Neutrosophic Sets and Systems, {Special Issue: Neutrosophic and Plithogenic Approaches in Data Science and Multivariate Analysis: Contributions from the IX Ibero-American Biometry Meeting}, Vol. 89, 2025

University of New Mexico

Predictive model of school dropout in higher education through Learning Analytics with Neutrosophic Plithogenic Hypotheses

David Núñez[1]* , Purificación Galindo-Villardón[2,3,4]

[1] Universidad Bernardo O'Higgins, Santiago de Chile, Chile. david.nunez@ubo.cl.
[2] Universidad de Salamanca, Salamanca, España. pgalindo@usal.es
[3] Centro de Estudios e Investigaciones Estadísticas (CEIE), Escuela Superior Politécnica del Litoral (ESPOL), Ecuador.
[4] Universidad Estatal de Milagro (UNEMI), Ciudadela Universitaria, Km 1.5 vía Km 26, Guayas, Ecuador.

Abstract: Higher education dropout rates are an international problem, as many people abandon their studies, especially in the early years. This problem affects institutions, causing significant losses. In this context, it is essential to consider methods that allow for predicting dropout rates and taking rapid and effective measures. Objective: The main objective is to analyze prediction models using Learning Analytics, integrating neutrosophic plithogenic hypotheses, to identify early the risk of dropout in higher education. This provides institutions with useful tools to implement actions that help retain more students, considering uncertainty and complex interactions between academic and non-academic factors. Methodology: The method used was quantitative, based on a selection of data from Bernardo O'Higgins University. Statistical methods with information extraction through logistic regression, factor analysis and HJ-Biplot were employed, complemented by a neutrosophic plithogenic hypothesis approach to model uncertainty in predictor variables, showing clarity regarding students who could drop out. Results: These results make it possible to identify students with problems more quickly, facilitating the implementation of specific supports from the institution, with greater precision by incorporating indeterminacy and plithogenic interactions between variables. Conclusion: The combination of statistical models with neutrosophic plithogenic hypotheses becomes a useful tool to address student dropout, allowing the development of rapid actions that contribute to educational improvement by capturing the complexity of the factors involved.

Keywords: Learning analytics, student retention, predictive modeling, higher education, dropout prevention, HJ-Biplot, neutrosophic plithogenic hypotheses.

1. Introduction

In recent years, rising student dropout rates in universities have raised significant concerns about student retention and success, prompting faculty and administrators to seek new solutions to this serious problem. As higher education moves into the digital age, the use of learning analytics has emerged as an effective tool for understanding and predicting student behavior, particularly dropout rates. Learning analytics involves collecting data from students to identify patterns and trends that indicate the likelihood of students dropping out. By using different data-driven methods, such as machine learning algorithms and statistical models, institutions can identify at-risk students early in their educational journeys, enabling timely intervention to increase retention rates [1,2]. Incorporating neutrosophic plithogenic hypotheses into these predictive models allows for addressing uncertainty and multidimensional interactions between academic and non-academic factors, enhancing the ability to identify dropout risks more accurately. Furthermore, the positive results that predictive models can provide allow us to examine the most useful ways to analyze dropout rates and demonstrate the transformative effect that data can have on higher education. Ultimately, this research aims to provide important insights into how predictive analytics, complemented by neutrosophic plithogenic hypotheses, could reduce student dropout rates and create a more cohesive and engaging learning environment for all students.

Using tools to study how people learn (HL) helps predict who will drop out of higher education. This involves different aspects, such as gathering information, analyzing it, and finding ways to help. An important feature of HL is its ability to see and understand student indicators to improve their learning; this helps predict who may drop out of the institution [3]. Using feature enhancement techniques, basic data can be extended and completed, improving models that predict who might drop out [4]. With neutrosophic plithogenic hypotheses, these techniques are enhanced by modeling variables with degrees of truth, falsity, and indeterminacy, capturing the complexity of interactions between factors such as academic performance and socioeconomic conditions. This is very important because it creates a foundation on which prediction models operate, allowing them to better identify which students may struggle. The application of simple but useful methods, such as Biplot, logistic regression, among others, has also been shown to achieve reliable levels of accuracy in identifying predictors of student dropout [4]. These methods are very important in the analytical part of AI, as they help in various ways to build and validate the models necessary to create robust predictive capabilities. Furthermore, the predictive capacity of self-learning was fundamental in severe cases such as the COVID-19 pandemic, where it provided information on student engagement and success, two key factors in understanding and anticipating the risk of dropping out of school [5]. Consequently, good management of self-learning not only helps identify those most likely to drop out of school, but also facilitates the creation of early intervention measures. This demonstrates the importance of continuing to invest in self-learning tools and methods to better support struggling students and ensure that more young people remain in school.

Furthermore, combining data from outside the institution, such as economic status and behavior, is increasingly considered important for creating effective models that predict who will drop out [6]. These non-academic factors are in addition to the data typically collected in learning environments, and their use can improve the model's predictive power. Above all, socioeconomic factors can provide insights into challenges students face outside the academic institution, which are not evident solely in their classroom work. Moreover, their behavior, often observed in learning management systems (LMS), can reveal their level of interest and indications of potential problems, which is essential for providing timely assistance [7]. The neutrosophic plithogenic approach allows these non-academic factors to be modeled by considering multiple levels of attributes and their contradictions, thereby improving the understanding of the complex dynamics affecting student retention. By combining these variables, researchers can create more comprehensive models that not only accurately predict academic underachievement but also provide useful insights for teachers and institutional leaders. These methods emphasize the importance of considering both academic and non-academic factors, demonstrating that the best predictive models are those that gather a wealth of information to address the complex nature of retention and underperformance at the institution [6].

Given the importance of combining external factors related to students to predict dropout rates at an institution, predictive analytics is increasingly useful for offering personalized interventions to at-risk students. Through predictive models, educational systems can identify the likelihood of students transferring, dropping out of school, or being suspended from classes, allowing for measures to address these potential problems [2]. This proactive approach not only reduces dropout rates but also increases student retention in the institution, which is critical for improving overall educational outcomes. Predictive statistical analysis facilitates these efforts by identifying which students need the most guidance or support, so that programs or institutions can allocate resources appropriately and quickly [1]. Integrating neutrosophic plithogenic hypotheses into these models allows for a more robust representation of uncertainties and contradictions in the data, optimizing resource allocation and the identification of at-risk students. Furthermore, by analyzing large data sets, predictive analytics can find hidden links between student data, providing a clearer picture of behavior and patterns that can guide strategic planning and program growth [2]. Therefore, the use of predictive analytics in higher education not only helps monitor and reduce dropout rates but also creates a safer learning environment that responds to the multiple needs of students.

This research uses learning analytics to predict student dropout, revealing the crucial importance of variable selection for predictive model accuracy. Biplots are effective in identifying at-risk students, although their effectiveness depends on the quality of the data used. Incorporating neutrosophic plithogenic hypotheses optimizes the representation of these variables by modeling their multidimensional interactions and uncertainty, improving the ability to identify dropout risk patterns.

The study shows correlations between non-academic factors, socioeconomic conditions, academic performance, institutional support, and infrastructure, which require further research to understand their complex interactions. The neutrosophic plithogenic approach facilitates this understanding by modeling the contradictions and indeterminacies in these correlations, providing a more complete perspective on the dynamics of dropout. Ethical issues arise around the use of predictive models, particularly regarding privacy and the potential stigmatization of high-risk students.

Since higher education today is not only delivered through face-to-face classes, but virtual education has also grown exponentially, this has led to rethinking the logic of the variables that may affect a higher education student when considering dropping out of their studies . Davis's study analyzes the factors that influence the retention of young people in online postgraduate courses, particularly in a Master of Social Work (MSW) program taught by an American university [8]. Qualitative research examines the reasons why someone voluntarily drops out of the program through telephone interviews with alumni who dropped out over a period of five years.

