features, which includes only those features that have a direct impact on the response. The final model represents the simplest possible version, which includes only the variables that are statistically significant in the analysis.

Through RFE by MLR, practitioners can shrink the feature space, increase model interpretability, and, if attention is focused on the most informative features, perhaps the performance of the model could be improved. It is particularly appropriate for the situation when the datasets are high-dimensional and numerous aspects are involved, and it helps to avoid the problem of the curse of dimensionality and also enhances the model's generalization ability. The feature selection method and the algorithm are depicted in Figure 1 and Figure 2 below.

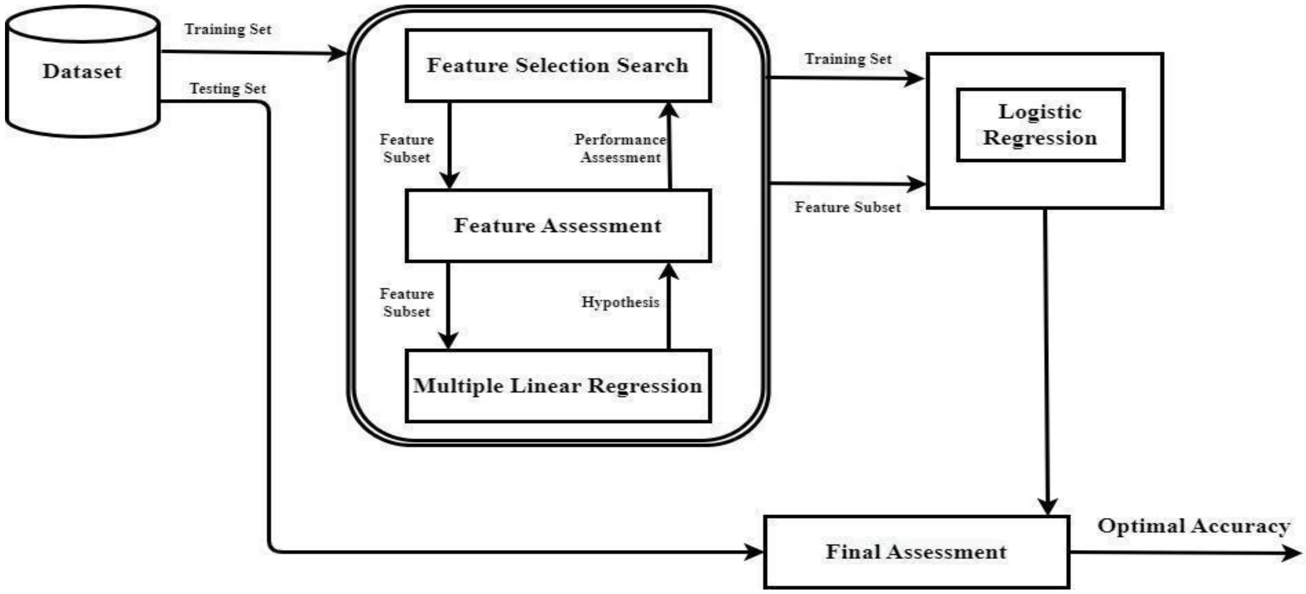


Figure 1. Wrapper algorithm approach for FSS Using MLR

Load feature set F^{Set} = {1, 2, . . . , n} at Set = n

Repeat

• Select best feature f_{best}ÎF^{Set} to eliminate the minimum significant cost increase

• F^{Set-1}=F^{Set} - {f_{best}} and Set - -

Figure 2. Backward Greedy Algorithm Using MLR

3.2 Multiple Linear Regression

Multiple Linear Regression (MLR) is used to determine the relationship between one variable and several other variables. It is used extensively in feature selection to find the relevant predictors for the outcome variable. Feature selection plays an important role in developing accurate predictors, especially for high dimensional data(He & Zheng, 2021; Aziz et al., 2022; Pina-Sánchez et al., 2023) .

1. **Multiple Regression**: This is an extension of simple linear regression method that allow incorporation of more than one independent variable. The general form of a multiple regression model is presented in Eq. (i).

