DOI: 10.24996/ijs.2025.66.8.29

ISSN: 0067-2904

Statistical Comparison of Some Machine Learning Techniques: A Case Study for Classifying Domestic Violence Crimes in Iraq

Amer F.A.H. ALNUAIMI *, Tasnim H.K. ALBALDAWI
Dept. of Mathematics, College of Science, University of Baghdad, Baghdad, Iraq

Received: 21/4/2024 Accepted: 17/8/2024 Published: 30/8/2024

Abstract
Crime classification and prediction represents a contemporary societal trend aimed at reducing or preventing criminal activities. One such crime is domestic violence, a global phenomenon known as the shadow pandemic or the hidden pandemic. This study aims to analyse data related to domestic violence crimes in Iraq and compare the effectiveness of the Random Forest (RF), Decision Tree (DT) and Naïve Bayes (NB) models using different metrics to develop an accurate model for classifying or predicting the type of domestic violence (physical or non-physical). The research employs a methodological framework that includes several stages. Subsequently, the three classifiers RF, DT, and NB are applied to the dataset to facilitate classification and prediction. The experiment results showed the superior performance of the RF classifier, achieving an accuracy score of (99.77%), compared to the DT classifier (99.07%) and NB (97.69%). For validation, different classification metrics were used. RF exhibited superior performance in all metrics compared to DT and NB algorithms, whose performance capabilities varied.

Keywords: Domestic violence, Machine learning, Classification, Random Forest, Decision Tree, Naive Bayes.

مقارنة إحصائية لبعض تقنيات التعلم الآلي: دراسة حالة لتصنيف جرائم العنف الأسري في العراق

عامر فائق عبد هاشم النعيمي*, تسنيم حسن كاظم البلداوي
قسم الرياضيات, كلية العلوم, جامعة بغداد, بغداد, العراق

الخلاصة
يمثل تصنيف الجريمة والتنبؤ بها اتجاها مجتمعيا معاصرا يهدف إلى الحد من الأنشطة الإجرامية أو منعها. احدى هذه الجرائم هي العنف الأسري وهي ظاهرة عالمية تعرف بجائحة الظل أو الجائحة الخفية. تهدف هذه الدراسة إلى تحليل البيانات المتعلقة بجريمة العنف المنزلي في العراق ومقارنة فعالية نماذج الغابة العشوائية وشجرة القراروبايز البسيط باستخدام مقاييس مختلفة لتطوير نموذج دقيق لتصنيف أو التنبؤ بنوع العنف المنزلي (جسدي أو غير جسدي). بعد ذلك، يتم تطبيق المصنفات الثلاثة الغابة العشوائية وشجرة القراروبايز البسيط على مجموعة البيانات لتسهيل التصنيف والتنبؤ. أظهرت نتائج التجربة الأداء المتفوق لمصنف الغابة العشوائية، حيث حقق درجة دقة بلغت (99.77%)، مقارنة بمصنف شجرة القرار (99.07%)ومصنف بايز البسيط (97.69%). و للتحقق من الصحة، تم استخدام مقاييس تصنيف

* Email: Aamer.Faeq1103a@sc.uobaghdad.edu.iq

مختلفة. أظهرت الغابة العشوائية أداءً فائقًا في جميع المقاييس مقارنة بخوارزميات شجرة القرار و بايز البسيط التي تباينت قدرات أدائها.

1. Introduction

A crime is an act punishable by law that harms both the individual and society [1]. Like other countries, Iraq faces enormous challenges related to crime, with dire consequences that threaten the safety and stability of its population [2], [3]. One of the most significant challenges is the phenomenon of domestic violence, known as the shadow pandemic or the hidden pandemic [4], defined as the occurrence of physical, sexual, or psychological abuse within relationships [5], [6]. Recently, crime prediction has gained significant popularity in the academic studies [7], [8], [9], [10].

Drawing from a review of the global literature on crime in the field of machine learning, classical statistical techniques such as decision trees (DT) and Naïve Bayes (NB) algorithms, along with modern statistical methods such as random forest (RF), have proven their effectiveness in crime prediction and classification [7], [8], [9], [10].

In the realm of literature, several machine learning approaches have been employed to examine crime data, with supervised machine learning (SL) being the most commonly utilized [8], [11], [12].

In a study conducted by Shojaee et al. (2013) [13], an experiment was carried out to enhance supervised classification learning algorithms for predicting crime status by employing two different feature selection methods tested on a real dataset. The comparisons in terms of area under curve (AUC) indicated that Naïve Bayes (0.898), K-nearest neighbor (KNN) (0.895), and neural networks (multilayer perceptron) (0.892) outperformed decision tree (J48) (0.727) and support vector machine (SVM) (0.678). Furthermore, the performance of mining results was enhanced by using the Chi-square feature selection technique. Abdul Jalil et al. (2017) [14] introduced a comparative study to evaluate the performance of filtering methods for selecting features of crime data using NB, SVM, LR, RF, random tree, multi-class classifier, and decision stump. The goal was to find a subset of features and classify crimes into three different categories; low, medium and high. Kim et al. (2018) [15] in their work, they analyzed crime data in Vancouver for 15 years (2003–2018) using two different data processing methods. K-nearest neighbor and boosted decision tree were implemented, and the decision tree outperformed, and other works, Alves et al. (2018) [16], Jha et al. (2019) [17], Aldossari et al. (2020) [18], Zhang et al. (2020) [19], Shah et al. (2021) [20], Hussain and Aljuboori (2022) [1], AlAbdouli et al. (2023) [21], and Alsubayhin et al. (2024) [22].

