

University of New Mexico

Plithogenic Analysis in the Optimization of Educational Technologies

Javier Gamboa-Cruzado¹ , Augusto Hidalgo-Sánchez² , Alfonso Tesén Arroyo³ , Angel Núñez Meza⁴ , Edwin Collazos Paucar⁵ , and Dante Manuel Macazana Fernández⁶

¹ National University of San Marcos, Lima. Peru. E-mail: jgamboac@unmsm.edu.pe

² National University of San Marcos, Lima. Peru. E-mail: ahidalgos@unmsm.edu.pe

³ National University Pedro Ruiz Gallo, Chiclayo. Peru. E-mail[: atesen@unprg.edu.pe](mailto:atesen@unprg.edu.pe)

⁴ National University Daniel Alcides Carrión, Pasco. Peru. E-mail: **anunezm@undac.edu.pe**

⁵ National University of San Marcos, Lima. Peru. E-mail: ecollazosp@unmsm.edu.pe

⁶ National University of San Marcos, Lima. Peru. E-mail: **dmacazanaf_af@unmsm.edu.pe**

Abstract. This study has investigated the impact of machine learning techniques on the prediction of academic performance, with the objective of analyzing their integration into teaching and their effects on educational quality, costs, and effectiveness of teaching strategies. To this end, neutrosophic and plithogenic statistical analyses were applied to evaluate the relationship between the independent and dependent variables. The results have shown that the combination of machine learning techniques with active teaching participation not only improves academic performance but also optimizes costs and raises educational quality. Consequently, it is concluded that the integration of technology and teaching participation is essential for improving academic results. This evidence provides a path for future research and applications in the educational field, by highlighting the need to balance technology and humanization in educational strategies.

Keywords: Educational technology, Plithogenic statistics, Teacher participation, Educational quality.

1 Introduction

The use of machine learning techniques to predict students' academic performance has become an essential tool in educational management, facilitating informed decision-making that promotes continuous academic progress. This tool not only aims to anticipate students' future performance but also to identify the most effective methodologies that highlight key variables influencing teaching and motivation, thereby contributing to the reduction of school dropout rates [1]. Since academic performance is a critical indicator of educational quality, the implementation of machine learning is positioned as a crucial component for the continuous improvement of academic management [2].

However, despite advancements in using machine learning for predicting academic performance in this field, numerous challenges remain unaddressed [3]. Specifically, understanding learning patterns in contexts with low literacy levels is fundamental to improving academic performance.

Furthermore, although there has been a growing interest in developing predictive models of academic performance, challenges such as low performance and high dropout rates in higher education persist [4, 5]. In this context, machine learning techniques have been intensively adopted, facilitating the processing and analysis of large volumes of data to support educational decision-making [6].

Therefore, the present study aims to analyze the impact of machine learning techniques on predicting students' academic performance and how their integration into teaching practices influences educational quality[7]. Additionally, this study seeks to determine the effect of these tools on academic outcomes, the operational costs of educational institutions, and the effectiveness of teaching intervention strategies.

2 Plithogenic statistics

Plithogenic statistics are applied to analyze complex data in education, particularly in the evaluation of machine learning techniques. This approach allows for the investigation of how these techniques influence various educational factors, such as academic performance, teacher participation, and the quality of teaching. By employing plithogenic statistics, a detailed understanding of the interactions between these factors and how they simultaneously affect the educational environment is obtained. To implement this method, the plithogenic dynamic must be defined in the context of academic performance prediction [8].

Plithogeny is the dynamic of different types of opposites, and/or their neutrals, and/or non-opposites, and their

organic fusion. Plithogeny is a generalization of dialectics (the dynamic of one type of opposites: $\langle A \rangle$ and \langle antiA>), neutrosophy (the dynamic of one type of opposites and their neutrals: $\langle A \rangle$ and $\langle \text{ant} A \rangle$ and $\langle \text{neut} A \rangle$), as Plithogeny studies the dynamic of many types of opposites and their neutrals and non-opposites $\langle \langle A \rangle$ and $\langle \langle \rangle$ and $\langle \rangle$ and \langle neutA \rangle , \langle B \rangle and \langle antiB \rangle and \langle neutB \rangle , etc.), and many non-opposites (\langle C \rangle , \langle D \rangle , etc.) all together. As an application and particular case derived from Plithogeny, the plithogenic set is an extension of the classical set, fuzzy set, intuitionistic fuzzy set, and neutrosophic set [9], and it has multiple scientific applications [10].

