RESEARCH ARTICLE

Revista Mexicana de Ingeniería Biomédica

VOL. 47 | NO. 1 | **JANUARY – APRIL** | PP 1-15

https://dx.doi.org/10.17488/RMIB.47.1.1550 E-LOCATION ID: e1550

Predicting Executive Function Impairments in Young Adults Using Machine Learning and Lifestyle Data

Predicción del deterioro de las funciones ejecutivas en adultos jóvenes utilizando aprendizaje automático y datos de estilo de vida

Alejandro Jarillo Silva[1], *Elizabeth Muñoz Ortiz*[2], *Olaf Montes de Oca Juárez*[1], *José Alberto Cruz Torentino*[1]

[1]Universidad de la Sierra Sur, Miahuatlán de Porfirio Díaz, Oaxaca - México
[2]Universidad Veracruzana, Xalapa, Veracruz - México

ABSTRACT

The development of executive function (EF) impairments in young individuals, such as difficulties with attention, memory, and problem-solving, is influenced by biological, social, and lifestyle factors. However, research on predicting these impairments remains limited due to a lack of reliable tools. This study analyzed 90 university students using EF tests, lifestyle, and sociodemographic questionnaires. Five machine learning models were evaluated: Decision Trees (DT), k-Nearest Neighbors (KNN), Support Vector Machines (SVM), Logistic Regression (LR), and Random Forest (RF), with cross-validation applied for model assessment. The results indicated a 62% incidence of EF impairments. Maternal education and nutrition were identified as key influencing factors. Among the models, DT performed best, achieving a recall of 61.9%, an F1-score of 62.1%, and an AUC of 66.54%, while RF had the lowest performance. Limitations include the cross-sectional nature of the data, which restricts causal inference, and the reliance on self-reported responses from participants, which may reduce data reliability. Despite these limitations, this study demonstrates the feasibility of using machine learning to predict EF impairments based on easily collected sociodemographic and lifestyle data. Sociodemographic and lifestyle variables are valuable predictors of EF impairments in young individuals. Machine learning tools offer a practical approach to assessing population-level EF health using accessible data.

KEYWORDS: cognitive impairments prediction, machine learning, neuropsychological tests

RESUMEN

El desarrollo de deterioros en las funciones ejecutivas (FE) en jóvenes, como dificultades en la atención, la memoria y la resolución de problemas, está influenciado por factores biológicos, sociales y de estilo de vida. Sin embargo, la investigación sobre la predicción de estos deterioros sigue siendo limitada debido a la falta de herramientas confiables. Este estudio analizó a 90 estudiantes universitarios mediante pruebas de FE y cuestionarios sobre estilo de vida y factores sociodemográficos. Se evaluaron cinco modelos de aprendizaje automático: Árboles de Decisión (DT), k-Nearest Neighbors (KNN), Máquinas de Soporte Vectorial (SVM), Regresión Logística (LR) y Bosques Aleatorios (RF), aplicando validación cruzada para la evaluación de los modelos. Los resultados indicaron una incidencia del 62% en deterioros de las FE. Se identificaron la educación materna y la nutrición como factores clave influyentes. Entre los modelos, DT obtuvo el mejor desempeño, con una sensibilidad del 61.9%, un F1-score de 62.1% y un AUC de 66.54%, mientras que RF tuvo el peor rendimiento. Las limitaciones incluyen la naturaleza transversal de los datos, lo que restringe la inferencia causal, y la dependencia de respuestas autoinformadas por los participantes, lo que podría afectar la fiabilidad de los datos. A pesar de esto, el estudio demuestra la viabilidad del uso de aprendizaje automático para predecir deterioros en las FE con datos sociodemográficos y de estilo de vida fácilmente recopilables. Las variables sociodemográficas y de estilo de vida son valiosos predictores de deterioros en las FE en jóvenes. Las herramientas de aprendizaje automático ofrecen un enfoque práctico para evaluar la salud de las FE a nivel poblacional utilizando datos accesibles.

PALABRAS CLAVE: aprendizaje automático, predicción de deterioro cognitive, pruebas neuropsicológica

Corresponding author

TO: **ALEJANDRO JARILLO SILVA**
INSTITUTION: **UNIVERSIDAD DE LA SIERRA SUR**
ADDRESS: **GUILLERMO ROJAS MIJANGOS S/N, CD UNIVERSITARIA, MIAHUATLÁN DE PORFIRIO DÍAZ, CP 70800, OAXACA, MÉXICO**
EMAIL: **ajarillo@unsis.edu.mx**

Received:

10 April 2025

Accepted:

2 October 2025

Published:

30 January 2026

INTRODUCTION

Neuropsychological tests are essential for assessing executive functions (EF), which comprise high-level cognitive processes such as planning, inhibitory control, working memory, and cognitive flexibility. These functions play a critical role in academic, emotional, and social development during adolescence and early adulthood, when the prefrontal cortex reaches its peak maturation[1][2][3]. Proper evaluation of EF allows for the early identification of difficulties that may affect learning, decision-making, and social adaptation[4][5][6].

Beyond their cognitive relevance, EF are also shaped by lifestyle factors such as nutrition, physical activity, sleep, stress management, and social support. Evidence shows that healthy habits can enhance domains like working memory, impulse control, and planning, whereas poor habits may hinder them[7][8][9]. For instance, aerobic exercise has been associated with improvements in attention and emotional regulation, while inadequate sleep and chronic stress are linked to impairments in decision-making[8][17].