This research aims to analyze data and develop an algorithm that provides high accuracy for classifying and predicting domestic violence crimes in Iraq (physical or non-physical). This will be achieved by evaluating the effectiveness of the RF, DT, and NB models using datasets of domestic violence reports in Iraq. Evaluation is a crucial aspect of this research, and the effectiveness of these models will be comprehensively assessed using various evaluation metrics (confusion matrix, recall, precision, F1-score, specificity, log-loss, ROC, AUC, and a five-fold cross-validation approach) to ensure their efficacy in classifying and predicting domestic violence crimes during the specified period. This study focuses on fifteen Iraqi governorates, excluding the Kurdistan region, which includes Erbil, Sulaymaniyah, and Dohuk. The findings of this study can inform public policies and guide efforts towards more effective measures to combat domestic violence in Iraq. Finally, this research is the first of its kind to utilize fifteen Iraqi governorates as a case study in this field.

2. Classification Algorithm

Among the classical and modern statistical methods, the following classification algorithms were applied in this study:

2.1 Naive Bayes (NB)

Naive Bayes learning (NB) [23], [24], [25] refers to the construction of a Bayesian probabilistic model that assigns a posterior class probability to an instance$P(Y=y|\boldsymbol{X}=\boldsymbol{x})$. A simple naive Bayes classier uses these probabilities to assign an instance to a class. By applying Bayes' theorem and simplifying the notation a little, we obtain the relationship given in Eq. (1).

$$P(y|\boldsymbol{x})=\frac{P(\boldsymbol{x}|y)P(y)}{p(x)}, \tag{1}$$

where $P(y|\boldsymbol{x})$ is the posterior probability, $p(\boldsymbol{x}|y)$ is the likelihood, $P(y)$ is the class prior probability, and $p(\boldsymbol{x})$ is the predictive probability density function.

2.2 Decision trees (DT)

Decision trees (DT) [23], [25], [26] or classification trees are a versatile machine learning algorithm that is used to solve classification and regression problems. It is one of the very powerful algorithms. It is a flowchart-like tree structure where each internal node indicates a feature, branches indicate rules, and leaf nodes indicate the result of the algorithm, which can be understood by the following measures:

- **Entropy**

Entropy measures the level of disorder or uncertainty in a given dataset or system, mathematically as in Eq. (2).

$$Entropy=-\sum_{i=1}^{n} p_i \log(p_i), \tag{2}$$

where *n* is the number of classes, p_i is the probability of occurrence class *i.* Log represents the natural logarithm.

- **Gini index (Gini impurity)**

The Gini index serves as a measure of inequality in sample as shown in Eq. (3) [27].

$$Gini\ index=1-\sum_{i=1}^{n} p_i^2 \tag{3}$$

where n is the number of classes and p_i^2 is the square of the probability p_i of occurrence class *i.*

2.3 Random Forest (RF)

Random Forest (RF) [25], [28] is an ensemble learning method and supervised machine learning algorithm based on a decision tree approach. In a RF, multiple decision trees are generated using training data, which is similar to a forest containing many trees. Each tree is built using a random subset of features sampled independently from the training data, ensuring diversity between trees. The predictions of all trees are combined through a majority voting process, where the most common prediction becomes the final prediction of the RF model. This approach is based on boosting and bagging principles.

3. Case Study: Domestic Violence Crimes in Iraq

The workflow used to construct a model for classifying the categories of domestic violence crimes, which are categorized as physical and non-physical violence in Iraq is depicted in Fig.

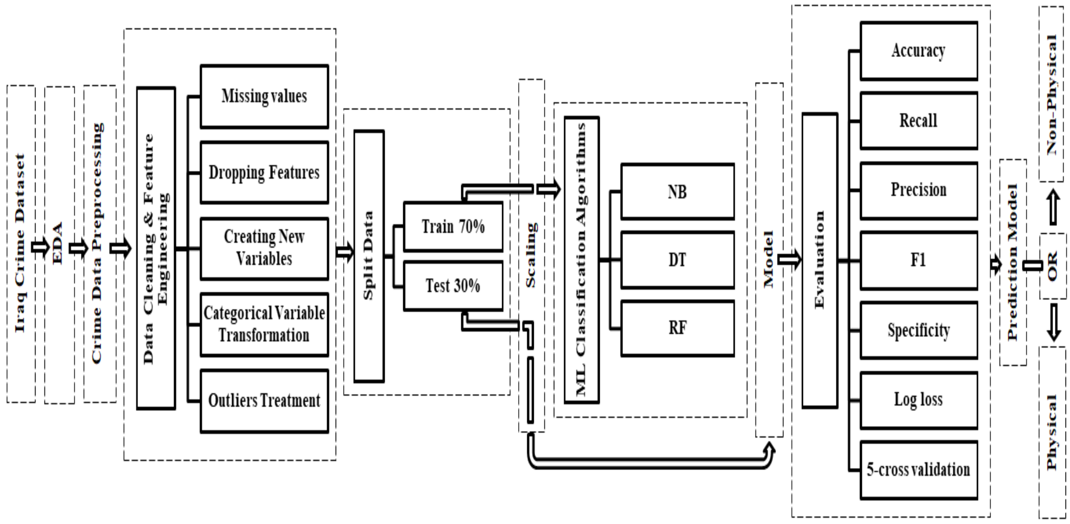


Figure 1: Workflow model

3.1 IRAQ-Domestic Violence Crimes Dataset

In this research, datasets in the context of criminology were utilized to build the classifier. These datasets comprise records of incidents related to domestic violence reported through the 497 hotline operated by the Iraqi Ministry of Interior, Department of Relations and Media department, and Department of Community Police.

The Iraqi dataset consists of 20 different features and 1439 tuples of categorical and numerical nature. These incidents occurred in 15 Iraqi governorates during the period from January 1, 2022, to December 31, 2022. Each entry in our dataset corresponds to a specific crime, described as in Table 1.

Table 1. Feature and classes of the crime dataset in Iraq

Features	Classes	Features	Classes	Features	Classes
Date	Continuous X1	**Marital Status**	Categorical X4	**Governorate**	Categorical X7
Day of Week	Categorical X2	**Educational Level**	Categorical X5	**Place of Residence**	Categorical X8
Sex	Categorical X3	**Profession**	Categorical X6	**City**	Categorical X9
Age	Continuous X10	**If the type of violence is physical, what type of violence?**	Categorical X14	**Reason for the violence**	Categorical X16
Housing Type	Categorical X11	**Perpetrator of Violence**	Categorical X15	**Type of Violence**	Categorical **Y**
Longitude & Latitude	Continuous X12 & X13				

To ensure the protection of the privacy of domestic violence victims, the disclosed information maintains a certain level of anonymity without specifying exact locations.