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_3+\cdots+\beta_nx_n+\epsilon \tag{i}$$

Where:

y is the dependent variable

β_0 is the intercept

$\beta_1,\beta_2,\beta_3,\ldots,\beta_n$ are the coefficients for the independent variables $x_1,x_2\ldots\ldots x_n$ respectively.

x_1, x_2 …… x_n are the independent variables.

ϵ is the error term, representing the difference between the predicted and actual values of y.

2. **Feature Selection**: This involves identifying a few of the most important features (variables) to be used in creating the model. The aim is to enhance the performance of models, minimize overfitting, enhance the precision of the model, and decrease computational time.

Techniques for Feature Selection using MLR

1. **Stepwise Regression:** These include the forward selection, backward elimination, and the bidirectional elimination procedures. These methods build or prune predictors in a stepwise fashion based on criteria including AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), or p-values of the predictors.
2. **Lasso Regression (L1 Regularization):** Lasso (Least Absolute Shrinkage and Selection Operator) imposes a penalty, which is the absolute value of the magnitude of coefficients. This can make some of the coefficients equal to zero, so can do variable selection exercise and the expression of Lasso Regression is presented in Eq. (ii).

$$\text{Minimize} \sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij}\right)^2 + \lambda \sum_{j=1}^{p}\left|\beta_j\right| \quad \text{(ii)}$$

3. **Ridge Regression (L2 Regularization)**: Unlike Lasso, Ridge regression adds a penalty of the type of the square of the magnitude of coefficients which minimizes multicollinearity and the mathematical expression of Ridge Regression is presented by Eq. (iii).

$$\text{Minimize} \sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij}\right)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \quad \text{(iii)}$$

4. **Elastic Net:** This is the combination of L1 and L2 regularization and it is useful when the predictor variables are multiple and correlated. The expression for Elastic Net is provided in Eq. (iv).

$$\text{Minimize} \sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij}\right)^2 + \lambda_1 \sum_{j=1}^{p}\left|\beta_j\right| + \lambda_2 \sum_{j=1}^{p} \beta_j^2 \quad \text{(iv)}$$

5. **Variance Inflation Factor (VIF):** This indicates the extent to which the variance of a regression coefficient is increased due to multicollinearity.

Logistic Regression in Crime Prediction

Logistic regression is a popular statistical approach that is commonly used for binary classification issues and has been successfully introduced into crime prediction to analyze and forecast the propensity of crime occurrence. LR estimates the likelihood that an input will fall into some specific class. In crime prediction, it can be used to predict the chances of finding a crime zone and the possibility of committing a crime. The model finds estimates of the coefficients of a logistic function that yield a value between 0 and 1, representing the probability(Chaudhuri et al., 2024; Chaudhuri & Das, 2024).

Application of LR in Crime Prediction:

- **Predicting Crime Hotspots:** Taking into consideration the crime statistics, the approach of logistic regression can detect areas with probabilities of future criminal occurrences.

- **Offender Profiling:** It can be applied in the evaluation of people with parameters that include age, gender, socio-economic status, and criminal history to protons reoffending.

- **Recidivism Prediction:** The findings also show that the results of logistic regression models can estimate the likelihood of a prior offender being involved in more crimes.

Practical Considerations

Data Quality: Logistic regression models are very sensitive to the data used hence the quality and completeness of data used will greatly affect the outcomes. In crime prediction, this includes past crime trends, population characteristics, and other social and economic factors.

Interpretability: Logistic regression can be considered advantageous in its interpretability. The coefficients of the model can give an insight as to what factors are most closely associated with crime.

Ethical Concerns: When applying logistic regression in crime prediction, ethical concerns such as bias in data, equal representation, and perpetuating social injustice have to be addressed.

Logistic regression remains an important tool in crime prediction because it maintains a balance between simplicity, interpretability, and effectiveness.