These findings highlight that adopting healthy lifestyles not only protects but may also enhance EF, supporting academic success and adaptive behavior. However, most studies have focused on children or older adults, leaving a gap in the understanding of these relationships in university students. This population is particularly relevant, as both EF and lifestyle factors are still in development, and difficulties at this stage may have lasting consequences on academic performance, social adaptation, and long-term well-being.

Neuropsychological tests, such as the BANFE-2 (Neuropsychological Battery of EF and Frontal Lobes, Second Edition), are valuable tools for evaluating EF in young people. These functions include cognitive processes such as planning, inhibitory control, cognitive flexibility, and working memory, which are fundamental for learning, adaptive behavior, and emotional regulation[10]. The BANFE-2 offers several advantages: it provides updated norms tailored to different age groups, ensuring accurate and comparable results; its design covers a wide range of executive subprocesses, making it a comprehensive tool for neuropsychological analysis; and, unlike other batteries, the BANFE-2 is specifically adapted for Spanish-speaking populations, reducing cultural biases.

On the other hand, the Lifestyle Profile Questionnaire (PEPS-I) is a tool designed to assess behaviors, habits, and attitudes related to a healthy lifestyle. It was developed to help identify specific areas where people may need to improve their health and overall well-being. This instrument evaluates dimensions such as nutrition, exercise, stress management, interpersonal support, self-actualization, and health responsibility using a Likert scale. Its applicability to young people and adolescents has been validated for reliability and validity[11]. In educational and clinical contexts, the PEPS-I contributes to the design of personalized programs that address factors such as family and social support, as well as self-care[12].

The relationship between EF and lifestyles in young people is a growing area of interest in neuropsychology and health sciences. EF, which includes skills such as inhibitory control, working memory, planning, and cognitive flexibility, are directly related to the ability of young people to adopt and maintain healthy habits[7]. For example, regular physical activity has been associated with improvements in working memory, attention, and emotional regulation. Aerobic exercise, in particular, has shown a positive impact on the development of the prefrontal cortex, where the EF reside[8].

There is evidence that adherence to healthy lifestyles positively impacts long-term EF preservation, with nutrition being a widely studied factor. In 2015, Zhu et al. conducted a 25-year follow-up study, periodically evaluating the diets of participants by categorizing and scoring foods as good, neutral or bad based on scientific evidence. These scores were then associated with cognitive and EF. A higher score, indicating a balanced diet, was associated with a more diverse and fluent vocabulary and speech, better mathematical processing, and overall improved cognitive performance[9].

On the other hand, supervised classification models are key tools in the analysis of neuropsychological and lifestyle data, particularly to identify patterns associated with EF impairments, classifying performance levels, and predicting clinical outcomes. For example, RF has been used to classify patients in different stages of mild cognitive impairment or Alzheimer's disease based on neuropsychological variables such as memory, attention, and language[13]. Comparative studies have been conducted between models such as Naïve Bayes (NB), Random Forest (RF), Decision Trees (DT), k-Nearest Neighbors (KNN), Support Vector Machines (SVM), AdaBoost, and Linear Discriminant Analysis (LDA) to optimize cognitive assessment related to EF, improve diagnostic accuracy, and reduce missed diagnoses, achieving results above 90\%[14]. The SVM model has been applied to classify the level of cognitive impairment based on the assessment of EF in children with Down syndrome[15]. The RF model has also been used for the prevention of cardiovascular disease based on lifestyle variables[16]. In 2024, Zhang et al. conducted research to identify sleep problems in students using lifestyle analysis, using models such as LR, Extreme Gradient Boosting Machine (XGBM), NB, SVM, DT, and CatBoosting Machine (CatBM), achieving classification results above 84\% in general[17].

Despite advances in the application of machine learning to cognitive assessment, limited research has addressed its integration with neuropsychological, lifestyle, and sociodemographic data in young adults. Alterations in EF during this developmental stage are clinically relevant, as they are associated with increased vulnerability to academic difficulties, maladaptive behaviors, and long-term risks for psychiatric or neurological conditions. Early identification of EF impairments may therefore contribute to timely educational, therapeutic, and preventive interventions that promote both academic success and mental health.

Most studies have focused on children, adolescents, or older adults, leaving a gap in the analysis of university students, a population in which EF are still developing and are sensitive to environmental factors and lifestyle habits[7][9][17]. This gap in the literature justifies the need to explore supervised classification models that can identify patterns and predict potential EF impairments, thereby contributing to the planning of targeted educational and preventive interventions for this group.

The aim of this study is to develop and evaluate supervised classification models capable of predicting EF impairments in university students based on lifestyle (PEPS-I) and sociodemographic variables. We hypothesize that these factors will significantly contribute to the accurate classification of students with and without EF impairments, and that machine learning models will achieve reliable predictive performance.

MATERIALS AND METHODS

Participants and Study Design

This project was approved by the Bioethics Committee of the Universidad de la Sierra Sur under reference number CEI-03/2022. The study involved 90 university students from an institution located in the state of Oaxaca, Mexico, during the period of January to December 2023. The participants were randomly selected from different academic disciplines (social sciences, health, and technology), and all gave their informed written consent, as approved by the Bioethics Committee of the Universidad de la Sierra Sur.

