Machine Learning Models Applied to the Analysis of Results from the Educational Quality Assessment Operation: "Aprender-2018"

Andrés Francisco Farías[1], Germán Antonio Montejano[2],

Ana Gabriela Garis[3], Andrés Alejandro Farías[4]

Sebastián Javier Farías[5]

National University of La Rioja, La Rioja, Argentina[1, 4]
National University of San Luis, San Luis, Argentina[2, 3]
afarias665@yahoo.com.ar[1], gmonte@unsl.edu.ar[2],
agaris@gmail.com[3], andresaf86@hotmail.com[4]
sebajavfarias@gmail.com[5]

Abstract. This paper proposes the use of machine learning models in the "Aprender" standardized assessment tests implemented in Argentina. These tests measure language and mathematics performance in primary and secondary school. The proposed study used data from the 2018 edition of the sixth-grade primary education assessment. During the research phase, language and mathematics performance were analyzed, the results of which are presented in this article. To this end, a preliminary feature selection was performed, followed by a preselection of some of the models used in this experiment, belonging to the Python library Scikit-Learn (Sklearn). The following classifier methods were considered: Extra Tree Classifier, Decision Tree Classifier, Random Forest Classifier, Gradient Boosting Classifier, and Kneighbors Classifier. Of these, the model that achieved the highest level of accuracy was identified. In addition, the datasets used underwent preliminary processing, during which missing and negative data were filled using the median of each column. Finally, the most significant features that lead to the best results were identified.

Keywords: Machine Learning Models, Selection Feature, Selection Models, Education

1 Introduction

The The procedure consisted of analyzing the Standardized Assessments of Educational Quality in Argentina, the National Assessment "Aprender 2018," using various machine learning models to determine the most notable characteristics in each case [1], [2].

Data frames from the National Operation "Aprender 2018" were processed, covering performance in Language and Mathematics in sixth grade of primary schools at the census level.

A feature preselection process was then carried out to determine whether any needed to be discarded.

Subsequently, a preselection of some of the models used in this experiment was made, belonging to the Python library Scikit-Learn (Sklearn). The following ensemble methods were considered: Extra Tree Classifier, Decision Tree Classifier, Random Forest Classifier, Gradient Boosting Classifier, and Kneighbors Classifier; the one that achieved the highest accuracy was chosen [3]. The data frames were then processed with the selected model. The results were tabulated and graphed for interpretation.

2 National Educational Quality Assessment Operation: "Aprender-2018"

The "Aprender" National Operation, applied in both primary and compulsory secondary education, is a learning assessment instrument implemented in Argentina since 2016.

This learning assessment dates back to the National Evaluation Operations (ONE), implemented between 1993 and 2013.

As mentioned above, the "Aprender" assessment began in 2016 and has been implemented in various editions, with a primary focus on assessing language and mathematics skills. However, in some editions, additional subjects such as Social Sciences, Natural Sciences, Citizenship Education, and Written Production have also been included.

The test's design varies in terms of coverage, as its implementation is based on a census in some cases and on a sample basis in others.

These standardized assessments facilitate not only the analysis of student performance in language and mathematics, but also, through the implementation of complementary booklets, the collection of information on the contexts in which this learning occurs. Research is also conducted on the influence of various factors on learning, including family characteristics, the educational context and attributes of each student, the educational level of the parents, extracurricular activities that students engage in outside of school hours, household assets, and the student's educational trajectory.

These factors determining learning have been analyzed since the first edition of "Aprender," and according to the Ministry of Education itself, the most appropriate model to identify all the factors determining learning would be the following: student learning or classroom average; unobservable family factors of the student; child development, a trait that seeks to capture, primarily, the nutrition and stimulation received in preschool; the socioeconomic level of the home; the educational level attained by the mother and father; the student's human capital; school capital, subdivided into physical, human, and social; classroom effect, also referred to in this report as the “mystery classroom,” a characteristic that attempts to capture the effectiveness of each teacher and their relationship with students; institutional factors, including school organization.

The model used in Argentina was reduced to five of these classes of independent characteristics: the socioeconomic status of the household, the educational level attained by the mother and father, the human capital of the students, the capital of the schools, and institutional factors [4], [5].

However, due to methodological issues, such as the young age of the respondents, not all of these characteristics can be administered in this type of assessment instrument. Consequently, a "forced omission" of these characteristics must be carried out, including those that can be effectively measured and yield valid results.

