Development of a Student Performance Prediction System with Deep Learning

Alexander EJ Villegas-Espinoza [1], Jorge Isaac Necochea-Chamorro [1]

[1] *Universidad César Vallejo, Lima -norte, Perú*

Abstract **– The present study focuses on the development of deep learning (DL) model to predict student performance and address the issue of low academic achievement among secondary school students. Utilizing a quasi-experimental design, the research process involved the enhancement of data with generative adversarial networks (GAN) and synthetic minority over-sampling techniques (SMOTE), followed by the training of a deep neural network (DNN). The model demonstrates high accuracy, achieving 96.81% and a Cohen's Kappa index of 0.866, indicating strong reliability.**

A comprehensive investigation is conducted to identify the key variables influencing student performance. These variables include self-concept, attitude towards subjects, parental satisfaction, and the use of additional learning resources. These factors are critical in building a robust predictive model capable of detecting students at risk of poor academic outcomes.

The findings underscore the efficacy of the model in not only identifying at-risk students but also in facilitating the implementation of personalized intervention strategies. The objective of these strategies is twofold: to reduce the rate of school failure and to enhance the overall quality of education.

Keywords **– Deep learning, student performance, prediction, deep neural networks, education.**

DOI: 10.18421/TEM144-64
https://doi.org/10.18421/TEM144-64

Corresponding author: Jorge Isaac Necochea Chamorro,
Universidad César Vallejo, Lima -norte, Perú
Email: jnecochea@ucvvirtual.edu.pe

Received: 04 December 2024.
Revised: 08 June 2025.
Accepted: 02 September 2025.
Published: 27 November 2025.

1. Introduction

Education is a constantly evolving pillar of modern society and plays a crucial role in preparing individuals for future challenges.

Nevertheless, a pervasive issue in educational institutions is the high rate of students who do not successfully complete the academic year [1], [2].

This phenomenon not only impacts individual academic performance but also exerts adverse effects on the quality of education and social development. It hinders student progress and interferes with their educational objectives. The aforementioned challenges are primarily attributable to the following factors: the instructor's engagement with students in their academic endeavors, the students' performance in the subject matter, their emotional state during the course, and the prevailing social context. As previously mentioned, it is imperative to implement innovative solutions that can predict and determine the student performance of each individual student.

Machine learning encompasses a subfield known as deep learning, which is distinguished by its capacity to process information through artificial neural networks. These networks are comprised of multiple layers that are capable of identifying complex patterns in voluminous data sets. The process can be likened to the human brain, which functions as an information repository, learning through sensory experiences and actions. It retains these inputs based on various situations, and, according to these inputs, it determines its point of inspiration.

In this study, this technology is applied to identify patterns [4] that are not easily detectable by humans. Specifically, one of its applications involves techniques capable of recognizing common characteristics in the data that would otherwise go unnoticed [3].

Student performance is a critical metric in educational assessment, reflecting learners' achievement of both short-term and long-term academic goals. This comprehensive measure generally comprises several evaluation components, including class participation, assignment completion, and assessment outcomes [5], [6].

In contemporary educational contexts, there is growing recognition of the need for advanced predictive systems that can identify at-risk students and enable timely interventions.

The present study endeavors to address a pivotal research question: The central question guiding this study is how deep learning (DL) models can effectively predict student academic performance. The investigation aligns with sustainable development goal (SDG) 9, which pertains to Industry, Innovation and Infrastructure [7], through its development of an artificial intelligence system employing deep neural networks (DNNs). The central hypothesis of this study posits that a deep neural network (DNN)-based model can predict academic performance with a success rate exceeding 90% by analyzing behavioral characteristics, learning dispositions, and parental involvement metrics.

The impetus for this research stems from the documented shortcomings of traditional student support systems. Current approaches often prove ineffective in identifying students who are experiencing academic difficulties until such difficulties reach a severe stage [8], [9]. DL-based predictive systems have the potential to effect transformative change in educational institutions by enabling the early identification of at-risk students, facilitating targeted interventions, and ultimately reducing failure rates while enhancing institutional educational quality [10].