Then, it is called a plithogenic set (P,a,V,d,c):

Where "P" is a set, "a" is an attribute (generally multidimensional), "V" is the range of values of the attribute, "d" is the degree of membership of the attribute value of each element x to the set P for some given criteria ($x \in$ P), and "d" means " d_F " or " d_{IF} " or " d_N ", when it is a fuzzy membership degree, an intuitionistic fuzzy membership, or a neutrosophic membership degree, respectively, of an element x to the plithogenic set P;

"c" means " c_F " or " c_{IF} " or " c_N ", when it is a fuzzy attribute value contradiction degree function, an intuitionistic fuzzy attribute value contradiction degree function, or a neutrosophic attribute value contradiction degree function, respectively [11].

Functions are defined according to the applications that experts need to address. d(∙,∙) and c(∙,∙) then use the following notation: $x(d(x,V))$, where $d(x,V) = \{d(x,v),$ for all $v \in V\}$, $\forall x \in P$. Thus, plithogenic statistical analysis allows for addressing the complexity of the perceptions of the analyzed sample [12, 13].

3 Material and Methods

The systematic literature review was conducted using an approach that allowed for an analytical response to the research questions related to the use of machine learning in academic performance [7]. The initial search in academic databases such as Science Direct and Scopus yielded 717,389 articles. These were filtered through exclusion criteria, resulting in the selection of 88 relevant studies (study sample). The selected articles were evaluated based on criteria, including the clarity of objectives and the organization of the document. The data extracted from these sources formed the basis for the modeling of plithogenic statistics.

A linguistic evaluation system was adapted to the plithogenic model to accurately capture the experts' opinions (see Table 1).

Scale	Plithogenic scale	S(T,I,F)	Using ML ma- chine learning techniques (VI1)	Teacher par- ticipation (VI2)	Academic per- formance (VD1)	Operational costs (VD2)	Ouality of edu- cation (VD3)
7	(0.95, 0.05. 0.05)	0.90	High Extremely (EH)	High Very (VH)	Extremely High (EH)	High Very (VH)	Extremely High (EH)
6	(0.80, 0.15. 0.10)	0.75	Very High (VH)	High(H)	Very High (VH)	High(H)	Very High (VH)
5	0.25, (0.65, 0.20)	0.65	High(H)	Moderately High(MH)	High(H)	Moderately High(MH)	High(H)
$\overline{4}$	0.35. (0.50, (0.30)	0.50	Moderately High (MH)	Medium (M)	Moderately High (MH)	Medium (M)	Moderately High (MH)
3	(0.35, 0.45, (0.40)	0.45	Medium (M)	Moderately Low (ML)	Medium (M)	Moderately Low (ML)	Medium (M)
\overline{c}	0.60, (0.20, 0.50)	0.35	Moderately Low (ML)	Low (L)	Moderately Low (ML)	Low (L)	Moderately Low (ML)
-1	0.75, (0.10, 0.65)	0.25	Low (L)	Very Low (V _L)	Low (L)	Very Low (V _L)	Low (L)

Table 1: Plithogenic scales to evaluate the educational impact of machine learning.

Consequently, the dataset is evaluated, which is formed totally or partially by data with some degree of indeterminacy and contradiction. The plithogenic statistical method is used to interpret and organize plithogenic data to reveal underlying patterns [14, 15].

