

Article

Not peer-reviewed version

Machine Learning Applications to Identify Young Offenders Using Data from Cognitive Function Test

María Claudia Bonfante , [Juan Contreras Montes](#) ^{*} , [Mariana Pino](#) , Ronald Ruiz , Gabriel González

Posted Date: 28 September 2023

doi: 10.20944/preprints202309.1845.v1

Keywords: cognitive functions; machine learning; feature selection; violence risk assessment



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Machine Learning Applications to Identify Young Offenders Using Data from Cognitive Function Test

María Claudia Bonfante ¹, Juan Contreras Montes ^{1*}, Mariana Pino ², Ronald Ruiz ² and Gabriel González ³

¹ Faculty of Engineering, Institución Universitaria de Barranquilla; Barranquilla 080002, Colombia, maria.bonfante@unisnu.edu.co; juancontrerasm@itsa.edu.co.

² Department of Psychology, Universidad Autónoma del Caribe; Barranquilla 080020, Colombia, mariana.pino@uac.edu.co; ronaldruzp@gmail.com.

³ Program Director, Fundación Hogares Claret; Barranquilla 080003, Colombia, gabriel.gonzalez@fhclaret.org.

* Correspondence: juancontrerasm@itsa.edu.co.

Abstract: Machine Learning techniques can be used to identify whether deficits in cognitive functions contribute to antisocial and aggressive behavior. This paper initially presents the results of tests conducted on delinquent and non-delinquent youths to assess their cognitive functions. The dataset extracted from these assessments, consisting of 37 predictor variables and one target, was used to train three algorithms that aim to predict whether the data corresponds to that of a young offender or a non-offending youth. Prior to this, statistical tests were conducted on the data to identify characteristics that exhibited significant differences in order to select the most relevant features and optimize the prediction results. Additionally, other feature selection methods, such as Boruta, RFE, and Filter, were applied, and their effects on the accuracy of each of the three machine learning models used (SVM, RF, and KNN) were compared. 80% of the data were utilized for training, while the remaining 20% were used for validation. The best result was achieved by the K-NN model trained with 19 features selected by the Boruta method, followed by the SVM model trained with 24 features selected by the filter method.

Keywords: cognitive functions; machine learning; feature selection; violence risk assessment

1. Introduction

In several developed countries, such as the United States, youth violence has become the third leading cause of death, and the leading cause among the African American community, among young people aged 10 to 24. For this reason, the Centers for Disease Control and Prevention (CDC) funds Youth Violence Prevention Centers (YVPCs), which use surveillance data to monitor youth violence and assess the impact of their interventions. Access to data is becoming increasingly possible as public health surveillance has mandated the systematic collection and analysis of data [1], which has generated large volumes of data available to researchers. The use and analysis of data have been crucial in trying to understand the alarming rates of violent exposure among low-income youth [2].

The assessment of violence risk has become a key element of the criminal justice system. Three benefits can be generated in the application of violence risk assessment tools: 1) Results can be used to identify more suitable treatments for the individual; 2) It can generate the necessary evidence for the early release of an individual; 3) It can determine the non-early release of an individual, thereby preventing recidivism [3].

One of the most widely used tools for assessing violence risk is the HCR-20 (Historical Clinical Risk Management-20), which includes 20 risk factors grouped into 3 categories: Historical (10 factors), Clinical (5 factors), and Future Risk (5 factors), and generates 3 risk ratings that can be summarized as low, moderate, or high [4], [5].

Another tool that has been used for assessing the risk of violence in hospitalized children and adolescents is the BRACHA method (Brief Rating of Aggression by Children and Adolescents), which has also been used to assess violence risks in schools [6]. This method allows for categorizing children and adolescents into lower or higher risk groups for aggression and violence.

Despite there currently being more than 200 violence risk assessment tools [7], the accuracy of these tools is still below expectations, with a significant presence of false positives and false negatives, with false positives being more common than false negatives [8]. Variables or risk factors associated with a higher probability of an individual acting violently or aggressively include, among others, criminogenic needs, demographic aspects, socioeconomic status, and intelligence. These factors are divided into two categories: static and dynamic. Static factors cannot be changed (age at the time of the first arrest, criminal history, neighborhood, abuse, etc.). Dynamic factors could be modifiable (impulsivity, job skills, drug use, etc.).

Forensic psychiatry is a branch of psychiatry that studies individuals with mental disorders who pose a risk to the public. Forensic psychiatrists support the criminal justice system by investigating the correlation between mental disorders and criminal behaviors. They assess the violence risks of offenders in prisons or secure hospitals and of individuals in the community with mental disorders. Studies conducted in 24 countries, both low and high-income, showed that one in eight men and one in sixteen women will subsequently commit a serious crime after leaving a psychiatric facility [9]. Due to the high prevalence of criminal acts committed by individuals with severe mental illnesses, the effort to predict the risk of criminal acts by individuals after being discharged from psychiatric facilities is becoming increasingly significant [10].

Fields of forensic psychiatry, such as the interaction of psychopathology, offense, and aggression, have not been sufficiently investigated, so this interaction has not been able to be fully understood. What is clear to the scientific community is that psychiatric illnesses and pathological behavior disorders are not influenced by single and independent factors and cannot be modeled as a linear process. Therefore, and because a large part of psychiatric phenomena do not have a monocausal origin but develop from the interaction of various factors, the application of Artificial Intelligence, supervised learning, and advanced statistics opens up new possibilities in the field of forensic psychiatry. Researchers have studied and compared models to identify the factors that distinguish offender patients with schizophrenia spectrum disorders (SSD), using machine learning algorithms applied to a dataset of 370 patients, achieving the highest accuracy with support vector machines (SVM), balanced accuracy of 77.6%, and an AUC of 0.87 [11].