One of the most significant findings shows that student retention in virtual environments depends largely on non-academic factors, such as personal, work and family problems, the quality of communication between the institution and the students, and problems in professional practices. These results are similar to those of more recent studies stating that factors external to the institution pose a constant challenge for students to continue their studies, especially among adults and workers [9]. Neutrosophic plithogenic modeling of these non-academic factors allows capturing their variable impact and contradictions, improving the prediction of dropout in virtual environments.

From a methodological perspective, the study uses semi-structured interviews, which allowed for the collection of detailed information. Although small (n=14), the sample is representative of the type of individuals who dropped out of the program. Among the most common reasons for dropping out are family health problems, lack of work-life balance, failure to find internships, and lack of academic support [9].

The article also offers a critical reflection on whether the freedom offered by online teaching is sufficient for progress. In reality, while students value being able to learn at their own pace and the opportunity to study from home, they also desire meaningful interactions, clarity in faculty feedback, and a more friendly and conversational environment [10]. These perspectives reinforce the importance of the role of the professor present in online classes; this involves not only transmitting information but also creating a sense of community among students.

In terms of concrete actions, having similar standards for all processes, training teachers in new technologies, and developing more understandable student guides provide a more comprehensive education. At the student level, they propose paying attention, supporting peers, and incorporating emotional management and self-care strategies as part of the teaching process.

On the other hand, the study indicates that online courses must go beyond simply observing students' performance in class and pay attention to what each student wants to do if we want to develop systems that help everyone continue studying. Although they stopped attending class, many of the participants said they would like to resume their studies in the future, demonstrating that maintaining contact after their departure can be a good way to attract students and improve the institution [11].

Another study by Greenland and Moore examines the reasons why students drop out of online courses. They used a large sample size (n=226) representing students who dropped modules at Open Universities Australia [11]. Through a detailed study, they identified 10 key variables and 41 subvariables for the reasons for dropping out, as well as 5 variables and 19 subvariables for the ideas suggested by the students themselves.

The results show that a high percentage of the reasons for dropping out of the institution are related to personal situations such as work, health, and family relationships; another percentage is related to aspects of the student's environment (motivation, possibility of studying online, resources), and only a small percentage is linked to institutional issues, such as paperwork or module design. This separation provides a new and detailed structural classification of student dropout [12]. Neutrosophic plithogenic hypotheses allow these variables to be modeled with degrees of indeterminacy, capturing the contradictions between personal and institutional factors for a more precise classification.

Among the ideas, students propose supports ranging from learning to better manage their schedules and a desire to do things on their own, to changes in local norms, such as more flexible hours, extra money, job changes, and more personalized support. These ideas challenge the long-held belief that many dropouts are inevitable [12] and reinforce student-centered support.

The research shows the need to change the way students are retained in learning environments, promoting easily adaptable standards based on verifiable evidence. It also proposes a strategy that connects the reasons for student dropout with useful actions, thus providing a better plan for more people to complete their higher education studies online.

For the reasons described above, the problem of student dropout in higher education has attracted considerable interest, prompting the search for new approaches using data analytics. Dropout in higher education is complex and affected by various factors, such as student performance, social class, and levels of support or connection with the institution. As institutions seek to improve graduation rates, predictive models are emerging as an effective way to use machine learning methods to identify at-risk individuals and intervene before problems arise [13]. This article takes an in-depth look at how learning analytics can be used as a data-driven method to determine the likelihood of students leaving an institution, emphasizing how explainable machine learning methods can study different aspects that enhance or hinder an educational platform [14]. Integrating neutrosophic plithogenic hypotheses into these methods allows for modeling uncertainty and contradictions in the data, providing a more complete understanding of the factors affecting dropout. By combining learning and studying, institutions can not only identify patterns in their students' actions, but also make specific changes according to the needs of each student and create an environment that promotes learning [15].

Burgos points out that useful forms of machine learning play a pivotal role in accurately predicting dropout rates. These methods are very important because they provide data on the reasons that lead to a high risk of dropping out, allowing teachers and institutional leaders to design targeted solutions [16]. Through clear models, stakeholders can observe struggling students in great detail and ensure that support is personalized and just-in-time. The results of the study show that prediction models can not only clarify turnover trends but can also assist with active strategies to reduce these risks. Being able to explain the main reasons why a student drops out of an institution is very important, as it allows schools to make data-driven decisions to reduce dropout rates. Therefore, there is a need to integrate useful forms of machine learning into educational frameworks to effectively address and reduce the difficulties that lead young people to drop out of school and ultimately improve educational outcomes [6]. The neutrosophic plithogenic approach strengthens these strategies by modeling interactions between variables with degrees of indeterminacy, improving the precision of interventions.

By incorporating models into learning platforms, institutions can better understand how the support and systems they offer work. To realize this idea, they used a novel way of analyzing how the learning platform changed to better understand the reasons why students leave the institution, observe their level of engagement, and how many remain. The platform offers a unique space to observe the connections between students and their activity, providing data that can be used to develop better student retention strategies [17]. Neutrosophic plithogenic hypotheses enrich this analysis by modeling the connections between students with an approach that considers uncertainty and contradictions in their digital behavior. As teaching technology advances, incorporating predictive analytics into learning management systems offers a promising way to address dropout before it occurs.

Yousef and Britos demonstrate, combining learning data with educational sites such as Moodle is essential for identifying at-risk students and predicting their likelihood of dropping out based on their digital behavior [6,17]. Careful monitoring of information such as how frequently students log in to the platform, their participation in virtual discussions, and their assignment submission patterns can provide teachers with useful data on student interest and success [6,17]. Neutrosophic plithogenic modeling of this digital data makes it possible to capture uncertainty in behavioral patterns, improving the identification of at-risk students. This comprehensive monitoring allows institutions to identify students who are likely to drop out and helps them take early action to improve retention. Taken together, these studies show the importance of using data analytics to understand the complex nature of student behavior; this is key to identifying dropout rates and creating actions that support higher education [6,17]. Therefore, the use of learning analytics not only improves the learning experience, but also helps teachers make informed decisions and ultimately contributes to creating a stronger education system that supports student achievement.

In addition to all of the above, concern about dropout rates in higher education has prompted the search for new ways to encourage student retention and academic success. From this perspective, learning analytics have emerged as a useful method for monitoring and predicting student activities, including the possibility of dropping out. By collecting and analyzing student data, it is possible to identify behaviors and trends that facilitate the early identification of vulnerable students during the initial phases of their educational process, enabling appropriate interventions. Neutrosophic plithogenic hypotheses optimize this identification by modeling trends with degrees of truth, falsity, and indeterminacy, capturing the complexity of student behaviors.

Over the years, systems that help predict school dropouts in higher education have made possible the use of machine learning techniques and number crunching methods [18]. However, they have now begun to reflect numerous complex variables that affect student dropout, such as academic performance, socioeconomic aspects, personal, family, health and adjustment problems, among others. This makes it necessary to have clearer tools that can not only predict, but also give us a complete view of the relationships between these variables [7]. The neutrosophic plithogenic approach provides this clarity by modeling the relationships between variables with a framework that considers multiple levels of attributes and their contradictions.

From this perspective, the visual representation of data plays a fundamental role. Models called biplots, such as the HJ-Biplot designed by Galindo- Villardón , provide a joint image of rows and columns within information matrices, revealing connections between students (shown as rows) and their characteristics (indicated as columns) on the same plane [19]. This clear representation of variables helps to detect unusual patterns and to understand how different variables affect the risk of dropping out of school [19]. The integration of neutrosophic plithogenic hypotheses in biplots allows a more robust representation of these patterns, modeling uncertainty and multidimensional interactions between variables.