The inclusion criteria required participants to be active university students, functionally independent, without physical limitations that affect their ability to perform the tests, and with normal or corrected vision and hearing. The exclusion criteria included individuals who refused to sign informed consent, had physical limitations preventing test completion, and/or self-reported psychiatric or neurological disorders. Finally, participants who did not complete the tests or withdraw from the study during the procedures were excluded from the final analysis. Data collection was carried out in a controlled environment at the Information Technology Center, specifically in the Human-Computer Interaction Laboratory and the Gesell Chamber, both facilities belonging to the Universidad de la Sierra Sur.

In Figure 1, the proposed framework is presented. Once the participants were approved, they signed the informed consent form. Subsequently, neuropsychological tests from the BANFE-2 battery, which are based on the dorsolateral prefrontal cortex, were administered in the Gessell chamber. The tests were administered by professionals with expertise in clinical research

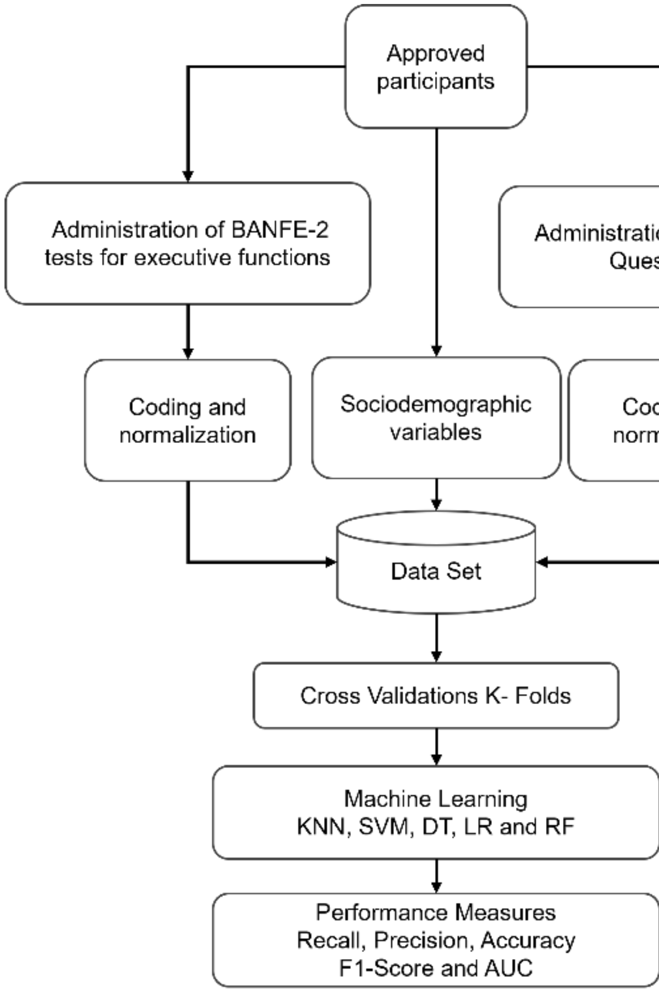

FIGURE 1. **Framework of the proposed machine learning prediction model**

These tests are described in the following:

1. Self-directed Pointing: Assesses the ability to use visuospatial working memory to self-direct the pointing to a series of figures.
2. Visuospatial Working Memory: Measures the ability to retain and actively reproduce the sequential visuospatial order of a series of figures.
3. Alphabetical Word Ordering: Evaluates the capacity to mentally manipulate and arrange verbal information stored in working memory.
4. Card Sorting: Primarily assesses EF, particularly cognitive flexibility, abstract reasoning, and the ability to adapt to new and changing rules.
5. Mazes: Evaluates the ability to systematically anticipate (plan) visuospatial behavior.
6. Tower of Hanoi: Measures the ability to sequentially anticipate actions, both in progressive and regressive order (sequential planning).
7. Consecutive Addition and Subtraction: Assesses the capacity to perform sequences in reverse order (inverse sequencing).

8. Verbal Fluency: Estimates the ability to fluently produce the greatest number of verbs within a limited time frame.

After administering the BANFE-2 and PEPS-I tests, the raw scores obtained from each participant underwent a coding and normalization process, which is a key step depicted in Figure 1. This procedure involved converting raw scores into standardized values based on normative data, taking into account participants' age and years of schooling [10]. Specifically, for the BANFE-2, coding was performed using Table A-7 of the manual (ages 16–30, 10–24 years of education), and normalization was conducted according to Table B-8. The process ensures that the scores are comparable across participants and accurately reflect individual performance relative to normative expectations.

Similarly, for the PEPS-I questionnaire, each subscale was coded considering the 48 variables across the dimensions of nutrition, exercise, health responsibility, stress management, interpersonal support, and self-actualization [9]. This standardized coding and normalization step allows for reliable integration of neuropsychological and lifestyle data into the subsequent machine learning prediction models, ensuring that the input data are consistent and suitable for analysis.

Following the administration of the BANFE-2 test, participants were categorized into two groups according to their performance: 'subjects without EF impairments' (n = 34) and 'subjects with EF impairments' (n = 56). These categories were used as classes for training machine learning models. Subsequently, the PEPS-I survey was administered to collect information on various lifestyle aspects, including nutrition, physical activity, health responsibility, stress management, interpersonal support, and professional development. To ensure accessibility for students, the survey was delivered in a digital format. Finally, participants completed a sociodemographic questionnaire. All data collected were securely stored for further processing and application of machine learning techniques.