Recent advancements in the field of educational data mining have substantiated the efficacy of machine learning methodologies. Notable studies include artificial neural network (ANN) implementations achieving 73% accuracy in pass/fail classification using demographic and institutional data [10], hybrid models combining Random Forest, SVM, and Decision Tree algorithms [11], and SMOTE-based approaches attaining prediction accuracies of 0.932 and 0.964 for mathematics and Portuguese performance, respectively [12]. Large-scale analyses employing 1D CNN-LSTM architectures have processed 3.8 million records with a mean absolute error of 0.593 [13], while optimized CNN implementations have achieved 100% accuracy in controlled datasets [14]. ANN models analyzing Nigerian student populations have demonstrated 97.36% prediction accuracy [15].

The methodological foundation of this research incorporates several key technical components. DNN architectures leverage multiple hidden layers with iterative weight optimization during training [16], while addressing common data limitations through generative adversarial networks (GANs) for synthetic data generation [17] and SMOTE techniques for class imbalance mitigation [18].

The validation framework utilizes k-fold cross-validation to ensure model robustness [19], with predictive systems analyzing historical patterns to forecast future academic performance [20]. This approach is particularly relevant in educational contexts where student outcomes are measured through graded assessments, participation metrics, and assignment completion—all of which are significantly influenced by familial, social, and demographic factors [21].

The significance of this study extends beyond the realm of technical innovation, offering practical solutions for educational institutions. The proposed system has the potential to transform student outcomes by enabling the early identification of academic risk factors and facilitating targeted support strategies. Furthermore, it could contribute to broader educational quality improvement and sustainable development goals.

2. Methodology

The development and training of the DL (deep learning) model in this research will be facilitated by advanced tools such as Google Colab and TensorFlow. Colab offers an accessible and collaborative environment that enables the execution of code in the cloud, thereby facilitating access to computing and storage resources [22]. The utilization of this technology is advantageous in the domains of data science and artificial intelligence, particularly in the context of machine and deep learning applications. TensorFlow, a robust and flexible library developed by Google, is ideal for building and evaluating neural networks, as it handles large volumes of data and allows the implementation of various types of neural models [23]. The creation of neural network models to identify students exhibiting poor academic performance will be facilitated by the utilization of TensorFlow and Keras modules. In addition, libraries such as pandas and seaborn will be employed for data manipulation and data processing techniques. These techniques integrate precision, recall, and F1 score metrics to evaluate the effectiveness of the predictions.

The implementation of DL models for the purpose of predicting student performance has a significant impact at both the educational and social levels. The identification of at-risk students and the provision of the requisite support have been demonstrated to improve not only their academic performance but also their mental well-being, thereby enhancing their confidence in overcoming adversities and achieving their goals. The findings of this study support the reduction of the school failure rate, promote educational equity, and strengthen the social fabric by preparing students to be productive and engaged members of society [24].

In the development of the study, the applied type was considered, given its purpose of solving specific problems and improving existing processes [25]. This line of research is of paramount importance as it aims to generate practical knowledge that can be implemented to address and solve educational challenges in an effective manner. The approach is quantitative in nature, as it relies on numerical data during the data processing stage to facilitate statistical interpretation that can be generalized to other analogous situations. The present study is considered quasi-experimental, a design in which the researcher manipulates the independent variable (IV) to observe its effect on the dependent variable (DV), but without randomly assigning participants to groups [27], [26].

For the purpose of data collection, the researchers employed a questionnaire, an instrument that has been demonstrated to facilitate the acquisition of both quantitative and qualitative data [28]. These data were utilized for the model dataset, and the instrument applied was digital forms [29], which are highly useful for obtaining data in a systematized and easily processable manner. The applied research endeavors to develop a tool for predicting student performance, thereby facilitating effective interventions to reduce school failure. Utilizing a quantitative approach, employs substantial volumes of data necessitated by DL. The quasi-experimental design enabled a comparison of the model's accuracy with real data from report cards, which were limited to high school students from a secondary educational institution.

The scope of the study is explanatory, analyzing how the DL system predicts student performance. The limitations of the study included the unavailability of data and the inaccurate interpretation of results. The independent variable in this study was the DL-based prediction system, while the dependent variable was student performance, which was assessed using a dataset processed in the model. The dimensions of student achievement centered on the student and the home environment. At the student level, the following factors were considered: attitudes, dispositions, demographic characteristics, and behavioral characteristics. These factors were included in order to reflect the students' commitment to learning. In the domestic environment, factors such as socioeconomic status and parental involvement have been demonstrated to exert influence on academic outcomes [30].

