

Explaining Factors of Student Attrition at Higher Education

Iara Alcauter, Lourdes Martinez-Villaseñor,
Hiram Ponce

Universidad Panamericana, Mexico City,
Mexico

{ialcaute, lmartine, hponce}@up.edu.mx

Abstract. The examination of student attrition within higher education is a dynamic field that seeks to tackle the complex task of preventing dropout occurrences and formulating effective retention strategies. This challenge becomes particularly pertinent within the realm of Science, Technology, Engineering, and Mathematics (STEM) disciplines. In the pursuit of these objectives, this research endeavors to assess prevailing data mining methodologies, specifically focusing on Decision Trees (DT), Random Forest (RF), Support Vector Machine (SVM), and Artificial Neural Networks (ANN) – all of which are widely employed for the prediction of student attrition. The study is conducted on a comprehensive dataset encompassing engineering students from a prominent Mexican university, with a specific emphasis on the application of variable selection through Recursive Feature Elimination (RFE) and addressing class imbalance via Synthetic Minority Over-sampling Technique (SMOTE). The outcomes of this investigation conspicuously identify Random Forest as the most optimal predictive model, yielding an impressive accuracy rate of 98%. Additionally, the research underscores the effectiveness of RFE in discerning influential variables. Furthermore, to provide complex insights and decision support, the study harnesses the Local Interpretable Model-Agnostic Explanations (LIME) technique to expound upon the factors that wield significant impact. This multifaceted analysis contributes to the advancement of strategies for enhancing student retention within STEM disciplines.

Keywords. Student attrition, machine learning, XAI, explainable artificial intelligence, higher education.

1 Introduction

Higher education degrees hold significant value in Mexico, like most OECD countries, as they lead to improved labor market outcomes

compared to lower educational levels [22]. Higher education institutions (HEIs) play a crucial role in a country's economic and social development, contributing to UN Agenda 2030 for Sustainable Development [32].

In Mexico the number of enrolled students in higher education has increased, with over three thousand institutions that offer more than thirty-five thousand educational programs. 35% of the Universities in Mexico are private [4].

In Latin America, access to university grew dramatically in the early 2000s, and in particular for those students from the low and middle income segments [13]. Most of these "new students" enrolled in new private universities, based on recent growth in middle-class household income, student loans, and scholarships [12].

Enrollment management and student retention have become priorities in universities in the United States and other developed countries worldwide. University dropout, understood as the discontinuation of studies without returning within a specified period, is a global phenomenon occurring in both public and private institutions.

Mexico is not exempt from this problem, which causes institutional, familial, and personal economic losses, as well as psychological issues and other negative social impacts [28].

In Higher education institution, the student attrition rate is one of the most commonly used indicators internationally to evaluate the internal efficiency of teaching and learning processes in tertiary education institutions [1]. Besides student dropout typically results in overall financial loss, lower graduation rates, and an inferior school reputation in the eyes of all stakeholders [14].

Defining school dropout is complex because there are no clear theoretical parameters that delimit it [18]. The term "at risk student" is commonly used in the field of education to describe a student who is at high risk of academic failure and who often requires the support and intervention of instructors to achieve academic success. Addressing this issue is essential for improving student retention and the societal impact of universities [33].

The primary purpose of this article is to present a comparison between different machine learning models to estimate the dropout of students in the engineering faculty of a private university in Mexico, seeking to decrease the dropout rate. The study encompasses four models: Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM) and Artificial Neural Networks (ANN).

Variable selection techniques and data balancing techniques were applied to enhance model performance. Additionally, we utilize Local Interpretable Model-Agnostic Explanations (LIME) to provide comprehensive insights into prediction factors.

This study makes significant contributions in several aspects. Firstly, it involves an in-depth analysis and identification of the key factors that influence student dropout within the field of engineering at one of the premier private universities in Mexico. Secondly, it devises a comprehensive model applicable to the entire school emphasizing the impact of academic performance specifically within the realm of mathematics-related subjects.

Thirdly, the study thoroughly evaluates the efficacy of widely used machine learning models for predicting dropout, ensuring optimal precision through a meticulous tuning of hyper parameters. Lastly, the study employs the elucidating capabilities of LIME to provide detailed explanations for the factors contributing to dropout. The organization of this study is structured into distinct sections.

The initial section serves as an introduction, delineating the salient aspects of the problem, outlining the path to its resolution, and emphasizing the contributions made.

In Section 2, an exploration of prior research within this domain is presented, underscoring its significance. Section 3 expounds upon the methodology employed, elucidating each technique and model utilized. The experimental setup is elaborated upon in Section 4, while the conclusions drawn from the study are encapsulated in Section 5.