3.2 Exploratory Data Analysis (EDA), Data Pre-processing, Data cleaning and Transformation, Training and testing dataset

The process of analysing a dataset helps to minimize the risks inveterate in the decision-making procedure by providing an understanding of statistical values and numbers such as graphs, pictures, tables, and charts [29] as in Table 2 and Fig. 2 (a) and (b).

Table 2: Description of the crime dataset in Iraq

	Governorate	No. of Violence Crimes	Percentage		Governorate	No. of Violence Crimes	Percentage
1	Baghdad	747	51.911	9	Karbala	45	3.1271
2	Nineveh	106	7.3662	10	Babil	41	2.8492
3	Salah al-Din	82	5.6984	11	Al-Najaf	38	2.6407
4	Diyala	77	5.3509	12	Wasit	25	1.7373
5	Al-Anbar	67	4.6560	13	Al-Qadisiyah	20	1.3898
6	Dhi-Qar	64	4.4475	14	Maysan	13	0.9034
7	Basra	60	4.1695	15	Al-Muthanna	5	0.3474
8	Kirkuk	49	3.4051				

(a) (b)

Figure 2 : Count of violence crime w.r.t. (a). type of violence, sex, and age. (b). marital status, education level, and profession.

By exploring the data presented in Table 2, it is observed that the rate of domestic violence crimes varies across different governorates. Analysis of the distribution of domestic violence crime categories provides a clear indication that physical violence, as shown in Fig. 2 (a), exhibits the highest recurrence rates in Iraq, indicating uneven distribution. Regarding gender, as shown in Fig. 2 (a), it is noteworthy that females were the highest reporters of domestic violence crimes. Concerning age, the histogram displayed in Fig. 2 (a) illustrates the distribution of data on the ages of petitioners in relation to domestic violence. It indicates an increase at certain ages and a decrease at others, with a more pronounced distribution towards higher values. As for marital status, as shown in Fig. 2 (b), the majority of reported cases of physical domestic violence occurred among married individuals, followed by single individuals, respectively. In terms of education level, as shown in Fig. 2 (b), reports of domestic violence cases were most prevalent among individuals with secondary education, followed by the graduates of universities or institutes, then those with primary education, and finally individuals with reading and writing. Additionally, regarding occupation, as depicted in Fig. 2 (b), the majority of reported of cases of domestic violence were among non-employees compared to the total number of employees, students, private sector employers, and retirees. Figure 3 represents the number of reports of domestic violence crimes for the top 80 regions, where the Doura region received the largest share with 46 reports, followed by Sadr City with 38 reports.

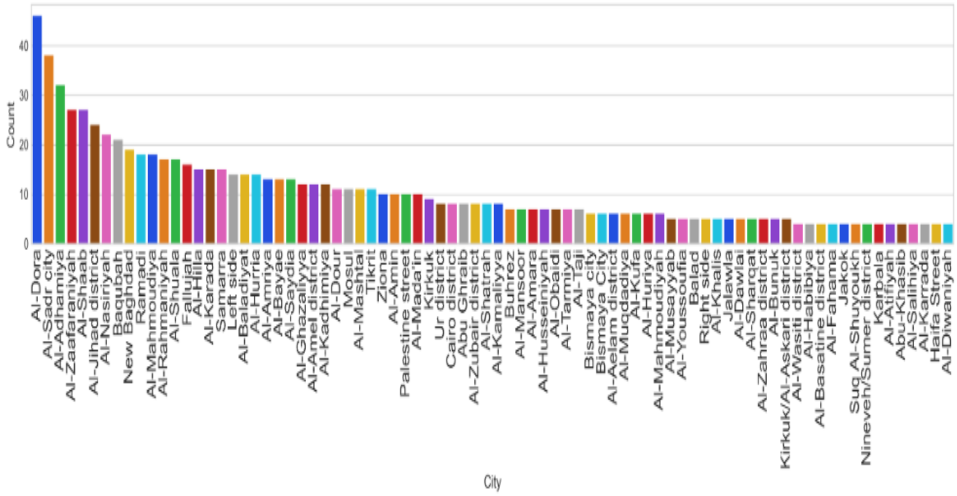


Figure 3: Distribution of domestic violence crimes in each region

Also, Fig. 4 represents the distribution of place of residence and housing type. Most of the reports were from city residents and private property, respectively.

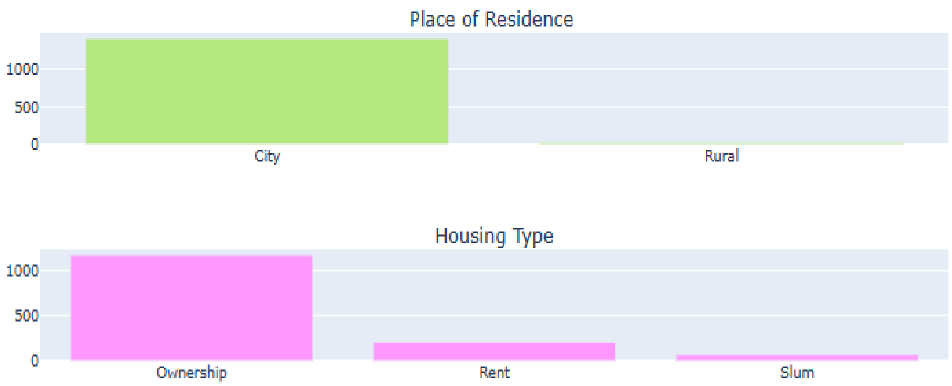


Figure 4: Distribution of domestic violence crimes w.r.t. place of residence and housing type

Figure 5 shows the top 10 perpetrators of violence. It is clear that the perpetrator of violence in the first place is the husband followed by the father respectively. As for any type of violence, Fig. 6 (a) displays the first six categories, and the most common type was hitting with the hand. Regarding the top 10 causes of violence, Fig. 6 (b) shows that family problems are the most common, with a total of 791 cases, which should be interpreted by psychiatrists and specialists.

