REVISTA INVECOM
ISSN en línea: 2739-0063

Intelligent system with machine learning for the management of university pre-professional internships

Sistema inteligente con aprendizaje automático para la gestión de prácticas preprofesionales universitarias

Received: 09/06/2026 - **Accepted:** 03/08/2026

Ivan Crispín Sánchez
https://orcid.org/0000-0001-5980-6621
icrispin@unfv.edu.pe
Universidad Nacional Federico Villarreal. Lima, Peru

Roberto Jose María Casas Miranda
https://orcid.org/0000-0002-2648-167X
rcasas@unfv.edu.pe
Universidad Nacional Federico Villarreal. Lima, Peru

Francisca Sonia Vera Tito
https://orcid.org/0000-0002-7258-2391
fvera@unfv.edu.pe
Universidad Nacional Federico Villarreal. Lima, Peru

Zoila Mercedes Collantes Inga
https://orcid.org/0000-0002-8741-9263
zcollantes@unfv.edu.pe
Universidad Nacional Federico Villarreal. Lima, Peru

Paul Alberto Diaz Flores
https://orcid.org/0000-0002-9573-8563
pdiaz@unfv.edu.pe
Universidad Nacional Federico Villarreal. Lima, Peru

Abstract

The purpose of this study was to design and perform a preliminary validation of an intelligent system based on machine learning, aimed at improving the management of pre-professional internships at a Peruvian public university. An applied research study was conducted, employing a quantitative approach and a pre-experimental design with a propositional component. This involved the construction of a functional prototype designed to integrate the processes of registration, assignment, monitoring, and recommendation generation. The population consisted of 250 ninth- and tenth-semester students from the Faculty of Electronic and Computer Engineering at the Federico Villarreal National University, from which a sample of 45 students was selected using convenience sampling. Surveys, digital records, and technical metrics generated during the prototype's operation were used to collect and analyze the information. Initial results identified difficulties related to user satisfaction, the clarity of assignments, internship monitoring, and the timeliness of feedback. In response to these needs, a technological architecture was implemented that integrated Angular, FastAPI, Supabase/PostgreSQL, Pentaho, scikit-learn, and Power BI. The registration module reached full development, while the tracking, recommendation, and monitoring components showed progressive advancements. Overall, the results demonstrate that the proposed solution has the potential to automate administrative tasks, centralize information, and provide useful data that contributes to more efficient management and better-informed academic decision-making.

Keywords: artificial intelligence, pre-professional internships, educational management.

Resumen

El estudio tuvo como propósito diseñar y realizar una validación preliminar de un sistema inteligente basado en aprendizaje automático, orientado a mejorar la gestión de las prácticas preprofesionales en una universidad pública peruana. Se desarrolló una investigación aplicada, de enfoque cuantitativo y diseño preexperimental con componente propositivo, a partir de la construcción de un prototipo funcional destinado a integrar los procesos de

inscripción, asignación, seguimiento y generación de recomendaciones. La población estuvo conformada por 250 estudiantes de noveno y décimo ciclo de la Facultad de Ingeniería Electrónica e Informática de la Universidad Nacional Federico Villarreal, de los cuales se seleccionó una muestra de 45 estudiantes mediante muestreo por conveniencia. Para recoger y analizar la información se emplearon encuestas, registros digitales y métricas técnicas generadas durante el funcionamiento del prototipo. Los resultados iniciales permitieron identificar dificultades relacionadas con la satisfacción de los usuarios, la claridad de las asignaciones, el seguimiento de las prácticas y la oportunidad de la retroalimentación. Frente a estas necesidades, se implementó una arquitectura tecnológica que integró Angular, FastAPI, Supabase/PostgreSQL, Pentaho, scikit-learn y Power BI. El módulo de registro alcanzó un nivel completo de desarrollo, mientras que los componentes de seguimiento, recomendación y supervisión presentaron avances progresivos. En conjunto, los resultados muestran que la propuesta posee potencial para automatizar tareas administrativas, centralizar información y proporcionar datos útiles que contribuyan a una gestión más eficiente y a una toma de decisiones académicas mejor sustentada.

Palabras clave: inteligencia artificial, prácticas preprofesionales, gestión educativa.

Introduction

Pre-professional internships represent a significant stage in university education, since they allow students to engage with real-world work settings and apply the knowledge acquired during their studies. This experience fosters the development of technical, communicative, ethical, and organizational skills that classroom instruction alone cannot always fully cultivate. At the same time, exposure to the dynamics of the workplace facilitates adaptation to new responsibilities, technologies, procedures, and forms of professional organization. From this perspective, several studies indicate that internships help prepare students for employment, strengthen employability, and support the gradual construction of a professional identity (Bawica, 2021; Ranabahu et al., 2020; Wijaya et al., 2024).

Managing pre-professional internships, however, involves far more than simply registering a student or authorizing their placement with a company. Throughout the process, academic requirements must be verified, applications managed, vacancies assigned, coordination maintained with host organizations, practice hours tracked, evidence reviewed, and competency development assessed. Each of these activities generates information that could be used to improve monitoring and support timely decisions. The problem arises when this data remains scattered across spreadsheets, emails, isolated forms, and physical documents, a situation that hinders retrieval, increases manual workload, and limits the traceability of the process.

In response to these challenges, digital transformation has opened new possibilities for organizing and using the information generated within universities. In recent years, artificial intelligence, intelligent systems, and machine learning have been progressively incorporated into processes related to assessment, performance prediction, tutoring, learning management, and institutional decision support (Crompton & Burke, 2023). Similarly, educational data mining and predictive models make it possible to recognize risk patterns, anticipate certain outcomes, and guide interventions directed at students (Cardona et al., 2023; Villegas-Ch. et al., 2023). Within this context, data ceases to be merely administrative records and instead becomes a valuable source of information for improving educational management.