2 Related Work

Vincent Tinto is an influential sociologist, known for his work on student retention and dropout in higher education. In [29], Tinto asserts that the lack of integration of students into the academic and social environment stands as one of the most influential factors contributing to student attrition.

He highlights the presence of various causes, encompassing personal, familial, economic, political, cultural, and institutional aspects, that either weaken or bolster a student's engagement.

Tinto further underscores the significance of implementing retention programs that provide support throughout students' university journey [30]. With Tinto's work as a background we observe in [17] that conceptualizing dropout is a matter more complex than most people think: The common description refers to students leaving their university studies before having completed their study program and obtained a degree.

Dropout definitions vary, including both voluntary and involuntary withdrawals. From another theoretical perspective, Astin's theory of student engagement propose a behavioral approach to understanding student attrition.

These theories accentuate the importance of student engagement in purposeful activities tied to enhanced learning outcomes. This perspective concludes that active student engagement plays a pivotal role in reducing university attrition rates [27].

The continuous flux of information generated by students upon entering university has spurred the development of educational data mining in various ways. In [8] that reviews the 50 most cited articles on the use of artificial intelligence in higher education, 46% of the articles are focused on the profiling and prediction of students with a focus on the conclusion of their studies.

As can be seen a pivotal application in this realm is the prediction of student performance, specifically aimed at identifying those who might be at risk of discontinuing their college journey.

Numerous scholars have harnessed data mining and machine learning techniques to prognosticate the determinants wielding the most influence on student retention and academic fulfillment [5].

In these eight comparative studies [23, 21, 3, 5, 10, 7, 16] and [19] associated with the prediction of withdrawals in higher education, four major concepts are evaluated: 1) The machine learning (ML) methods used, 2) The data used, 3) The metrics used to evaluate the performance of the model and in some of them 4) The size of the data set.

1) ML Methods: In all these articles it is established that the most used methods are classification methods. They all include as prediction methods: Logistic/linear Regression (LR), Decision Trees (DT), Random Forest (RF), Artificial Neural Networks (ANN), Support Vector Machine (SVM), K-Nearest Neighbors (KNN) and Naïve Bayes (NB). The most used are DT, RF, ANN and SVM. Notably, certain techniques exhibiting superior efficacy involve ensemble approaches like RF.

2) Data used: In general, the data is classified into: socio-demographic data, academic background, current academic data, characteristics of the program or university, behavioral characteristics, financial data and some add family background, behavior in learning management systems and activity on social networks. The most used data are current academic, socio-demographic and academic background data. Notably, a critical domain of focus lies within the realm of freshmen, as it represents the stage wherein a substantial proportion of dropout incidents materialize.

3) The most used measure is Accuracy, followed by Precision, Recall and f1-score. Additional metrics like area under the curve, mean absolute error and specificity are included.

4) The size of the data set is a characteristic that some of the reviews compare. No uniformity in size is observed, as there are from less than 100

records to more than 10,000. No observation is concluded in this regard.

The articles analyzed refer to prediction, presenting different approaches such as drop out prediction, prediction of reaching the end of the year, prediction of graduation time, prediction of leaving in the first year, etc.

It is noteworthy that a universal model encompassing all institutions is elusive; rather, model customization is contingent upon the unique array of variables inherent to each educational entity.

Given that the majority of Artificial Intelligence (AI) algorithms employed in student performance prediction rely on so-called black box techniques—where predictions are generated without explicating their origins—this study introduces two frameworks for Explainable AI, catering to both local and general explanations.

This distinction is pivotal for enabling precise actions based on predictions while also engendering a sense of trust by elucidating the origins of predictions to ensure impartiality and ethical reliability. Explainable AI (xAI), as a comprehensive concept, aims to construct and employ models that users can interpret and comprehend.

One avenue involves developing robust and fully explainable models, such as the deep k-nearest neighbors' approach and teaching explanations for decisions, as outlined by Dieber et al. [9].

The Local Interpretable Model-Agnostic Explanations (LIME) framework emerges as a prominent tool within the literature, particularly noted for its efficacy in explaining image-related matters.

The overarching objective of an Explainable AI (XAI) system is to render its behavior intelligible to human users by furnishing comprehensive explanations [15].

A proficient XAI system should elucidate its capabilities and comprehensions, delineate its past and present actions, forecast its subsequent steps, and unveil the crucial information shaping its decision-making process.

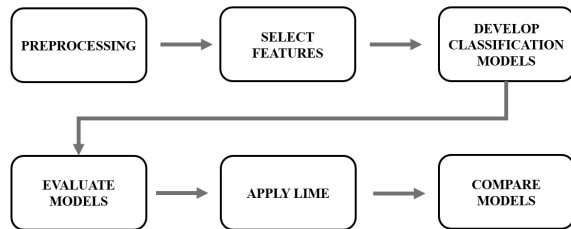


Fig. 1. Research Methodology

3 Methodology

In this section, we delineate the research methodology. The process is illustrated in Figure 1 which outlines the sequential stages undertaken to fulfill our objectives.