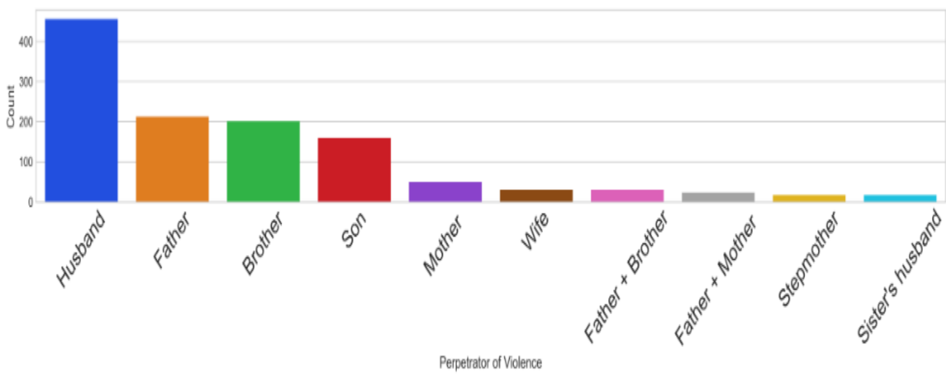
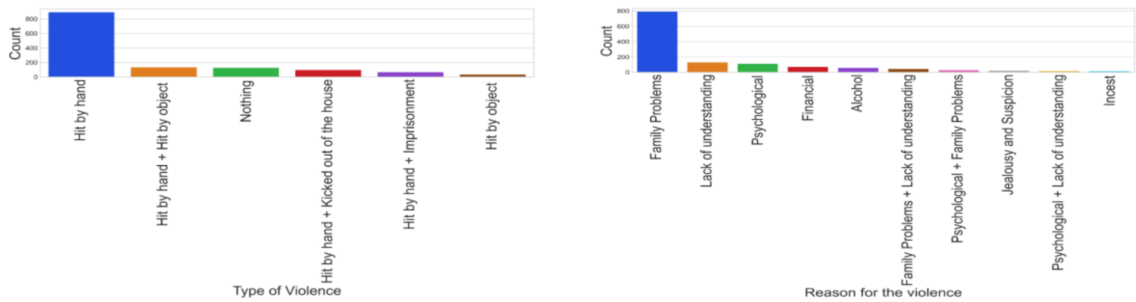


Figure 5: Distribution of perpetrators of violence



(a) (b)

Figure 6: (a) Distribution of any type of violence, (b) Distribution of reason for the violence

In order to further analyze the dataset, it is necessary to preprocess the dataset by splitting the attributes in the date feature, dividing the timestamp into date, month, year, hour, minute, time period, hour type, weekend, working hour, and season as in Fig. 7.

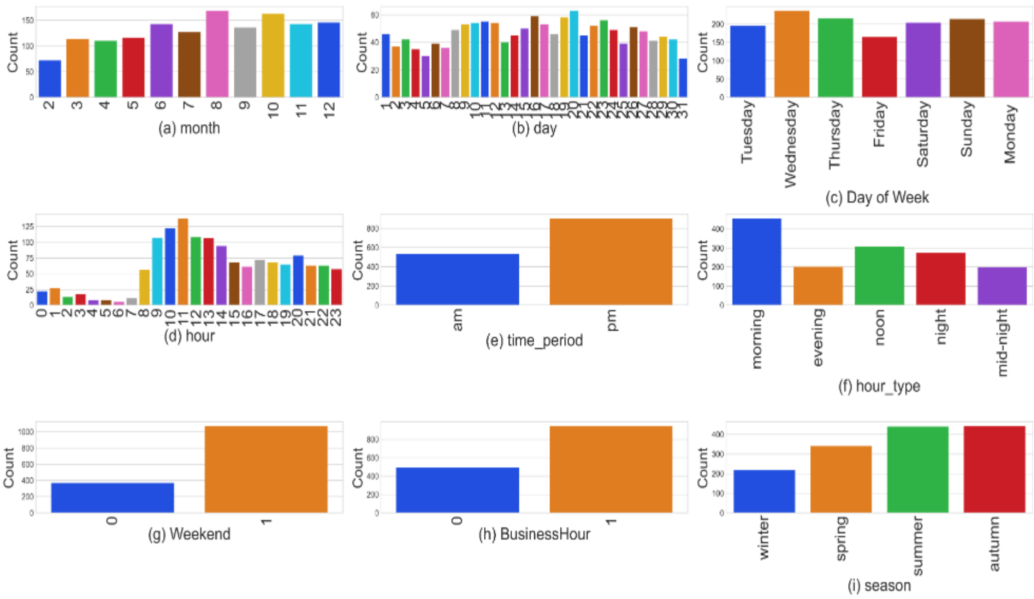


Figure 7: Distribution of domestic violence reports by month, day, days of week, hour, time period, hour type, weekend, business hour, and season

In Fig. 7 (a), August was the month with the highest number of reports, followed by November and December. February was the month with the lowest number of crime incidents, and similarly, from Fig. 7 (b) and (c), it appears that reports of domestic violence occur on all days of the month as well as on all days of the week, Fig. 7 (d) showing that from 8 am to 23:59, the most reports were reported, with the peak being from 9 to 14, compared to other hours. Most of the reports occurred at 11 noon and began to stabilize from 15 noon until 23:59, where the reports were somewhat similar. Figure 7 (e) shows that the morning period is the most reported. In addition, regarding the type of time also, Fig. 7 (f) represented the largest number of domestic violence reports from the morning period, followed by noon and night. As for weekends, Fig. 7 (g) shows that reports of domestic violence are concentrated during the days of the week, based on the fact that the weekend in Iraq is Friday and Saturday. As for the working hours, Fig. 7 (h) shows that reports are concentrated during typical work hours compared to off-hours. Regarding seasons, Figure 7 (i) shows the distribution of reported domestic violence cases by season. It was noted that reporting increased from winter, spring, summer, and fall, respectively, and summer and fall had similar reporting rates.