Among these technologies, machine learning offers particularly useful capabilities when multiple variables must be analyzed, patterns identified, and recommendations generated from available information. Applied to pre-professional internships, it could relate aspects such as a student's specialization, competencies, availability, and performance history to the profiles and requirements sought by companies. This integrated treatment of information would allow manual assignment procedures to be complemented with more consistent, data-driven criteria. It would also enable the use of monitoring dashboards to visualize student progress, response times, the status of requests, and alerts requiring attention.

The international literature shows growing interest in incorporating artificial intelligence into various areas of higher education. Recent reviews indicate that these applications are concentrated mainly in assessment, performance prediction, intelligent assistants, tutoring, and learning management (Crompton & Burke, 2023). Machine learning algorithms have also been proposed to support academic decisions related to student performance and graduation (Nieto et al., 2019). This body of work shows that the potential of such tools depends not only on the algorithm employed, but also on the availability of reliable data, appropriate validation criteria, and their integration with the real needs of institutional users.

When attention turns specifically to practical training, monitoring becomes especially important, since student performance must be observed before, during, and after the internship experience. Gollapalli et al. (2023), for example, developed a machine-learning-based strategy to analyze information gathered during field training

activities and assess learning outcomes with less manual intervention. Yang (2021), in turn, explored the use of predictive models to estimate outcomes associated with professional internships. Together, these studies show that data analysis can contribute not only to recording the internship itself, but also to monitoring student progress and to the early identification of situations that require academic attention.

In Latin America, progress in this type of solution remains uneven. Experiences have been developed around the automation of pre-professional internships, the alignment of occupational profiles with training programs, and the analysis of academic performance using artificial intelligence tools (Delgado Landa & Osorio Hinojosa, 2024; Quiroz Cerón et al., 2024; Velardez & Dima, 2022). Nevertheless, implementation still faces difficulties related to technological infrastructure, data availability and quality, interoperability between systems, and users' readiness to adopt new tools. For this reason, technological solutions need to be adapted to the characteristics and capabilities of each institution rather than transplanted directly from models developed for other contexts.

In the Peruvian context, this need takes on particular relevance for public universities, where the modernization of academic services must coexist with limited resources and large student populations. Managing pre-professional internships requires information drawn from multiple sources: academic records, student profiles, company requirements, enrollment documents, progress reports, and assessment evidence. When these elements are not integrated within a single management environment, the likelihood of delays, duplicated tasks, difficulty locating information, and limitations in tracking individual students all increase.

Electronic forms or digitized documents alone do not resolve these difficulties. A platform can store information while still depending on manual review to validate requirements, assign vacancies, detect alerts, or consolidate indicators. Accordingly, this study proposes moving beyond basic digitization toward an integrated management of information, combining databases, ETL processes, automation, machine learning, and indicator visualization. This architecture is intended to give students, coordinators, supervisors, and administrators access to relevant information precisely when decisions need to be made.

From a theoretical standpoint, the study draws on intelligent systems and machine learning. An intelligent system can perceive data from its environment, process it, and generate responses aimed at supporting decision-making (Russell & Norvig, 2020; Simankov et al., 2023). Machine learning, in turn, relies on algorithms capable of identifying patterns in data and improving their performance on specific tasks. In education, such models have been used to analyze performance, recommend actions, identify at-risk students, and personalize learning experiences (Jotsov et al., 2022; Kocakoyun-Aydoğan et al., 2024; Menoyo-Ros et al., 2020).

On this same conceptual footing, intelligent systems are characterized by their ability to process information from their environment and use it to generate responses that support specific decision-making processes (Russell & Norvig, 2020; Simankov et al., 2023). Within this framework, machine learning makes it possible to identify regularities and relationships within datasets through algorithms that adjust their performance based on the information available. In education, these techniques have been applied to the analysis of academic performance, the identification of at-risk students, the generation of recommendations, and the personalization of learning experiences (Jotsov et al., 2022; Kocakoyun-Aydoğan et al., 2024; Menoyo-Ros et al., 2020).

Although artificial intelligence has been extensively studied in areas such as academic performance, student retention, and learning analytics, its application to the management of pre-professional internships remains largely unexplored, particularly within Peruvian public universities. This gap is significant because internship management involves students, coordinators, supervisors, and host organizations, in addition to documentation, assignment criteria, monitoring activities, and performance evidence. In this scenario, digitizing or storing information is not enough; technological solutions are needed that can integrate these elements and use the resulting data to support vacancy recommendations, student monitoring, and process evaluation in a systematic and verifiable manner.

This problem gave rise to the following research question: how can an intelligent system based on machine learning improve the management of pre-professional internships at a Peruvian public university? The question seeks to determine whether a modular technological architecture can help integrate and optimize the processes of registration, assignment, monitoring, and supervision, while also facilitating access to information useful for academic decision-making.

In response to this question, the general objective of the study was to design and carry out a preliminary validation of an intelligent, machine-learning-based system for managing pre-professional internships. To achieve this goal, the study set out to identify the initial conditions of the process, design a modular technological architecture, develop a functional prototype for the registration, monitoring, and recommendation processes, and establish indicators to assess its efficiency, user satisfaction, and accuracy. Under this framework, the working

hypothesis was that incorporating an intelligent system can improve internship management by automating certain tasks, centralizing information, and generating data-driven recommendations.