These stages align with the conventional steps inherent in the implementation of a machine learning model, grounded in the principles of Knowledge Discovery in Databases (KDD) as expounded in [11]. Each of the steps will be explained in the following sections.

3.1 Data

The dataset utilized in this study is authentic and exclusively constructed for the purpose of this research. It encompasses comprehensive data pertaining to students who have been enrolled in various academic programs within the School of Engineering at Private Mexican University. The information is predominantly sourced from the student information system, complemented by insights from diverse unstructured sources.

Encompassing 43 distinct features, the data set encapsulates a wide spectrum of student details, encompassing demographic attributes (age, gender, residence, nationality), academic history (high school background, GPA), particulars of the admission process (enrollment term, prerequisites), financial assistance, fiscal transactions (collections), academic attainment (overall GPA, GPA in Mathematics subjects - particularly mathematics during the first year), engagement with tutoring sessions, and a pivotal indicator denoting whether the student withdrew from studies or remained enrolled.

With a total of 4709 records, the data set encompasses a cohort of engineering students who commenced their academic journey since the year 2003. This compilation spans individuals who either discontinued their studies, successfully completed their university tenure, or are currently in the progression of their educational pathway.

3.1.1 Preprocessing

The objective of preprocessing is to render the raw data amenable for utilization in data mining techniques. Various activities were done to predictive data mining, as outlined in [2], were taken into account:

Data Cleaning - In the initial phase, outliers within each feature were eliminated, with replacements based on averages and densities. Erroneous type values were substituted with values derived from densities.

Students who re-enrolled after 2003 were excluded. This step was taken as some of these students had a study duration of up to 20 years, potentially impacting data set dynamics. We eliminate high school academic average feature because only 38 % of the students have it.

Discretization and Scaling - Continuous features such as debt indicators and the count of tutoring sessions were transformed into range values. For features like financial indicators and tutor sessions, we devised consistent event range groups, with exceptions for cases with 0 events. Grade averages were retained with a couple of decimal places, as these values lie between 0 and 10, making scaling unnecessary.

Handling Missing Features - Null fields were addressed by substituting them with average values in certain instances and considering densities and other feature values for consistency in other cases. Reason-based imputation was employed for missing or blank variables.

Age values were assigned in a manner that preserved the distribution proportions of students across different age groups. Additionally, missing values in high school GPA were replaced using the mean. This imputation considered the student applied major and the originating high school's rank.

Encoding - This technique was employed to convert categorical variables into numerical form. A common approach is one-hot encoding, which generates binary columns for each category in the original variable. The data set includes diverse explanations for student attrition.

We posited that these variables could harbor significant insights for preventing dropout. Types and reasons for dropout were transformed into variables with binary values of 0 and 1, subsequently reclassified into categories such as economic, academic, focus, health, engagement, etc. An additional variable indicating whether studies were concluded was introduced.

Observing that students within the School frequently switch majors within the engineering domain and may extend study duration, we appended binary-coded variables (0 and 1) to capture this behavior. Finally, the engineering programs were encoded to discern potential complexity variations among them.

3.1.2 Data Balancing

As we can read in [20] one of the challenging problems on predicting student attrition is imbalance data set, because the number of students completing their studies far outweighs those who dropout.

To counteract this, we employed techniques aimed at balancing the data set to mitigate its influence on results. One such widely used technique is SMOTE (Synthetic Minority Over-sampling Technique), along with its variant, SMOTE-Tomek.

SMOTE involves oversampling the minority class by generating synthetic instances rather than mere replication. Synthetic examples are introduced along line segments connecting any subset of the k nearest neighbors of the minority class sample [6].

SMOTE-Tomek integrates SMOTE and Tomek links. Tomek links, outlined in [31], serve as either an under-sampling or data cleaning method. When employed as under-sampling, only majority class examples are removed, while as a data cleaning method, instances from both classes can be discarded.

Although SMOTE effectively balances class distribution by oversampling the minority class, certain issues typical in skewed data sets remain unresolved.

In our study, both SMOTE and SMOTE-Tomek were applied as balancing methods, enabling a comparative analysis of outcomes. This approach allows us to explore the suitability of oversimplification versus a combination of oversimplification and under simplification techniques in addressing an imbalanced data set.

3.1.3 Features Selection (FS)

The Recursive Feature Elimination (RFE) method was employed to identify and retain the most significant variables. In [25] it is indicated that adequate selection of features may improve accuracy and efficiency of classifier methods.