The dataset underwent cleaning by removing some unnecessary features that did not add marginal value. Optimal classification models necessitate continuous data. To accomplish this objective, each categorical feature was converted into a single fast encoding matrix using the Label-Encoder function within the “Pre-processing” module of sklearn in Python. Table 3 displays the correlation coefficient between our target and other features, using Spearman’s correlation coefficient.

Table 3: The correlation coefficient values between target and other features

Features	correlation coefficient	Features	correlation coefficient	Features	correlation coefficient
Year	NaN	**Marital Status**	-0.001216	**Governorate**	-0.056826
Day of Week	0.002805	**Educational Level**	0.045776	**Place of Residence**	NaN
Sex	0.001773	**Profession**	0.019905	**City**	-0.031956
Age	-0.117556	**Physical Violence Type**	-0.731119	**Reason for the violence**	-0.024379
Housing Type	-0.072595	**Perpetrator of Violence**	-0.058000	**Type of Violence**	1.000000
Longitude	-0.002979	**Minute**	0.018734	**Month**	0.091500
Day	0.012626	**Latitude**	-0.017966	**Business Hour**	-0.007217
Hour	-0.008549	**Weekend**	-0.008830	**Time Period**	-0.015737
Hour Type	-0.036304	**Season**	-0.054715		

NaN in Table 3 indicates that it was not possible to calculate the correlation coefficient for the target and year because the data used are for one year and for the target and place of residence because the data are not diverse enough.

The dataset is divided into two parts, the test and training data sets at 30% and 70% respectively.

3.3 *Performance Measures*

The model’s classification performance on Iraqi domestic violence crime data can be assessed using a confusion matrix, depicted in Table 4. This matrix summarizes how accurately the model categorizes data, aligning with both actual and predicted classifications in the dataset.

Table 4: Confusion matrix

PRIDECT VALUE		ACTUAL VALUE	
		Positive (1)	**Negative (0)**
	Positive (1)	True Po. TP	False Po. FP
	Negative (0)	False Ne. FN	True Ne. TN

Through the confusion matrix, various verification measures can be derived, such as accuracy, precision, recall, f1, and specificity as appeared in Eqs. (4) to (8) respectively.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{4}$$

$$Precision = \frac{TP}{TP+FP} \tag{5}$$

$$Recall = \frac{TP}{TP+FN} \tag{6}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision+Recall} \tag{7}$$

$$Specificity = \frac{TN}{TN+FP} \tag{8}$$

For an unbalanced data set, such as the current Iraq domestic violence crimes data set with an unbalanced distribution of classes, it is necessary to compare accuracy using other measures such as the log loss measure represented by Eq. (9).

$$Logloss = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}\log(p_{ij}), \tag{9}$$

where y_{ij} indicates whether i^{th} sample belongs to j^{th} class or not and p_{ij} indicates the probability of i^{th} sample belongs to i^{th} class. M is the total number of attributes and N is the total number of samples.

Finally, for all models, a 5-fold cross validation technique was employed [30], [31], [32], [33], [34] as illustrated in Fig. 8.

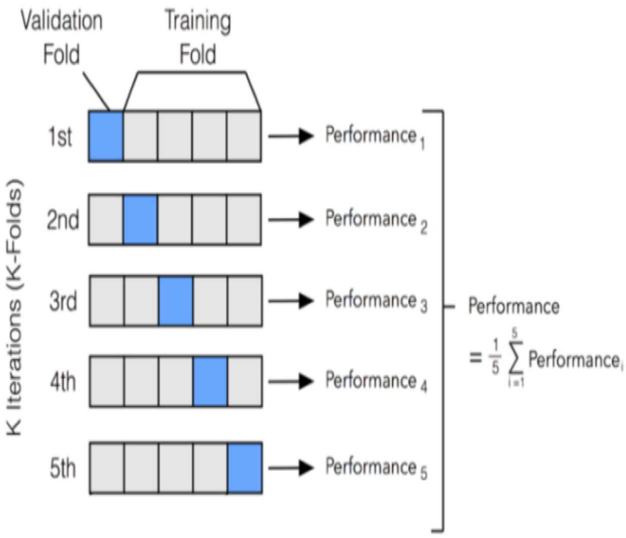


Figure 8: 5-fold cross-validations [35]

4. Results and Discussions

This section discusses the obtained results. Regarding the confusion matrix, Table 5 illustrates the outcome of the proposed model for three methods. Based on this analysis, it can be inferred that the RF algorithm outperformed the other two methods in this scenario, demonstrating excellent performance by correctly classifying the samples without any false negatives (FNs).

When evaluating accuracy compared to the proposed metrics, the results are depicted in Table 6. It is observed that the RF algorithm attained the highest accuracy compared to other algorithms.

Table 5: Confusion matrix values obtained using the proposed algorithms

Algorithm	TP	FP	TN	FN
DT	383	1	42	6
RF	389	1	42	0
NB	381	2	41	8

Table 6: The result of accuracy on test, precision, recall, specificity and F1 score

	Model	Test Acc.	Precision	Recall	F1	Specificity
1	**DT**	99.07	99.74	99.23	99.48	97.67
2	**RF**	**99.77**	**99.74**	**100.0**	**99.87**	**97.67**
3	**NB**	97.69	99.48	97.94	98.70	95.35

Furthermore, among the algorithms, RF offered a good precision of 99.74%, meaning that it produces results that are more relevant to identifying domestic violence crimes than irrelevant results.
When calculating the recall of the algorithms, the RF obtained an estimate of 100% for the model's ability to effectively identify and target actual domestic violence crimes, which indicates that it does not contain false negative results. RF also obtained a high F1 score compared to other algorithms, with a high value of 99.87%, which means that the model achieves low false positives and low false negatives. Finally, the RF achieved a high ability of 97.67% in measuring specificity to correctly distinguish between true negative results and avoid false positive results and the Fig. 9 shows the results graphically.