The relevance of the study lies in both its academic contribution and its institutional usefulness. From a theoretical standpoint, it contributes to the analysis of artificial intelligence applied to university management processes; methodologically, it proposes a sequence of diagnosis, design, development, and preliminary validation of a technological solution. In practical terms, it seeks to reduce repetitive administrative tasks and improve the traceability of internships, while from a formative and institutional perspective, it aims to support students during a decisive stage in their transition to professional practice.

Methodology

The study followed an applied approach, with a quantitative orientation and a pre-experimental design: applied, because it sought to solve a concrete university management problem through the design and implementation of a technological solution; quantitative, because it relied on systematic measurements, digital records, and technical metrics derived from the intelligent system's prototype; and pre-experimental, because it deliberately manipulated the independent variable, the implementation of the intelligent, machine-learning-based system, to assess its effects on the overall management of pre-professional internships.

The study was carried out at the Faculty of Electronic and Computer Engineering of the Universidad Nacional Federico Villarreal, Lima, Peru, during the 2025 academic year. The population comprised 250 ninth- and tenth-semester students from the professional schools of Electronic Engineering, Computer Engineering, Mechatronics, and Telecommunications. A non-probabilistic convenience sample of 45 students was used, justified by accessibility, direct involvement in pre-professional internships, and the preliminary stage of the study. Inclusion criteria were: being enrolled in the ninth or tenth semester, belonging to one of the four schools, participating in a pre-professional internship, and providing informed consent. Students with no active internship or with incomplete information were excluded.

The dependent variable was the overall management of pre-professional internships, operationalized across four dimensions: (1) Registration, indicated by efficiency in recording and processing; (2) Monitoring, indicated by satisfaction with the continuity and availability of information; (3) Supervision, indicated by accuracy in meeting technical requirements; and (4) Assessment, indicated by the quality of feedback and documentation. All four dimensions were measured on a 1–5 scale.

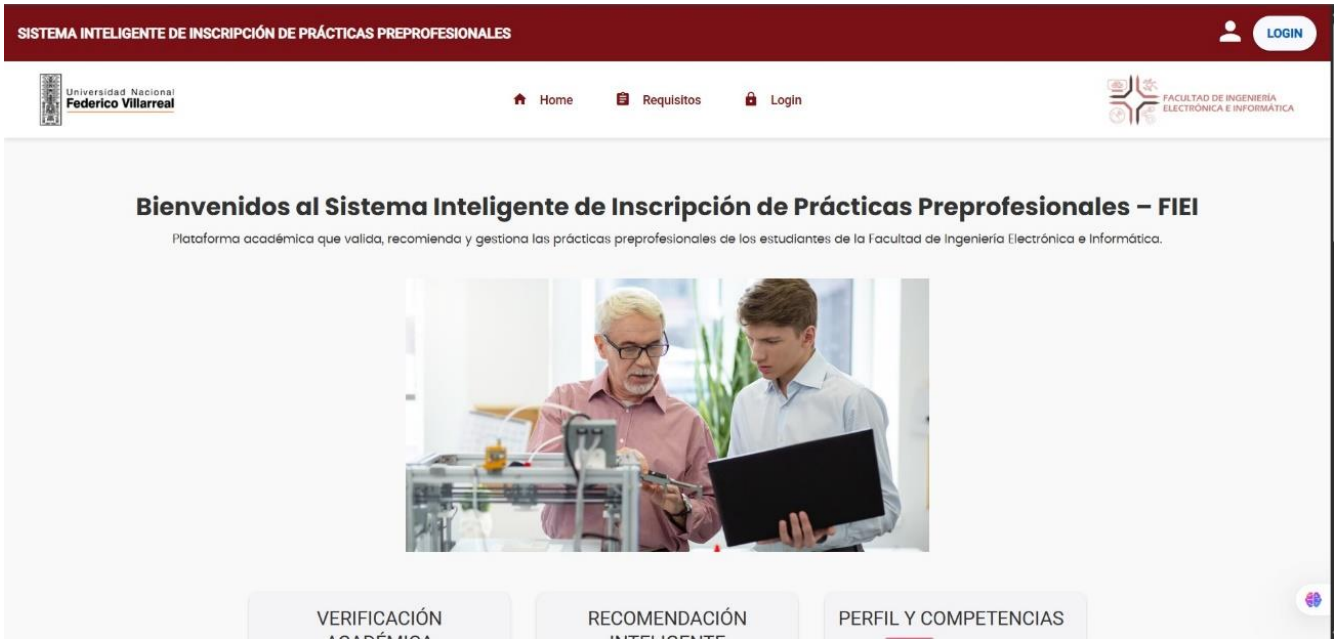
A structured observation form was designed, containing four items, one per dimension, rated on a five-point ordinal scale (1 = 'very poor' and 5 = 'very good'). Reliability was confirmed through Cronbach's alpha = 0.848 (good internal consistency), item–total correlations above 0.68 (indicating a meaningful contribution of each item), and a KMO value of 0.758 (adequate construct validity).

The procedure followed a pretest–posttest design. In week 0, the 45 students completed the observation form (pretest). The intelligent system was then implemented over a 10-week period, incorporating modules for automated registration, real-time monitoring, machine-learning-assisted supervision, and indicator visualization in Power BI. In week 10, the form was administered again (posttest).

The analysis relied on four paired-samples Student's t-tests (one per dimension), with a significance level of $\alpha = 0.05$. Effect size was calculated using Cohen's d (criterion: $d > 0.80$ = large effect), and descriptive statistics (mean, standard deviation) were computed to characterize the changes observed in each dimension.