Mathematical Expression of Logistic Regression

$$b0 = y000 - (b015 \times avg015 + b037 \times avg037 + b039 \times avg039 \times b040 \times avg040 + b071 \times avg071 + b073 \times avg073 + b086 \times avg086 + b89 \times avg089) \quad \text{(v)}$$

$$y = b0 + b15 \times x015[i] + b37 \times x037[i] + b39 \times x039[i] + b40 \times x040[i] + b71 \times x071[i] \times b73 \times x073[i] \times b86 \times x086[i] + b89 \times x089[i] \quad \text{(vi)}$$

Equation (v) and (vi) finds the value of the intercept and dependent variable y based on several independent variables x_i. Here's a breakdown of the components of the equation:

- y: The dependent variable or the outcome that is to be predicted.
- b15, b37, b39, b40, b71, b73, b86, b89 are the coefficients represent the change in the dependent variable y for a one-unit change in the respective independent variable, holding all other variables constant.

- avg015, avg037,avg039, avg040, avg071, avg073, avg086, avg089 are the average values of the respective independent variables across all observations in the dataset.
- x015= variable representing pctWInvInc.
- x037= variable representing MalePctDivorce.
- x039= variable representing FemalePctDiv.
- x040= variable representing TotalPctDiv.
- x071= variable representing PctHousOwnOcc.
- x073= variable representing PctVacMore6Mos.
- x086= variable representing MedOwnCostPctInc.
- x089= variable representing PctBornSameState.

- $\bar{y}$= Average of ViolentCrimesPerPop variable.
- $\bar{x}$= Average of all independent fields particularly.

$$b_i=\frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2} \tag{vii}$$

Where i=1, 2....n number of independent features

b_i=correlation coefficient

x_i=ith independent feature

y_i= dependent feature of ith row

$$p=\frac{e^y}{1+e^y} \tag{viii}$$

Where p=Totality of violent crimes per 100K population(dependent feature)

Equation (vii) and (viii) finds the value of individual correlation coefficient and totality of violent crimes per 100K population.

Assessment of Model Performance

Table 3: Confusion matrix and performance evaluation metrics and statistical tests

<table>
<tr><th>S/N</th><th>Metrics</th><th>Formula/ Description</th></tr>
<tr><td>1</td><td>Confusion Matrix</td><td>
<table>
<tr><td></td><td></td><td colspan="2">Actual</td></tr>
<tr><td rowspan="3">Predicted</td><td></td><td>Positive</td><td>Negative</td></tr>
<tr><td>(Positive)</td><td>True Positive, T_P</td><td>False Positive, F_P</td></tr>
<tr><td>(Negative)</td><td>False Negative, F_N</td><td>True Negative, T_N</td></tr>
<tr><td></td><td></td><td>Sensitivity= $\frac{T_P}{(T_P + F_N)}$</td><td>Specificity= $\frac{T_N}{(F_P + T_N)}$</td></tr>
</table>
</td></tr>
<tr><td>2</td><td>Accuracy</td><td>$\frac{T_P + T_N}{(T_P + F_N) + (F_P + T_N)}$</td></tr>
<tr><td>3</td><td>Precision</td><td>$\frac{T_P}{T_P + F_P}$</td></tr>
<tr><td>4</td><td>AUC
(Area under the curve)</td><td>A curve plotted between sensitivity and (1-specificity) is called receiver operating characteristic (ROC). AUC measures the degree to which the curve is up in the north-west corner.</td></tr>
<tr><td>5</td><td>Kappa Statistic</td><td>$\left(P_c - P_b\right) \ / \ \left(1 - P_b\right)$
P_c is complete agreement probability, and P_b represents likelihood ‘by chance’. Its range is (-1, 1).</td></tr>
</table>

Results and discussions

Confusion matrix is a very important measure for assessing the classifications and contains the actual, as well as the predicted classification made by different classifiers. It provides a framework for presenting performance indicators that are important in evaluating the effectiveness of these systems. It is necessary to emphasize measures of sensitivity, specificity, and accuracy when comparing the proposed model with other methods. Sensitivity or true positive rate refers to the probability of these actual positive samples being correctly classified. Specificity is defined as the number of true negatives out of the total negatives in the population as classified by the model.

Accuracy gives a general measure of how accurate the model is, including the true positive and the true negatives.

By reviewing the data from the confusion matrix, the researchers are provided with the opportunity to have an absolutely unbiased perspective at the proposed model, as well as other approaches, to enshrine their virtues and vices. The probability models provide different solution recommendations concerning the choice of the classification action, while the sensitivity, specificity, and accuracy tests represent different views of the models to enhance the decision-making process.