In [12], the enormous potential of applying machine learning (ML) and Deep Learning in forensic psychiatry and violence risk assessments is recognized; however, it also raises significant risks that can arise from the use of these technologies in decisions to hospitalize or medicate individuals, in sentencing recommendations, or specific police surveillance, as the value systems guiding these technologies are defined by those who design them, so it recommends analyzing the ethical implications that may be involved in the use of these techniques.

Researchers in Denmark developed a supervised ML model designed to assess the risk of people in the general psychiatric system, according to their future probability of requiring treatment in a forensic institution. The goal is to predict whether people will commit a crime during or after outpatient treatment or after discharge from hospital care, leading to a court-ordered psychiatric admission in Denmark, through which criminals with mental illnesses are referred for forensic psychiatric treatment. This study used a sample of 45,720 psychiatric patients, of whom 1% committed a crime. The dataset comprises 39 predictors, which were divided into three groups: socioeconomic, psychiatric history, and criminal history. Four ML models were used: Logistic Regression, Random Forest (RF), XGBoost; and LightGBM. The performance metrics for each model are presented, including Recall, Accuracy, and F1-Score [13].

In [14], a series of HARM (Hamilton Anatomy of Risk Assessment) models using machine learning techniques are proposed to predict longitudinal physical aggression in patients with schizophrenia in forensic settings. Data from 151 patients were used, with follow-ups at 4, 12, and 18 months. The R language was used, and the following nine machine learning algorithms were

implemented: Boosted Logistic Regression, Elastic Net, Lasso Regression, k-nearest neighbors, Adaptive Boosting, Extreme Gradient Boosting, Random Forest, Bagged CART, Conditional Forest. The best performance was achieved with the Random Forest model, with a Balanced Accuracy of 86.60%, Accuracy = 87.33%, and an AUC of 0.914.

Some researchers believe that the results obtained with ML models focus more on predicting the general probability of a crime occurring in a group sample rather than individually, suggesting better statistical strategies. They have developed proposals to predict the type of crime committed by psychiatric patients on an individual level. They conducted tests with models such as Random Forest (RF), Elastic Net, and Support Vector Machine (SVM), applied to a representative and diverse sample of 1240 patients from the forensic mental health system. Considering clinical, historical, and sociodemographic variables as predictors, they developed separate models for each type of criminal offense, using feature selection methods. The models showed the following performance: 1) For the prediction of sexual and violent crimes at an individual level, 20 predictor variables were considered, achieving a sensitivity of 83.26% and specificity of 77.42%; For the prediction of sexual and non-violent crimes at an individual level, 30 predictor variables were considered, achieving a sensitivity of 74.60% and specificity of 80.65%; 3) For the prediction of sexual, violent, and non-violent crimes at an individual level, 36 predictor variables were considered, achieving a sensitivity of 82.44% and specificity of 60% [15].

A study by [16] focused on defining ML algorithms to obtain models that would identify the factors that distinguish between homicide/involuntary manslaughter and all other crimes committed by individuals with schizophrenia spectrum disorders (SSD), as well as the factors that distinguish between completed homicide/unpremeditated homicide and other violent crimes. The study concluded that variables related to criminal, psychiatric, and clinical factors have a high influence. They analyzed 358 offender patients with schizophrenia spectrum disorders admitted to the Forensic Inpatient Therapy Center at the Zurich Psychiatry Hospital between 1982 and 2016. Of the total 358 patients, 36.6% (131) experienced one or more direct coercive measures, while 63.4% (227) did not experience coercion. The database contained a total of 569 variables. To counteract possible overfitting and improve the model's quality, they reduced the number of variables using the chi-squared test, ultimately obtaining the following ten most significant variables: 1) threat of violence; 2) (actual) violence towards others; 3) application of direct coercive measures during previous psychiatric hospital treatments; 4) Positive and Negative Syndrome Scale (PANSS): poor impulse control; 5) Positive and Negative Syndrome Scale (PANSS): lack of cooperation; 6) prescription of haloperidol during hospital treatment; 7) total PANSS score at admission; 8) cumulative daily equivalent antipsychotic dose of olanzapine at discharge; 9) Positive and Negative Syndrome Scale (PANSS): hostility; and 10) legal prognosis estimated by a team of licensed forensic psychiatrists. Eight different ML models were trained, with the final Naïve Bayes model achieving the best results, with a balanced accuracy of 73.28%, an AUC of 0.8468, sensitivity of 72.87%, and specificity close to 73.68%.

The objective of the study conducted by [17] was to identify key differences in offender patients and non-offender patients with schizophrenia spectrum disorders (SSD) regarding aggressive behavior using various supervised ML algorithms. They used a dataset with 740 patients, 370 offenders, and 370 non-offenders, and 39 predictor variables: sociodemographic data, psychiatric data, pharmacotherapy, adverse events during hospitalization, childhood or youth abuse data, and other physical or neurological illnesses. The best result was achieved with the "gradient boosting" algorithm, achieving an accuracy of 79.9%, sensitivity of 77.3%, specificity of 82.5%, and an AUC of 0.87.

However, since the brain and behavior are highly complex systems involving multiple levels of temporal and spatial granularity and millions of nonlinear feedback loops, some researchers argue that a better understanding of the common and distinct pathophysiological mechanisms underlying psychiatric disorders is needed to provide more effective and personalized treatments. They also suggest that the analysis of "small" experimental samples using conventional statistical approaches has largely failed to capture the heterogeneity underlying psychiatric phenotypes, so they

recommend modern ML algorithms and approaches, such as deep learning, with large volumes of data to achieve better results and a greater understanding of psychiatric phenotypes [18].