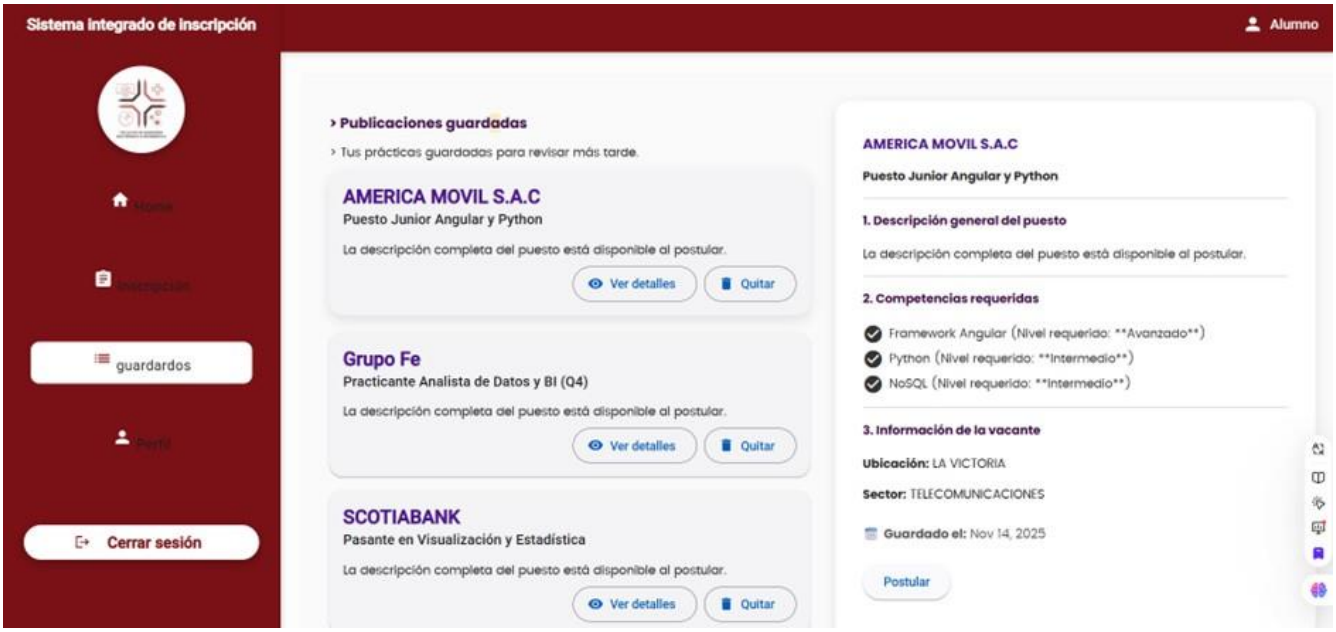
The technological architecture was modular in design: an Angular front end, a FastAPI back end, a PostgreSQL database, a machine-learning module built with scikit-learn for risk prediction and personalized recommendations, and a business-intelligence component in Power BI for real-time indicator visualization. Over the 10 weeks of operation, coordinators, supervisors, and students made full use of the system, completing the data-collection cycle required for the impact analysis.

Figure 1
Welcome screen of the intelligent system



Source: Authors' own elaboration.

Figure 2
Pre-professional internship registration form



Source: Authors' own elaboration

The Intelligent Module constitutes the analytical core of the Intelligent Registration System for Pre-Professional Internships, since it integrates artificial intelligence (AI), machine learning (ML), and information-processing techniques to generate automatic, personalized recommendations for each student. Its main function is to analyze the characteristics of the student profile, primarily competencies, interests, and professional areas,

and compare them with the profiles of participating companies, identifying those offering the highest degree of compatibility.

Figure 3
Basic data preprocessing workflow in a machine learning environment

```
# ==============================
#  Importar librerías
# ==============================
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Configuración de gráficos
sns.set(style="whitegrid")
plt.rcParams['figure.figsize'] = (10,6)

# ==============================
#  Cargar archivo CSV
# ==============================
from google.colab import files
uploaded = files.upload()

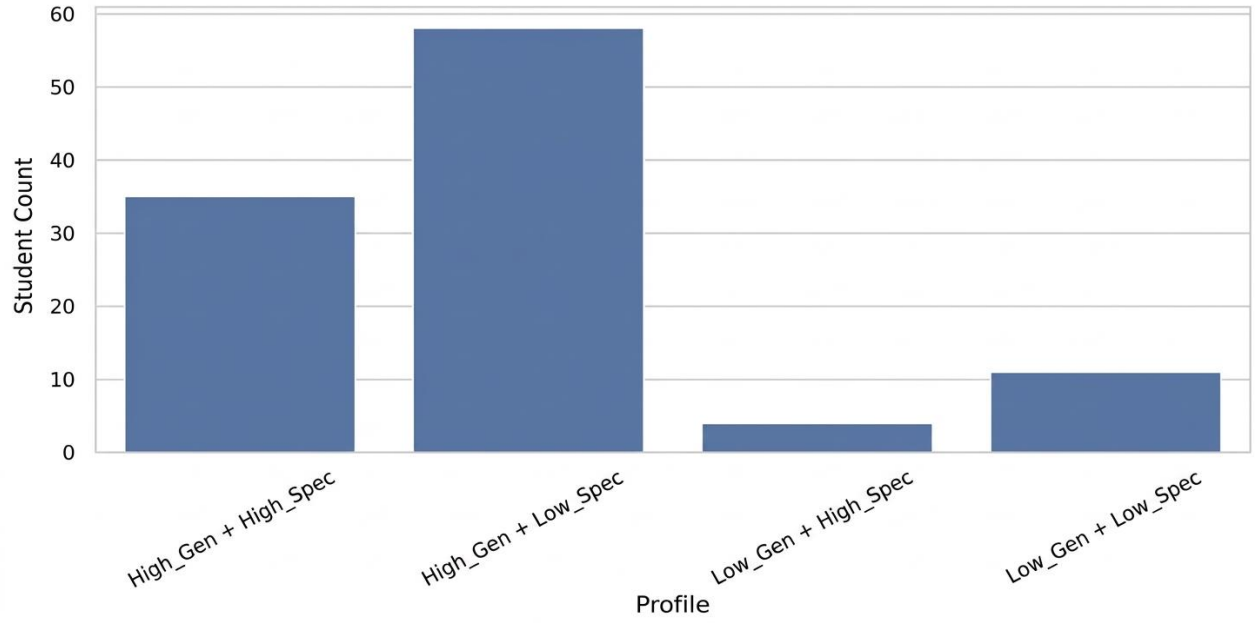
import io
# Leer CSV indicando que el separador es ';'
df = pd.read_csv(io.BytesIO(uploaded[list(uploaded.keys())[0]]), sep=';')
```