For the plithogenic statistical modeling in this study, a random variable P is referenced, representing the lower and upper levels that the studied plithogenic variable can reach, within an indeterminate and contradictory interval. Thus, it follows the plithogenic mean of the variable (\overline{P}) when formulating [16]:

$$
\overline{P} = \frac{1}{n_P} \sum_{i=1}^{n_P} P_i \tag{1}
$$

Javier G. Cruzado, Augusto H. Sánchez, Alfonso T. Arroyo, Angel N. Meza, Edwin C. Paucar, Dante M. Macazana F. Plithogenic Analysis in the Optimization of Educational Technologies

Where n_p is a plithogenic random sample from the studied population. Once the mean is defined, the next step is to calculate the variance of the plithogenic sample. To do this, it is necessary to convert a plithogenic number to a scalar number according to the methodology analyzed in the study materials. Subsequently, the following equation is used to calculate S_P^2 [16]:

$$
S_P^2 = \frac{\sum_{i=1}^{n_P} (P_i - \bar{P}_i)^2}{n_P} \tag{2}
$$

Subsequently, the plithogenic coefficient (CV_p) is calculated, which measures the consistency of the variable. The lower the value of CV_P , the more consistent the performance of the analyzed element is compared to the others studied. The following equation is proposed [16]:

$$
CV_P = \sqrt{S_P^2} \times 100
$$
 (3)

4 Results

The following presents an analysis of the implications of applying machine learning techniques in education. The research has addressed how these techniques can influence various aspects of the educational system, from the overall quality of education to the impact on institutional costs. In this context, several hypotheses about the effectiveness and benefits of machine learning in different educational areas are explored (see Table 2).

Table 2: Neutrosophic hypothesis of the study.

To explore how machine learning techniques influence the educational system using a plithogenic approach, it is essential to clearly define the variables to be modeled. The following includes the independent and dependent variables for statistical modeling, as well as the evaluation of each variable in the presented sample (see Tables 3 and 4), and the representation of determining and indeterminate elements (see Figure 1).

Table 3: Variables to measure machine learning.

Javier G. Cruzado, Augusto H. Sánchez, Alfonso T. Arroyo, Angel N. Meza, Edwin C. Paucar, Dante M. Macazana F. Plithogenic Analysis in the Optimization of Educational Technologies

Table 4: Sample analyzed for variables VI1, VI2, VD1, VD2 and VD3.

Interpretation of Results:

Use of machine learning techniques (VI1): Values for this variable range from 0.61 to 0.99, covering a spectrum from low to extremely high levels. This indicates that educational institutions apply machine learning techniques to varying degrees. Institutions using advanced techniques achieve greater accuracy in predicting academic performance and better resource optimization, contributing to higher educational quality.

Teacher participation in predictive analytics (VI2): Values range from 0.81 to 1, showing high to extremely high levels of teacher involvement in educational data analysis. Intense participation allows teachers to better adapt pedagogical strategies, leading to significant improvements in student academic performance.

Academic performance (VD1): The variable shows values between 0.7 and 1, reflecting generally high academic performance. This suggests that machine learning techniques and effective teaching practices are associated with good academic results. Consistency in high performance validates the effectiveness of the applied educational strategies.

Operational costs (VD2): Operational costs range from 0.71 to 0.99, from medium to extremely high levels. This variability reflects how different institutions manage their expenses related to the implementation of machine

Javier G. Cruzado, Augusto H. Sánchez, Alfonso T. Arroyo, Angel N. Meza, Edwin C. Paucar, Dante M. Macazana F. Plithogenic Analysis in the Optimization of Educational Technologies

learning techniques. Institutions with high costs invest in advanced technologies, while others optimize resources and reduce expenses in the long term.

Quality of education (VD3): Values for this variable range from 0.69 to 0.99, indicating high educational quality. This shows that effective implementation of predictive techniques and active teacher participation contribute to a positive educational experience. High educational quality signifies that machine learning techniques are well-applied and adapted to student needs.

Figure 1: Indeterminacy of the analyzed sample. Source: Own elaboration.

In Figure 1, the levels of indeterminacy for each variable are observed. Among them:

Usage of Machine Learning Techniques (VI1): This section represents the level of application of machine learning techniques, which is clearly defined and measured. The determined part shows that, generally, the use of these techniques is quite high and well understood. However, the undetermined part reflects challenges such as the lack of standardization and variability in the data, introducing uncertainty into the results and the effectiveness of these techniques.

Teacher Participation in Predictive Analysis (VI2): Teacher participation is high and well established, suggesting a good degree of involvement in predictive analysis. Nonetheless, the undetermined part represents difficulties in accurately measuring this participation, including differences in teacher training and motivation.