Specific violence within psychiatric hospitalization wards is another globally significant problem that has been studied by the scientific community. In [19], a multivariable prediction model was developed to assess the risk of violence in hospitalized patients using patient medical record data. The model predicts which patients will exhibit violent behavior during the first 4 weeks after admission. AUC values in the range of 0.797 to 0.764 were achieved.

In [20], a systematic review was conducted, including 182 studies and 8 articles, on the use of ML techniques to predict the risk of violence in psychiatric patients in clinical and forensic settings. They analyzed the machine learning methods used in each study, sample size, model performance parameters such as AUC, accuracy, specificity, sensitivity, and the predictors used in each case. The analyzed studies showed AUC values in the range of 0.63 to 0.95.

In (Sonnweber et al., 2022) [21], ML methods were used to identify the factors that distinguish between hospitalized patients with schizophrenia spectrum disorders who have confirmed alcohol or illicit substance use during forensic psychiatric hospitalization and those who do not. The database consisted of 364 cases, with demographic, clinical, and criminal data from residents at the Zurich Forensic Psychiatry Hospital from 1982 to 2016. Different ML methods were used: logistic regression, decision tree, RF, gradient boosting, k-nearest neighbors, SVM, and Naive Bayes. The best performance was achieved with the gradient boosting model, with an AUC of 0.735, sensitivity of 81.48%, and specificity of 57.58%.

This work aims to identify the knowledge obtained from databases related to social history, cognitive testing, and risk assessment for young people referred for delinquent behavior, using ML techniques, which are more effective at identifying whether deficits in cognitive functions contribute to antisocial and aggressive behaviors. This study is motivated by the importance of prevention and promotion of mental health services and tools that enable the effective treatment of individuals so that they can reintegrate into community life, reduce the risk of recidivism, and reduce the substantial costs associated with incarceration.

2. Materials and Methods

2.1. Dataset

Our study sample included adolescent offenders ($n = 66$) and adolescent non-offenders ($n = 62$), aged 14–18 years of the male gender. The offender group was selected from “Centro de Reeducción el Oasis” through the foundation “Hogares Claret in Barranquilla, Colombia, where they are imprisoned for violation of punishable offenses, like sexual abuse, homicide, theft, among others. On the other side, the control group was selected from different educative institution located in the same city and had to meet the following criteria: a) male gender b) Age between 14 and 18 years old, c) not more than 12 years of education, d) absence of criminal background, e) absence of neurological, psychiatric or physical diagnosis. The parents or legal representatives of both group (adolescent offenders and adolescent no offenders) were due informed about the research by means of the informed consent which was filled out voluntarily or in case of participants under 18 years of age by parents or guardians. This study was approved by the Caribbean Ethics Committee and followed the ethical principles of the Declaration of Helsinki.

The dataset was built using the following cognitive instruments: Osterrieth Complex Figure; Ineco Frontal Screening; Montreal Cognitive Assessment; Stroop Color-Word Test; WAIS; and Symbol Digit Modalities Test. The final dataset has 138 observations and 39 features as shown in Table 1.

Table 1. Features.

Cognitive tests	Features
Osterrieth Complex Figure	Copy Short-term memory

	Long-term memory
	Motor programming
	Interference resistance
	Motor inhibitory control
	Verbal inhibitory control
Ineco Frontal Screening (IFS)	Verbal working memory
	Numerical working memory
	Visual working memory
	Abstraction capacity
	IFS Total
	Executive Visuospatial
	Identification
	Memory
Montreal Cognitive	Attention
Assessment (MOCA)	Language
	Abstraction
	Orientation
	MOCA Total
STROOP	Word
	Color
	Word-Color
VFT (Verbal Fluency Test)	Phonological Fluency Success
	Phonological Fluency Repetitions
	Phonological Fluency Distortions
	Semantic Fluency Success
	Semantic Fluency Repetitions
	Semantic Fluency Distortions
	Exclusive Fluency Success
	Exclusive Fluency Repetitions
	Exclusive Fluency Distortions
WAIS (Wechsler Adult	Matrices
Intelligence Scale)	Similarities
	Vocabulary
SDMT (Symbol Digit	Total correct answers
Modalities Test)	Total answers

The dependent variable is called "Group" and has two classes: adolescent offenders (SG) and adolescent non-offenders (CG), as showed in Figure 1.

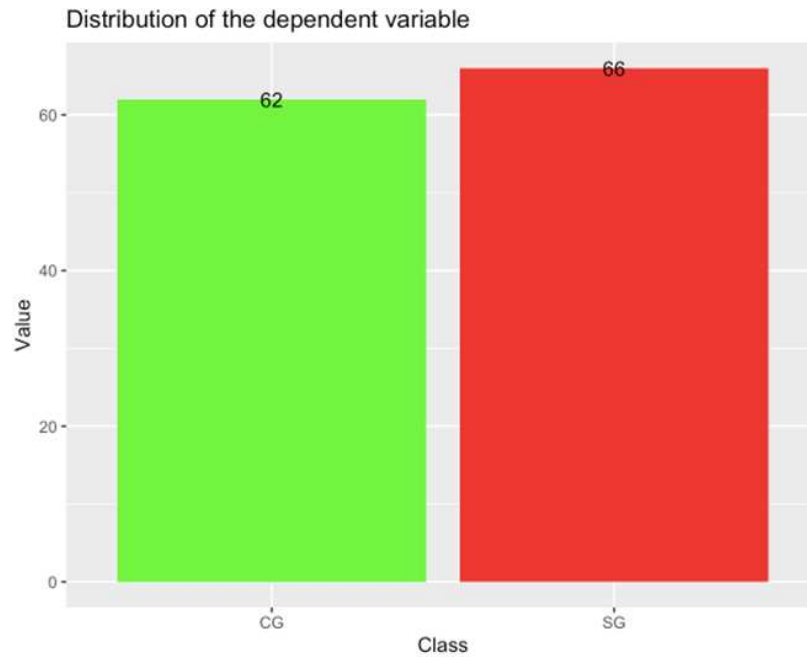


Figure 1. Distribution of classes (groups).