The extant research identified that student performance is influenced by household factors such as socioeconomic status and parental involvement in the study [30]. The population comprised 155 high school students in Canta, of whom 112 participated in the study. The data presented herein were obtained through the utilization of a questionnaire that had previously undergone validation [30], [31].

The dataset encompassed nominal, ordinal, and interval measurement scales. Subsequent to the process of data cleansing, 2,000 synthetic data points were generated through the employment of GAN networks utilizing k-fold cross-validation (with five folds). This process culminated in a total of 10,000 data points allocated for training (80%) and testing (20%), respectively. The DNN model implemented was a multi-categorical classification model, incorporating dense layers, dropout, and an output of four neurons. To address class imbalance, SMOTE was employed in conjunction with k-fold cross-validation, thereby enhancing the model's robustness and ensuring the accuracy of its predictions.

As an ethical imperative in this research, the transparency of the results was ensured, precluding any fabrication in the methods to maintain scientific integrity [32]. With regard to data management, informed consent was obtained from the students, with the purpose of using their data for the research being explained [33]. These data were protected to ensure student privacy and stored securely to prevent unauthorized access [34].

3. Results

The model demonstrated its capacity to identify students who were at risk, thereby highlighting key areas for improvement, including teamwork, the topics covered, and self-concept.

A series of key patterns were identified, including self-concept (0.347) and parental satisfaction (0.337), which were determined to be the most relevant indicators of student performance. In addition to these factors, attitudes towards subjects, topics covered, and teamwork were also considered. However, factors such as economy and parental involvement were found to have a lower impact (<0.05) as shown in Table 1.

Table 1. Pearson correlation of the variables

Days attended	Pearson correlation	,198
	Next (bilateral)	,036
	N	112
Interventions active	Pearson correlation	,128
	Next (bilateral)	,179
	N	112
Job in equipment	Pearson correlation	,286
	Next (bilateral)	,002
	N	112
Topics viewed	Pearson correlation	,289
	Next (bilateral)	,002
	N	112
Resources used	Pearson correlation	,272
	Next (bilateral)	,004
	N	112
Socioeconomic level	Pearson correlation	,063
	Next (bilateral)	,508
	N	112
Stake active parents	Pearson correlation	-,119
	Next (bilateral)	,213
	N	112
Parents' satisfaction	Pearson correlation	,337
	Next (bilateral)	,000
	N	112
Self-concept	Pearson correlation	,347
	Next (bilateral)	,000
	N	112
Attitude towards the subjects	Pearson correlation	,299
	Next (bilateral)	,001
	N	112
Exercise	Pearson correlation	-,101
	Next (bilateral)	,290
	N	112
Responsible parents	Pearson correlation	,051
	Next (bilateral)	,592
	N	112

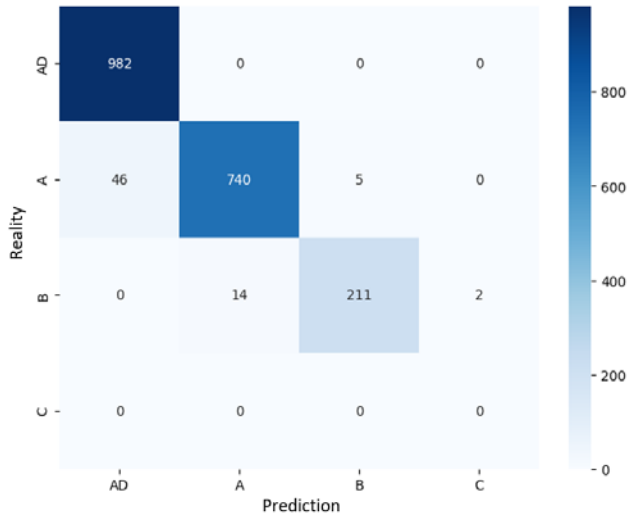

Figure 1. Confusion matrix

The development of the deep neural network (DNN) entailed the incorporation of 11 inputs, the implementation of hidden layers with L2 regularization, the integration of dropout with a 40% probability, and the construction of an output comprising 4 softmax neurons. The model, which was trained with 10,000 synthetic data points generated using a generative adversarial network (GAN) and SMOTE, achieved an accuracy of 97.50% in the training phase and 96.65% in the testing phase as shown in Figure 1.

The accuracy, F1 score, and recall metrics exhibited a 95% success rate or higher, thereby substantiating their reliability (see Table 2 for details).