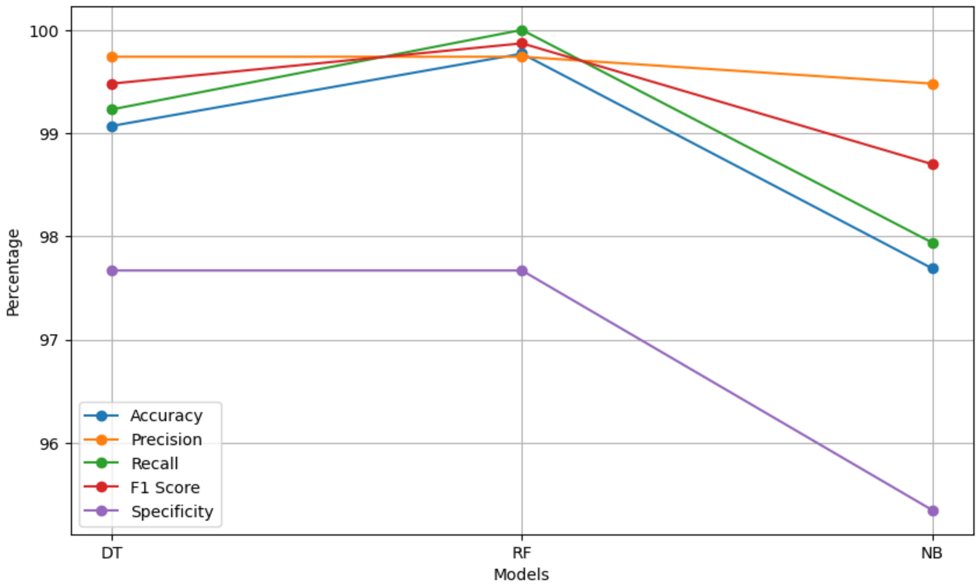
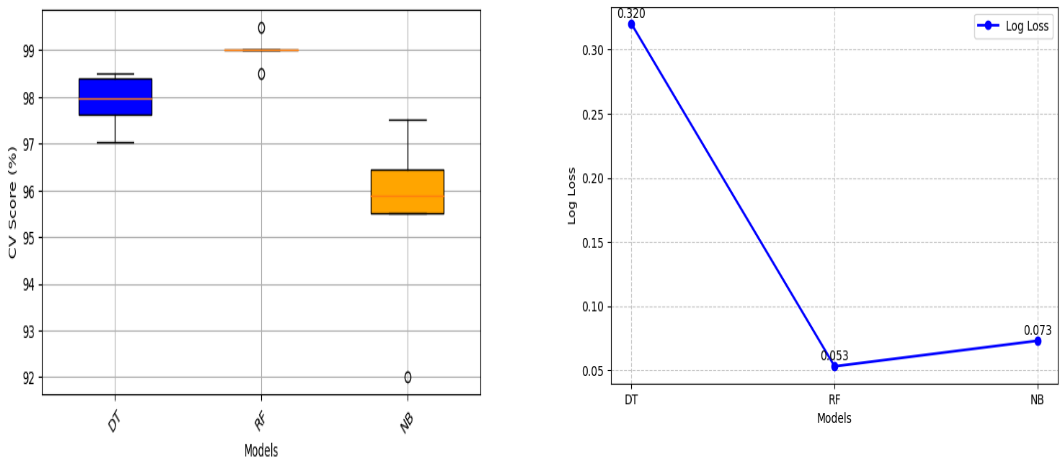


Figure 9: Performance metrics for different models

Table 7 presents the results of the five-fold cross-validation measure, providing an objective assessment of the performance of the three selected models, utilizing all data for training and evaluation. RF demonstrated the ability to outperform DT and NB in all five verification runs and also on average, as depicted in Fig. 10 (a). Through the log loss measurement values for the three algorithms, it is observed that the RF and NB algorithms outperform with scores of 0.053 and 0.073, respectively. This indicates that these models exhibited better performance in predicting the categories as shown in Fig. 10 (b), compared to the DT algorithm which scored 0.320.

Table 7: The results of the 5-fold cross-validation measure (DT, RF, and NB)

	Model	CV Score 1	CV Score2	CV Score 3	CV Score 4	CV Score 5	Average CV
1	**DT**	98.51%	97.03%	98.01%	98.51%	97.51%	97.91%
2	**RF**	**98.51%**	**99.01%**	**99.50%**	**99.00%**	**99.00%**	**99.01%**
3	**NB**	95.52%	97.52%	92.02%	96.52%	95.52%	96.23%



(a) (b)

Figure 10: (a) 5-fold cross-validation scores for proposed algorithms, (b) Log Loss for proposed algorithms

Arriving at the AUC results (area under the ROC curve) for the three algorithms in binary classification, as in Fig. 11. Which shows the superiority of the RF algorithm with AUC = 1.0, which means that the RF model is able to completely distinguish between the positive and negative classes without any ambiguity, and this is a very good and ideal performance. It is followed by NB and DT respectively.