2.2. Feature selection

In order to reduce the number of input variables, choose the most relevant features, prevent and reduce overfitting in the machine learning training process, four methods of feature selection were considered.

2.2.1. Features that are significantly different among groups

By identifying the most important features that are significantly different among each group, it is possible to develop a more accurate model that can accurately predict the class labels for new data. To select features, the procedure consisted of identifying those characteristics that presented significant differences among each of the groups (classes). The homogeneity of variance test method used was the Levene test. As a result, 23 features with $p < 0.05$ were extracted from the original 36 features. These 22 features are shown in Figure 2.

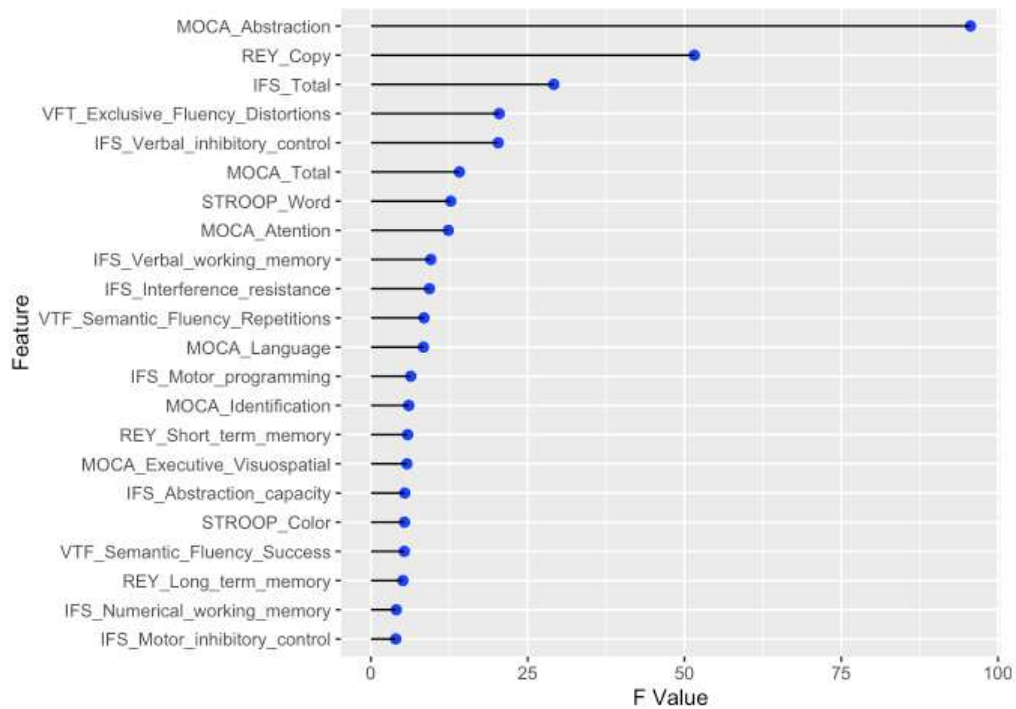


Figure 2. Features that are significantly different among groups.

2.2.2. Boruta algorithm

Boruta is a feature selection algorithm that works as a wrapper algorithm around Random Forest. This algorithm helps to select and rank variables that are statistically significant. The strictness of the algorithm can be adjusted modifying the p values and the number of times the algorithm is run. The 19 features selected are shown in Figure 3.

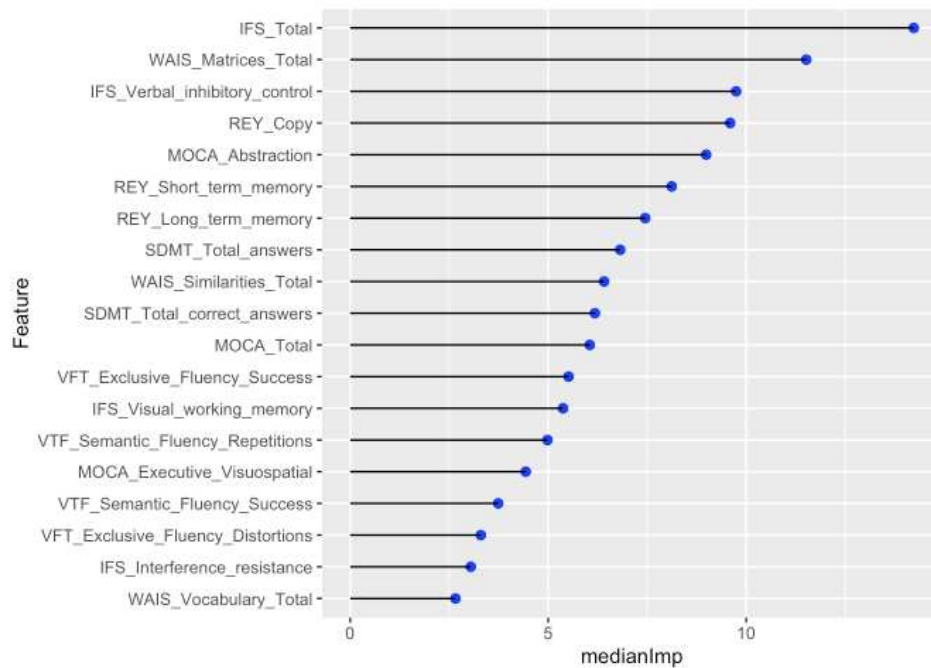


Figure 3. Features selected by Boruta algorithm.