ML-based models

In this study, various machine learning models were evaluated, including KNN, SVM, DT, LR, and RF. Neuropsychological, lifestyle, and sociodemographic data often exhibit high dimensionality and, in this case, a limited sample size. Cross-validation was applied to assess the ability of the models to generalize unseen data, thereby reducing the risk of overfitting, where models become too tailored to training data and lose accuracy in new datasets. Furthermore, cross-validation minimizes the variability associated with random data sampling[18]. As shown in Figure 1, a five-fold cross-validation procedure was implemented to ensure robust model evaluation. In this approach, 80% of the dataset was allocated to the training phase, while the remaining 20% was reserved for testing, allowing for reproducible and unbiased performance estimation.

SVM models are effective when working with datasets where the number of features (variables) is high relative to the number of observations, as is often the case with neuropsychological tests, lifestyle habits, and sociodemographic variables. By focusing on finding an optimal hyperplane that maximizes the margin between classes, SVMs perform well in terms of generalization, even with small or unbalanced data sets, which are common in clinical and population studies[19].

The KNN algorithm is particularly suitable for neuropsychological, lifestyle, and sociodemographic variables due to its flexible and non-parametric nature. These characteristics make it an effective tool in contexts where relationships between variables may be complex and nonlinear[20].

DT models do not require linear relationships between variables, making them ideal for neuropsychological data. These data often involve complex and non-linear interactions between factors such as lifestyle, sociodemographic background, and neurocognitive conditions related to EF[21]. The LR model is a widely used parametric approach for predicting binary or categorical outcomes. It is especially valuable in neuroscience studies that aim to identify risk factors for conditions such as EF impairments based on predictive variables[22].

Finally, the RF model is a non-parametric machine learning approach based on multiple decision trees. It is well suited for heterogeneous and complex datasets, such as those found in neuropsychological and sociological studies. By combining multiple trees, RF reduces the risk of overfitting and improves generalization [23].

Performance Measures

To evaluate supervised classification models in studies involving neuropsychological, lifestyle, and sociodemographic variables, it is essential to use performance metrics that reflect both the precision of the model and its generalizability. These measures should account for the characteristics of the problem, such as class imbalances and the importance of minimizing errors in clinical or social contexts [24]. The performance metrics used in this study are:

1. Sensitivity or True Positive Rate (Recall):

$$Recall = \frac{TP}{TP + FN} \tag{1}$$

Where TP represents true positives and FN represents false negatives. This metric measures the model's ability to correctly identify positive cases, which is critical in medical contexts where false negatives can have severe consequences.

2. Precision:

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

Where FP represents false positives. Precision indicates the proportion of positive predictions that are correct, which is important in scenarios where false positives incur high costs.

3. Accuracy: This is a widely used performance metric in supervised classification models. Measures the proportion of correct predictions made by the model relative to the total number of evaluated cases.

$$Accuracy = \frac{TP + TN}{TN + TP + FP + FN} \tag{3}$$

Where TN represents true negatives. This metric evaluates the model's ability to correctly detect negative cases, which is particularly useful in studies where minimizing false positives is important.

4. F1-Score:

$$F1 - Score = 2x\frac{Precision\ x\ Recall}{Precison + Recall} \tag{4}$$

This metric combines precision and recall into a single value, making it particularly useful when dealing with imbalanced classes.

5. AUC (Area Under the Curve): This metric measures the model's ability to distinguish between classes. A value close to 1 indicates excellent discrimination.

As a complementary tool to evaluate the proposed models, the probability density curve is used. This curve allows for the exploration and analysis of the model's behavior in terms of confidence and the quality of the probabilities assigned to its predictions. Facilitates a more precise interpretation and promotes effective optimization of the results[25].

Statistical Analysis

In this study, continuous variables were summarized using mean and standard deviation (SD), while categorical variables were presented as proportions. The chi-square test was used to compare categorical distributions. Independent means were compared using the Mann–Whitney U test due to violations of normality and homogeneity of variance assumptions. All statistical analyzes were performed with JASP version 0.18.1.

RESULTS AND DISCUSSION

Results

Of the 95 students who initially participated, 5 were excluded because they withdrew from the study. A total of 90 students were therefore included in the analysis. According to the BANFE-2 results, 56 students (62%) were classified as having cognitive process impairments, while 34 students (38%) were classified as normal. Of the participants, 36 (40%) were male and 54 (60%) were female. The distribution by academic level was as follows: 42% first year, 9% second year, 12.2% third year, 27.8% fourth year, and 6.7% fifth year. Regarding the academic area, 73 students (81%) belonged to health sciences, and 17 students (19%) to social sciences (see Table 1).

TABLE 1. **Participant characteristics and stratified comparison based on the presence of EF impairments in university students**

Statistics Results			
Variable	Alterations in EF		P-value
	Yes	No	
Participants n(%)	56(22)	34(38)	0.287
Gender(Male/Female%)	20.36(36.64)	16.18(47.53)	0.47
Years(Mean, SD)	20.32(1.84)	20.73(3.04)	0.375
Grade n(%)			
First	22(39.3)	16(47.1)	
Second	8(14.3)	2(5.9)	
Third	9(16.1)	2(5.9)	
Fourth	14(25)	11(32.3)	
Fifth	3(5.4)	3(8.8)	
Academic profile n(%)			0.748
Social	10(17.9)	7(20.6)	
Health	46(82.1)	27(79.4)	
Father's education level n(%)			0.25
Less than 6 years	14(25)	5(15)	
More than 6 years	20(75)	32(85)	
Mother's education level n(%)			**0.001**
Less than 6 years	36(25)	2(15)	
More than 6 years	20(75)	32(85)	
PEPS-I Results (mean,SD)			
Nutrition Score	14.57(3.7)	14.14(2.93)	0.75
Exerciser score	11.55(3.45)	10.73(3.48)	0.28
Health responsibility	21.86(4.38)	20.24(3.99)	0.08
Stress management	16.02(3.37)	16.68(3.67)	0.39
Interpersonal support	20.04(4.09)	20.97(3.80)	0.27
Update	39.89(6.94)	39.59(6.86)	0.87
BANFE-2 Results (mean,SD)			
Work memory	91.7(11.75)	106.32(8.13)	
EF	86.2(13.4)	108.3(7.3)	