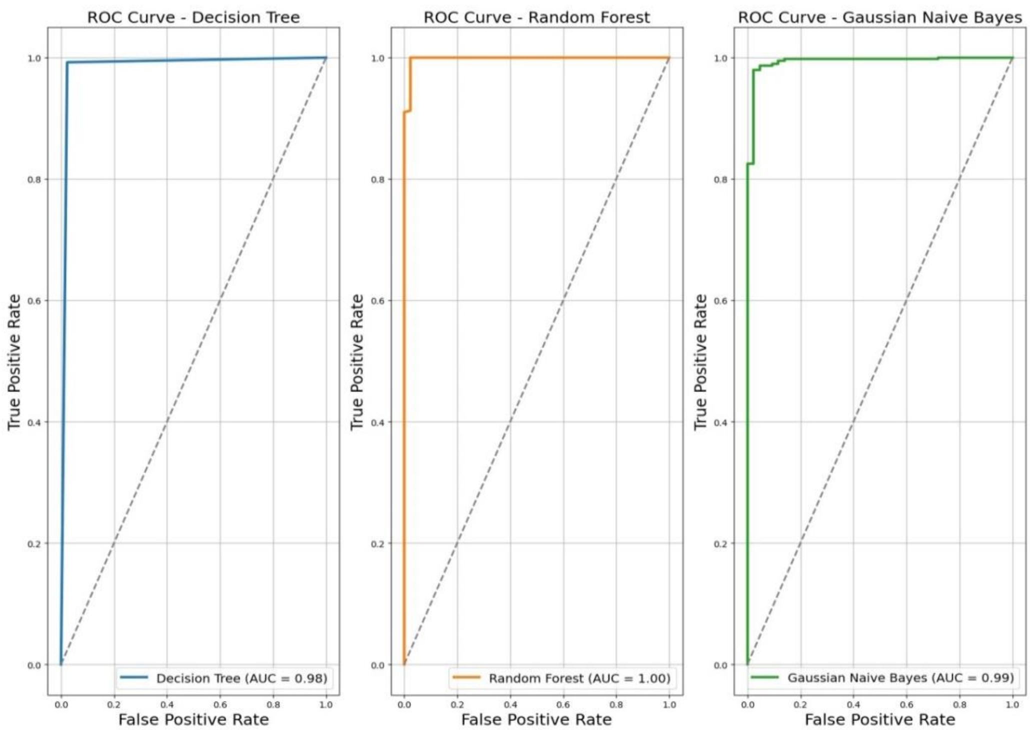


Figure 11: Receivers operating characteristic (ROC) curve proposed algorithms

5. Conclusions

This study compared the accuracy degrees of the DT, RF and NB algorithms in classifying and predicting the type of domestic violence (physical and non-physical) with different metrics. The result of this study indicates the superior performance of the RF classifier as it achieved an accuracy score of (99.77) compared to the DT classifier (99.07) and NB (97.69).

For validation, different classification metrics were used, RF showed superior performance in all metrics compared to DT and NB algorithms whose performance capabilities varied.
The experiment was performed on a desktop computer (all-in-one), running Windows 10 pro, Intel (R) Core (TM) i5-7400T CPU @ 2.40 GHz processor, with 8 GB RAM and OS 64-bit.
These positive findings highlight the need for enhanced collaboration among researchers and specialists in combating domestic violence. They emphasize the importance of exchanging knowledge, experiences, and developing comprehensive strategies to address this persistent issue. Moreover, this study suggests future advancements in criminal and intelligence analysis using mathematical machine learning algorithms.
Looking to the future, there is a plan to continue applying other classification algorithms to the crime dataset to evaluate their performance.

References

[1] F. S. Hussain and A. F. Aljuboori, "A Crime Data Analysis of Prediction Based on Classification Approaches," *Baghdad Science Journal,* vol. 19, no. 5, pp. 1073-1077, 2022.

[2] H. H. Hussein and A. T. Abdulameer, "Crime Prediction Using Big Data Analysis," in *Proceedings of 2nd International Multi-Disciplinary Conference Theme: Integrated Sciences and Technologies, IMDC-IST* , Sakarya, Turkey, 2021, doi: 10.4108/eai.7-9-2021.2314943.

[3] L. Elluri, V. Mandalapu and N. Roy, "Developing Machine Learning Based Predictive Models for Smart Policing," in *2019 IEEE International Conference on Smart Computing (SMARTCOMP)*, Washington, DC, USA, 2019, pp. 198-204.

[4] A. J. Rashid, "Domestic Violence in Iraq in Light of the Repercussions of the Corona Crisis," Bayan Center.org, 2020. [Online]. Available: https://www.bayancenter.org/wp-content/uploads/2020/10/8767589.pdf.

[5] UN Women , "Violence Against Women in the Iraqi Society: A Legal, Judicial and Statistic Study," Iraq.UN Women.org, 2021. [Online]. Available: https://iraq.unwomen.org/sites/default/files/Field%20Office%20Iraq/Attachments/Publications/2021/SVAW/Violence%20Against%20WomenAbstract%20Englishcompressed%201.pdf.

[6] L. E. P. Center, " Domestic Violence," International Association of Chiefs of Police, Alexandria, Virginia, U.S.A. , 2021. [Online]. Available: https://www.theiacp.org/sites/default/files/2021-07/Domestic%20Violence%20FULL%20-%2006292020.pdf.

[7] V. Mandalapu, L. Elluri, P. Vyas and N. Roy, "Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions," *IEEE Access,* vol. 11, pp. 60153-60170, 2023.

[8] K. Jenga, C. Catal and G. Kar, "Machine Learning in Crime Prediction," *Journal of Ambient Intelligence and Humanized Computing,* vol. 14, p. 2887–2913, 2023.

[9] J. Yin, "Crime Prediction Methods Based on Machine Learning: A Survey," *Computers, Materials & Continua,* vol. 74, no. 2, pp. 4601-4629, 2023.

[10] U. M. Butt, S. Letchmunan, F. H. Hassan, M. Ali, A. Baqir and H. H. R. Sherazi, "Spatio-Temporal Crime HotSpot Detection and Prediction: A Systematic Literature Review," *IEEE Access,* vol. 8, pp. 166553-166574, 2020.

[11] N. Esquivel, O. Nicolis, B. Peralta and J. Mateu, "Spatio-Temporal Prediction of Baltimore Crime Events Using CLSTM Neural Networks," *IEEE Access,* vol. 8, pp. 209101-209112, 2020.

[12] E. F. Khalaf and A. H. Taresh, "Survey: Crime Prediction Using Machine Learning Approach," *Journal of Al-Qadisiyah for Computer Science and Mathematics,* vol. 14, no. 3, p. 51–57, 2022.