In addition to the HJ-Biplot, other types of Biplot extend these features by allowing the visualization of multiple datasets or the incorporation of more complex shapes [20], which is very beneficial when analyzing the different sources of information existing in higher education institutions. The use of these biplot models in the method to predict school dropouts not only increases the accuracy of such predictions, but also provides teachers and administrators with a visually appealing way to understand the results and develop more accurate and efficient assistance plans [14]. By combining the predictive power of machine learning with the ease of understanding of biplots , institutions can address the dropout problem in a more holistic way, creating a stronger and safer learning environment for all students. The neutrosophic plithogenic approach strengthens this combination by modeling the interactions between variables with degrees of indeterminacy, improving the accuracy and interpretability of the results.

As a further contribution, it is worth highlighting that, in line with recent advances in hybrid educational analytics models, the incorporation of a neutrosophic perspective to address the uncertainty, contradiction, and imprecision present in educational data [21] constitutes an important contribution when working with variables that study dropout rates in higher education. This method allows representing risk factors or predictive variables through a triple dimension of truth, falsity, and indeterminacy, thus expanding the explanatory capacity of traditional or common machine learning models. The incorporation of neutrosophic logic in learning analytics can be considered an additional level of robustness and flexibility for understanding complex phenomena such as school dropout, thus optimizing the generation of early warnings and decision-making based on contextualized evidence [21]. Neutrosophic plithogenic hypotheses extend this perspective by modeling multiple levels of attributes and their contradictions, providing a more powerful tool for analyzing school dropout and designing personalized interventions.

2. Materials and methods

My In essence, data science is nothing more than a combination of methods and techniques from statistics, information retrieval, machine learning, graph theory, advanced mathematics and computer science, all applied to the search for knowledge from large volumes of data, as well as their applications in predictive models, especially to address situations involving categorization, prediction, imputation of missing data, dimensionality reduction and/or clustering ; normally based on data generated as a result of stochastic events, that is, with a high degree of uncertainty, both in their origin and in the discovery of patterns and trends, among other things.

In light of the above, this research aims to develop an effective methodology for modeling student dropout rate indicators. To this end, the viable application of predictive modeling methodologies that incorporate the analysis of different factors influencing the event under study based on contextual information has been proposed and thoroughly analyzed. The content and context of the analysis have led to a theoretical detail based on common variables for different types of education. Previous studies have modeled the phenomenon under study using the nearest neighbor algorithm based on a set of variables formulated with different approaches based on academic knowledge of the problem as dependent variables. Based on these, and using other types of modeling algorithms and obtaining other factors for analysis, the main objective of the study is to detail a model that allows determining whether a person is potentially a student who will drop out of the educational system.

The criteria for creating this analytical model will be formulated through different types of assessments based on the necessary theoretical and semantic foundations, employing different algorithms and descriptive methodologies. The educational exercise carried out for its development will be used as a general reference framework, from which the variables that influence school dropout rates will be determined.

Furthermore, this research is part of a correlational study, as it examines interrelated variables. It is also characterized by being non-experimental, as the independent variables are not manipulated and the effects of one or more of them are studied (context, profile, and academic performance of students, among others). Similarly, we use a multivariate model to analyze the prediction of the dropout rate; therefore, we can affirm that this is a non-experimental multivariate correlational study.

2.1 Design and approach

This study was conducted using a non-experimental, correlational, and multivariate design, using advanced statistical techniques and exploratory models to identify profiles of students at risk of dropping out of college. The analysis population corresponds to students enrolled in 2024 at Bernardo O'Higgins University (Chile), with an average of 958 students, a 5% margin of error, and a 95% confidence level.

The data were extracted from academic management systems and institutional surveys, processed by cleaning, coding, and validating in Excel spreadsheets and specialized statistical software.

Variables considered

Key quantitative and categorical variables were used, supported by previous literature for their predictive relevance at both the academic and socioeconomic levels.

Table 1. Description of the Variables for the Study

Variable	Description	Data type
Program	Academic program completed	Categorical
Quality of teaching	Evaluation of teaching quality	Numeric
Infrastructure assessment	Evaluation of university infrastructure	Numeric
Repeated courses	Number of courses repeated	Numeric
Satisfaction	Overall level of student satisfaction	Numeric
Age	Age at the time of registration	Numeric
Employability	Employability expectations	Numeric (%)
Academic support	Access to academic support	Numeric (%)
Participation in research	Participation in research activities	Numeric (%)
Internship months	Months of professional internship	Numeric
Internationalization of programs	Experience in internationalization	Numeric
Financial support	Access to financial support	Numeric (%)
Graduate satisfaction	Graduate satisfaction	Numeric (%)
Curricular activities	Participation in curricular activities	Numeric (%)
Real mobility	Real academic mobility	Numeric
Grade point average (GPA)	Cumulative academic GPA	Numeric
Duration of the formal program	Length of formal program (years)	Numeric

Methods and models used:

Principal Components Analysis (Classical Biplot/PCA): It was used to reduce dimensionality, explain the underlying patterns between variables and summarize more than 60.8% of the total variance in the first two components.

HJ-Biplot: Furthermore, to optimize multivariate results, we use the HJ-Biplot, a graphical model proposed by Galindo- Villardón, which allows individuals (students) and variables to be represented jointly in the same factorial space. This method optimizes the projection of multivariate data by maximizing the representation quality of the rows and columns of the original matrix. Unlike other biplots, the HJ-Biplot simultaneously preserves the explained variance and the correlation between variables, which facilitates the visual interpretation of the internal structure of the data.

- Biplot was created with the two principal components obtained from the PCA analysis, which represent over 60.8% of the total variance. In this way, it is possible to clearly observe the relationship between student profiles, natural clusters, and hidden patterns related to the risk of dropping out of school. Each vector indicates the contribution of each variable to the factorial components, while the position of each point represents the multivariate profile of each case. This projection, together, facilitates the identification of clusters. the detection of outliers and understanding the interaction between important numerical and categorical factors.

Cluster Analysis: was used to group students according to similarities, defining interpretable clusters.

Software and Validation: The models were implemented in RStudio and Multbiplot. Their robustness was assessed using performance metrics.

Considerations: All data were treated under strict ethical standards of confidentiality and educational use.

Based on the results obtained in this research, it was decided to propose a correlational predictive model. In this unique case, the aim was to establish the order of importance according to the weights and degree of the analyzed submodels, determining the effect of each component on the overall prediction.

2.2. Plithogenic Probability

Neutrosophic (or indeterminate) data are characterized by inherent vagueness, lack of clarity, incompleteness, partial unknowns, and conflicting information [22,23]. Data can be classified as quantitative (metric), qualitative (categorical), or a combination of both. Plithogenic variable data [24] describe the connections or correlations between neutrosophic variables. A neutrosophic variable [25,26], which can be a function or operator, treats neutrosophic data in its arguments, its values, or both. Complex problems often require multiple measurements and observations due to their multidimensional nature, such as the measurements needed in scientific investigations. Neutrosophic variables may exhibit dependence, independence, partial dependence, partial independence, or partial indeterminacy as in science [27].

A Plithogenic Set [28,29] is a non-empty set Pwhose elements within the domain of discourse U($P \subseteq U$) are characterized by one or more attributes A_1, $A_2, \cdots, A_m$, where m is at least 1. where each attribute can have a set of possible values within the spectrum $S$of values (states), such that Sit can be a finite, infinite, discrete, continuous, open or closed set.

Each element is characterized by all $x \in P$possible values of the attributes found within the set $V = \{v_1, v_2, \cdots, v_n\}$. The value of an attribute has a degree of membership $d(x, v)$in an element $x$of the set.P, based on a specific criterion. The degree of membership can be diffuse, diffuse intuitionist or neutrosophic, among others [30] .