As part of the statistical analysis, comparisons between the “with impairments” and “without impairments” groups were conducted to determine whether significant differences existed in gender, age, academic profile, parental education level, nutrition, physical activity, health responsibility, stress management, interpersonal support, and self-actualization.

Statistical tests (Mann-Whitney U) were applied and it was observed that among the six domains assessed by the PEPS-I test, there is apparently insufficient statistical evidence to conclude a significant difference. However, evaluating each variable in isolation does not allow the identification of factor combinations that may be relevant to classifying or predicting an outcome, such as the presence of EF impairments.

Predictive performance was assessed using the metrics Recall, Precision, Accuracy, F1-Score, and AUC (see Table 2 and Figure 2). The DT model achieved the best result in Recall (61.9), followed by the SVM model with 60.8. In terms of precision, the SVM model scored the highest (66.6), followed by DT with 62.8. For accuracy, SVM led with 67.8, followed by DT with 65.6. In the F1-Score metric, DT achieved 62.1, followed by the LR model with 60.6. Finally, for AUC, DT also led with 66.54, followed by SVM with 62.23. In summary, the models that showed the best overall performance according to the evaluated metrics were DT and SVM.

TABLE 2. **Predictive performance of models to estimate the risk of EF impairments among university students**

Results					
Metrics (%)	Models				
	SVM	KNN	RF	LR	DT
Recall	60.8	54.3	50.2	60.5	61.9
Precision	66.6	56.1	50.2	61.5	62.8
Accuracy	67.8	61.1	56.7	64.4	65.6
F1-Score	60.3	53.2	48.9	60.6	62.1
AUC	62.23	50.63	54.35	59.61	66.6