Table 2. Model metrics

Model metrics (DNN)	
Precision	96.81% (0.9681)
Accuracy (Recall)	96.65% (0.9665)
F1 Score	96.67% (0.9667)
Accuracy	96.65% (0.9665)

The findings underscore the efficacy of the model in predicting student performance. To validate the hypothesis, Cohen's Kappa index was employed, yielding a value of 0.866, signifying substantial concordance with the actual data (see Table 3).

Table 3. Cohen's Kappa

	Worth	Standard error asymptotic	Approximate T
Kappa	,866	,051	13,737
Number of cases valid	112		

According to the predictions illustrated in Figure 2, the model identified that 10.71% of students have a high probability of failing, while 76.79% have a low probability.

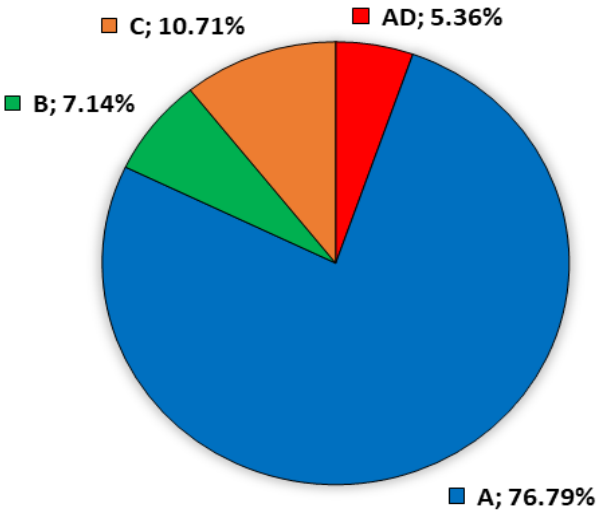

Figure 2. Prediction results

Furthermore, an examination of the model's application to the student data set reveals the following findings: Among the 112 respondents, 6 (5.36%) exhibited an extremely low probability of failing, 86 (76.79%) demonstrated a low probability of failing, 8 (7.14%) exhibited a moderate probability of failing, and finally, 8 (10.71%) exhibited a high probability of failing during the school year.

As illustrated in Figure 3, low-performing students exhibit several academic disadvantages, including a lack of active interventions, limited engagement in teamwork, insufficient topic coverage, fewer resources utilized, lower parental satisfaction, weaker academic self-concept, and less favorable attitudes toward school subjects, along with reduced parental responsibility.

Interestingly, however, they engage in significantly more physical exercise than their higher-performing peers.

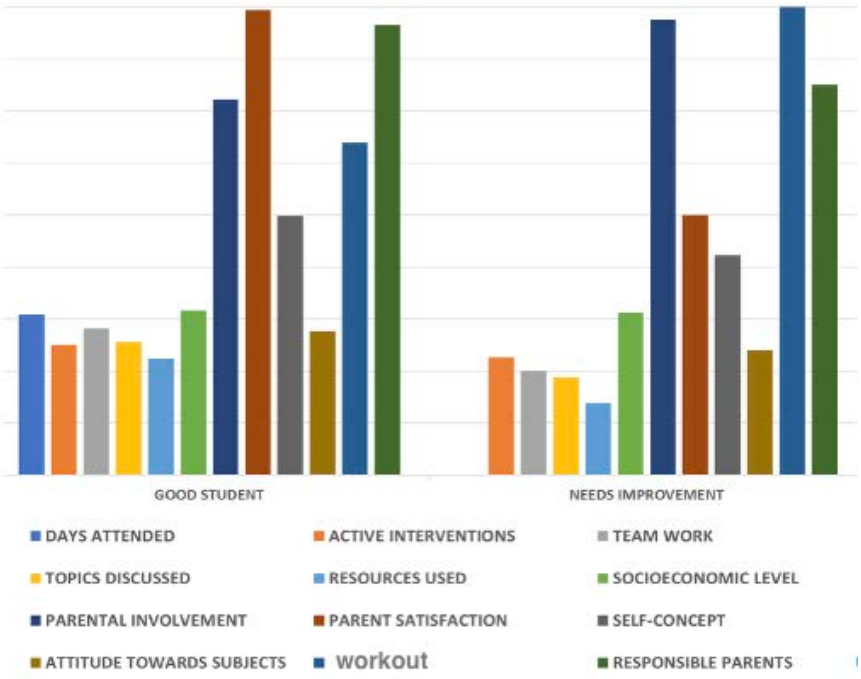

Figure 3. Comparison of indicators between students disapproved and approved

4. Discussion

The findings of this study provide valuable insights into factors influencing academic performance and the development of a predictive model using advanced machine learning techniques.