[13] S. Shojaee, A. Mustapha, F. Sidi and M. A. Jabar, "A Study on Classification Learning Algorithms to Predict Crime Status," *International Journal of Digital Content Technology and its Applications,* vol. 7, no. 9, pp. 361-369, 2013.

[14] M. Abdul Jalil, F. Mohd and . N. M. M. Noor, "A Comparative Study to Evaluate Filtering Methods for Crime Data Feature Selection," *Procedia Computer Science,* vol. 116, pp. 113-120, 2017.

[15] S. Kim, P. Joshi, P. . S. Kalsi and P. Taheri, "Crime Analysis Through Machine Learning," in *2018 IEEE 9th Annual Information Technology,Electronics and Mobile Communication Conference (IEMCON)*, Vancouver, BC, Canada, 2018, pp. 415-420.

[16] L. G. Alves, H. V. Ribeiro and F. A. Rodrigues, "Crime Prediction Through Urban Metrics and Statistical Learning," *Physica A: Statistical Mechanics and its Applications,* vol. 505, p. 435–443, 2018.

[17] P. Jha, R. Jha and A. Sharma, "Behavior Analysis and Crime Prediction Using Big Data and Machine Learning," *International Journal of Recent Technology and Engineering (IJRTE),* vol. 8, no. 1, pp. 461-468, 2019.

[18] B. S. Aldossari, F. M. Alqahtani, N. S. Alshahrani, M. M. Alhammam, R. M. Alzamanan and N. Aslam, "A Comparative Study of Decision Tree and Naive Bayes Machine Learning Model for Crime Category Prediction in Chicago," in *Proceedings of 2020 6th International Conference on Computing and Data Engineering*, Sanya, China, January, PP. 34 - 38, 2020.

[19] X. Zhang, L. Liu, L. Xiao and J. Ji, "Comparison of Machine Learning Algorithms for Predicting Crime Hotspots," *IEEE Access,* vol. 8, pp. 181302-181310, 2020.

[20] N. Shah, N. Bhagat and M. Shah, "Crime Forecasting: A machine Learning and Computer Vision Approach to Crime Prediction and Prevention," *Visual Computing for Industry, Biomedicine, and Art,* vol. 4, no. 9, p. 1–14, 2021.

[21] S. K. AlAbdouli, . A. F. Alomosh, A. B. Nassif and Q. Nasir, "Comparison of Machine Learning Algorithms for Crime Prediction in Dubai," *International Journal of Advanced Computer Science and Applications(IJACSA),* vol. 14, no. 9, pp. 169-173, 2023.

[22] A. Alsubayhin, M. S. Ramzan and B. Alzahrani, "Crime Prediction Model Using Three Classification Techniques: Random Forest, Logistic Regression, and LightGBM," *(IJACSA) International Journal of Advanced Computer Science and Applications,* vol. 15, no. 1, pp. 240-251, 2024.

[23] M. A. Kashmoola, M. K. Ahmed and N. . Y. A. Alsaleem, "Network Traffic Prediction Based on Boosting Learning," *Iraqi Journal of Science,* vol. 63, no. 9, pp. 4047-4056, 2022.

[24] F. H. Fadel and S. F. Behadili, "A Comparative Study for Supervised Learning Algorithms to Analyze Sentiment Tweets," *Iraqi Journal of Science,* vol. 63, no. 6, pp. 2712-2724, 2022.

[25] A. F. A. H. Alnuaimi and . T. H. Albaldawi , "An Overview of Machine Learning Classification Techniques," in *Fifth International Scientific Conference of Alkafeel University (ISCKU)*, Najaf, Iraq, BIO Web Conf., Vol. 97, Art. no. 00133, pp.1-24, Apr. 2024.

[26] A. Ram and Meenakshi, "Short Review on Machine Learning and its Application," *Specialusis Ugdymas / Special Education ,* vol. 1, no. 43, pp. 9894-9902, 2022.

[27] A. Saud, S. Shakya and B. Neupane, "Analysis of Depth of Entropy and Gini Index Based Decision Trees for Predicting Diabetes," *Indian Journal of Computer Science,* vol. 6, no. 6, pp. 19-28, 20 Feb 2022.

[28] A. Cutler, D. R. Cutler and J. R. Stevens, "Random Forests," in *Ensemble Machine Learning: Methods and Applications*, New York, Springer , 2012, pp. 157-176.

[29] I. Pradhan, K. Potika, M. Eirinaki and P. Potikas, "Exploratory Data Analysis and Crime Prediction for Smart Cities," in *the 23rd International Database Applications & Engineering Symposium (IDEAS)*, Athens Greece, Art. no. 4, pp. 1-9, June 2019.

[30] U. Saeed, M. Sarim, A. Usmani and A. Mukhtar, "Application of Machine learning Algorithms in Crime Classification and Classification Rule Mining," *Research Journal of Recent Sciences,* vol. 4, no. 3, pp. 106-114, 2015.

[31] M. Khan, A. Ali and Y. Alharbi, "Predicting and Preventing Crime: A Crime Prediction Model Using San Francisco Crime Data by Classification Techniques," *Complexity,* vol. 2022, no. 1, pp. 1-13, 2022.

[32] D. Sardana, S. Marwaha and R. Bhatnagar, "Supervised and Unsupervised Machine Learning

Methodologies for Crime Patterns Analysis," *International Journal of Artificial Intelligence and Applications (IJAIA),* vol. 12, no. 1, pp. 83-99, 2021.

[33] G. H. Larios, "Case Study Report, San Francisco Crime Classification," Università degli Studi del Piemonte Orientale, Italy, 2016.

[34] D. Berrar, "Cross-validation," in *Encyclopedia of Bioinformatics and Computational Biology*, Oxford, Academic Press, 2018, p. 542–545.

[35] H. B. Rezaei, A. Amjadian, M. V. Sebt, R. Askari and A. Gharaei, "An Ensemble Method of the Machine Learning to Prognosticate the Gastric Cancer," *Annals of Operations Research,* vol. 328, no. 10, pp. 151-192, 13 Jan. 2023.