2.2.3. Recursive Feature Elimination (RFE) Algorithm

This algorithm selects features sequentially. It starts building a model using all the features and removes those with lowest score at each iteration. It is a wrapper method. The 14 features selected are shown in Table 2.

Table 2. Features selected by RFE algorithm.

Cognitive tests	Features	Ranking
Osterrieth Complex Figure	Copy	4
	Short-term memory	7
	Long-term memory	9
Ineco Frontal Screening (IFS)	Verbal inhibitory control	3
	Visual working memory	13
	IFS Total	1
Montreal Cognitive Assessment (MOCA)	Abstraction	5
	MOCA Total	12
VFT (Verbal Fluency Test)	Semantic Fluency Repetitions	11
	Exclusive Fluency Success	14
WAIS (Wechsler Adult Intelligence Scale)	Matrices_Total	2
	Similarities_Total	10
SDMT (Symbol Digit Modalities Test)	Total correct answers	8
	Total answers	6

2.2.4. Filter Algorithm

In this case we use the filterVarImp() function in R (from Caret library), which uses ROC curve analysis on each predictor and use area under the curve as scores. This function can be used to evaluate features without the need of a specific model. It is, the selection of features is independent of any machine learning algorithms. The score of each feature is shown in Figure 4.

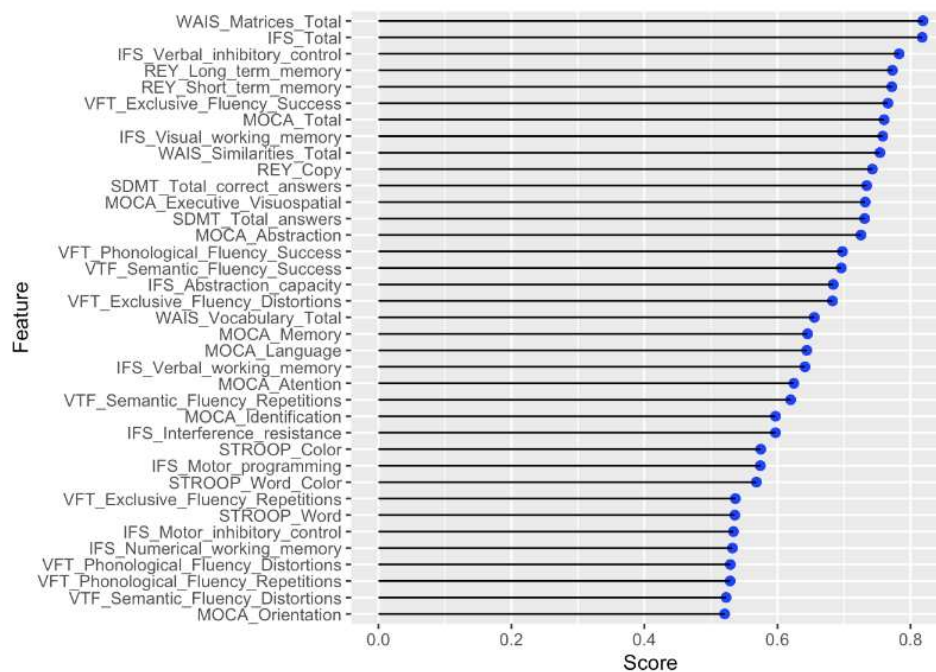


Figure 4. Score of features according to by filterVarImp() function in R.

In this case, the features with score higher than 0.6 were selected.

3. Results

The next step was training and testing 3 machine-learning models using the features selected by each feature selection method. Results are shown in Table 3.

Table 3. Metric results by machine-learning models.

Model	Metric	Features selected by				
		All Features	Levene test	Boruta	RFE	Filter
SVM	Sensitivity/Recall	1.0000	1.0000	1.0000	0.9167	0.9167
	Specificity	0.6923	0.6154	0.5385	0.7692	0.8462
	F1	0.8571	0.8276	0.8000	0.8452	0.8800
	Balance Accuracy	0.8462	0.8077	0.7692	0.8429	0.8814
RF	Sensitivity/Recall	0.8333	0.6667	1.0000	0.7500	0.8333
	Specificity	0.8462	0.7692	0.6923	0.6923	0.7692
	F1	0.8333	0.6957	0.8571	0.7200	0.8000
	Balance Accuracy	0.8397	0.7179	0.8462	0.7212	0.8013
KNN	Sensitivity/Recall	0.9167	1.0000	1.0000	0.9167	0.9167
	Specificity	0.6923	0.4615	0.7692	0.8462	0.6923
	F1	0.8148	0.7742	0.8889	0.8800	0.8148
	Balance Accuracy	0.8045	0.7308	0.8846	0.8814	0.8045

The best results were achieved with the K-NN model trained with features selected by the Boruta method and the SVM model trained with features selected by the filter method.

4. Discussion

The main objective of this study was to identify young law offenders from young non-law offenders based on a dataset extracted from cognitive tests, such as the Osterrieth Complex Figure (REY), Ineco Frontal Screening (IFS), Montreal Cognitive Assessment (MOCA), Stroop Color-Word Test, WAIS; and Symbol Digit Modalities Test (SDMT). These cognitive tests were administered to a group of young individuals consisting of 66 law offenders and 62 non-law offenders.

In total, the dataset contains 39 features and 138 observations. Different methods were applied to reduce the number of features. Initially, a total of 22 features that showed significant differences between both groups were extracted. Subsequently, feature selection methods such as Boruta (19 features), RFE (14 features), and Filter (24 features) were applied.