The research identified key variables affecting student performance, including self-concept, parental satisfaction, attitude toward subjects, class attendance, teamwork, and use of additional resources. These results align with the findings of [15], which underscores the pivotal role of emotional support and self-efficacy in academic achievement. Furthermore, [10] underscores the pivotal role of student engagement, emphasizing its impact on active participation in the learning process.

A salient finding was that students with lower academic performance exhibited higher levels of physical activity, contrasting with studies associating exercise with cognitive improvements [35].

The developed model employed a deep neural network (DNN) with generative adversarial networks (GAN) for synthetic data generation, synthetic minority over-sampling technique (SMOTE) for class balancing, and techniques including dropout and L2 regularization to prevent overfitting. This approach is corroborated by the findings of [12], which advocates for the implementation of these techniques to enhance model generalization. Furthermore, [13] demonstrated the efficacy of GANs in expanding limited datasets, validating their application in this study. In comparison with previous studies, our model demonstrated superior performance metrics, with precision, accuracy, and F1-score values of 96.81%, 96.65%, and 96.67%, respectively.

These results surpass the 73% accuracy reported in [10]. While [15] achieved slightly higher precision (97.36%), the study distinguishes itself through its utilization of Cohen's Kappa (0.866), signifying nearly perfect agreement between predictions and actual data. This finding indicates that the model is both accurate and robust under conditions of limited data.

The model identified 10.71% of students as being at risk of failing during the academic year, which is consistent with the findings reported in [14]. That study used predictions to implement personalized support strategies. The critical variables associated with poor performance (teamwork, resource utilization, self-concept) align with the recommendations made by [11], which advocate for collaborative environments to enhance academic outcomes.

The present study corroborates the following conclusion: The following conclusions can be drawn from the extant literature:

Socio-emotional factors (self-concept, parental satisfaction) and study habits (attendance, resource use) are key determinants, consistent with the findings of [10] and [15].

Advanced techniques (GAN, SMOTE, regularization) significantly improve predictive model accuracy, as demonstrated by the research of [12] and [13].

Early identification of at-risk students enables targeted interventions, as proposed by [14]. Future research should examine the role of intrinsic motivation in the exercise-performance relationship and the impact of AI-based interventions in real educational settings.

5. Conclusions

The present study successfully identified key predictors of academic performance through quantitative analysis of cognitive and non-cognitive factors, including motivation, concentration, subject attitudes, and behavioral characteristics. Despite the exclusive use of quantitative instruments, such as Likert-scale questionnaires, the study revealed an unexpected finding: higher sports participation among students with lower academic performance. This observation highlights the potential value of incorporating qualitative methodologies in future investigations to enhance the understanding of such complex behavioral patterns.

The deep neural network (DNN) model exhibited strong predictive validity, with an accuracy of 96.81% in academic risk classification, thereby substantiating the efficacy of the quantitative approach. However, the limitations of relying solely on quantitative data became evident when interpreting nuanced findings, such as the sports-performance relationship. This finding indicates that while quantitative methods offer substantial predictive capacity, it is also possible to be further augmented by qualitative insights to provide a comprehensive explanation of the underlying mechanisms.

From a pragmatic standpoint, the model effectively identifies students who are at risk of problematic academic domains, such as teamwork skills, academic self-concept, and class participation. The findings regarding sports participation, as derived from the quantitative data, suggest that future intervention strategies would benefit from the incorporation of more comprehensive assessment tools that encompass both academic and lifestyle factors.

For future research directions, methodological expansion is recommended in order to include qualitative instruments such as structured interviews and ethnographic observations. This mixed-methods approach would preserve the quantitative model's predictive strengths while providing a more profound understanding of contextual factors. Subsequent studies could develop new quantitative metrics informed by qualitative discoveries, particularly regarding extracurricular impacts. The statistical validity of the model must be subjected to rigorous evaluation through established quantitative measures, while incorporating qualitative dimensions.

This research establishes that quantitative models can effectively predict academic risk, while simultaneously highlighting how qualitative methods could provide essential explanatory power for complex behavioral phenomena. The extant findings provide a validated predictive framework and a clear methodological pathway for more comprehensive future studies that bridge quantitative and qualitative approaches.