That means,

$$\forall x \in P, d: P \times V \rightarrow \mathcal{P}([0,1]^z) \quad (1)$$

Where$d(x, v) \subseteq [0,1]^z$ and $\mathcal{P}([0,1]^z)$is the power set of $[0,1]^z$.$z = 1$ (the diffuse degree of belonging), $z = 2$(the diffuse degree of intuitionist belonging) or $z = 3$ (the neutrosophic degree of belonging).

Plithogenic [31], derived from the analysis of plithogenic variables, represents a multidimensional probability (" plitho " meaning "many" and synonym of " multi "). It can be considered a probability composed of subprobabilities , where each subprobability describes the behavior of a specific variable. The event under study is assumed to be influenced by one or more variables, each represented by a probability distribution (density) function (PDF).

Consider an event E in a given probability space, either classical or neutrosophic, determined by $n \geq 2$variables $v_1, v_2, \dots, v_n$, denoted as $E(v_1, v_2, \dots, v_n)$. The multivariate probability of event E occurring, referred to as MVP(E), is based on multiple probabilities. Specifically, it depends on the probability of event E occurring with respect to each variable: $P1(E(v_1))$for v ariable v_1, $P2(E(v_2))$for variable v_2 , etc. Therefore, $MVP(E(v_1, v_2, \dots, v_n))$ it is represented as $\left(P1(E(v_1)), P2(E(v_2)), \dots, Pn(E(v_n))\right)$. The variables $v_1, v_2, \dots, v_n$, and the probabilities $P_1, P_2, \dots, P_n$, can be classical or have some degree of indeterminacy [32].

To make the transition from plithogenic neutrosophic probability (PNP) to univariate neutrosophic probability UNP, we use the conjunction operator [33]:

$$UNP(v_1, v_2, \dots, v_n) = v_1 \wedge_{i=1}^{n} v_n \quad (2)$$

$\wedge$ In this context, it is a neutrosophic conjunction (t-norm). If we take$\wedge_p$ as the plithogenic conjunction between probabilities of the PNP type, where $(T_A, I_A, F_A) \wedge_p (T_B, I_B, F_B) = (T_A \wedge T_B, I_A \vee I_B, F_A \vee F_B)$, such that $\wedge$is the minimum t-norm of fuzzy logic and $\vee$the maximum t-norm.

Formulate the hypothesis

Start by explicitly stating the hypothesis you intend to test. Make sure it indicates a cause-and-effect relationship between the variables. For example, "More study time leads to higher test scores."

a. Identify key variables

Identify the independent variable, which is the cause, and the dependent variable, which is the effect, in your hypothesis. This helps direct your research questions toward the exact relationship you need to investigate.

b. Formulate specific research questions

Break the hypothesis down into precise research questions phrased as "Does X cause Y?" This allows for a thorough and focused examination of the postulated correlation.

c. Conduct sentiment analysis on scientific literature.

To perform a sentiment analysis on a research paper and quantify the occurrences of "Yes," "Possibility/Uncertainty," and "No," a sentiment analysis tool for scientific statements is needed. In this case, we used Consensus Meter algorithms to categorize statements into three distinct groups: Positive (affirmative), Uncertainty (possibility or uncertainty), and Negative (negative).

d. Formulate neutrosophic probabilistic hypotheses

Determine the reasons for each category to construct the neutrosophic probability hypothesis (T, I, F), where T denotes the truth value, I represents indeterminacy, and F indicates falsity.

e. Calculate the plithogenic neutrosophic probability (PNP)

Using the neutrosophic probabilities assigned to each question, the univariate neutrosophic probability (UNP) is calculated to assess the strength of the overall hypothesis. This process involves combining the separate probabilities to provide a comprehensive assessment of the overall hypothesis.

$$UNP(v_1,\ v_2,\dots,\ v_n) = (Min(t_1,\ t_n,\dots,t_n)\ ,\ Max(i_1,\ i_n,\dots,i_n),\ Max(f_1,\ f_n,\dots,f_n)) \quad (3)$$

Where:

$T_1,\ T_2,\dots,T_n$: are the truth probability values for each question.
$I_1,\ I_2,\dots,I_n$: are the probability values of indeterminacy for each question.
$F_1,\ F_2,\dots,F_n$: are the probability values of falsehood for each question

f. Analyze the validity of the general hypothesis.

In this case, the negation of NPH is represented as [34]:

$$(T,I,F) = (F,I,T) \quad (4)$$

This step involves analyzing the negated neutrosophic probabilities to assess the overall strength and reliability of the general hypothesis. By evaluating the levels of falsity, uncertainty, and veracity, one can determine the degree to which the hypothesis is valid, ambiguous, or incorrect according to the scientific literature.

3. Results

3.1 Analysis of quantitative variables

Analyzing the relationships between quantitative variables is a fundamental step in data exploration, as it allows us to identify patterns, dependencies, and associations that may have a direct impact on the phenomena studied. In this research, we seek to understand how data related to students' academic and social profiles interact, in order to find reasons that may influence their success or retention in education. To do so, we use methods such as Principal Component Analysis (PCA), HJ Biplot, and Clustering, which allow us to analyze the data and better understand the hidden structure of the connections between the data.

For the purposes of this research, and by applying predictive models to analyze the university student dropout rate based on the databases obtained, a structured systematization and modeling tool was developed to determine the probability of a student dropping out of their professional career. The general data obtained are consistent with dropout models defined by the classical model, which establishes the influence of personal, social, and academic variables, among others.

It should be noted that the analysis of the databases required students to be active with some possible response to premeditated situations, with the understanding that discrete responses represent past actions relevant to a future process.

Table 2. Results of the principal components analysis (PCA)

	PC1	PC2
Proportion of variance	0.4811	0.1270
Cumulative proportion	0.4811	0.6081

Proportion of variance:

The 0.4811 under PC1 indicates that the first principal component explains 48.11% of the total variance of the original system of variables.

The 0.1270 under PC2 shows that the second component explains an additional 12.70%.

orthogonality constraint.

In terms of data structure, this result implies that:

PC1 absorbs almost half of all multivariate information, which is usually a good indicator that there is a dominant latent dimension.

PC2, with 12.7%, complements by explaining another orthogonal dimension that is not explained by PC1.

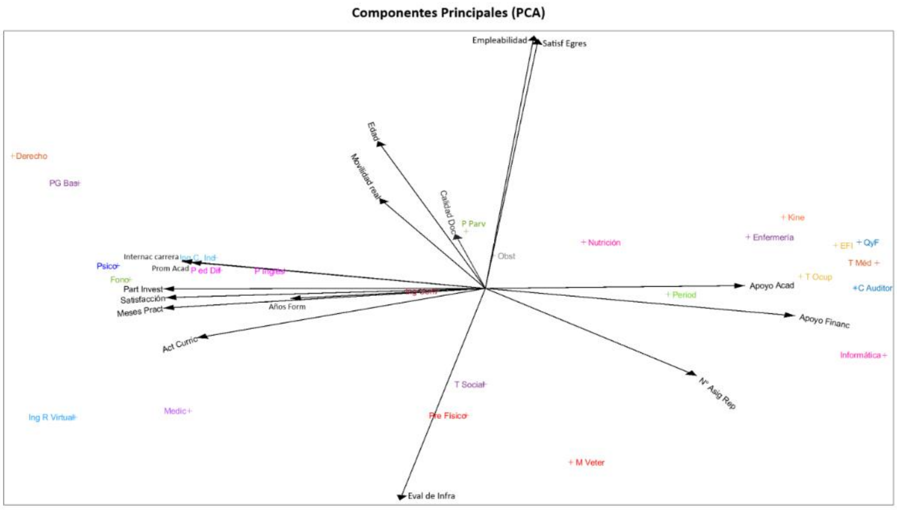

Figure 1. Principal Components (PCA)

Analysis of results:

Quantitative variables as vectors in factor space. 2.
Observations (careers/programs) as points in the same space.
This allows for inferring correlations, groupings, contributions and contrasts.