The figure 5 shows the number of times each feature was selected by each of the methods employed:

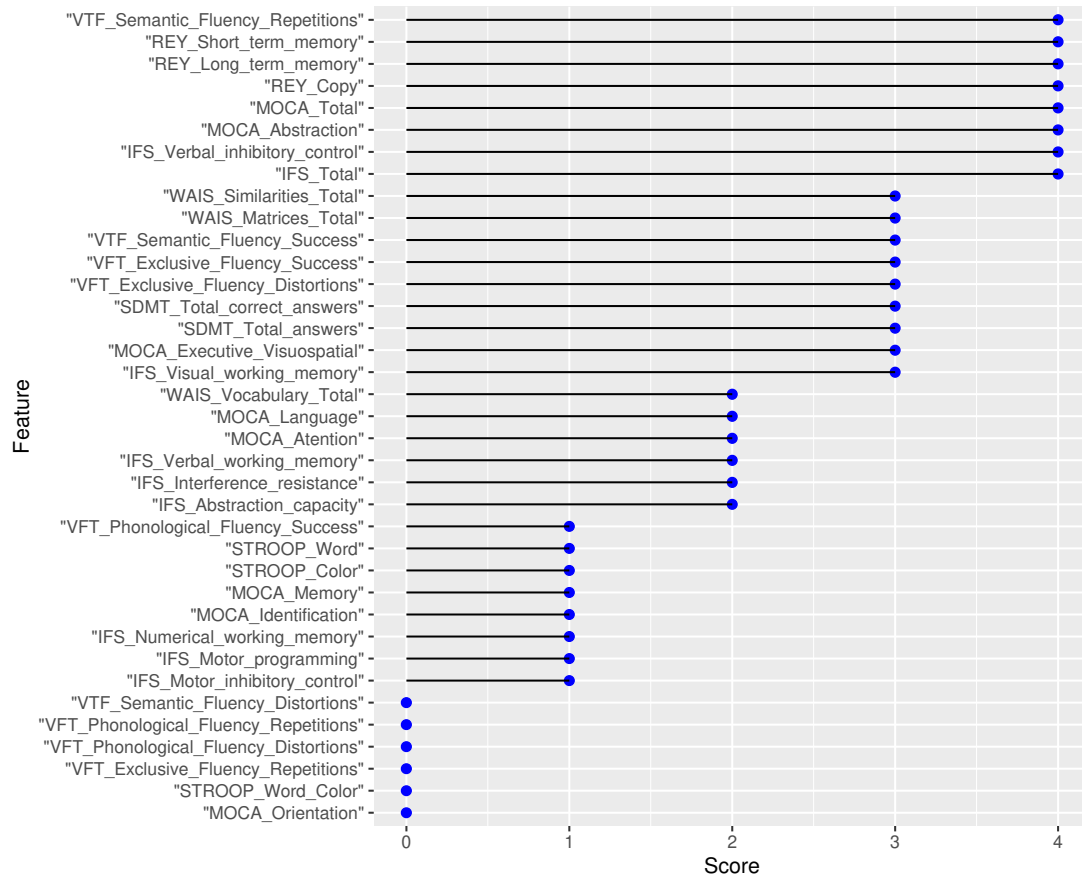


Figure 5. Number of times each feature was selected by all methods.

In the table 4, the 8 common features that were selected by the four aforementioned methods are presented

Table 4. Eight (8) common features selected by each of the 4 methods.

Cognitive tests	Features
Osterrieth Complex Figure	Copy Short-term memory Long-term memory
Ineco Frontal Screening (IFS)	Verbal inhibitory control IFS Total
Montreal Cognitive Assessment (MOCA)	Abstraction MOCA Total
VFT (Verbal Fluency Test)	Semantic Fluency Repetitions

In the table 5, the 14 common features that were selected by the three feature selection methods (Boruta, RFE, and Filter) are presented

Table 5. 14 common features selected by each of the 3 features selection methods.

Cognitive tests	Features
Osterrieth Complex Figure	Copy Short-term memory Long-term memory
Ineco Frontal Screening (IFS)	Verbal inhibitory control

	Visual working memory
	IFS Total
Montreal Cognitive Assessment (MOCA)	Abstraction MOCA Total
VFT (Verbal Fluency Test)	Semantic Fluency Repetitions Exclusive Fluency Success
WAIS (Wechsler Adult Intelligence Scale)	Matrices Similarities
SDMT (Symbol Digit Modalities Test)	Total correct answers Total answers

According to the employed feature selection methods, the three most relevant features are: IFS_Total, WAIS_Matrices_Total, and IFS_Verbal_Inhibitory_control

The summary of the results obtained for each model is presented in Figure 6 and Figure 7.

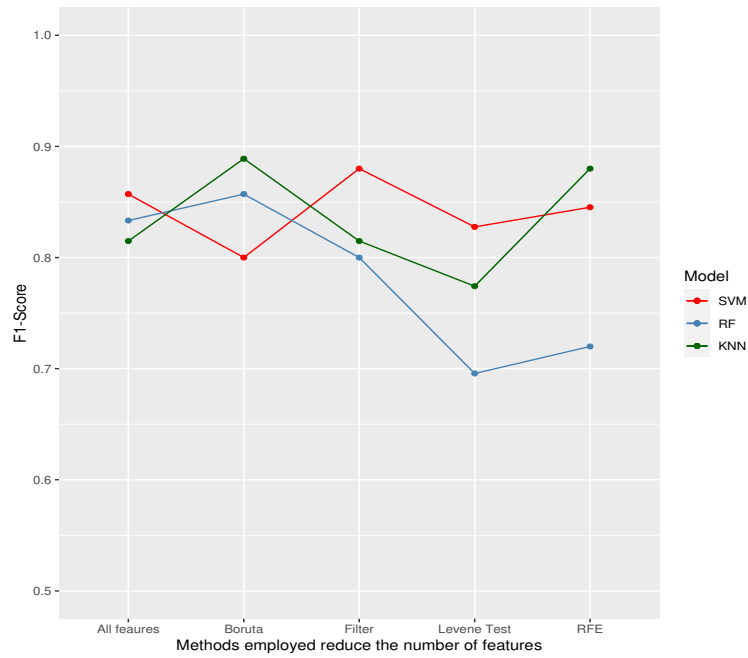


Figure 6. F1 score obtained for each model on the testing data.