The logistic regression model was tested using a UCI dataset with three different training-testing ratios(50-50%, 66-34% and 80-20%) and 10 fold cross validations, and their accuracies are shown in the Table 4, Table 5, Table 6 and Table 7 respectively. A split ratio of 80:20 and 50:50 for training and testing yielded the highest accuracy of 100% in all performance matrices.

Table 4, Table 5, Table 6 and Table 7 provides the classification report, including sensitivity, specificity, precision, recall, F1-score, and accuracy for the logistic regression classifier using the UCI dataset with 50-50, 66-34 and 80-20 training-testing split and 10 fold cross validations. The model achieved 100% score in accuracy, sensitivity, specificity, precision, recall and F1-score.

The LR model also performs remarkable in 10 fold cross validations and reach 100% accuracy in all performance measures in most of the folds.

Table 8 shows the ROC (Receiver Operating Characteristic) curve, used to further investigate the model's performance. The ROC curve visualizes the trade-off between the true positive rate (TPR) and false positive rate (FPR), ranging from 0 to 1. The area under the ROC curve indicates the model's ability to distinguish between classes, with a curve closer to 1 indicating higher classification capability.

Table 4. Comparison of Accuracies for 80-20% train-test split

ATTRIBUTES	VALUE RANGE (%)
Confusion Matrix	152 0 0 241
Accuracy	100.0
Sensitivity	100.0
Specificity	100.0

Precision	100.0
Recall	100.0
F1_Score	100.0

The model correctly classified all instances, resulting in no false positives or false negatives in 80-20 % split. The perfect scores across all metrics indicate the model's ability to distinguish between classes flawlessly in this train-test split.

Table 5. Comparison of Accuracies for 66-34% train-test split

Attributes	Value range (%)
Confusion Matrix	254 0 11 388
Accuracy	98.315
Sensitivity	100.0
Specificity	97.24
Precision	95.84
Recall	100.0
F1_Score	97.88

The model misclassified 11 instances as false positives in 66-34% split. While the sensitivity and recall remain perfect, the presence of false positives slightly reduces the specificity and precision.

Table 6. Comparison of Accuracies for 50-50% train-test split

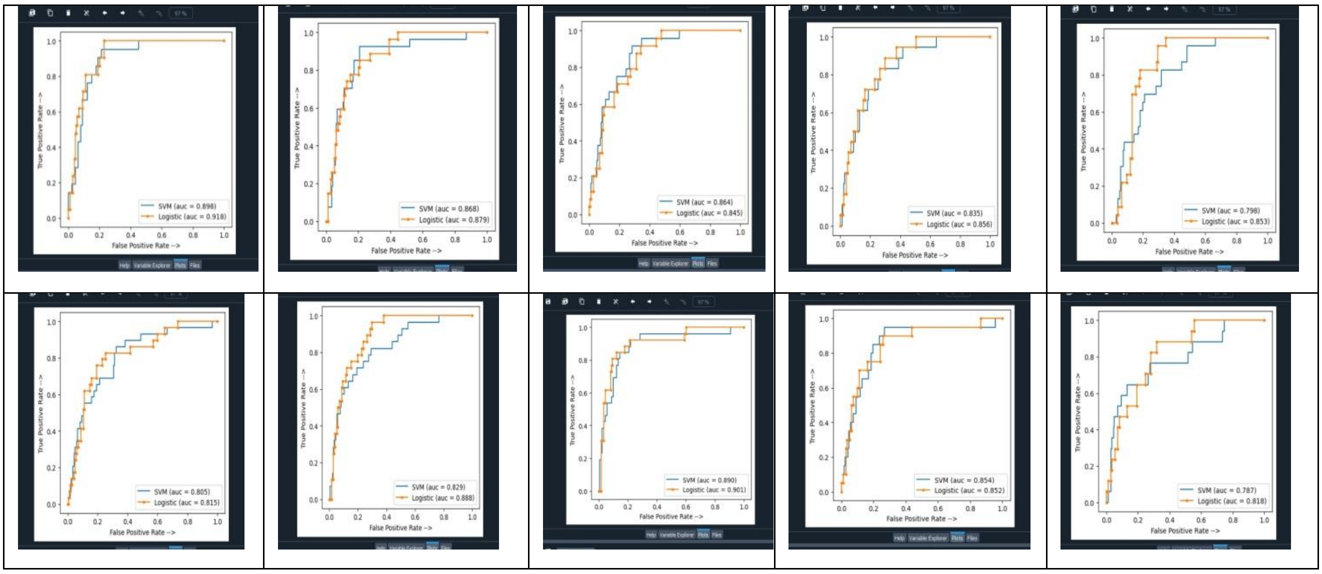
Attributes	Value range (%)
Confusion Matrix	388 0 0 611
Accuracy	100.0
Sensitivity	100.0
Specificity	100.0
Precision	100.0
Recall	100.0
F1_Score	100.0