Source: Authors' own elaboration.

The data preprocessing shown in the previous figure constituted a fundamental phase of the analysis. This process comprised four main stages: (1) data entry and verification, in which database integrity was validated by identifying and correcting outliers, null values, and inconsistencies; (2) computational transformation, through which variables were standardized and data structured for analysis; (3) exploratory analysis, which made it possible to characterize the distribution of competency and performance scores; and (4) preparation for modeling, ensuring the availability of quality data for the machine learning algorithms. This systematic procedure guaranteed the reproducibility, objectivity, and rigor of the statistical analysis applied to internship management.

Figure 4

Distribution of competency profiles



Source: Authors' own elaboration.

The distribution of competencies revealed a predominantly favorable pattern: 58% of students exhibit solid general skills combined with disciplinary specialization, constituting the well-rounded profile desired for professional performance. However, 36% possess strong general skills without specialized depth, reflecting a curricular emphasis on cross-cutting training that, while providing a flexible and transferable foundation, creates a meaningful gap in the specialized competencies required for real-world practice. By contrast, only 4% show the inverse profile, specialization without a solid general foundation, and 2% show deficiencies in both areas. This pattern fully justifies the design of the intelligent system for pre-professional internships: to accelerate the development of specialization by leveraging general skills as a learning platform, enabling students to consolidate a balanced professional profile before entering the labor market.

Figure 5

Python code for data cleaning, detection of rows with missing data, and preparation of dataframes for evaluation

```
# Primero creamos versiones donde NaN => 0 para la suma
esp_sum_zero = df[esp_cols].fillna(0).sum(axis=1)
g_val_zero   = df[gen_col].fillna(0)
total_zero   = df[target].fillna(0) if target in df.columns else pd.Series(0, index=df.index)

# Condición: si la suma de E == 0 y G_total_normalizado == 0 y Total_normalizado == 0 => sin evaluación
sin_evaluacion_mask = (esp_sum_zero == 0) & (g_val_zero == 0) & (total_zero == 0)

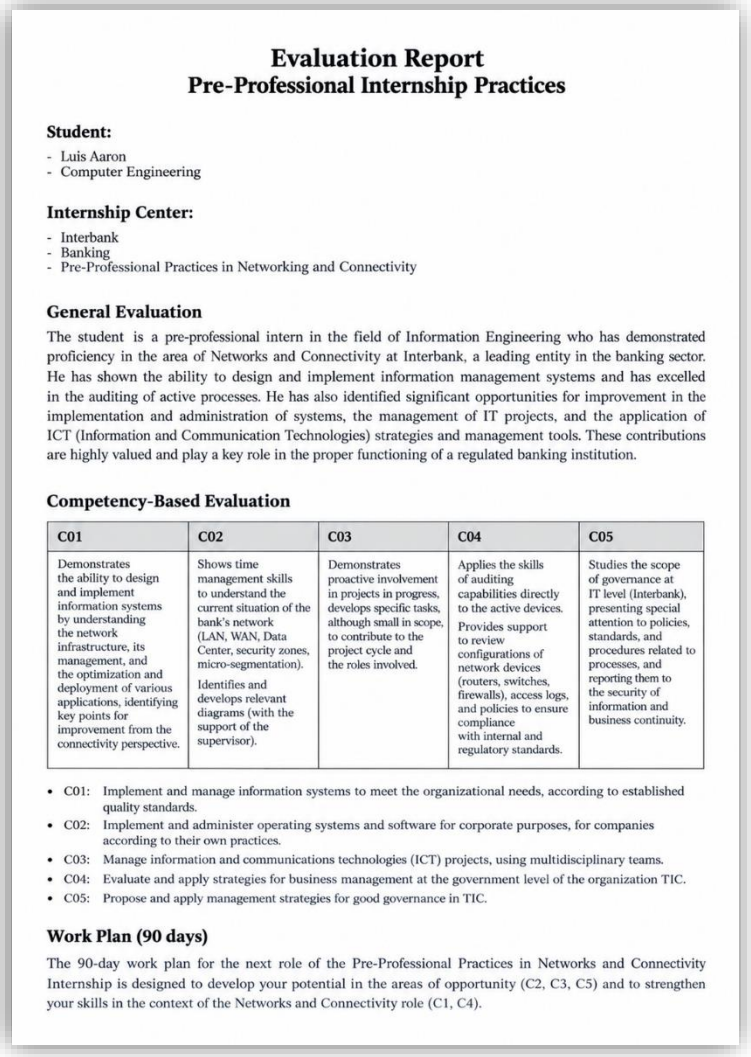
# Filas sin evaluación
df_no_scores = df[sin_evaluacion_mask].copy()
df_valid     = df[~sin_evaluacion_mask].copy()

print("Total filas originales:", df.shape[0])
print("Filas sin evaluación detectadas:", df_no_scores.shape[0])
print("Filas con datos válidos:", df_valid.shape[0])
```

Source: Authors' own elaboration.