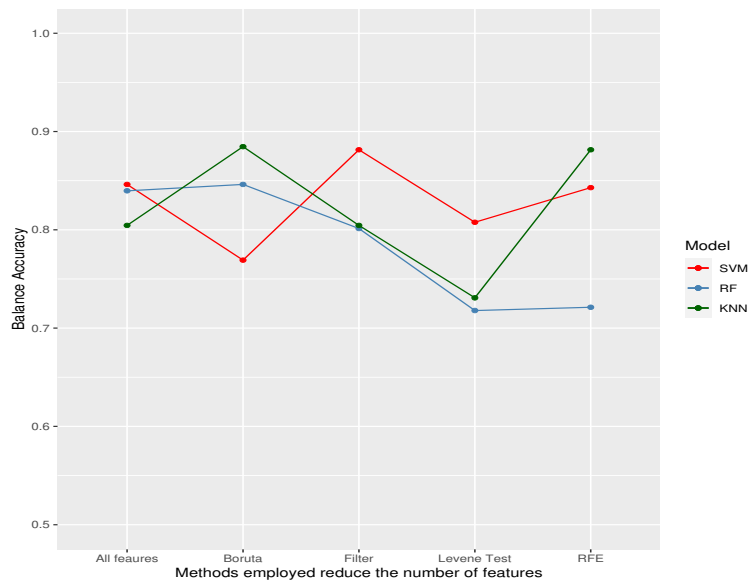


Figure 7. Balance accuracy obtained for each model on the testing data.

5. Conclusions

In this study, a database containing the results of cognitive tests conducted on a study group (SG) consisting of 66 young offenders in a reform school in the city of Barranquilla, northern Colombia, and a control group (CG) consisting of 62 non-offender high school students was used. All the young individuals were male and aged between 14 and 18 years old. The database has a total of 37 predictor variables.

The objective was to find a model to predict which group each young individual belongs to based on the variables resulting from the cognitive tests. Three models were employed: Support Vector Machine (SVM), Random Forest (RF), and the K-Nearest Neighbors (K-NN) model. Each of these models was trained using a different number of variables, as follows: 1) All 37 variables; 2) Only 22 variables that showed significant differences (Levene Test); 3) Only 19 variables selected by the Boruta method; 4) Only 14 variables selected by the RFE method; 5) Only 24 variables selected by the Filter method.

The best result was achieved by the K-NN model trained with 19 features selected by the Boruta method, followed by the SVM model trained with 24 features selected by the filter method.

Author Contributions: Conceptualization, M.B., J.C. and M.P; methodology, J.C. and M.B; software, J.C.; validation, M.P. and R.R.; formal analysis, R.R.; investigation, M.B., J.C. and M.P.; resources, M.P. and R.R.; data curation, J.C.; writing—original draft preparation, M.B. and J.C.; writing—review and editing, M.P. and R.R.; visualization, M.B. and J.C.; supervision, M.P and G.G.; project administration, M.P and G.G.; funding acquisition, G.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research and APC was funded by Institución Universitaria de Barranquilla and Ministerio de Ciencia Tecnología e Innovación in Colombia, grant number 68229 BPIN 20200000100006.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are openly available in Kaggle at <https://www.kaggle.com/datasets/mariaclaudiabonfante/funcionescognitivas>

Acknowledgments: We would like to extend our gratitude for the support received from Ministerio de Ciencia Tecnología e Innovación – MinCiencias Colombia, and the Fundación Clarets, the Centro de Reeducción el Oasis, and the Fundación Luz Esperanza in Barranquilla, Colombia, Their invaluable contributions significantly facilitated the realization of this research

Conflicts of Interest: The authors declare no conflict of interest

References

1. Masho, S. W., Schoeny, M. E., Webster, D., & Sigel, E., Outcomes, Data, and Indicators of Violence at the Community Level. *Journal of Primary Prevention*, 2016, 37(2), 121–139. <https://doi.org/10.1007/s10935-016-0429-4>
2. Voith, L., Salas, M., Sorensen, A., Thomas, T. J., Coulton, C., & Barksdale, E., Identifying Risk Factors and Advancing Services for Violently Injured Low-Income Black Youth, *Journal of Racial and Ethnic Health Disparities*, 2022. <https://doi.org/10.1007/s40615-022-01365-9>
3. Douglas, T., Pugh, J., Singh, I., Savulescu, J., & Fazel, S., Risk assessment tools in criminal justice and forensic psychiatry : The need for better data, *European Psychiatry*, 2017, 42, 134–137.
4. Rolin, S.A., Bareis, N., Bradford, J.M., Rotter, M., Rosenfeld, B., Pauselli, L., Compton, M.T., Stroup, T.S., Appelbaum, P.S., Dixon, L.B., Violence Risk Assessment for Young Adults Receiving Treatment for Early Psychosis, 2021, 76: 101701. doi:10.1016/j.ijlp.2021.101701.
5. De Vogel, V., De Beuf, T., Shepherd, S., Schneider, R.D., Violence Risk Assessment with the HCR-20V3 in Legal Contexts: A Critical Reflection, *Journal of Personality Assessment*, 2022, 104:2, 252-264, DOI: 10.1080/00223891.2021.2021925
6. Barzman, D., Ni, Y., Griffey, M., Bachtel, A., Lin, K., Jackson, H., Sorter, M., DelBello, M., Automated Risk Assessment for School Violence: a Pilot Study, *Psychiatric Quarterly*, 2018, Vol. 89, pp. 817–828. <https://doi.org/10.1007/s11126-018-9581-8>
7. Singh, J. P., Desmarais, S. L., Hurducas, C., Arbach-lucioni, K., Condemarin, C., Dean, K, Doyle, M., Folino, J. O., Godoy-cervera, V., Grann, M., Mei, R., Ho, Y., Matthew, M., Nielsen, L. H., Pham, T. H., Rebocho, M. F., Reeves, K. A., Rettenberger, M., Ruitter, C. De, Ruitter, C. De., International Perspectives on the Practical