Similar to the 80-20 split, the model correctly classified all instances with no errors. Perfect scores across all metrics confirm the model's outstanding performance.

Table 7: Comparison of Accuracies for 10-fold cross-validation

TEST CASES	Accuracy	Sensitivity	Specificity	Recall	Precision	F1_score
01	95.65	100.0	90.16	100.0	92.77	96.24
02	99.46	100.0	99.09	100.0	98.70	99.34
03	97.31	100.0	94.59	100.0	94.93	97.40
04	100.0	100.0	100.0	100.0	100.0	100.0
05	100.0	100.0	100.0	100.0	100.0	100.0
06	99.39	100.0	98.92	100.0	98.63	99.31
07	100.0	100.0	100.0	100.0	100.0	100.0
08	98.14	100.0	96.34	100.0	96.38	98.15
09	98.61	100.0	97.26	100.0	97.26	98.61
10	99.29	100.0	98.57	100.0	98.63	99.31

Table 8. Comparison of ROC Curve and AUC values (10-fold cross-validations)



While some folds achieved perfect scores, others showed slight variations in accuracy, specificity, and precision. The overall high performance across folds indicates good generalization ability.

Conclusion

The proposed work is intended to not only adopt but also expand the utilization of explainable AI techniques in the context of crime rate prediction. Most of the current approaches in machine learning that have been designed with an emphasis on the risk assessment of criminal occurrences are mainly 'black box' models which means that the underlying process occurring within the model and the reasons behind a particular prognosis are not clearly definable. Such opacity can lead to dilution of trust and confidence among users, stakeholders and the public at large who may not believe in the prediction without understanding the models behind it.

To address this issue, this study has come up with model that are precise, possessing high predictive accuracies, but more importantly, comprehensible. This new model offers similar interpretability as the earlier regression models that simplify the relation between the input variables and the output.

Key contributions of the study include:

1. **Enhanced Predictive Accuracy:** The new model not only have higher ability in crime prediction to traditional machine learning models but also give more information about the crime prediction process.

2. Interpretability: The model shows the influence that each variable has on the prediction from a global as well as from a local standpoint.

> Global Perspective: It displays the aggregated contribution of each variable toward all the predictions. It gives an overall picture of which factors tend to matter most in the determination of the model's decisions.
>
> Local Perspective: Under this perspective, this study determines the contribution of each variable for each individual prediction. It facilitates the identification of the features that contributed to a given forecast.

3. Identification of important features using Explainable AI: The analysis of crime rates with the help of socio-economic and demographical factors can help identify possible factors, which might impact the crime rate in a particular region. This analysis included the following variables:

Total number of violent crimes per 100K population: An independent variable which is a measure of the rate of crime is among the variable of interest.

Percentage of households with investment or rent income in 1989: Shows the economic status and the level of economic welfare within a certain society.

Percentage of males who are divorced: Looks at stability and potential stressors in a socially constructed male entity.

Percentage of females who are divorced: Likely, similar to the above, it might indicate social stability and potential stressors within a female population.

Percentage of the population who are divorced: An aggregate of social stability of the whole population.

Percent of households owner-occupied: Indicates how stable a community is and how much it has invested in properties and this can be a good measure of the level of crime in an area.

Percent of vacant housing that has been vacant more than 6 months: Reflects on the overall economic status and possible areas of weakness or deterioration.

Median owner's cost as a percentage of household income for owners with a mortgage: Indicates the cost aspect of homeownership, which would point towards the strain in an area.