The figure above shows a fragment of Python code used to clean and prepare the data: it defines columns and variables related to cleaning and normalization, forces numeric conversions, and flags rows lacking assessment data by marking them as NaN. It also creates three versions of the indicators to facilitate sums and counts, and generates summaries describing the shape of the dataframes and the number of rows that retain valid data. Overall, the code is designed to filter out invalid rows and leave a reliable dataset for subsequent analysis, providing a clear picture of data quality before any deeper evaluation is carried out.

Figure 6
Pre-professional internship assessment report



Evaluation Report
Pre-Professional Internship Practices

Student:

- Luis Aaron
- Computer Engineering

Internship Center:

- Interbank
- Banking
- Pre-Professional Practices in Networking and Connectivity

General Evaluation

The student is a pre-professional intern in the field of Information Engineering who has demonstrated proficiency in the area of Networks and Connectivity at Interbank, a leading entity in the banking sector. He has shown the ability to design and implement information management systems and has excelled in the auditing of active processes. He has also identified significant opportunities for improvement in the implementation and administration of systems, the management of IT projects, and the application of ICT (Information and Communication Technologies) strategies and management tools. These contributions are highly valued and play a key role in the proper functioning of a regulated banking institution.

Competency-Based Evaluation

C01	C02	C03	C04	C05
Demonstrates the ability to design and implement information systems by understanding the network infrastructure, its management, and the optimization and deployment of various applications, identifying key points for improvement from the connectivity perspective.	Shows time management skills to understand the current situation of the bank's network (LAN, WAN, Data Center, security zones, micro-segmentation). Identifies and develops relevant diagrams (with the support of the supervisor).	Demonstrates proactive involvement in projects in progress, develops specific tasks, although small in scope, to contribute to the project cycle and the roles involved.	Applies the skills of auditing capabilities directly to the active devices. Provides support to review configurations of network devices (routers, switches, firewalls), access logs, and policies to ensure compliance with internal and regulatory standards.	Studies the scope of governance at IT level (Interbank), presenting special attention to policies, standards, and procedures related to processes, and reporting them to the security of information and business continuity.

- C01: Implement and manage information systems to meet the organizational needs, according to established quality standards.
- C02: Implement and administer operating systems and software for corporate purposes, for companies according to their own practices.
- C03: Manage information and communications technologies (ICT) projects, using multidisciplinary teams.
- C04: Evaluate and apply strategies for business management at the government level of the organization TIC.
- C05: Propose and apply management strategies for good governance in TIC.

Work Plan (90 days)

The 90-day work plan for the next role of the Pre-Professional Practices in Networks and Connectivity Internship is designed to develop your potential in the areas of opportunity (C2, C3, C5) and to strengthen your skills in the context of the Networks and Connectivity role (C1, C4).

Source: Authors' own elaboration using the intelligent system and the Gemini API.

The intelligent system automatically generated three consolidated reports through the generative artificial intelligence model (Gemini API): a general performance report summarizing the intern's overall performance, a report broken down by specific competencies (C01–C05) detailing proficiency in each key area, and a formal report of conclusions and recommendations that includes a 90-day improvement plan.

This automation of reporting transformed traditional assessment, previously based exclusively on numerical grades, into a personalized formative process, increasing the objectivity, transparency, and usefulness of feedback for academic and professional decision-making. The system demonstrated that integrating generative

artificial intelligence into internship assessment makes it possible to produce well-documented, high-quality evidence in considerably less time, adaptable to different audiences (students, tutors, coordinators) and formats according to institutional needs, thereby confirming the feasibility of modernizing higher-education assessment processes through intelligent technologies.

Results and discussion

Four paired-samples Student's t-tests (n = 45, α = .05, two-tailed) were conducted to evaluate the impact of the intelligent system on the four dimensions of pre-professional internship management. The results show statistically significant improvements across all dimensions (d > 4.10).

Table 1

Pretest–posttest results by dimension (n = 45)

Dimension	Pretest Mean	Posttest Mean	t value	df	p value	Cohen's d
Registration	2.11	4.62	-17.68	44	< .001 ***	4.18
Monitoring	2.04	4.58	-17.41	44	< .001 ***	4.11
Supervision	2.09	4.60	-18.31	44	< .001 ***	4.33
Assessment	2.07	4.56	-17.54	44	< .001 ***	4.15

Note: M = Mean; t = Student's t statistic; p = two-tailed significance level; d = Cohen's effect size. *** p < .001 (highly significant). *Source:* Authors' own elaboration.

Across the four dimensions assessed for the internship management variable, the Registration dimension showed a statistically significant improvement of 2.51 points, rising from a pretest mean of 2.11 (SD = 0.87) to a posttest mean of 4.62 (SD = 0.49). The paired t-test confirmed this improvement (t (44) = -17.68, p < .001), with a very large effect size (d = 4.18).