The assessment process is structured around four distinct performance levels: these levels are defined as "Below Basic," "Basic," "Satisfactory," and "Advanced," as described in the technical notes section of the national report. The analysis of the data for Language and Mathematics is carried out with respect to the variations presented by the associated factors [6].

3 Description of the procedure

To analyze the data from the National Operation "Aprender 2018", applying machine learning models, the following procedure was implemented [7],[8]:

- Data Debugging and analyzing the data contained in the dataframe
- Pre-selection of features from dataframes
- Pre-selection of Machine Learning Models
- Application of the selected Machine Learning model to the dataframes and visualization.
- Interpretation of the results

3.1 Data Debugging and analyzing the data

The Language Performance data frame contains 562,214 rows, while the Mathematics Performance data frame contains 556,802. In both cases, a significant amount of missing data and negative values were identified. To address this problem, the median completion method was used within each column to replace the missing values.

Columns with data that were not significant for the analysis were removed.

3.2 Pre-selection of features from dataframes

Language Performance. The Sequential Forward Feature Selection technique was used for the preliminary selection, considering the average score of the different feature groupings as a comparison value. Initially, an exploratory analysis was performed to determine the optimal grouping of the features [9], [10]. Subsequently, the results table was analyzed to identify the features that constituted the best performing group, as shown in Table 1 and Figure 1.

Table 1. Sequential Forward Feature Selection, Language Performance. Highlighted in yellow, the best result

Numbers features	avg_score
1	0,469428
2	0,470373
3	0,471763
…..	……..
23	0,504336
24	**0,504339**
25	0,504166
26	0,504080
27	0,504141
28	0,504189
29	0,503991

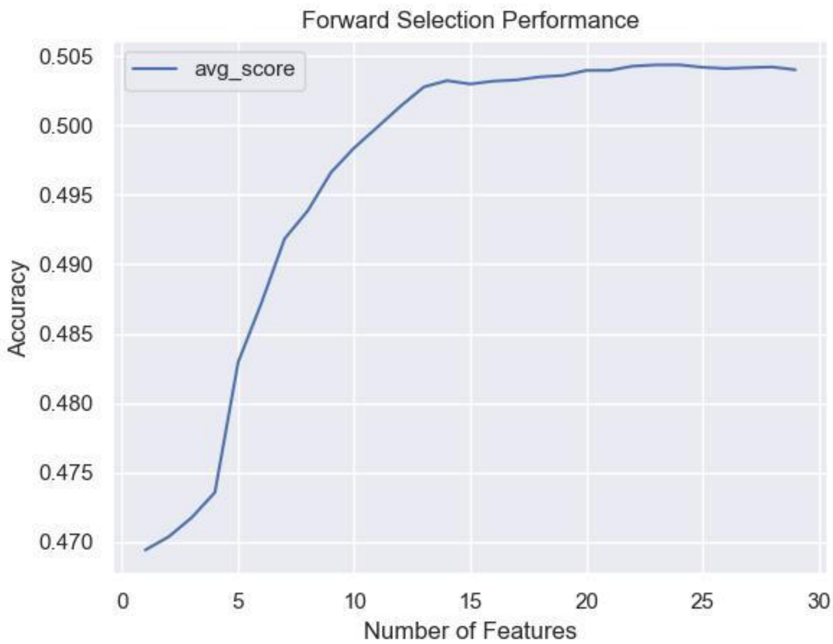

Fig. 1. Graph of the grouping of features for Language Performance

Performance Mathematics. The Sequential Forward Feature Selection technique was used for the preliminary selection, considering the average score of the different feature groupings as a comparison value. Initially, an exploratory analysis was performed to determine the optimal grouping of the features [9], [10]. Subsequently, the results table was analyzed to identify the features that constituted the best performing group, as shown in Table 2 and Figure 2.

Table 2. Sequential Forward Feature Selection, Mathematics Performance. Highlighted in yellow, the best result

Numbers features	avg_score
1	0,389046
2	0,394201
3	0,410464
…..	……..
22	0,425720
23	**0,425733**
24	0,425661
25	0,425635
26	0,425563
27	0,425422
28	0,425551