Academic Performance (VD1): Academic performance is relatively stable and measurable, indicating that student grades are clearly recorded. However, the undetermined part reflects issues such as variability in assessment and uncontrolled external factors, introducing uncertainty into the interpretation of performance data.

Operational Costs (VD2): Operational costs are relatively clear and well documented, facilitating financial management. Despite this, the undetermined part suggests challenges in accurately forecasting costs, such as fluctuations in expenses and changes in budgets, which affect planning and the evaluation of economic efficiency.

Quality of Education (VD3): The quality of education is evaluated quite stably and consistently, reflecting a high level of satisfaction among students and teachers. The undetermined part is low, indicating that uncertainties in this variable are minor. However, variations in the perception of quality or contextual factors not fully reflected in the evaluation should be considered.

Therefore, the impact of the variables in the research is evaluated based on the analysis of the weighted average value, the range of standard deviation, and the weighted coefficient of variation for each variable. These data detail the behavior and consistency of the analyzed variables (see Table 5) [17, 18].

Table 5: Obtaining \bar{x}_P , S_P and CV_P of the variables VI1, VI2, VD1, VD2 and VD3.

Variable	$\overline{\mathbf{X}}$ _p	Sp	\bf{CV}_{P}
VI1	$0.71 + 0.85$ I	$0.059 + 0.072$ I	$0.083 + 0.085$ I
VI ₂	$0.85 + 0.93$ I	$0.033 + 0.037$ I	$0.039 + 0.04$ I
V _D 1	$0.77 + 0.9$ I	$0.05 + 0.052$ J	$0.065 + 0.058$ I
V _D 2	$0.76 + 0.93$ I	$0.03 + 0.035$ I	$0.039 + 0.038$ I
VD3	$0.84 + 0.89$ I	$0.04 + 0.055$ I	$0.048 + 0.062$ I

Javier G. Cruzado, Augusto H. Sánchez, Alfonso T. Arroyo, Angel N. Meza, Edwin C. Paucar, Dante M. Macazana F. Plithogenic Analysis in the Optimization of Educational Technologies

Usage of Machine Learning Techniques (VI1): This variable shows a high average in the use of machine learning techniques but presents moderate uncertainty. Compared to other variables, its high variability makes predictions of academic performance less consistent. However, intensive use of techniques enhances the prediction and reliability of predictive models.

Teacher Participation in Predictive Analysis (VI2): Teacher participation is high and its uncertainty is low, indicating greater consistency in implementing predictive analysis. Compared to VI1, this variable offers greater stability, facilitating the accuracy of predictions. High teacher participation helps make machine learning models more effective and reliable by contributing to more stable predictions of academic performance.

Academic Performance (VD1): Academic performance shows a high average but with considerable uncertainty. Its variability is moderate, affecting the stability of predictive models. Compared to VI2, which has lower variability, VD1 may present additional challenges in prediction due to its lack of consistency. Variability in academic performance complicates the task of modeling and predicting student success with precision.

Operational Costs (VD2): Operational costs show a high mean with low uncertainty, indicating that cost data is stable. Compared to VD1, operational costs are more predictable, which facilitates resource optimization without significantly affecting the accuracy of academic performance predictions. Stability in costs allows machine learning models to adjust more efficiently, contributing to better resource management and intervention strategies.

Quality of Education (VD3): The quality of education has a high average with moderate uncertainty. Its variability, although lower than that of VD1, can still impact the stability of predictions. Compared to VI2 and VD2, the quality of education is less consistent, which may influence the accuracy of predictive models. Improving the stability of this variable is crucial for enhancing the accuracy of predictions related to academic performance.

After performing the comparative plithogenic analysis [18, 19, 20] , the correlation between independent and dependent variables is evaluated to identify the strongest relationships and determine which hypothesis is most viable (see Table 6). The following relationships are proposed:

- ❖ H1: Correlation between VI1 and VD3.
- ❖ H2: Correlation between VI1 and VD1.
- ❖ H3: Correlation between VI1 and VD2.
- ❖ H4: Correlation between VI2 and VD1.

Table 6: Pearson correlation coefficients.