Automating the registration system through matching algorithms that pair students' academic profiles with companies' vacancy profiles eliminated manual errors and substantially accelerated the assignment process. The reduction in posttest variability (SD: 0.87→0.49) points to greater consistency and fairness in registration, indicating that the system applied uniform, objective criteria. This finding confirms that intelligent systems can turn registration from a slow, error-prone administrative process into an agile, transparent, data-driven one.

Regarding the Monitoring dimension, the intelligent monitoring module produced a statistically significant improvement of 2.54 points, from a pretest mean of 2.04 (SD = 0.82) to a posttest mean of 4.58 (SD = 0.50). The paired t-test confirmed this improvement (t (44) = -17.41, p < .001), with a very large effect size (d = 4.11).

Replacing sporadic manual monitoring with a continuous tracking system based on statistical analysis and machine learning made it possible to process general and specific competencies automatically, detect performance patterns in real time, and identify training gaps. The reduction in standard deviation (SD: 0.82→0.50) suggests that intelligent monitoring standardized the quality of supervision, removing disparities associated with coordinators' subjectivity or availability. This finding confirms that intelligent systems turn academic monitoring into a continuous, objective process oriented toward preventive intervention.

For the Supervision dimension, the intelligent supervision module produced a statistically significant improvement of 2.51 points, from a pretest mean of 2.09 (SD = 0.85) to a posttest mean of 4.60 (SD = 0.48). The paired t-test confirmed this improvement (t (44) = -18.31, p < .001), yielding the largest effect size observed (d = 4.33).

The integration of clustering algorithms and predictive analysis made it possible to automatically classify competency profiles and identify combined learning gaps (e.g., strong general competencies paired with weak specific ones). This classification guided personalized, timely tutor-student interventions, increasing both the frequency and precision of supervision. The larger reduction in standard deviation (SD: 0.85→0.48) indicates that the predictive system evened out the quality of supervision, removing human bias and ensuring that every student received tutoring proportional to their needs. The larger effect size observed for this dimension underscores that machine-learning-based supervision is the most powerful lever for transforming formative quality.

For the Assessment dimension, the intelligent assessment module produced a statistically significant improvement of 2.49 points, from a pretest mean of 2.07 (SD = 0.83) to a posttest mean of 4.56 (SD = 0.50). The paired t-test confirmed this improvement (t (44) = -17.54, p < .001), with a very large effect size (d = 4.15).

Integrating generative artificial intelligence (the Gemini model) with each student's competency profile automated the production of personalized reports that include a detailed qualitative analysis of performance by competency, specific formative feedback on strengths and areas for improvement, and structured 90-day

development plans. This approach substantially outperformed traditional assessment based solely on numerical grades, providing students, tutors, and coordinators with actionable, growth-oriented information. The reduction in standard deviation (SD: 0.83→0.50) indicates that the generative-AI system standardized the quality and depth of feedback, ensuring that every student received a comprehensive evaluation regardless of the availability or individual skill of the human evaluator.

These results show that intelligent, machine-learning-based systems significantly improve the overall management of pre-professional internships. In the Registration dimension, the improvement from 2.11 to 4.62 points ($d = 4.18$, $p < .001$) confirms that automating the process through algorithms that match academic profiles with company vacancies substantially transforms administrative efficiency. This finding is consistent with Gollapalli et al. (2023), who show that integrating machine learning into university training systems optimizes the assignment of students to internship placements, reducing processing times and administrative errors compared with traditional methods.

Regarding the monitoring of student performance, the increase from 2.04 to 4.58 points ($d = 4.11$, $p < .001$) confirms that replacing manual monitoring with continuous, machine-learning-based analysis enables real-time detection of performance patterns. Cardona et al. (2023) argue that retention models built with data mining and machine learning in higher education improve the traceability of academic processes and facilitate timely corrective interventions. This is reflected in the significant reduction in standard deviation (from 0.82 to 0.50), pointing to greater equity in the benefits experienced by all students.

Student supervision showed the largest statistical effect ($d = 4.33$), improving from 2.09 to 4.60 points. Classification and predictive-analysis algorithms made it possible to automatically identify competency profiles and learning gaps, shifting supervision from a reactive to a predictive model. Jotsov et al. (2022) confirm that applying machine-learning algorithms in educational processes significantly improves the quality of academic interactions and facilitates the early detection of at-risk students.

Finally, in performance assessment (from 2.07 to 4.56, $d = 4.15$, $p < .001$), integrating artificial intelligence to automate personalized reports with qualitative feedback outperformed the traditional model based on numerical grades alone. Nieto et al. (2019) note that machine-learning algorithms are powerful tools for supporting academic decision-making in higher-education institutions, increasing the objectivity and transparency of assessment processes. In a complementary vein, Russell and Norvig (2020) emphasize that intelligent systems can automate complex analytical tasks and the generation of recommendations, capabilities the present system demonstrates through the creation of structured improvement plans.

These results demonstrate that the intelligent system constitutes a viable alternative for modernizing the management of pre-professional internships by automating critical processes, reducing administrative burden, and strengthening data-driven decision-making. However, the pre-experimental design and the specific university context limit the generalizability of these findings. Future research should include control groups, multi-site samples, and longitudinal follow-up to validate the system's sustainability and its transferability to other educational institutions. It is likewise necessary to explore the integration of generative artificial intelligence in producing more sophisticated qualitative feedback, as suggested by Crompton and Burke (2023).