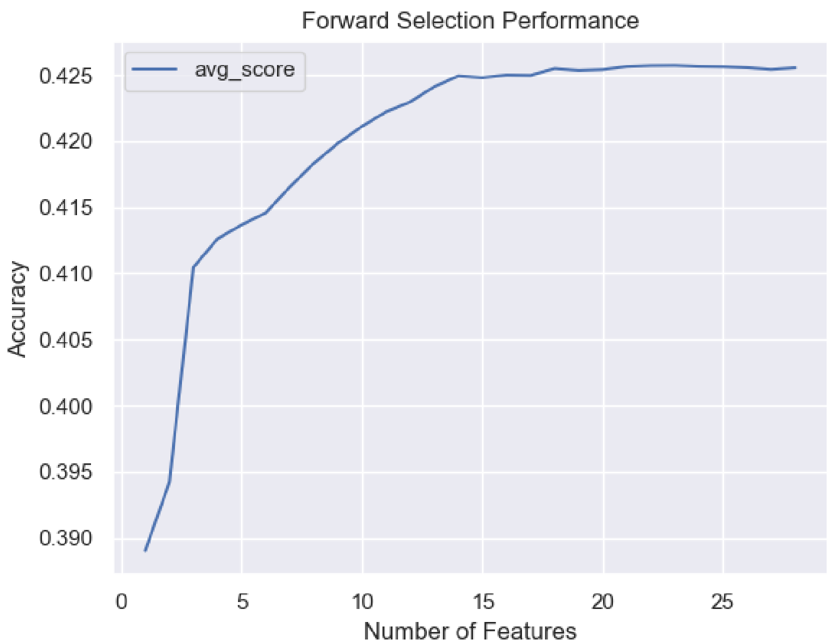

Fig. 2. Graph of the grouping of features for Mathematics Performance

3.3 Pre-selection of Machine Learning Models

Language Performance and Performance Mathematics. In this research, certain models from the Scikit-Learn (Sklearn) Python library were employed on dataframes, with the optimal grouping of features. The model that attains the highest score is selected, as indicated in Table 3 and Table 4, the Gradient Boosting Classifier model presents the highest level of accuracy [11], [12].

Table 3. Language Performance : Accuracy of the models of Machine Learning. Highlighted in yellow, the best result

Models	Accuracy
ExtraTreesClassifier	0,3638
Decision Tree Classifier	0,3589
Random ForestClassifier	0,3627
GradientBoostingClassifier	**0,3671**
KNeighborsClassifier	0,3186

Table 4. Mathematics Performance : Accuracy of the models of Machine Learning. Highlighted in yellow, the best result

Models	Accuracy
ExtraTreesClassifier	0,3926
Decision Tree Classifier	0,3820
Random ForestClassifier	0,3940
GradientBoostingClassifier	**0,3966**
KNeighborsClassifier	0,3428

3.4 Application of the selected Machine Learning model to the dataframes and visualization

Language Performance. The model that obtained the highest score in the preselection is used: Gradient Boosting Classifier, using optimal feature pooling. The results are shown in Table 5, ordered from highest to lowest.

Table 5. Importance of Features for Language Performance

Features	Description of the features	Importance
sector	Management Sector	0,2749
ap20	Are you doing well in your language class?	0,0954
isocioa	Student Socioeconomic Index	0,0947
ap4	How many people do you live with?	0,0900
ap8	Approximately how many books are there where you live?	0,0794
ap14	Besides attending school, do you help your parents or relatives with their work?	0,0696
ap21a	In your opinion, how well do you read?	0,0686

Features	Description of the features	Importance
ap17	Have you ever repeated a grade?	0,0467
ap15	Do you work outside the home for someone other than your family?	0,0356
ap19	Do you get along well with your classmates?	0,0257
ap21b	In your opinion, how well do you write?	0,0200
ap9	What is your mother's highest educational level?	0,0197
ap31	Do you look for information or discuss these topics on the internet (social media, websites, forums, etc.)	0,0175
ap6	How many rooms does the place where you live have, not counting the kitchen and bathroom?	0,0116
ap16	Did you go to kindergarten?	0,0105
ap11	Are your mother or father members of Indigenous communities or descendants of Indigenous communities?	0,0090
ap10	What is your father's highest educational level?	0,0088
ap24	When you're working in class... Do your teachers explain things to you again if you don't understand?	0,0073
ambito	Setting	0,0064
ap12	Do they speak any indigenous languages at home?	0,0040
migración	Immigration status	0,0036
ap38	How many days did you miss school for this reason?	0,0009
ap18	Do you like going to school?	0,0001
ap42	You used a notebook with enlarged type because...	0,0000

The columns chart shows the data from the previous table, Figure 3.