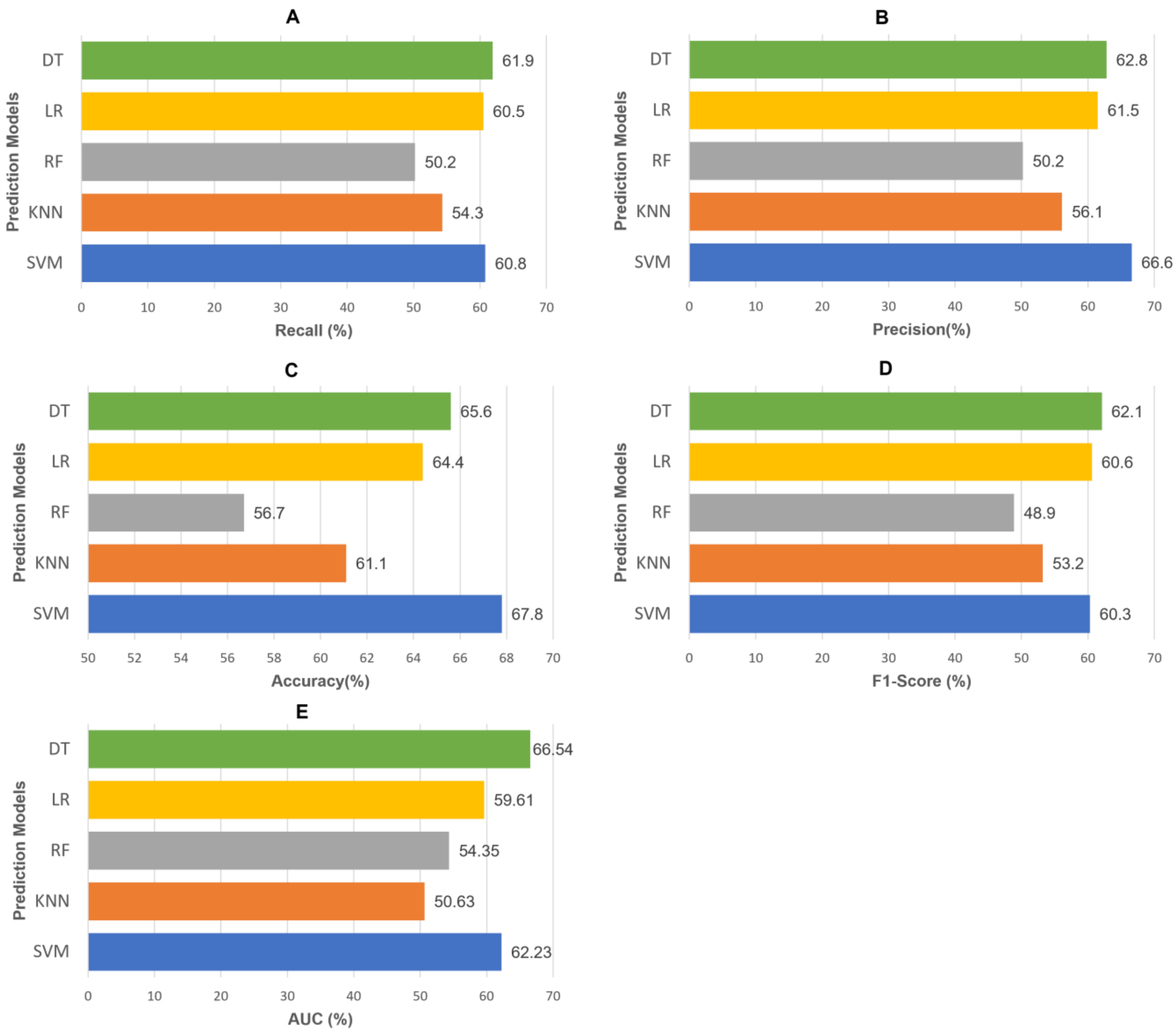

FIGURE 2. **Evaluation metrics of prediction performance for all developed models. (A) Recall; (B) Precision; (C) Accuracy; (D) F1-Score; (E) AUC**

A smaller overlap was identified between participants with and without cognitive process impairments, particularly in the DT and RF models (Figure 3). This result indicated that a clear separation of the predicted risk was achieved in both models, distinguishing between participants with cognitive impairments and those without. In contrast, there was a relatively large overlap in the LR, KNN, and SVM models.

Figure 3. Density curves for all developed models. (A) Logistic regression; (B) Decision Tree; (C) KNN; (D) Random Forest (E) Support Vector Machine. The red curve represents subjects without impairments in EF processes, while the blue curve represents subjects with impairments in EF processes

Discussion

From statistical analysis, it was identified that mother's level of education is a relevant factor (p-value < 0.01) in the development of cognitive processes in university students, as it influences the quality of the educational environment at home, the access to cultural and economic resources, and the promotion of academic expectations. Mothers with higher educational levels tend to provide enriched cognitive stimuli and knowledge about health and nutrition, which supports proper neurological development[25][26][27]. However, this factor is not deterministic, as cognitive development also depends on factors such as educational quality, sociocultural context, and public policies[28]. In regions of Mexico such as Oaxaca, mothers with lower formal educational levels may transmit cultural values and practical skills that contribute to the cognitive development of their children, highlighting the importance of inclusive and culturally adapted educational approaches[29].

Finally, nutrition is considered a risk factor for the presence of cognitive process impairments in young people. The relationship between nutrition and cognitive processes in young people has been extensively studied, identifying both risk factors and protective elements associated with diet. Malnutrition, especially during critical stages of development, can cause brain damage that affects EF. The severity and timing of malnutrition are key determinants of its impact on cognitive development[30]. Although nutrition is an essential component, cognitive development is also influenced by genetic factors, life experiences, socioeconomic environment, and educational quality. It is the interaction of these elements that shapes human thinking and learning[31]. Some studies suggest that despite adverse nutritional conditions, certain individuals develop resilience mechanisms that allow them to maintain adequate EF, implying that nutrition is not the only determinant of cognition[32].

The study developed and validated several machine learning models to predict the risk of cognitive process impairments in young people, using data from multiple domains, such as sociodemographic variables and lifestyle factors (physical activity, nutrition, sleep, stress, emotional management, and social environment). Among the most notable findings, it was identified that the DT and SVM models performed better in key metrics such as recall, precision, accuracy, F1 score, and AUC. However, the DT model stood out as the most effective in prediction, as evidenced by the probability density plots, which showed a greater capacity to identify risk patterns. DT and SVM machine learning models enable an efficient and accurate analysis of the data of the PEPS-I questionnaire and sociodemographic variables to identify risk patterns. Among the advantages of using these models is their ability to detect non-linear interactions between variables, such as the time spent on physical activity and sleep quality, which might not be evident in traditional analyzes.

Despite the progress made, there are few studies that have explored the integration of predictive models to predict EF impairments using categorical and continuous data simultaneously[33][34]. This combination allows for a comprehensive analysis of factors related to different dimensions, such as the social environment, physical activities, eating habits, and other characteristics associated with lifestyle. The implementation of these approaches represents an innovation in the use of artificial intelligence to address complex public health issues, with the potential to improve early detection and guide personalized interventions in at-risk groups.

This work not only offers new perspectives on the application of predictive models, but also underscores the importance of validating their performance in specific contexts and assessing their ability to handle heterogeneous data. In particular, the performance of the DT model highlights its practical usefulness due to its interpretability and ability to generate clear rules that could be easily applied in clinical or educational setting It is necessary to consider several limitations when interpreting the current findings. The data used in this study were cross-sectional, which is a limitation. In addition, there are various factors that can alter executive and cognitive functions; the use of the PEPS-I questionnaire simplified the process of evaluating these covariates. However, the reliability of the data is reduced as the responses were self-reported by the study participants.

CONCLUSIONS

This study highlights the relevance of adopting an integrated and interdisciplinary approach to understanding EF alterations in university students from Oaxaca. While maternal education level may influence cognitive outcomes, it should not be viewed as an isolated determinant. EF alterations stem from a complex interplay of structural, cultural, and socioeconomic variables. Addressing these factors through culturally sensitive and evidence-based strategies is essential for designing inclusive interventions in public health and education.

From a biomedical engineering perspective, the integration of machine learning techniques particularly DT models demonstrate significant potential in identifying and interpreting patterns within multifactorial data. These tools enable the translation of lifestyle and sociodemographic variables into actionable insights, contributing to the development of personalized and preventive approaches in cognitive health.