The primary objective of this procedure is to pinpoint and eliminate irrelevant or redundant features, thus diminishing the data set's dimensionality and enhancing the efficiency of learning algorithms.

Feature Selection (FS) algorithms encompass two main components: (i) a selection algorithm that generates potential feature subsets to identify an optimal arrangement, and (ii) an evolutionary algorithm that assesses the quality of the suggested feature subset by providing a 'measure of goodness' to the selection algorithm.

Our study uses Recursive Feature Elimination (RFE) as feature selection method to evaluate model performance with different subsets of features and select those that result in the best performance.

3.2 Classification Models

Our experimentation is conducted using four distinct classifiers: Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Networks (ANN).

These classification methods were selected because they are the most used in review articles as we explain later, considering both conventional and deep learning approaches.

3.2.1 Decision Trees (DT)

The DT algorithm is selected as it is well known for predictive modeling of education-based data. In this review [16] on machine learning application of determining the attributes influencing academic performance is indicated that 14 of the 84 publications that were examined, employed the DT method.

The DT algorithms were able to outperform all other algorithms when accuracy is considered. In the realm of education, some researchers have harnessed decision tree algorithms to illustrate the impact of data mining technology, particularly in predicting student dropouts, segmenting students based on performance, managing student retention, and projecting student attrition.

Notably, in a specific predictive study, bagged trees, adaptive boosting trees, and random forests achieved respective accuracy of 88.7%, 95.7%, and 96.1% [19].

3.2.2 Support Vector Machine (SVM)

In [16] it is established that SVM algorithm is used in education for tracking learner involvement and engagement in courses online. In the majority of applications of machine learning, it has been acknowledged as among the most trustworthy and effective algorithms [26].

This algorithm offers noteworthy accuracy and excel particularly with small data sets. Their proficiency extends to predicting at-risk and marginal students [19].

3.2.3 Random Forest (RF)

As indicated in [16] RF is one of the most the supervised ensemble machine learning algorithm most used to predict student at risk. RF operate by constructing a number of decision trees during the training time and producing the output of the class, which is the mode of the classes of the individual trees.

This review [21] sets that RF algorithm have the highest accuracy beating other algorithms in the prediction of students at risk and students' dropout.

3.2.4 Artificial Neural Networks (ANN)

Neural networks prominently feature among widely employed algorithms in the education domain for predicting student performance as we can validated in [23, 21] and [5].

ANN hold particular appeal due to their ability to classify patterns without requiring explicit training. Inherent parallelism bestows ANN with the capability to expedite computational processes, rendering them suitable for predictive tasks in the educational data mining realm [19].

3.3 Model Evaluation

To compare different machine learning methods to predict students at risk of dropout, we use the following performance metrics:

1) Accuracy: It is a measure of how often the model's predictions are correct, compared to the actual outcomes in the data set. In other words, accuracy measures the percentage of correct predictions made by the model. The formula to calculate is:

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \quad (1)$$

2) Precision: It is a measure used particularly in classification tasks. It measures the ability of a model to correctly identify the positives instances (or the instances of a specific class) among all the positives instances (correct or incorrect):

$$Precision = \frac{\text{True Positives}}{(\text{True Positives} + \text{False Positives})} \quad (2)$$

where: True positives are instances correctly predicted as positive. False positives are instances incorrectly predicted as positive.

3) Recall: Also known as true positive rate this measure evaluates the ability of a model to correctly identify positive instances among all the actual positive instances in the data set:

$$Recall = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (3)$$

where False negatives are instances incorrectly predicted as negative by the model, but they are actually positive.

4) F1 score: It is a measure commonly used in binary classification tasks. It is a value that balance the trade-off between precision and recall:

$$F1Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall}. \quad (4)$$

In the eight reviews analyzed previously, the most used metric is Accuracy.

Unfortunately, if we have a imbalanced data sets it will tend to be high, even when a correct prediction is not made. Because of this reason it is integrated additional metrics.

First, the cost of losing a student is very high, so we seek to minimize false negatives which is what Recall measures.

Second, intervention initiatives also require a high and focused effort, so minimizing false positives would help us avoid work, that's why we use Precision.

And finally we use F1-Score because it give us a balance between precision and recall in imbalanced data sets.

3.4 LIME

LIME, or Local Interpretable Model-Agnostic Explanations, stands as one of the most prominent model-agnostic frameworks within the literature, particularly emphasizing its efficacy in enhancing interpretability for tabular models [24].

Functioning as an algorithm capable of faithfully elucidating predictions from any classifier or regressor, LIME achieves this by creating a local approximation using an interpretable model.