The study ultimately demonstrates that while artificial intelligence tools provide powerful predictive capabilities, their educational application benefits from incorporating multiple methodological perspectives to fully address the complexities of student performance.

The findings regarding sports participation serve as a compelling example of how ostensibly quantitative data can reveal limitations in purely statistical interpretations, thereby naturally leading to qualitative methodological expansion. Future research endeavors should prioritize the development of integrated assessment frameworks that maintain the rigor of predictive analysis while achieving greater explanatory depth through methodological diversity.

References:

[1]. Díaz Barajas, D., & Ruiz Olvera, A. (2018). Reprobación escolar en el nivel medio superior y su relación con el autoconcepto en la adolescencia. *Revista Latinoamericana de Estudios Educativos, 48*(2), 125–142. Doi: 10.48102/rlee.2018.48.2.49

[2]. Medina Esparza, F., & Martínez Iñíguez, J. E. (2022). El fenómeno de la reprobación en la educación superior. *RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo, 13*(25).

[3]. Alzubaidi, L., et al. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data, 8*, 1-74.

[4]. Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN computer science, 2*(6), 1-20. Doi: 10.1007/s42979-021-00815-1

[5]. Tadese, M., Yeshaneh, A., & Mulu, G. B. (2022). Determinants of good academic performance among university students in Ethiopia: a cross-sectional study. *BMC medical education, 22*(1), 395. Doi: 10.1186/s12909-022-03461-0

[6]. Kassarnig, V., et al. (2018). Academic performance and behavioral patterns. *EPJ Data Science, 7*, 1-16. Doi: 10.1140/epjds/s13688-018-0138-8

[7]. Naciones Unidas. (2023). *Objetivo 9: Construir infraestructuras resilientes, promover la industrialización sostenible y fomentar la innovación.* Naciones Unidas. Retrieved from: https://www.un.org/sustainabledevelopment/es/infrastructure/ [accessed: 10 November 2024]

[8]. Porcelli, A. M. (2020). Artificial Intelligence and Robotics: Its social, ethical and legal dilemmas. *Derecho global. Estudios sobre derecho y justicia*, *6*(16), 49-105. Doi: 10.32870/dgedj.v6i16.286

[9]. Gutiérrez-Monsalve, J. A., Garzón, J., & Segura-Cardona, A. M. (2021). Factores asociados al rendimiento académico en estudiantes universitarios. *Formación universitaria*, *14*(1), 13-24. Doi: 10.4067/S0718-50062021000100013

[10]. Gil-Vera, V. D., & Quintero-López, C. (2021). Predicción del rendimiento académico estudiantil con redes neuronales artificiales. *Información tecnológica*, *32*(6), 221-228. Doi: 10.4067/s0718-07642021000600221

[11]. Sultana, J., Rani, M. U., & Farquad, M. A. H. (2019). Student's performance prediction using deep learning and data mining methods. *Int. J. Recent Technol. Eng*, *8*(1), 1018-1021.

[12]. Aslam, N., et al. (2021). An improved early student's academic performance prediction using deep learning. *International Journal of Emerging Technologies in Learning (iJET)*, *16*(12), 108-122. Doi: 10.3991/ijet.v16i12.20699

[13]. Li, S., & Liu, T. (2021). Performance prediction for higher education students using deep learning. *Complexity*, *2021*(1), 9958203. Doi: 10.1155/2021/9958203

[14]. Akour, M., Alsghaier, H., & Al Qasem, O. (2020). The effectiveness of using deep learning algorithms in predicting students achievements. *Indonesian Journal of Electrical Engineering and Computer Science*, *19*(1), 387-393.

[15]. Yauri, R. A., et al. (2024). A Machine Learning Approach in Predicting Student's Academic Performance Using Artificial Neural Network. *Journal of Computational and Cognitive Engineering*, *3*(2), 203-212. Doi: 10.47852/bonviewJCCE3202470

[16]. Talaei Khoei, T., Ould Slimane, H., & Kaabouch, N. (2023). Deep learning: systematic review, models, challenges, and research directions. *Neural Computing and Applications*, *35*(31), 23103-23124. Doi: 10.1007/s00521-023-08957-4

[17]. Lee, M. (2023). Recent advances in generative adversarial networks for gene expression data: a comprehensive review. *Mathematics*, *11*(14), 3055. Doi: 10.3390/math11143055

[18]. Nemade, B., et al. (2023). A comprehensive review: SMOTE-based oversampling methods for imbalanced classification techniques, evaluation, and result comparisons. *International Journal of Intelligent Systems and Applications in Engineering*, *11*(9s), 790-803.