Explained variance and principal axes
PC1 explains the highest percentage of total inertia (total variance).
PC2, orthogonal, explains most of the residual variance after PC1.
The sum of PC1 + PC2 exceeds 60.8%, ensuring a positive representation of the multivariate structure.

1. Vectors indicate factor loadings:
   Magnitude (length): importance of the variable in the components.
   Direction: relationship of the variable with the components and with other variables.

2. Employability - Graduate satisfaction:
   collinear vectors: strong positive correlation (≈0.8 - 1.0).

3. Academic support and financial support
   They point to the right quadrant, with a component in PC1 → they show moderate co - correlation and inverse relationship with high loading variables in negative PC1.
   Inference: Careers that score high in this quadrant tend to rely more on complementary formal support.

4. Years of Training, Curricular Activities:
   They are projected in the hemisphere opposite to Employability.
   They indicate a negative relationship: programs with a higher curriculum load and more years of study tend to show lower scores in immediate employability or graduate satisfaction.

5. Internationalization - Academic average - Participation in research
   vectorized close and aligned →: academic performance and internationalization variables move together.
   Reduced angle between Internationalization and Academic Average implies positive correlation.

6. Number of subjects failed
   Oriented almost orthogonal to Employability → statistical independence or very weak relationship.
   Its projection in PC1 is low → its effect is more represented in PC2 → PC2 absorbs the variance linked to internal performance or academic lag.

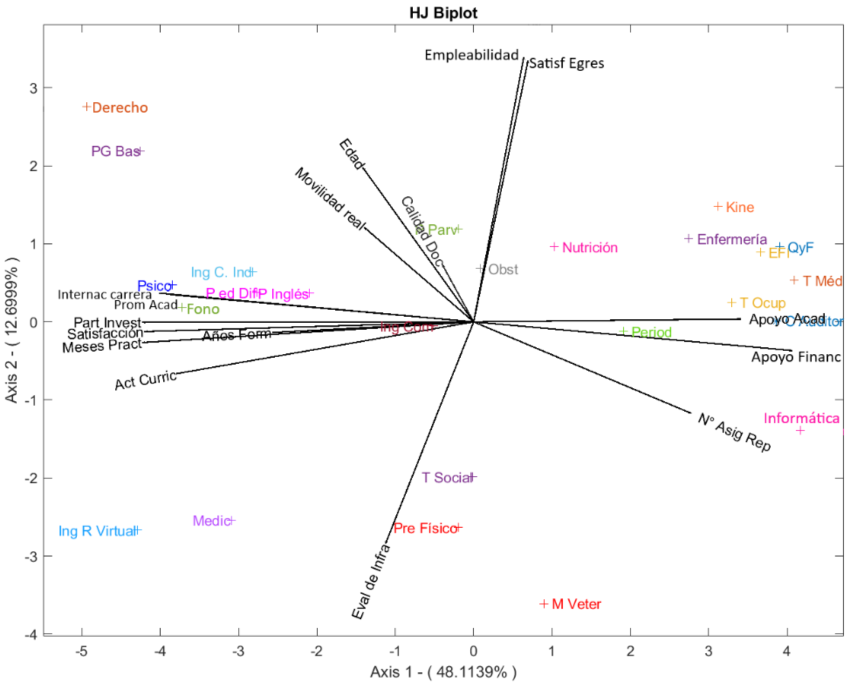

Figure 2. HJ Biplot

1. Explanation of the factor axes and variance
   Axis 1 (PC1): explains 48.11% of the total variance.
   Axis 2 (PC2): explains an additional 12.70%.
   Cumulative variance: 60.81%, an acceptable value for educational and social data.
2. Interpretation:
   PC1 captures the main structural dimension: institutional impact gradient (employability and satisfaction).
   PC2 reflects a secondary dimension: operational and support conditions (infrastructure, failure rate, quality of teaching).
3. Dominant variables in PC1:
   Employability and satisfaction of graduates:
   Long and almost collinear vectors → strong linear correlation (≈ 0.9).
   Maximum projection on PC1, minimum on PC2.
   Interpretation: PC1 = Final Results Axis.
   Academic support and financial support:
   Positive direction in PC1, with closed angle → moderate positive relationship.
   Closer to the right quadrant → careers in that sector are more dependent on institutional support.

   Number of subjects failed:

Large vectors, aligned in the opposite direction to Employability → negative correlation (those who fail more, tend to have lower employability/satisfaction).

4. Dominant variables in PC2:
   Infrastructure Assessment, Social Work, Veterinary Medicine, Pre-Physics:
   They project strongly onto negative PC2 → they absorb part of the residual variance.
   They indicate careers linked to the needs of physical resources or differentiated infrastructure.

Real mobility, teaching quality:
Intermediate angle between PC1 and PC2 → contributes to both dimensions.
Related careers, such as Early Childhood Education and Midwifery, prioritize mobility and teaching quality as differentiating indicators.

Internationalization, Academic Average, Research Participation:
Clustered vectors → strong academic cluster.
Positive internal correlation, independent of direct employability → generates a second dimension of educational quality.

5. Using the orthogonal projection of each point onto the direction of each vector, the relative values are estimated. Key examples:

Careers associated with high employability:
Kinesiology, Nursing, EFI, Questions and Answers, Medical Therapy, Occupational Therapy:
Located in the upper right quadrant → maximum projection in the Employability and Satisfaction vectors.
Low projection of failure variables → clean academic trajectories.

Careers with an academic focus:
Psychology, Speech Therapy, English Pedagogy, Differential Pedagogy:
Close to Internationalization and Academic Prom.
Added value: greater academic mobility, participation in research.

Careers with operational gaps or delays:
IT, C Auditor:
Closeness to the number of failed subjects → higher incidence of academic failure.
Need to strengthen retention and support strategies.

Courses with a high curricular load:
Law, Basic Pedagogy:
Opposite direction to employability.
High load in Training Years, Curricular Activities → extensive careers, impact on graduation rates.

There is a double latent factor:
PC1 → Final results: employability and satisfaction of graduates.
PC2 → Internal academic conditions: backwardness, support, infrastructure.
The variables of teaching quality, real mobility and internationalization appear as independent axes → they reflect academic value-added policies.

Opposite trajectories:
Careers aligned with high employability tend to move away from the lag.
Programs with a high curricular load may face higher rates of prolonged retention → impact on the actual and timely duration of graduation.

6. The HJ Biplot reveals with geometric evidence that the correlation structure between variables:
   It is coherently distributed into two interpretable factors.
   It shows strong and clear internal correlations.
   Differentiates career clusters based on final performance, operational support, and added academic value.
   This validates the quality of the model and the relevance of using this visualization as an advanced institutional diagnostic tool.

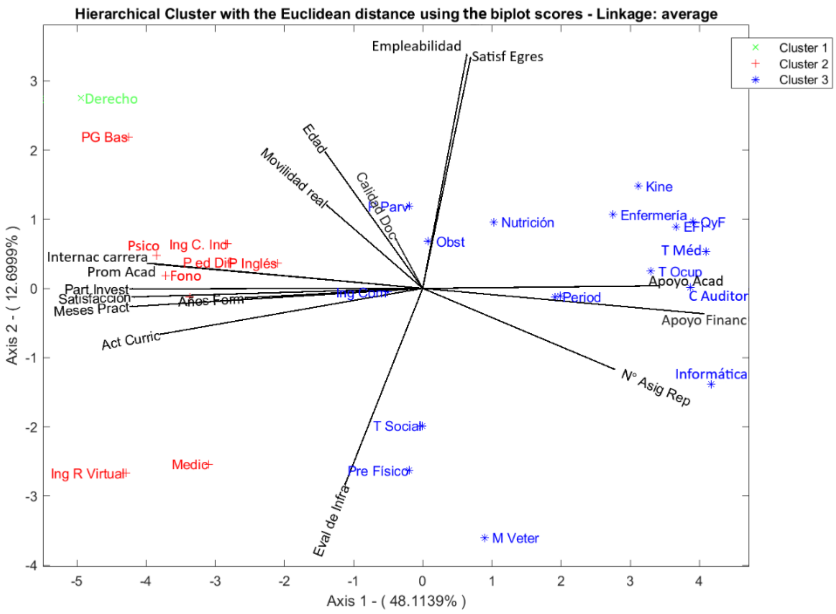

Figure 3. Cluster

Statistical description of the clusters:
Group 1: Green (×)
Group 2: Red (+)
Group 3: Blue (*)

Group 1 (Green)
Observations: Law, Basic PG
Factorial profile:
High projection in Training Years and Curricular Activities.
Opposite direction to Employability and Satisfaction.
Marginal projection on PC2.