The table shows the Pearson correlation coefficients, which indicate the strength and direction of the linear relationships between the variables. This analysis focuses on how the independent variables (VI1 and VI2) relate to the dependent variables (VD1, VD2, and VD3). The analysis reveals that teacher participation in predictive analysis (VI2) has a significant influence on academic performance (VD1), suggesting that the active involvement of teachers in using machine learning tools is crucial for improving student outcomes. On the other hand, the use of machine learning techniques (VI1) has a moderate impact on the quality of education (VD3) and a limited effect on operational costs (VD2) and academic performance (VD1). This indicates that while technology is a significant facilitator, the commitment and participation of teachers in the predictive analysis process are essential for achieving better academic performance results.

The acceptance of hypothesis H4 confirms that the active participation of teachers in predictive analysis positively affects student academic performance. This result highlights several key aspects that must be considered in the research:

- ❖ Importance of Teacher Intervention: The analysis explores how the integration of machine learning techniques, combined with teacher involvement in predictive analysis, affects critical variables such as academic performance, operational costs, and educational quality. Hypothesis H4 emphasizes the significant impact of teacher participation on improving academic performance, highlighting the crucial role within the analyzed neutrosophic context.
- ❖ Direct Connection with Dependent Variables: H4 establishes a clear relationship between teacher participation in predictive analysis and improvement in academic performance (VD1), one of the most relevant dependent variables. Academic performance, as a key indicator of educational process

effectiveness, is directly influenced by this participation, reinforcing the relevance of the hypothesis within the research framework.

❖ Applicability and Relevance: The analysis of the data and the involved variables suggests that H4 aligns with the idea that active teacher participation in the interpretation and application of predictive analysis has a significant impact on student academic outcomes.

The correlation between the variables allows for the identification of areas where different variables coincide or overlap, helping to mitigate specific challenges in predicting academic performance. Among those identified are:

- ❖ Correlation between VI1 (Use of Machine Learning Techniques) and VD1 (Academic Performance): The Pearson correlation coefficient is 0.0973, indicating a moderate intersection between these variables. This suggests that while appropriate use of machine learning techniques positively impacts academic performance, it is not a decisive factor. The intersection highlights an area where optimizing these techniques can help overcome challenges in accurately predicting academic performance.
- ❖ Correlation between VI2 (Teacher Participation in Predictive Analysis) and VD1 (Academic Performance): With a coefficient of 0.491, this intersection is significant. The strong correlation indicates that active teacher participation is crucial for improving academic outcomes. Here, the intersection reveals that involving teachers in the predictive process mitigates challenges associated with applying machine learning techniques without human supervision and adjustment.
- ❖ Correlation between VI1 (Use of Machine Learning Techniques) and VD3 (Quality of Education): The high intersection (0.3695) shows that the use of machine learning techniques also affects perceptions of educational quality. This suggests that optimizing these techniques not only improves performance but also enhances overall satisfaction among students and teachers. Addressing challenges in perceiving educational effectiveness accurately is crucial for improving quality.

Conclusion

The validation of hypothesis H4 confirms that active teacher participation in predictive analysis significantly enhances students' academic performance. Additionally, the use of plithogenic statistics has been crucial for identifying and understanding the complex relationships between independent and dependent variables. This approach has allowed for the detection of areas of indeterminacy and contradiction, providing a more comprehensive view of how interactions between technology and teacher participation impact educational outcomes.

The plithogenic analysis has demonstrated that combining machine learning techniques with strong teacher involvement is essential for improving academic performance. The areas of intersection and union indicated that the most effective educational strategies are those that balance technology with human involvement, while the levels of contradiction highlight the need for balanced implementation. These solid and applicable relationships suggest a clear path for designing educational policies and practices that optimize academic performance through machine learning.