[19]. Laura-Ochoa, L. (2019). Evaluation of classification algorithms using cross validation. *Proceedings of the LACCEI International Multi-Conference for Engineering, Education and Technology*. Latin American and Caribbean Consortium of Engineering Institutions. Doi: 10.18687/LACCEI2019.1.1.471

[20]. Nowak, S. (2022). *Modelos predictivos: ¿Qué son y para qué se usan?* Nuclio Digital School. Retrieved from: https://nuclio.school/blog/modelos-predictivos-que-son-y-usos/#Que-son-los-modelos-predictivos [accessed: 12 November 2024]

[21]. Rodríguez Jiménez, F. J., Pérez-Ochoa, M. E., & Ulloa-Guerra, Ó. (2021). Aula invertida y su impacto en el rendimiento académico: Una revisión sistematizada del período 2015-2020. *EDMETIC. Revista de Educación Mediática y TIC*, *10*(2), 1–25. Doi: 10.21071/edmetic.v10i2.13240

[22]. Google. (n.d.). *Google Colaboratory*. Google. Retrieved from: https://colab.research.google.com/ [accessed: 12 November 2024]

[23]. Abadi, M., et al. (2016). {TensorFlow}: a system for {Large-Scale} machine learning. *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, 265-283.

[24]. UNESCO. (n.d.). *Education for sustainable development*. UNESCO. Retrieved from: https://www.unesco.org/en/education/sustainable-development [accessed: 15 November 2024]

[25]. Castro Maldonado, J. J., Gómez Macho, L. K., & Camargo Casallas, E. (2023). Applied research and experimental development in strengthening the competences of the 21st century society. *Tecnura*, *27*(75), 140-174. Doi: 10.14483/22487638.19171

[26]. Hernández, R., Fernández, C., & Baptista, P. (2014). Definiciones de los enfoques cuantitativo y cualitativo, sus similitudes y diferencias. *Metodología de la investigación*, *6*, 2-21.

[27]. Galarza, C. A. R. (2021). Diseños de investigación experimental. *CienciAmérica: Revista de divulgación científica de la Universidad Tecnológica Indoamérica*, *10*(1), 1-7. Doi: 10.33210/ca.v10i1.356

[28]. Anguita, J. C., Labrador, J. R., & Campos, J. D. (2013). La encuesta como técnica de investigación. Elaboración de cuestionarios y tratamiento estadístico de los datos (I). *Atención primaria*, *31*(8), 527.

[29]. López, D. M. O., & Gómez, M. C. S. (2006). Técnicas de recolección de datos en entornos virtuales más usadas en la investigación cualitativa. *Revista de investigación educativa*, *24*(1), 205-222.

[30]. Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

[31]. Vijayalakshmi, V., & Venkatachalapathy, K. (2019). Comparison of predicting student's performance using machine learning algorithms. *International Journal of Intelligent Systems and Applications*, *10*(12), 34. Doi: 10.5815/ijisa.2019.12.04

[32]. Colegio de Ingenieros del Perú. (2018). *Código de ética del Colegio de Ingenieros del Perú*. CIP. Retrieved from: https://www.cip.org.pe/publicaciones/reglamentosCNCD2018/codigo_de_etica_del_cip.pdf [accessed: 20 November 2024]

[33]. Universidad César Vallejo. (2022). *Código de Ética en Investigación de la Universidad César Vallejo* (Versión 01, Código PP-DG-02.01). [Resolución Consejo Universitario RCUN °0470-2022-UCV].

[34]. Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica – Concytec. (2024). *Código Nacional de Integridad Científica*. Aprobado mediante Resolución de Presidencia N° 028-2024-CONCYTEC-P. Retrieved from: https://portal.concytec.gob.pe/images/publicaciones/Codigo-integridad-cientifica.pdf [accessed: 02 December 2024]

[35]. Coppola, S., Matrisciano, C., & Vastola, R. (2024). Exploring the relationship between physical activity and cognitive function in children. *Journal of Physical Education and Sport*, *24*(5), 1266-1274. Doi: 10.7752/jpes.2024.05144