Nutritional deficiencies, while confirmed as a critical risk factor, must be understood within a broader context of social determinants of health. The application of machine learning algorithms to data obtained through instruments like the PEPS-I questionnaire supports a systemic view of cognitive development and facilitates early detection of EF impairments.

In summary, this research validates the use of supervised learning models as effective, low-cost, and scalable tools for the biomedical monitoring of cognitive risk factors in young populations. The findings support the design of intelligent systems and digital health applications capable of aiding in decision-making processes, shaping targeted interventions, and guiding public policy toward improving mental health and academic outcomes.

ETHICAL STATEMENT

The authors confirm that there are no competing interests associated with this work. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. All necessary ethical approvals and informed consents were obtained in accordance with institutional guidelines. The experimental data and complementary information supporting the findings of this study are available on GitHub https://github.com/jarillo-silva/filesData?tab=readme-ov-file#filesdata

REFERENCES

[1] P. Anderson, "Assessment and development of executive function (ef) during childhood", Child neuropsychology, vol. 8, no. 2, 2002, doi: https://doi.org/10.1076/chin.8.2.71.8724

[2] J.R. Best and P.H. Miller, "A developmental perspective on executive function", Child development, vol. 81, no. 6, 15 Nov. 2010, doi: https://doi.org/10.1111/j.1467-8624.2010.01499.x

[3] A. Diamond, "Executive functions", Annual review of psychology, vol. 64, no. 1, January 2013, doi: https://doi.org/10.1146/annurev-psych-113011-143750

[4] C. Blair and C.C. Raver, "School readiness and self-regulation: A developmental psychobiological approach", Annual review of psychology, vol. 66, no. 1, January 2015, doi: https://doi.org/10.1146/annurev-psych-010814-015221

[5] S.J. Blakemore and S. Choudhury, "Development of the adolescent brain: implications for executive function and social cognition", Journal of child psychology and psychiatry, vol. 47, no. 3-4, May 2006, doi: https://doi.org/10.1111/j.1469-7610.2006.01611.x

[6] A.G. Gutiérrez-García and M. Landeros-Velázquez, "Evaluación de funciones ejecutivas en estudiantes universitarios con niveles de autoeficacia percibida baja", Revista Electrónica de Psicología Iztacala, vol. 20, no. 2, June 2017.

[7] J.R. Best, P.H. Miller and J.A. Naglieri, "Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample", Learning and individual differences, vol. 21, no. 4, August 2011, doi: https://doi.org/10.1016/j.lindif.2011.01.007

[8] A. Diamond and D.S. Ling, "Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not", Developmental cognitive neuroscience, vol. 18, no. 1, April 2016, doi: https://doi.org/10.1016/j.dcn.2015.11.005

[9] N. Zhu, D.R. Jacobs, K. Meyer, K. He, L. Launer, J. Reis, K. Yaffe, S. Sidney, R. Whitmer and L. Steffen, "Cognitive function in a middle aged cohort is related to higher quality dietary pattern 5 and 25 years earlier: the cardia study", The journal of nutrition, health & aging, vol. 19, no. 1, 2015, doi: https://doi.org/10.1007/s12603-014-0491-7

[10] J. Flores, F. Ostrosky and A. Lozano," Banfe-2 batería neuropsicológica de funciones ejecutivas y lóbulos frontales. México: Manual moderno", 2014.

[11] F.S. Mogollón García, L.D. Becerra Rojas and J.S. Adolfo Ancajima Mauriola, "Estilos de vida saludables en estudiantes de pregrado", Conrado, vol. 16, no. 75, July 2020. Available: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1990-86442020000400069&lng=es&tlng=es

[12] Y.E. Salazar-Granizo, C. Hueso-Montoro and R.A. Caparros-Gonzalez, "Lifestyles and Academic Stress in University Students of Health Sciences: A Mixed-Methodology Study", Healthcare, vol. 12, no. 1, January 2024, doi: https://doi.org/10.3390/healthcare12141384

[13] E.E. Bron, M. Smits, W.M. Van Der Flier, H. Vrenken, F. Barkhof, P. Scheltens, J.M. Papma, R.M. Steketee, C.M. Orellana, R. Meijboom, et al., "Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural mri: the caddementia challenge", NeuroImage, vol. 111, no. 1, 2015, doi: https://doi.org/10.1016/j.neuroimage.2015.01.048

[14] C. Simfukwe, S.S. A. An and Y.C. Youn, "Comparison of machine learning algorithms for predicting cognitive impairment using neuropsychological tests", Applied Neuropsychology: Adult, vol. 1, no. 1, Sep. 2024, doi: https://doi.org/10.1080/23279095.2024.2392282

[15] M.F. Jojoa-Acosta, S. Signo-Miguel, M.B. Garcia-Zapirain, M. Gimeno-Santos, A. Méndez-Zorrilla, C.J. Vaidya, M. Molins-Sauri, M. Guerra-Balic and O. Bruna-Rabassa, "Executive functioning in adults with down syndrome: machine-learning-based prediction of inhibitory capacity", International Journal of Environmental Research and Public Health, vol. 18, no. 20, sep. 2021, doi: https://doi.org/10.3390/ijerph182010785

[16] K. Modi, I. Singh, Y. Kumar, "A comprehensive analysis of artificial intelligence techniques for the prediction and prognosis of lifestyle diseases", Archives of Computational Methods in Engineering, vol. 30, no. 8, june 2023, doi: https://doi.org/10.1007/s11831-023-09957-2