While the LIME framework, especially renowned for its prowess in image interpretation, has garnered significant attention, its application to tabular data remains relatively understudied. Moreover, existing research predominantly employs LIME as a benchmark rather than critically assessing LIME's inherent usability.

To bridge this gap, our paper employs LIME on tabular machine learning models and comprehensively evaluates its performance across comparability, interpretability, and usability dimensions [9].

Initially introduced by Ribeiro et al. in 2016, LIME operates as an open-source framework designed

Table 1. Balance training data

Method	Withdrawn	Persisted
None	1375	2154
SMOTE	2154	2154
SMOTE-Tomek	2134	2134

to unveil the decision-making mechanisms of machine learning models and cultivate trust in their application.

The term "local" implies that the framework scrutinizes specific observations, offering insight into how a particular instance is classified rather than providing a holistic understanding of a model's overall behavior. "Interpretable" underscores the framework's aim to render a model's operations intelligible to users.

The term "Model-Agnostic" reflects LIME's adaptability to any present or future blackbox algorithm, disregarding whether the model is transparent or not.

LIME treats all models as black boxes, irrespective of their inherent transparency. The output generated by the LIME framework is denoted as "explanations" [24].

4 Experimentation

4.1 Data

The experimentation was done in a Python notebook in Google Colab. Colab provides a service with an Intel Xeon at 2.20 GHZ, 13 GB of Ram, Tesla K80 accelerator and 12 GB of VRam GDDR5.

This tool allows us to read our data set and apply the python libraries for machine learning. We use numpy, pandas, matplotlib, seaborn, imblearn, tensorflow, sklearn, and lime.

As we explained above, we apply the preprocessing techniques to the data set to later divide it into training and testing and start the experimentation process. 75% of the data set was assigned for training and from this set the SMOTE and SMOTE-Tomek technique was applied to balance the data set. Table1 shows how the proportion of the values remained.

4.2 Model Development

In our research article, we conducted a variable selection exercise for each of the four methods. Specifically, we opted to choose 10 variables out of the total of 42.

Regardless of the model selected, the consistently chosen variables were as follows: number of semesters with a scholarship, total average, total average of mathematics subjects, average of the last completed cycle, average of mathematics subjects in the last cycle, average of the first semester, average load of subjects, average of failed subjects, percentage progress, and debts.

The process of selecting the most crucial features using the Recursive Feature Elimination (RFE) method is influenced not only by the method itself but also by the training data. In our study, we employed three distinct training data sets: the original data set, the one on which SMOTE was applied, and the data set on which SMOTE was applied followed by Tomek.

As outlined previously, the determination of the optimal machine learning method for predicting students at risk of dropout involved the evaluation of four primary indicators: Accuracy, Precision, F1 score, and Recall.

Each of the methods underwent validation with the three distinct data sets. Within each validation process, we conducted hyper parameter optimization by testing a range of different values to observe their impact on performance improvement.

After fine-tuning the hyper parameters, the optimal performance values were as follows: In the Random Forest model, the values were n -estimators=70 and criterion=entropy. For the Decision Tree model, max-depth was set to 10, criterion to entropy, and class-weight to balanced. In the SVM model, the best-performing kernel was RBF, with C =10 and gamma=scale. In the case of ANN, the values were hidden layer sizes=200, activation=relu, and initial learning rate=0.005.

Our iterative process concluded once precision indicators ceased to exhibit further modifications. Upon completion of each method's execution, we proceeded to validate the influence of the variables across various instances within the test data set.

Table 2. Model Performance

Dataset	Accuracy	Precision	F1-Score	Recall
Random Forest				
Original	97.8%	95.1%	97.1%	97.3%
SMOTE	97.1%	95.1%	97%	97.3%
S-Tomec	96.8%	94.9%	96.7%	96.9%
Decision Tree				
Original	97.1%	95.1%	97%	97.3%
SMOTE	96.8%	94.9%	96.7%	96.9%
S-Tomec	87.4%	87.4%	77.25%	87.2%
SVM				
Original	96.8%	95.8%	96.7%	96.7%
SMOTE	96.84%	95.3%	96.7%	96.8%
S-Tomec	96.9%	95.5%	96.8%	96.9%
ANN				
Original	96.5%	96.5%	96.5%	96.5%
SMOTE	95.15%	95.2%	95.2%	95.2%
S-Tomec	95.6%	95.6%	95.6%	95.6%

5 Results

In Table 2, the outcomes of applying the model to this data set are presented. The indicators exhibit notably high values, a result anticipated due to the data set's exclusive inclusion of students who have completed the process, thereby maintaining consistency across variables.