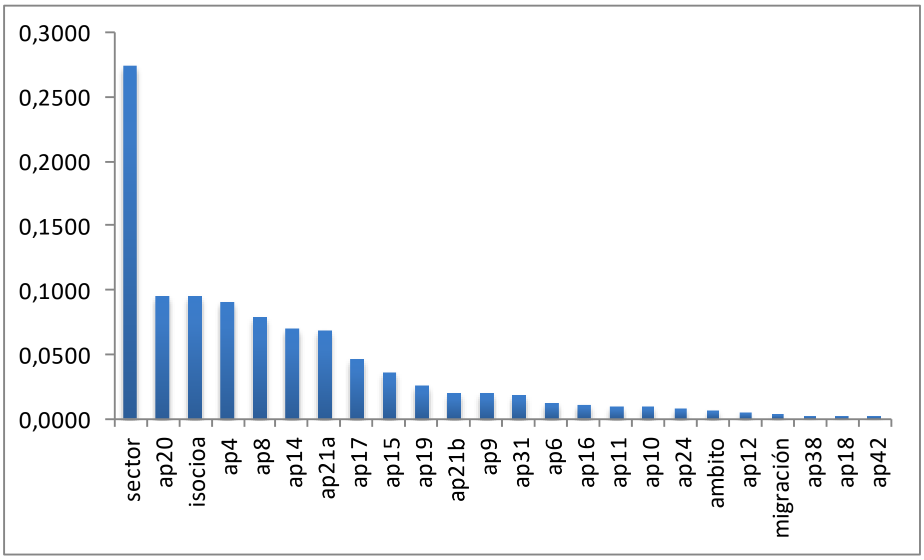

Fig. 3. Column chart of the importance of features in Language Performance

Performance Mathematics. The model that obtained the highest score in the preselection is used: Gradient Boosting Classifier, using optimal feature pooling. The results are shown in Table 6, ordered from highest to lowest.

Table 6. Importance of Features for Mathematics Performance

Features	Description of the features	Importance
sector	Management Sector	0,2097
ap23	In your opinion, how do you solve math problems?	0,1583
ap22	Are you doing well in your math class?	0,1271
isocioa	Student Socioeconomic Index	0,1238
ap8	Approximately how many books are there where you live?	0,0729
ap17	Have you ever repeated a grade?	0,0646
ap4	How many people do you live with?	0,0585
ap14	In addition to attending school, do you help your parents or relatives at work?	0,0391
ap15	Do you work outside the home for someone who is not part of your family?	0,0292
ap6	How many rooms does the place where you live have, not counting the kitchen and bathroom?	0,0231
ap31	Do you look for information or discuss these topics on the internet? (social media, websites, forums, etc.)	0,0176
ap9	What is your mother's highest level of education?	0,0152
ap16	Did you go to kindergarten?	0,0134
ap2	Gender	0,0104
ap10	What is your father's highest level of education?	0,0098
ap11	Do your mother or father belong to Indigenous groups or are they descendants of Indigenous groups?	0,0087
ap12	Do they speak any indigenous languages at home?	0,0068
migración	Immigration status	0,0044
ap1	How old are you?	0,0037
ap37	Did you have to miss school to accompany your family due to work-related relocations?	0,0026
ap38	How many days did you miss school for this reason?	0,0007
ap18	Do you like going to school?	0,0005
ap42	Did you use a notebook with enlarged type because...	0,0001

The columns chart shows the data from the previous table, Figure 4.

3.5 Interpretation of the results

In Language performance, the most important features are the management sector (sector), the student's perception of how they do in Language class (ap20), the student's socioeconomic index (isocioa), the indicative feature of overcrowding, how many people the student lives with (ap4), the number of books in the student's home (ap8), collaboration in family work (ap14), the perception of how well they read (ap21a), the level of repetition (ap17), work outside the home (ap15), how they get along with classmates (ap19), the perception of how the student writes (ap21b), and the mother's educational level (ap9). In Mathematics Performance, the most significant features are: management sector (sector), perception of how well students solve math problems (ap23), perception of how well they do in math class (ap22), student socioeconomic status (isocioa), how many liqueurs they have at home (ap8), grade repetition rate (ap17), how many people they live with (ap4), whether they work inside their home (ap14) or outside the home (ap15), how many bedrooms their home has (ap6), whether they search for information online (ap31), and their mother's educational level (ap9).

The results obtained do not differ from the official Ministry reports, which were processed using other resources.