Percent of people born in the same state as currently living: Indicates social integration and population's stability, migration etc.

Studying these variables will help identify more comprehensive socio-economic factors that may lead to increased or decreased rates of crime. For instance, higher density of buildings that are long-term vacant and higher percentage of people who are divorced may relate to higher incidence of crime. On the other hand, more homeownership and stability of the communities (people who live in birth state) are associated with a lower crime rate.

Thus, by incorporating these interpretability features into the study, it is possible to increase people's confidence in the model for making predictions. Such transparency may help increase the adoption and integration of machine learning tools in crime rate predictions, which ultimately contribute to policy-making and creation of efficient and realistic crime fighting measures.

References

Almaw, A., & Kadam, K. (2018, June). Crime data analysis and prediction using ensemble learning. In 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 1918-1923). IEEE.

Alves, L. G., Ribeiro, H. V., & Rodrigues, F. A. (2018). Crime prediction through urban metrics and statistical learning. Physica A: Statistical Mechanics and its Applications, 505, 435-443.

Apicella, A., Isgrò, F., Prevete, R., & Tamburrini, G. (2020). Middle-level features for the explanation of classification systems by sparse dictionary methods. International Journal of Neural Systems, 30(08), 2050040.

Aziz, R. M., Hussain, A., Sharma, P., & Kumar, P. (2022). Machine learning-based soft computing regression analysis approach for crime data prediction. Karbala International Journal of Modern Science, 8(1), 1-19.

Bandekar, S. R., & Vijayalakshmi, C. (2020). Design and analysis of machine learning algorithms for the reduction of crime rates in India. Procedia Computer Science, 172, 122-127.

Caplan, J. M., & Kennedy, L. W. (2010). Risk terrain modeling manual: Theoretical framework and technical steps of spatial risk assessment for crime analysis. Rutgers Center on Public Security.

Cesario, E., Catlett, C., & Talia, D. (2016, August). Forecasting crimes using autoregressive models. In 2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech) (pp. 795-802). IEEE.

Chainey, S. P. (2013). Examining the influence of cell size and bandwidth size on kernel density estimation crime hotspot maps for predicting spatial patterns of crime. Bulletin of the Geographical Society of Liege, 60, 7-19.

Chaudhuri, A. K., & Das, S. (2024). The Performance of Feature Selection Approaches on Boosted Random Forest Algorithms for Predicting Cardiovascular Disease. In Computer Vision and AI-Integrated IoT Technologies in the Medical Ecosystem (pp. 288-310). CRC Press.

Chaudhuri, A. K., Das, S., & Ray, A. (2024). An Improved Random Forest Model for Detecting Heart Disease. In Data-Centric AI Solutions and Emerging Technologies in the Healthcare Ecosystem (pp. 143-164). CRC Press.

Chaudhuri, A. K., Sinha, D., Banerjee, D. K., & Das, A. (2021). A novel enhanced decision tree model for detecting chronic kidney disease. Network Modeling Analysis in Health Informatics and Bioinformatics, 10, 1-22.

Choy, C. A., Robison, B. H., Gagne, T. O., Erwin, B., Firl, E., Halden, R. U., ... & S. Van Houtan, K. (2019). The vertical distribution and biological transport of marine microplastics across the epipelagic and mesopelagic water column. Scientific reports, 9(1), 7843.

Deng, L., Hinton, G., & Kingsbury, B. (2013, May). New types of deep neural network learning for speech recognition and related applications: An overview. In 2013 IEEE international conference on acoustics, speech and signal processing (pp. 8599-8603). IEEE.

Dubey, N., & Chaturvedi, S. K. (2014). A survey paper on crime prediction technique using data mining. Int. J. Eng. Res. Appl, 4(3), 396-400.

Gupta, A., Syed, A., Mohammad, A., & Halgamuge, M. N. (2016). A comparative study of classification algorithms using data mining: crime and accidents in Denver City the USA. International Journal of Advanced Computer Science and Applications, 7(7), 374-381.

Hassan, M. F., & Abdel-Qader, I. (2015, December). Performance analysis of majority vote combiner for multiple classifier systems. In 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA) (pp. 89-95). IEEE.