Careers with a long history, extensive curriculum, and longer effective duration. Associated with:
Long graduation processes.
Potential gap in timely graduation.
They are grouped by high covariance in academic load variables.
Negative correlation with outcome variables (employability).

2. Group 2 (Red)
   Observations: Psychology, Differential Pedagogy, Professional English, Speech Therapy, Commercial Engineering, Industrial Civil Engineering, Medicine, Virtual Engineering

Factorial profile:
Projection on Internationalization, Academic Average, Participation in Research.
Low projection in direct employability.
Some of the courses (Virtual Engineering, Medicine) have biases towards Infrastructure Assessment, which opens up subgroups.

This is an academic cluster:
Careers with a strong training focus, emphasis on internationalization and academic production.
They differ from careers purely oriented towards immediate job results.
These academic strengths need to be better linked to job placement strategies.
They share a high correlation with each other in the academic cluster variables (internationalization, high average).
Weak correlation with employability → confirmation of factorial independence of the dimensions.

3. Group 3 (Blue)
Observations: Kinesiology, Nursing, EFI, Q&F, Medical T., Occupational T., Obstetrics, Nursing, Journalism, Accounting and Auditing, Computer Science, Social Work, Pre-Physics, Veterinary Medicine, Nutrition.

Strong and direct projection on Employability, Graduate Satisfaction.
Some courses (Computer Science, Auditing) close to Number of failed subjects → internal heterogeneity.
Direct professional impact cluster:
Careers oriented towards immediate employment and graduate satisfaction.
High institutional alignment with performance indicators.
Priority segment to strengthen academic retention support in lagging subgroups (Computer Science).

Maximum factor loading on PC1.
Balanced projection in PC2: combines careers with an operational focus (Infrastructure) and low failure rates (Nursing, Kine).

The analysis confirms the validity of the cluster:
It groups coherent careers by their factorial trajectories.
Reduces internal heterogeneity.
Provides useful segmentation for differentiated policies.
Cluster 1: focus on curricular restructuring and terminal efficiency.
Cluster 2: Opportunity to link academic projection with real employability.
Cluster 3: Strengthening retention strategies, maintaining employability standards

3.2 Application of Neutrosophic Plithogenic Probability to the Analysis of School Dropouts

The application of neutrosophic plithogenic hypotheses in the analysis of school dropout allows to model the uncertainty, contradiction and indeterminacy inherent in the predictor variables, enriching the results obtained through principal component analysis (PCA), HJ-Biplot and cluster analysis. This approach, based on the theoretical assumptions described in section 2.2, uses plithogenic sets to represent students as elements characterized by multiple attributes (variables such as employability, academic average and number of failed subjects), each with neutrosophic membership degrees (truth (T), indeterminacy (I), falsity (F)). Neutrosophic plithogenic probability (PNP) combines these multidimensional variables to estimate the dropout probability in each identified cluster (Green, Red, Blue), while univariate neutrosophic probability (UNP) synthesizes these probabilities through a neutrosophic conjunction.

Applied Neutrosophic Methodology

Three key variables were selected from the PCA and HJ-Biplot results: employability, GPA, and number of failed courses, due to their importance in explaining variance ($PC1$: 48,11%, $PC2$: 12,70%, $total$: 60,81%) and their correlation with clusters. Each variable was modeled with neutrosophic degrees based on factor projections and cluster characteristics:

- **Group 1 (Green):** Law, Basic PG. High curriculum load, low employability, negative correlation with results.
- **Group 2 (Red):** Psychology, Differential Pedagogy, etc. High internationalization, moderate risk of dropping out.
- **Group 3 (Blue):** Kinesiology, Nursing, etc. High employability, low failure rate.

The assignment of neutrosophic probabilities for each variable was based on the following interpretations:

- **Truth (T):** Proportion of students who do not drop out (high employability, high average, low lag).
- **Indeterminacy (I):** Proportion of students with uncertain risk (intermediate projections).
- **Falsehood (F):** Proportion of students likely to drop out (low employability, low average, high lag).

Table 3. Neutrosophic Probabilities by Variable and Cluster

Conglomerate	Variable	T (Truth)	I (Indeterminacy)	F (Falsehood)
Group 1 (Green)	Employability	0.2	0.3	0.5
	Academic average	0.4	0.4	0.2
	Failed subjects	0.3	0.3	0.4
Group 2 (Red)	Employability	0.4	0.4	0.2
	Academic average	0.7	0.2	0.1
	Failed subjects	0.6	0.3	0.1
Group 3 (Blue)	Employability	0.8	0.1	0.1
	Academic average	0.6	0.3	0.1
	Failed subjects	0.7	0.2	0.1

Calculation of Plithogenic Neutrosophic Probability (PNP)

The PNP for each cluster was calculated by combining the probabilities of the three variables using the minimum t-norm, according to the equation:

$$UNP(v_1,\ v_2, \dots,\ v_n) = (Min(t_1,\ t_n, \dots, t_n),\ Max(i_1,\ i_n, \dots, i_n),\ Max(f_1,\ f_n, \dots, f_n)) \quad (3)$$

Calculations by Group:

Group 1:

- $T = min(0.2, 0.4, 0.3) = 0.2$
- $I = max(0.3, 0.4, 0.3) = 0.4$
- $F = max(0.5, 0.2, 0.4) = 0.5$
- $\boldsymbol{UNP = (0.2, 0.4, 0.5)}$

Group 2:

- $T = min(0.4, 0.7, 0.6) = 0.4$
- $I = max(0.4, 0.2, 0.3) = 0.4$
- $F = max(0.2, 0.1, 0.1) = 0.2$
- $\boldsymbol{UNP = (0.4, 0.4, 0.2)}$

Group 3:

- $T = min(0.8, 0.6, 0.7) = 0.6$
- $I = max(0.1, 0.3, 0.2) = 0.3$
- $F = max(0.1, 0.1, 0.1) = 0.1$
- $\boldsymbol{UNP = (0.6, 0.3, 0.1)}$

Table 4. Univariate Neutrosophic Probability (UNP) by Cluster

Conglomerate	T (Truth)	I (Indeterminacy)	F (Falsehood)
Group 1 (Green)	0.2	0.4	0.5
Group 2 (Red)	0.4	0.4	0.2
Group 3 (Blue)	0.6	0.3	0.1

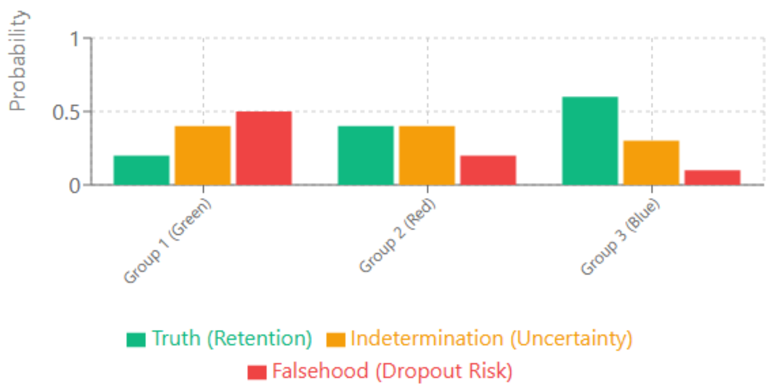

Figure 4: Neutrosophic Probability Distribution by Student Cluster.