Remarkably, the random forest method yielded the most favorable outcomes, closely followed by the decision trees which demonstrated commendable performance. Unexpectedly, the neural networks displayed comparatively lower performance, a deviation from their typically superior performance that could be related to imbalanced data or features selection.

The performance of Random Forest could be related to the variety of machine learning problems that ensemble methods have been successfully used like feature selection, missing features, imbalanced data, error correction, etc. as indicated in [34].

LIME provides an explanation for each instance by illustrating how individual feature values contribute to the prediction outcome. As depicted in Figure 2 we observe an instance where there is a 45% probability of the student being at risk of dropping out and a 55% likelihood of their continuing their studies.

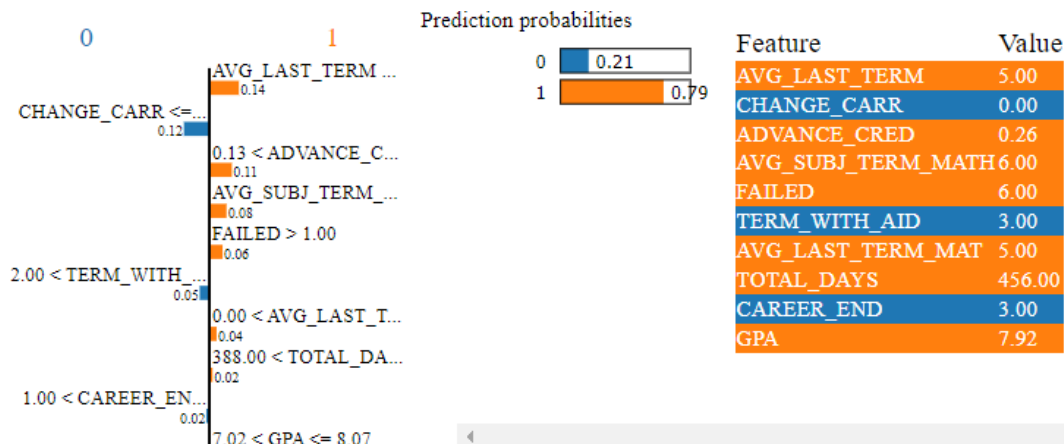


Fig. 2. LIME Graphic of a local instance

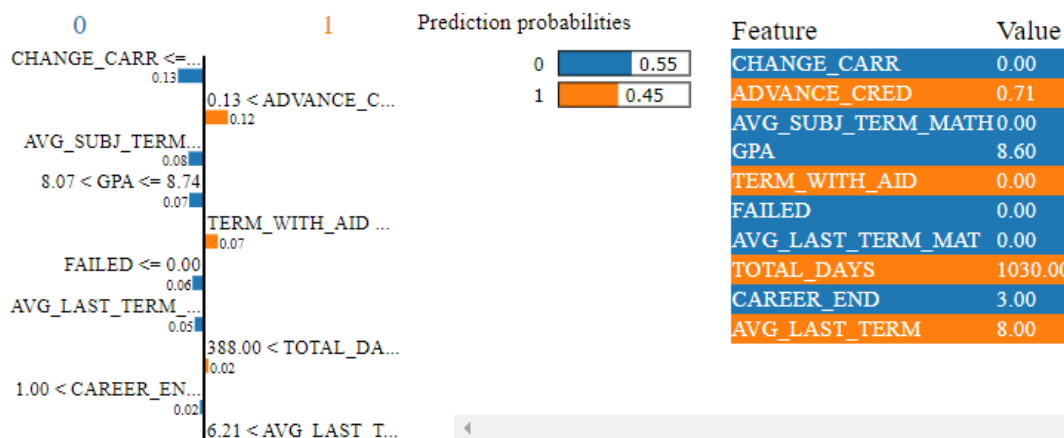


Fig. 3. LIME Graphic of a local instance

This distribution is elucidated by various indicators: the student maintains an GPA of 8.6, student has never changed his career, student has not failed any subject. The features that affect student success are the numbers with financial aid are 0, the last term average is 8 and he has studied 71% of the subjects.

These indicators hold the potential to facilitate a targeted evaluation of students in similar circumstances, enabling us to provide them with precise support and intervention strategies.

Another example is depicted in Figure 3, where a student has an 21% probability of completing their studies. The features that have a positive influence are that he has never changed his career, that he

has had 3 cycles with a scholarship and that the career he is studying is 3.

Features that influence negatively are the average obtained in the last cycle is 5, he only has an advance of 26% of credits, he has failed 6 subjects, he has spent 456 days in the degree and a GPA of 7.92. The influence of variables on a potentially dropping-out student is demonstrated.

Although these analyses are localized, it is possible to extrapolate student behaviors and identify those at higher risk. Moreover, explaining the conditions to tutors and supervisors for ensuring student success becomes straightforward and comprehensible.