He, J., & Zheng, H. (2021). Prediction of crime rate in urban neighborhoods based on machine learning. Engineering Applications of Artificial Intelligence, 106, 104460.

Hoermann, S., Bach, M., & Dietmayer, K. (2018, May). Dynamic occupancy grid prediction for urban autonomous driving: A deep learning approach with fully automatic labeling. In 2018 IEEE International Conference on Robotics and Automation (ICRA) (pp. 2056-2063). IEEE.

Hu, T., Zhu, X., Duan, L., & Guo, W. (2018). Urban crime prediction based on spatio-temporal Bayesian model. PloS one, 13(10), e0206215.

Joshi, N., & Srivastava, S. (2014). Improving classification accuracy using ensemble learning technique (using different decision trees). Int. J. Comput. Sci. Mob. Comput, 3(5), 727-732.

Kalinic, M., & Krisp, J. M. (2018). Kernel density estimation (KDE) vs. hot-spot analysis–detecting criminal hot spots in the City of San Francisco. Lund, Sweden.

Kennedy, L. W., Caplan, J. M., & Piza, E. (2011). Risk clusters, hotspots, and spatial intelligence: risk terrain modeling as an algorithm for police resource allocation strategies. Journal of quantitative criminology, 27, 339-362.

Lan, M., Liu, L., & Eck, J. E. (2021). A spatial analytical approach to assess the impact of a casino on crime: An example of JACK Casino in downtown Cincinnati. Cities, 111, 103003.

Law, J., Quick, M., & Chan, P. (2014). Bayesian spatio-temporal modeling for analysing local patterns of crime over time at the small-area level. Journal of quantitative criminology, 30, 57-78.

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. Artificial intelligence, 267, 1-38.

Nasridinov, A., & Park, Y. H. (2014). A study on performance evaluation of machine learning algorithms for crime dataset. Adv. Sci. Technol. Lett.-(Networking Commun. 2014), 66, 90-92.

Pina-Sánchez, J., Buil-Gil, D., Brunton-Smith, I., & Cernat, A. (2023). The impact of measurement error in regression models using police recorded crime rates. Journal of Quantitative Criminology, 39(4), 975-1002.

Retnowardhani, A., & Triana, Y. S. (2016, November). Classify interval range of crime forecasting for crime prevention decision making. In 2016 11th international Conference on knowledge, Information and creativity support systems (KICSS) (pp. 1-6). IEEE.

Rojarath, A., Songpan, W., & Pong-inwong, C. (2016, August). Improved ensemble learning for classification techniques based on majority voting. In 2016 7th IEEE international conference on software engineering and service science (ICSESS) (pp. 107-110). IEEE.

Rummens, A., & Hardyns, W. (2021). The effect of spatiotemporal resolution on predictive policing model performance. International Journal of Forecasting, 37(1), 125-133.

Safat, W., Asghar, S., & Gillani, S. A. (2021). Empirical analysis for crime prediction and forecasting using machine learning and deep learning techniques. IEEE access, 9, 70080-70094.

Sathyadevan, S., Devan, M. S., & Gangadharan, S. S. (2014, August). Crime analysis and prediction using data mining. In 2014 First international conference on networks & soft computing (ICNSC2014) (pp. 406-412). IEEE.

Sun, Y., Wang, X., & Tang, X. (2015). Deeply learned face representations are sparse, selective, and robust. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2892-2900).

Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1701-1708).

Verma, A., & Mehta, S. (2017, January). A comparative study of ensemble learning methods for classification in bioinformatics. In 2017 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence (pp. 155-158). IEEE.

Zhang, Q., Yuan, P., Zhou, Q., & Yang, Z. (2016, May). Mixed spatial-temporal characteristics based crime hot spots prediction. In 2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD) (pp. 97-101). IEEE.

Zhang, X., Liu, L., Lan, M., Song, G., Xiao, L., & Chen, J. (2022). Interpretable machine learning models for crime prediction. Computers, Environment and Urban Systems, 94, 101789.

Zhang, X., Liu, L., Xiao, L., & Ji, J. (2020). Comparison of machine learning algorithms for predicting crime hotspots. IEEE access, 8, 181302-181310.