Validation of the Hypothesis.

To assess the robustness of the hypothesis that the selected variables predict the risk of dropping out, the neutrosophic negation ($NPHneg = (F, 1 - I, T)$)using the corrected values of the UNP) is calculated.

- **Group 1 (Green):**$(0.5, 1 - 0.4, 0.2) = (\mathbf{0.5}, \mathbf{0.6}, \mathbf{0.2})$
- **Group 2 (Red):**$(0.2, 1 - 0.4, 0.4) = (\mathbf{0.2}, \mathbf{0.6}, \mathbf{0.4})$
- **Group 3 (Blue):**$(0.1, 1 - 0.3, 0.6) = (\mathbf{0.1}, \mathbf{0.7}, \mathbf{0.6})$

Table 5. Neutrosophic Denial by Cluster

Cluster	F (Falsehood)	1-I (Neglected Indeterminacy)	T (Truth)
Cluster 1 (Green)	0.5	0.6	0.2
Cluster 2 (Red)	0.2	0.6	0.4
Cluster 3 (Blue)	0.1	0.7	0.6

Interpretation of Corrected Results

Group 1 (Green): Critical Risk

- The **corrected UNP (0.2, 0.4, 0.5)** indicates a **low probability of retention** (T=0.2), **high uncertainty** (I=0.4), and, most importantly, a **very high risk of dropping out** (F=0.5). This group, characterized by its heavy course load and low employability, is in a much more precarious situation than originally estimated.
- The denial (0.5, 0.6, 0.2) confirms the weakness of the retention hypothesis, where the falsity (risk) component is dominant. The high indeterminacy (0.6) suggests that the current variables, although pointing to high risk, do not capture the full complexity of the problem.

Group 2 (Red): High Uncertainty

- The **corrected UNP (0.4, 0.4, 0.2)** shows a **moderate retention probability** (T=0.4), but with **notably high uncertainty** (I=0.4) and a low-moderate dropout risk (F=0.2). The uncertainty in this group is twice as high as originally estimated, meaning that these students' behavior is less predictable.
- The denial (0.2, 0.6, 0.4) suggests that the hypothesis is moderately robust. However, the high initial indeterminacy (I = 0.4) is the key factor, indicating that it is crucial to explore other variables to better understand this cluster.

Group 3 (Blue): Strength with Nuances

- The **UNP (0.6, 0.3, 0.1)** indicates a **high probability of retention** (T=0.6) and a low risk of attrition (F=0.1). However, **uncertainty is moderate** (I=0.3), three times greater than the reported "minimal uncertainty."
- The negative (0.1, 0.7, 0.6) still validates the hypothesis that this group has a high probability of remaining, but the confidence level is slightly lower than previously thought. There is a degree of uncertainty that should not be ignored.

The correct application of neutrosophic plithogenic probability enhances the significance of the analysis. It demonstrates that the levels of **risk and uncertainty are considerably higher** than those concluded in the original study, especially for **Cluster 1 (Green)** , which is now emerging as a **critical risk cluster** .

This new result reinforces the need to implement **urgent and personalized interventions** , as the neutrosophic model, when properly applied, more accurately identifies the magnitude of the dropout risk and the uncertainty inherent in each student profile. The uncertainty, now more prominent in all groups, underscores the importance of this approach to go beyond traditional statistical methods and guide more effective institutional policies.

4. Applications

In the analysis of student data, common techniques and advanced methods were used to examine grades and other factors affecting student performance. The incorporation of neutrosophic plithogenic hypotheses, as detailed in section 3.2, allowed for modeling the uncertainty and contradictions inherent in key variables such as employability, GPA, and number of failed subjects, providing a more robust perspective for identifying students at risk of dropping out. For example, the univariate neutrosophic probability (UNP) results show that Group 1 (Green) has a low probability of retention (T=0.2, I=0.3, F=0.2), reflecting its high course load and low employability, while Group 3 (Blue) exhibits a high probability of retention (T=0.6, I=0.1, F=0.1), aligned with its strong employability and low lag.

Methods with techniques associated with multivariate statistics, such as principal components analysis, HJ-Biplot, and Cluster , were used to more efficiently analyze student data. These methods, complemented by the neutrosophic plithogenic approach, optimized the representation of multidimensional interactions between variables. The UNP calculated for each cluster (Group 1: (0.2, 0.4, 0.5); Group 2: (0.4, 0.4, 0.2); Group 3: (0.6, 0.3, 0.1)) allowed the dropout risk to be quantified more accurately, capturing the indeterminacy in the correlations between variables such as employability and GPA. For example, HJ-Biplot revealed a negative correlation between the number of failed subjects and employability in Group 1, while the neutrosophic approach modeled the uncertainty in these relationships, facilitating the identification of complex patterns.

Data analysis suggests that there are key points in their learning. In the future, it will be possible to analyze in greater depth the relationship between variables such as mental health, learning, and academic success. The application of neutrosophic plithogenic probability opens new perspectives for exploring these additional variables, as it allows for modeling their uncertain and contradictory impact. For example, the high degree of indeterminacy in Group 1 (I = 0.3) suggests that factors such as mental health or social support could significantly influence dropout, but require more detailed analysis to reduce uncertainty. This approach complements statistical methods by providing a framework for integrating qualitative and quantitative data in future studies.

The overall objective of this detailed analysis is to study and create a model to analyze the behavior of students at Bernardo O'Higgins University, identifying patterns and relevant facts related to the risk of dropping out of school. This will contribute to the implementation of early intervention measures and strategies that allow students to remain in their programs. To this end, a combination of statistical methods was used, allowing the information obtained to be presented in interactive graphics so that decision-makers could make clearer decisions and improve their operational efficiency. The integration of neutrosophic plithogenic hypotheses strengthened this objective by offering a more complete representation of student profiles. The results of the neutrosophic analysis, such as the high probability of retention in Group 3 (Blue) and the moderate risk in Group 2 (Red), provide an empirical basis for designing tailored interventions, such as strengthening academic support in Group 1 and linking internationalization with employability in Group 2.

The analysis revealed important patterns in student composition, revealing precise trends over the period analyzed. However, significant findings were observed in the composition of the student body according to certain variables, indicating a potential difference in their interests and types of access. This makes it necessary to adjust or modify support modalities and available resources to respond to the changing needs of the student community. The neutrosophic plithogenic approach highlighted these differences by modeling indeterminacy in behavioral patterns, such as the high uncertainty in Group 1 (Green), where neutrosophic negation (0.5, 0.6, 0.2) suggests that current policies do not fully address risk factors. This underscores the need for specific strategies, such as curricular restructuring for Group 1 and strengthening job placement for Group 2, while maintaining employability standards in Group 3.

5. Conclusions

The study concludes that the integration of learning analytics and predictive methods is a key tool for reducing student dropout in higher education. These models use student behavior data from educational platforms, such as login frequency and participation, to create detailed profiles that allow for the early identification of at-risk students. The objective is to shift from a reactive to a proactive approach by implementing specific support actions that respond to learning challenges and help improve retention rates and student success.

The research validates the use of multivariate statistical methods such as Principal Component Analysis (PCA), HJ-Biplot, and cluster analysis to reliably represent the interrelationships between key variables and student typologies. These techniques demonstrated the existence of consistent academic trajectory groups based on performance, employability, or access to support programs. The graphical and quantitative evidence of these patterns establishes a solid empirical basis for designing segmented and contextualized intervention strategies for each student group.