- Application of Violence Risk Assessment: A Global Survey of 44 Countries, *International Journal of Forensic Mental Health*, 9013, 2014, <https://doi.org/10.1080/14999013.2014.922141>
8. Tortora, L., Meynen, G., Bijlsma, J., Tronci, E., & Ferracuti, S., Neuroprediction and A.I. in Forensic Psychiatry and Criminal Justice: A Neurolaw Perspective, *Frontiers in Psychology*, 11(March), 2020, 1–9, <https://doi.org/10.3389/fpsyg.2020.00220>
 9. Coid, J., Mickey, N., Kahtan, N., Zhang, T., & Yang, M., Patients discharged from medium secure forensic psychiatry services: Reconvictions and risk factors, *British Journal of Psychiatry*, 190(MAR.), 2007, 223–229. <https://doi.org/10.1192/bjp.bp.105.018788>
 10. Watts, D., Cardoso, T. D. A., Librenza-garcia, D., Ballester, P., & Passos, I. C., Predicting criminal and violent outcomes in psychiatry: a meta- analysis of diagnostic accuracy, *Translational Psychiatry*, 2022, September. <https://doi.org/10.1038/s41398-022-02214-3>
 11. Hofmann, L. A., Lau, S., & Kirchebner, J, Advantages of Machine Learning in Forensic Psychiatric Research—Uncovering the Complexities of Aggressive Behavior in Schizophrenia. *Applied Sciences (Switzerland)*, (2022), 12(2). <https://doi.org/10.3390/app12020819>
 12. Cockerill, R. G, Ethics implications of the use of artificial intelligence in violence risk assessment. *Journal of the American Academy of Psychiatry and the Law*, 2020, 48(3), 345–349. <https://doi.org/10.29158/JAAPL.003940-20>
 13. Trinhhammer, M. L., Merrild, A. C. H., Lotz, J. F., & Makransky, G, Predicting crime during or after psychiatric care: Evaluating machine learning for risk assessment using the Danish patient registries, *Journal of Psychiatric Research*, 2022, 152(June), 194–200. <https://doi.org/10.1016/j.jpsychires.2022.06.009>
 14. Watts, D., Mamak, M., Moulden, H. Upfold, C., De Azevedo Cardoso, T., Kapczinski, F., Chaimowitz, G., The HARM models: Predicting longitudinal physical aggression in patients with schizophrenia at an individual level, *Journal of Psychiatric Research*, 2023, Vol. 161, pp. 91-98, doi: <https://doi.org/10.1016/j.jpsychires.2023.02.030>
 15. Watts, D., Moulden, H., Mamak, M., Upfold, C., & Chaimowitz, G., Predicting offenses among individuals with psychiatric disorders - A machine learning approach, *Journal of Psychiatric Research*, 2021, 138(October 2020), 146–154
 16. Günther, M. P., Kirchebner, J., & Lau, S, Identifying Direct Coercion in a High Risk Subgroup of Offender Patients With Schizophrenia via Machine Learning Algorithms, *Frontiers in Psychiatry*, 2020, 11(May), 1–9.
 17. Kirchebner, J., Lau, S., & Machetanz, L, Offenders and non-offenders with schizophrenia spectrum disorders: Do they really differ in known risk factors for aggression?, *Frontiers in Psychiatry*, 2023, 14(April), 1–12. <https://doi.org/10.3389/fpsy.2023.1145644>
 18. Koppe, G., Meyer-Lindenberg, A., & Durstewitz, D, Deep learning for small and big data in psychiatry, *Neuropsychopharmacology*, 2021, 46(1), 176–190. <https://doi.org/10.1038/s41386-020-0767-z>
 19. Menger, V., Spruit, M., Est, R. Van, Nap, E., & Scheepers, F., Machine Learning Approach to Inpatient Violence Risk Assessment Using Routinely Collected Clinical Notes in Electronic Health Records. *JAME Network Open*, 2019, 2(7), 1–12. <https://doi.org/10.1001/jamanetworkopen.2019.6709>
 20. Parmigiani, G., Barchielli, B., Casale, S., Mancini, T., & Ferracuti, S., The impact of machine learning in predicting risk of violence: A systematic review, 2022, December. <https://doi.org/10.3389/fpsy.2022.1015914>
 21. Sonnweber, M., Kirchebner, J., Philipp, M., Kappes, J. R., & Lau, S., Exploring substance use as rule-violating behaviour during inpatient treatment of offender patients with schizophrenia, *Criminal Behaviour and Mental Health*, 2022, 10.1002/cbm.2245

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.