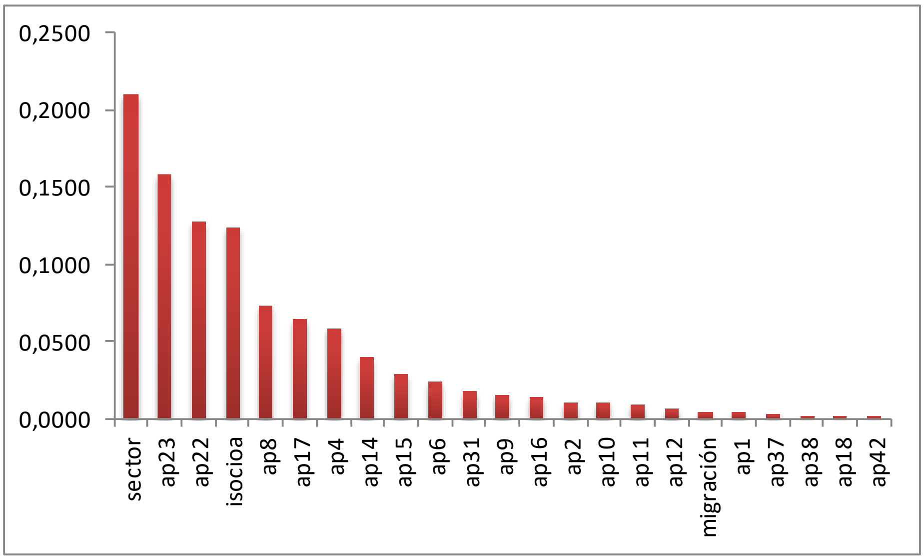

Fig. 4. Column chart of the importance of features in Mathematics Performance

4 Conclusions and Future Work

The results of this work are consistent with those of official reports obtained using other analytical techniques.

Interestingly, processing large volumes of data from assessment programs using machine learning models yields results that show the same trends as the statistical analyses performed on these same tests.

Clearly, the objective of this work has been successfully met: to demonstrate the viability of using machine learning tools on the results of standardized assessments of educational quality that handle large volumes of data for the analysis and interpretation of educational performance.

For future work, the techniques used in this article can be applied to the various "Aprender" and third-grade operational programs in primary and secondary school.

References

[1] Llach, J. J., & Cornejo, M. (2018). *Factores condicionantes de los aprendizajes en la escuela primaria y media.* In LIII Reunión Anual de la Asociación Argentina de Economía Política (La Plata, 14 al 16 de noviembre de 2018).

[2] Ministerio de Educación, Cultura, Ciencia y Tecnología. (2018) *Factores Condicionantes de los Aprendizajes. Primaria y Secundaria. Ciudad Autónoma de Buenos Aires*, Argentina. Recuperado de https://www.argentina.gob.ar/sites/default/files/factores_condicionantes_de_los_aprendizajes.pdf

[3] Abro, A. A., Khan, A. A., Talpur, M. S. H., Kayijuka, I., & Yaşar, E. (2021). Machine learning classifiers: a brief primer. *University of Sindh Journal of Information and Communication Technology*, *5*(2), 63-68

[4] Llach, J. J., Schumacher, F., & Llach, J. (2006). *La segregación social en la educación primaria argentina*. Buenos Aires: Granica.

[5] Ministerio de Educación y Deportes. Secretaría de Evaluación Educativa (2016) *Aprender 20216. Informe de Resultados*. Recuperado de https://www.argentina.gob.ar/sites/default/files/reporte_nacional.pdf

[6] Adrogué, C. (2014). *Calidad y equidad de la educación primaria pública argentina*. Pilquen-Sección Psicopedagogía, 11(1), 12.

[7] Chakrapani, P., & Chitradevi, D. (2022, April). *Academic performance prediction using machine learning: A comprehensive & systematic review*. In 2022 International Conference on Electronic Systems and Intelligent Computing (ICESIC*)* (pp. 335-340). IEEE.

[8] Adedeji, O. B., Olayinka, O. O., Adebare, A., Idowu, P. A., Ayoade, A. O., & Ademola, A. O. (2025). *A Machine Learning-Based Predictive Model for the Classification of Academic Performance of Students*. University of Ibadan Journal of Science and Logics in ICT Research, 13(1), 69-79.

[9] Ververidis, D., & Kotropoulos, C. (2005, September). *Sequential forward feature selection with low computational* cost. In 2005 13th European Signal Processing Conference (pp. 1-4). IEEE.

[10] Kaur, A., Guleria, K., & Trivedi, N. K. (2021, March). Feature selection in machine learning: Methods and comparison. In *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 789-795). IEEE.

[11] Konstantinov, A. V., & Utkin, L. V. (2021). *Interpretable machine learning with an ensemble of gradient boosting machines.* Knowledge-Based Systems, 222, 106993.

[12] Kunapuli, G. (2023). *Ensemble methods for machine learning*. Simon and Schuster.