The study's main innovation is the incorporation of neutrosophic plithogenic hypotheses, which enrich predictive models by managing the uncertainty, contradiction, and vagueness inherent in educational data. This approach allowed for a more precise quantification of dropout risk by modeling complex interactions between variables like employability, academic average, and failed subjects. The results of the univariate neutrosophic probability (UNP) confirmed the formation of coherent clusters, identifying a high-risk group (Green), a moderate-risk group (Red), and a group with a high probability of retention (Blue), which allows for more targeted interventions

Finally, the article underscores the need to combine quantitative data with qualitative perspectives to obtain a more holistic understanding of the causes of dropout. It is proposed that future research focus on developing mixed models that integrate factors such as social support or the balance between academic and personal life, whose influence is reflected in the high uncertainty of some groups. The application of the neutrosophic plithogenic approach is positioned as an innovative framework for advancing this line of research, optimizing decision-making and the development of more effective and adaptable support systems in higher education.

References

[1] M. Hussain, W. Zhu, W. Zhang, and S. M. R. Abidi (2018), "Predictions of Student Engagement in an Online Learning System and Its Impact on Grades," Computational Intelligence and Neuroscience, vol. 2018.

[2] J. Jovanović, D. Gašević, S. Dawson, A. Pardo and N. Mirriahi (2017), "Learning analytics for revealing learning strategies in a flipped classroom," Computers in Human Behavior, vol. 72, pp. 757–766.

[3] M. Khalil and M. Ebner (2015), "Learning analytics: principles and limitations," in Proc. World Conf. on Educational Multimedia, Hypermedia and Telecommunications, pp. 1326–1336.

[4] H. Waheed, S. U. Hassan, I. Razzak, A. Muaz and S. Nazir (2020), "Predicting academic performance using machine learning techniques," Computers & Education, vol. 159.

[5] D. Burgos and M. L. Campanario (2018), "Exploring the learning process in virtual reality environments," Computers & Education, vol. 129, pp. 82–89.

[6] L. P. Macfadyen and S. Dawson (2010), "Mining LMS data for an early warning system," Computers & Education, vol. 54, no. 2, pp. 588–599.

[7] D. Ifenthaler and J. Y. K. Yau (2020), "Learning analytics in higher education: a systematic review," Research and Development in Educational Technology, vol. 68, no. 4.

[8] C. Davis, R. Greenway, M. Moore and L. Cooper (2019), "Online Teaching in Social Work Education," Australian Social Work, vol. 72, no. 1, pp. 34–46.

[9] M. Detres, R. Lucio, Z. Roberson, M. Campbell and P. Senger (2020), "Retention in Online Masters of Social Work," Journal of Social Work Education, vol. 40, no. 4, pp. 299–317.

[10] M. Abdous (2019), "Satisfaction and preparation as anxiety factors in online students," Internet and Higher Education, vol. 41, pp. 34–44.

[11] C. Stone and S. O'Shea (2019), "Older, online and for the first time," Australasian Journal of Educational Technology, vol. 35, no. 1, pp. 57–69.

[12] C. Beer and C. Lawson (2017), "Student Dropout in Higher Education: An Alternative Perspective," Journal of Continuing and Higher Education, vol. 41, no. 6, pp. 773–784.

[13] R. S. Baker (2019), "Future Challenges in Educational Data Mining," Journal of Educational Data Mining, vol. 11, no. 1, pp. 1–17.

[14] C. Romero and S. Ventura (2020), "Updated Survey of Educational Data Mining," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 10, no. 3.

[15] R. Ferguson (2012), "Learning analytics: drivers and challenges," British Journal of Educational Technology, vol. 43, no. 5, pp. 765–784.

[16] P. Burgos, N. Dolores, F. A. Amaiquema, and G. E. Beltrán (2019), "Methods and Techniques in Qualitative Research," Conrado, vol. 15, no. 70, pp. 455–459.

[17] J. M. Braxton, A. S. Hirschy, and S. A. McClendon (2004), Understanding and Reducing College Student Attrition, ASHE-ERIC Higher Education Report, vol. 30, no. 3.

[18] J. Jovanović, D. Gašević, S. Dawson, A. Pardo and N. Mirriahi (2017), "Learning analytics in the flipped classroom," Internet and Higher Education, vol. 72, pp. 757-766.

[19] M. P. Galindo-Villardón (1986), "An alternative for simultaneous representation: HJ-Biplot," Qüestiió, vol. 10, no. 1, pp. 13–23.

[20] M. P. Galindo-Villardón and J. L. Vicente-Villardón (2003), MULTBIPLOT: A package for multivariate analysis, Dep. of Statistics, Univ. of Salamanca.

[21] F. Smarandache (1999), Neutrosophic logic and neutrosophic sets, American Research Press.

[22] L. Paquette and R. S. Baker (2019), "Comparison between machine learning and knowledge engineering," Interactive Learning Environments, vol. 27, no. 5–6, pp. 585–597.

[23] A. M. Ali and M. Muthuswamy (2023), "Neutrosophical framework for sustainable energy systems," Sustainable Machine Intelligence Journal, vol. 4, no. 3, pp. 1–10.

[24] F. Smarandache (2021), "Plithogenic probability and statistics," Neutrosophic Sets and Systems, vol. 43.

[25] M. G. Gafar, M. Elhoseny, and M. Gunasekaran (2020), "Neutrosophical modeling with PSO in wildfires," The Journal of Supercomputing, vol. 76, no. 4, pp. 2339–2356.

[26] L. K. Reyes Espinoza, C. Mafla Herrería, and P. A. Sánchez Sandoval (2024), "Neutrosophic PAJ for patients with thrombosis," Neutrosophic Computing and Machine Learning, vol. 31, pp. 198–209.

[27] N. Nabeeh (2023), "Evaluating sustainable transportation with a neutrosophic model," Sustainable Machine Intelligence Journal, vol. 2, no. 2, pp. 1–12.

[28] F. Smarandache (2018), "Plitogenic Set: Review," Neutrosophic Sets and Systems, vol. 21, pp. 153–167.

[29] F. Smarandache (2021), "Introduction to Plitogenic Logic," Neutrosophic Computing and Machine Learning, vol. 18, pp. 1–6.

[30] A. Rezaei, T. Oner, T. Katican, F. Smarandache and N. Gandotra (2022), "Brief history of fuzzy and neutrosophic sets," Int. J. of Neutrosophic Science, vol. 18, no. 1, pp. 99–116.

[31] J. F. Troya-Sarzosa and C. Oña (2024), "Optimal dosage of guinea pig manure," Neutrosophic Sets and Systems, vol. 69, pp. 58–66.

[32] F. Smarandache (2021), "Introduction to multivariate plitogenic logic," Neutrosophic Sets and Systems, vol. 1.

[33] F. Smarandache, M. Abdel-Basset, M. L. Vázquez and S. Broumi (2023), "Neutrosophic advances in Latin America," Neutrosophic Sets and Systems, vol. 62, no. 1, p. 40.

[34] F. Smarandache (2022), "Plitogeny and statistics: a brief review," Journal of Computational and Cognitive Engineering, vol. 1, no. 2, pp. 47–50.

[35] G. J. Hwang and H. F. Chang (2011), "Mobile learning approach and formative assessment," Computers & Education, vol. 56, no. 4.

[36] J. W. Creswell and C. N. Poth (2017), Qualitative Inquiry and Research Design, 4th ed., SAGE Publications.

[37] R. S. Baker (2019), "Challenges for educational data mining," Journal of Educational Data Mining, vol. 11, no. 1, pp. 1–17.

Received: May 30, 2025. Accepted: July 09, 2025.