6 Conclusions

As we discussed, it is critical for the University, particularly the School of Engineering, which has a higher dropout rate, to identify students at risk of leaving. As we observed in the model evaluation tables, the Random Forest model with the original data set performed the best.

Despite the data set having a 29% dropout rate versus 71% non-dropout rate, the models did not show any improvement when balanced. As anticipated, during the variable selection process, at least two of the indicators related to performance in the mathematics area appeared in all models.

Undoubtedly, the explanation provided by the LIME models is one of the most crucial aspects, as it enables effective communication of the model's behavior and, consequently, its outcomes. The level of confidence achieved through this explanation allows for proactive measures to be taken and focuses attention on students at lower risk.

One of the great challenges in the realization of this article was the generation of the data set since the quality of the information was refined step by step. It was not fully comprehensive, which means that features such as high school GPA could further enhance the performance of the models.

This suggests that refining the data set with additional relevant characteristics could yield even better results. We consider that it is important for future work to better shape this data set, to integrate other variables that allow predicting students at risk well in advance. Integrate the most recent information and also the most academic, such as what can be obtained from the LMS. It is also interesting to apply another description method that is global in its explanation, such as SHAP (Shapley Additive explanation).

References

1. **Albornoz, M., Osorio, L. (2018).** Rankings de universidades: Calidad global y contextos locales. *Revista Iberoamericana de Ciencia, Tecnología y Sociedad*, Vol. 13, No. 37, pp. 13–51.

2. **Alexandropoulos, S. A., Kotsiantis, S. B., Vrahatis, M. N. (2019).** Data preprocessing in predictive data mining. *The Knowledge Engineering Review*, Vol. 34, No. e1. DOI: 10.1017/S026988891800036X.
3. **Alwarthan, S. A., Aslam, N., Khan, I. U. (2022).** Predicting student academic performance at higher education using data mining: A systematic review. *Applied Computational Intelligence & Soft Computing*, Vol. 2022. DOI: 10.1155/2022/8924028.
4. **Asociación Nacional De Universidades e Instituciones De Educación Superior (2018).** *Visión y acción 2030: Una propuesta de la ANUIES para renovar la educación superior en México.* Consejo Reginal Sur-Sureste.
5. **Cardona, T., Cudney, E. A., Hoerl, R., Snyder, J. (2020).** Data mining and machine learning retention models in higher education. *Journal of College Student Retention: Research, Theory & Practice*, Vol. 25, No. 1. DOI: 10.1177/1521025120964920.
6. **Chawla, N. V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W. P. (2002).** SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, Vol. 16, pp. 321–357.
7. **Chen, Y., Zhai, L. (2023).** A comparative study on student performance prediction using machine learning. *Education and Information Technologies*, Vol. 28, pp. 12039–12057. DOI: 10.1007/s10639-023-11672-1.
8. **Chu, H. C., Hwang, G. H., Tu, Y. F., Yang, K. H. (2022).** Roles and research trends of artificial intelligence in higher education: A systematic review of the top 50 most-cited articles. *Australasian Journal of Educational Technology*, Vol. 38, No. 3, pp. 22–42. DOI: 10.14742/ajet.7526.
9. **Dieber, J., Kirrane, S. (2020).** Why model why? Assessing the strengths and limitations of LIME.
10. **Fahd, K., Venkatraman, S., Miah, S. J., Ahmed, K. (2022).** Application of