[17] L. Zhang, S. Zhao, W. Yang, Z. Yang, Z. Wu, H. Zheng, M. Lei, "Utilizing machine learning techniques to identify severe sleep disturbances in chinese adolescents: an analysis of lifestyle, physical activity, and psychological factors", Frontiers in Psychiatry, vol. 15, nov. 2024, doi: https://doi.org/10.3389/fpsyt.2024.1447281

[18] Z. John Lu, "The elements of statistical learning: data mining, inference, and prediction", Journal of the Royal Statistical Society Series A: Statistics in Society, vol. 173, no. 3, june 2010, doi: https://doi.org/10.1111/j.1467-985X.2010.00646_6.x

[19] S. Gornale, S. Kumar, R. Siddalingappa, P.S, " Hiremath, Survey on handwritten signature biometric data analysis for assessment of neurological disorder using machine learning techniques", Transactions on Machine Learning and Artificial Intelligence, vol. 10, no. 2, April 2022, doi: https://doi.org/10.14738/tmlai.102.12210

[20] S. Swati, M. Kumar, S. Namasudra, "Early prediction of cognitive impairments using physiological signal for enhanced socioeconomic status", Information Processing & Management, vol. 59, no. 2, March 2022, doi: https://doi.org/10.1016/j.ipm.2021.102845

[21] J. Hou, H. Jiang, Y. Han, R. Huang, X. Gao, W. Feng, Z. Guo, "Lifestyle influence on mild cognitive impairment progression: A decision tree prediction model study", Neuropsychiatric Disease and Treatment, vol. 20, no. 1, Feb. 2024, doi: https://doi.org/10.2147/NDT.S435464

[22] D.W. Hosmer Jr, S. Lemeshow, R.X, Sturdivant, Applied logistic regression, 3rd ed. USA: John Wiley & Sons, pp. 35.

[23] L. Breiman, "Random forests", Machine learning, vol. 45, no. 1, Jan. 2001, doi: https://doi.org/10.1023/A:1010933404324

[24] D.M. Powers, "Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation", International Journal of Machine Learning Technology, vol 2, no. 1, Jan. 2001, doi: https://doi.org/10.48550/arXiv.2010.16061

[25] G.W. Evans, P. Kim, "Childhood poverty, chronic stress, self-regulation, and coping", Child development perspectives, vol. 7, no. 1, Nov. 2013, doi: https://doi.org/10.1111/cdep.12013

[26] P.E. Davis-Kean, "The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment", Journal of family psychology, vol. 19, no. 2, Sep. 2005, doi: https://doi.org/10.1037/0893-3200.19.2.294

[27] P. Glewwe, E.A. Miguel, "The impact of child health and nutrition on education in less developed countries" Handbook of development economics, vol. 4, 2007, doi: https://doi.org/10.1016/S1573-4471(07)04056-9

[28] UNESCO, Global Education Monitoring Report 2019: Migration, Displacement and Education-Building Bridges, Not Walls ,UN, Mexico, 2018. Accessed: Feb. 25, 2025. [Online]. Available: https://hdl.handle.net/20.500.12799/6023

[29] K.Y.C. Cuellar, G.I. Díaz, R.A. Ortíz, N.M. Pérez, J.G. Ariza, "Genealogías culturales e historias de familia en oaxaca: Diseño epistémico, estrategia metodológica y reflexividad", Estudios sobre las culturas contemporáneas, vol. 20, no. 40, 2014. Accessed: Dec. 12, 2024. [Online]. Available: https://dialnet.unirioja.es/servlet/articulo?codigo=4990957

[30] J.R. Ocaña-Noriega, G.S. Sagñay-Llinin, "La malnutrición y su relación en el desarrollo cognitivo en niños de la primera infancia", Polo del conocimiento, vol. 5, No. 12, Sep. 2020, doi: https://doi.org/10.23857/pc.v5i12.2044

[31] J.A. Cardona Murillo, F.Z. Leidy Andrea, "Influencia de la alimentación en el desarrollo cognitivo de los estudiantes en el grado de transición de la institución educativa Luis Carlos Galán Sarmiento del municipio de Carepa", Corporación Universitaria Minuto de Dios, Urabá, Antioquia, Colombia, 2021. [Online]. Available: https://repository.uniminuto.edu/items/9c2b881d-8efc-4eee-941b-3901a995d529

[32] J. Bernal, "Resiliencia, infancia y nutrición: propuesta de indicadores para medición en niños y adolescentes", Revista Salud Bosque, vol. 7, no. 2, Feb. 2017, doi: https://doi.org/10.18270/rsb.v7i2.2192

[33] P. Zaninotto, G.D. Batty, M. Allerhand, I.J. Deary, "Cognitive function trajectories and their determinants in older people: 8 years of follow-up in the English longitudinal study of ageing", J Epidemiol Community Health, vol. 72, no. 8, Aug. 2018, doi: https://doi.org/10.1136/jech-2017-210116

[34] G.R. Poudel, A. Barnett, M. Akram, E. Martino, L.D. Knibbs, K.J. Anstey, J.E. Shaw, E. Cerin, "Machine learning for prediction of cognitive health in adults using sociodemographic, neighbourhood environmental, and lifestyle factors", Inter- national Journal of Environmental Research and Public Health, vol. 19, no. 17, 2022, doi: https://doi.org/10.3390/ijerph191710977