References

- [1] Pallathadka, H., Wenda, A., Ramirez-Asís, E., Asís-López, M., Flores-Albornoz, J., & Phasinam, K. (2023). Classification and prediction of student performance data using various machine learning algorithms. Materials Today: Proceedings, 80(Part 3), 3782-3785.
- [2] Dabhade, P., Agarwal, R., Alameen, K., Fathima, A., Sridharan, R., & Gopakumar, G. (2021). Educational data mining for predicting students' academic performance using machine learning algorithms. Materials Today: Proceedings, 47(Part 15), 5260-5267.
- [3] Yağcı, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. Smart Learning Environments, 9(1), 11.
- [4] Kuadey, N. A., Ankora, C., Tahiru, F., Bensah, L., Agbesi, C. C. M., & Bolatimi, S. O. (2024). Using machine learning algorithms to examine the impact of technostress creators on student learning burnout and perceived academic performance. International Journal of Information Technology, 16(4), 2467-2482.
- [5] Rastrollo-Guerrero, J. L., Gómez-Pulido, J. A., & Durán-Domínguez, A. (2020). Analyzing and predicting students' performance by means of machine learning: A review. Applied Sciences, 10(3), 1042.
- [6] Maehara, R., Benites, L., Talavera, A., Aybar-Flores, A., & Muñoz, M. (2024). Predicting financial inclusion in Peru: Application of machine learning algorithms. Journal of Risk and Financial Management, 17(1), 34.
- [7] Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student performance prediction using machine learning techniques. Education Sciences, 11(9), 552.
- [8] Gutiérrez, P., et al. (2023). Neutrosophic Sets and Systems Special Issue: Neutrosophic advancements Plithogenic Iadov model to study university teaching practices in the complexity of the educational process of comprehensive training by competencies. Neutrosophic Sets and Systems, 62(1), 79-83.
- [9] Cristina, E., et al. (2022). Neutrosophic statistical analysis of self-assessment of rehabilitation knowledge in university students. Neutrosophic Sets and Systems, 52(October), 2-5.
- [10] Smarandache, F. (2022). Plithogeny, plithogenic set, logic, probability and statistics: A short review. Journal of Computational and Cognitive Engineering, 1(2), 47-50.
- [11] González Caballero, E., Leyva Vázquez, M., & Smarandache, F. (2021). On neutrosophic uninorms. Neutrosophic Sets and Systems, 45(2021), 340-348.
- [12] La Cantera, D., Hernandez, D., Quiroz, R. C., Queija, M. L., Gonzalez, J. R., & Vazquez, M. Y. L. (2024). Analyzing interdisciplinary education in general medicine using Smarandache's multivalued logic hypothesis theory and plithogenic probability. Neutrosophic Sets and Systems, 70(1), 23.
- [13] Ali, M. A., & Muthuswamy, M. (2023). Neutrosophic multi-criteria decision-making framework for sustainable evaluation of power production systems in renewable energy sources. Sustainable Machine Intelligence Journal, 4(3), 1-10.
- [14] Sánchez, F. C., Almeida Blacio, J. H., Flores Bracho, M. G., Andrade Santamaria, D. R., & Casanova, R. S. (2021). Neutrosophic and plithogenic statistical analysis in educational development. Neutrosophic Sets & Systems, 44.
- [15] Smarandache, F. (2021). Introducción a la lógica plitogénica. Neutrosophic Computing and Machine Learning, 18, 1-6.
- [16] Troya-Sarzosa, J. F., & Oña, C. (2024). Using plithogenic statistics to determine optimal guinea pig manure dosing in prickly pear cultivation. Neutrosophic Sets and Systems, 69(1), 8.
- [17] Leyva Vázquez, M., & Smarandache, F. (2021). Resolución de sistemas de ecuaciones lineales neutrosóficas mediante computación simbólica. Neutrosophic Computing and Machine Learning, 15, 41-46.
- [18] Vázquez, M. L., & Torres, L. C. (2024). Integrating neutrosophic numbers in regression analysis for enhancing predictive modelling through uncertainty representation. Plithogenic Logic and Computation, 1, 120-126.
- [19] Marín-Eléspuru, C. U., Melgarejo-Mariño, D. M., Solsol-Hidalgo, E. A., Chiroque-Sernaqué, D., & Balbuena Hernández, J. R. (2024). Exploring the impact of performance audits on the management of public organizations through the analysis of plithogenic statistics. Neutrosophic Sets and Systems, 69(1), 21.
- [20] El-Douh, A. A., Lu, S., Abdelhafeez, A., & Aziz, A. S. (2023). Assessment of health sustainability using neutrosophic MCDM methodology: Case study COVID-19. Sustainable Machine Intelligence Journal, 3, 1-1.

Received: June 21, 2024. Accepted: August 12, 2024