- machine learning in higher education to assess student academic performance, at-risk, and attrition: A meta-analysis of literature. *Education and Information Technologies*, Vol. 27, pp. 3743–3775. DOI: 10.1007/s10639-021-10741-7.
11. **Fayyad, U., Piatetsky-Shapiro, G., Smyth, P. (1996).** The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, Vol. 39, No. 11, pp. 27–34. DOI: 10.1145/240455.240464.
 12. **Ferreira, F. H., Messina, J., Rigolini, J., López-Calva, L. F., Lugo, M. A., Vakis, R. (2012).** Economic mobility and the rise of the Latin American middle class. DOI: 10.1596/978-0-8213-9634-6.
 13. **Ferreira, M. M. (2017).** At a crossroads : Higher education in Latin America and the Caribbean. *Education for Global Development*.
 14. **Gansemmer-Topf, A. M., Schuh, J. H. (2006).** Institutional selectivity and institutional expenditures: Examining organizational factors that contribute to retention and graduation. *Research in Higher Education*, Vol. 47, pp. 613–642. DOI: 10.1007/s11162-006-9009-4.
 15. **Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., Yang, G. Z. (2019).** XAI—Explainable artificial intelligence. *Science Robotics*, Vol. 4, No. 37. DOI: 10.1126/scirobotics.aay7120.
 16. **Issah, I., Appiah, O., Appiahene, P., Inusah, F. (2023).** A systematic review of the literature on machine learning application of determining the attributes influencing academic performance. *Decision Analytics Journal*, Vol. 7. DOI: 10.1016/j.dajour.2023.100204.
 17. **Kehm, B. M., Larsen, M. R., Sommersel, H. B. (2019).** Student dropout from universities in Europe: A review of empirical literature. *Hungarian Educational Research Journal*, Vol. 9, No. 2, pp. 147–164. DOI: 10.1556/063.9.2019.1.18.
 18. **Lassibille, G., Navarro-Gómez, L. (2008).** Why do higher education students drop out? Evidence from Spain. *Education Economics*, Vol. 16, No. 1, pp. 89–105. DOI: 10.1080/09645290701523267.
 19. **Lynn, N. D., Emanuel, A. W. (2021).** Using data mining techniques to predict students' performance. a review. *IOP Conference Series: Materials Science and Engineering*, Vol. 1096. DOI: 10.1088/1757-899X/1096/1/012083.
 20. **Mduma, N. (2023).** Data balancing techniques for predicting student dropout using machine learning. *Data*, Vol. 8, No. 3, pp. 49. DOI: 10.3390/data8030049.
 21. **Nawang, H., Makhtar, M., Hamzah, W. M. (2021).** A systematic literature review on student performance predictions. *International Journal of Advanced Technology and Engineering Exploration*, Vol. 8, No. 84, pp. 1441–1453. DOI: 10.19101/IJATEE.2021.874521.
 22. **OECD (2019).** Reviews of national policies for education, the future of mexican higher education: Promoting quality and equity. DOI: 10.1787/9789264309371-en.
 23. **Oqaidi, K., Aouhassi, S., Mansouri, K. (2022).** Towards a students' dropout prediction model in higher education institutions using machine learning algorithms. *International Journal of Emerging Technologies in Learning*, Vol. 17, No. 18, pp. 103–117. DOI: 10.3991/ijet.v17i18.25567.
 24. **Ribeiro, M. T., Singh, S., Guestrin, C. (2016).** “Why should i trust you?” Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1135–1144. DOI: 10.1145/2939672.2939778.
 25. **Sánchez-Marño, N., Alonso-Betanzos, A., Tombilla-Sanromán, M. (2007).** Filter methods for feature selection—a comparative study. *International Conference on Intelligent Data Engineering and Automated Learning*, Vol. 4881, pp. 178–187. DOI: 10.1007/978-3-540-77226-2_19.

26. **Shahiri, A. M., Husain, W., Rashid, N. A. (2015).** A review on predicting student's performance using data mining techniques. *Procedia Computer Science*, Vol. 72, pp. 414–422. DOI: 10.1016/j.procs.2015.12.157.
27. **Shcheglova, I., Gorbunova, E., Chirikov, I. (2020).** The role of the first-year experience in student attrition. *Quality in Higher Education*, Vol. 26, No. 3, pp. 307–322. DOI: 10.1080/13538322.2020.1815285.
28. **Sosu, E. M., Pheunpha, P. (2019).** Trajectory of university dropout: Investigating the cumulative effect of academic vulnerability and proximity to family support. *Frontiers in Education*, Vol. 4. DOI: 10.3389/feduc.2019.00006.
29. **Tinto, V. (1975).** Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, Vol. 45, No. 1, pp. 89–125. DOI: 10.2307/1170024.
30. **Tinto, V. (1982).** Limits of theory and practice in student attrition. *The Journal of Higher Education*, Vol. 53, No. 6, pp. 687–700. DOI: 10.2307/1981525.
31. **Tomek, I. (1976).** Two modifications of CNN. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-6, No. 11, pp. 769–772. DOI: 10.1109/TSMC.1976.4309452.
32. **United Nations (2015).** Transforming our world: The 2030 agenda for sustainable development. Resolution adopted by the General Assembly on 25 September 2015.
33. **Veerassamy, A. K., D'Souza, D., Apiola, M. V., Laakso, M. J., Salakoski, T. (2020).** Using early assessment performance as early warning signs to identify at-risk students in programming courses. *2020 IEEE Frontiers in Education Conference (FIE)*, pp. 1–9. DOI: 10.1109/FIE44824.2020.9274277.
34. **Zhang, C., Ma, Y. (2012).** Ensemble machine learning: Methods and applications. Springer. DOI: 10.1007/978-1-4419-9326-7.

*Article received on 06/06/2023; accepted on 16/09/2023.
Corresponding author is Lourdes Martínez-Villaseñor.*