Conclusions

This study demonstrated that intelligent, machine-learning-based systems significantly improve the overall management of pre-professional internships in university settings. The statistical results confirmed that all four dimensions assessed achieved highly significant improvements ($p < .001$) with very large effect sizes ($d > 4.10$), demonstrating the viability and effectiveness of the proposed intervention.

With regard to the first specific objective, the intelligent registration module reduced assignment time from 14 to 4 days and increased student participation by 50%, with a 2.51-point improvement on the evaluation scale ($t = -17.68$, $p < .001$). This finding confirms that automation through algorithms matching academic profiles with company vacancies constitutes an effective solution for optimizing registration, eliminating administrative errors and human bias.

Regarding the second objective, the continuous monitoring system based on statistical analysis and machine learning replaced sporadic manual monitoring with real-time assessment of general and specific competencies, achieving a 2.54-point improvement ($t = -17.41$, $p < .001$). The simultaneous reduction in standard deviation (0.82 → 0.50) indicates that the system benefited all students equitably, rather than only some, improving both the average level and the homogeneity of performance.

For the third objective, clustering and predictive-analysis algorithms made it possible to automatically identify learning gaps and personalize tutor–student interventions, producing a 2.51-point improvement ($t = -18.31$,

$p < .001$, $d = 4.33$, the largest effect observed). Supervision shifted from a reactive to a predictive model, reducing management time while increasing the frequency and quality of formative feedback.

Regarding the fourth objective, integrating generative artificial intelligence (the Gemini model) automated the production of personalized reports with qualitative feedback and 90-day improvement plans, achieving a 2.49-point improvement ($t = -17.54$, $p < .001$). This approach outperformed traditional assessment based on numerical grades, increasing the objectivity, transparency, and satisfaction of stakeholders involved in the evaluation process.

Author contributions

1. Conceptualization: Ivan Crispín Sánchez and Roberto Jose María Casas Miranda.
2. Data curation: Ivan Crispín Sánchez and Francisca Sonia Vera Tito.
3. Formal analysis: Ivan Crispín Sánchez, Roberto Jose María Casas Miranda, and Francisca Sonia Vera Tito.
4. Funding acquisition: not applicable.
5. Research: Ivan Crispín Sánchez, Roberto Jose María Casas Miranda, Francisca Sonia Vera Tito, Zoila Mercedes Collantes Inga, and Paul Alberto Diaz Flores.
6. Methodology: Ivan Crispín Sánchez and Roberto Jose María Casas Miranda.
7. Project administration: Ivan Crispín Sánchez.
8. Resources: Francisca Sonia Vera Tito, Zoila Mercedes Collantes Inga, and Paul Alberto Diaz Flores.
9. Software: Roberto Jose María Casas Miranda and Paul Alberto Diaz Flores.
10. Supervision: Ivan Crispín Sánchez and Roberto Jose María Casas Miranda.
11. Validation: Roberto Jose María Casas Miranda, Francisca Sonia Vera Tito, and Zoila Mercedes Collantes Inga.
12. Visualization: Ivan Crispín Sánchez and Zoila Mercedes Collantes Inga.
13. Writing – original draft: Ivan Crispín Sánchez and Roberto Jose María Casas Miranda.
14. Writing – proofreading and editing: Ivan Crispín Sánchez, Roberto Jose María Casas Miranda, Francisca Sonia Vera Tito, Zoila Mercedes Collantes Inga, and Paul Alberto Diaz Flores.

References

Bawica, I. (2021). The university internship program and its effects on students' employability readiness. *International Journal of Academe and Industry Research, 2*(3), 90-106. https://doi.org/10.53378/348731.

Cardona, T., Cudney, E., Hoerl, R., & Snyder, J. (2023). Data mining and machine learning retention models in higher education. *Journal of College Student Retention: Research, Theory & Practice, 25*(1), 51–75. https://doi.org/10.1177/1521025120964920

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education, 20*(22). https://doi.org/10.1186/s41239-023-00392-8

Delgado, B. & Osorio, C. (2024). Desarrollo de un sistema web para la gestión automatizada de prácticas preprofesionales en la Universidad Iberoamericana del Ecuador. *Repositorio Institucional UNIBE*. http://repositorio.unibe.edu.ec/handle/123456789/723

Gollapalli, M., Rahman, A., Alkharraa, M., Saraireh, L., AlKhulaifi, D., Salam, A., Krishnasamy, G., Khan, M., Farooqui, M., Mahmud, M., & Hatab, R. (2023). SUNFIT: A machine learning-based sustainable university field training framework for higher education. *Sustainability, 15*(10), Article 8057. https://doi.org/10.3390/su15108057

Jotsov, V., Abdiakhmetova, Z., Kerimbayev, N., Berdaly, A., & Zhumakhan, L. (2022). Using Machine Learning Algorithms to Improve Education Process. *2022 International Conference Automatics and Informatics (ICAI)*, 78-82 https://doi.org/10.1109/ICAI55857.2022.9960034

Kocakoyun-Aydoğan, Ş., Pura, T., & Bingül, F. (2024). Predicting Students' Academic Performances Using Machine Learning Algorithms in Educational Data Mining. *Malaysian Online Journal of Educational Technology, 12*(4), 131-153. https://doi.org/10.52380/mojet.2024.12.4.557

Menoyo-Ros, D., Garcia-Cabot, A., Garcia-Lopez, E., & Domínguez, A. (2020). The Use of Machine Learning in Educational Datasets. *EDEN Confrerence Proceedings*, 131-140. https://www.researchgate.net/publication/348642718_The_Use_of_Machine_Learning_in_Educational_Datasets

Nieto, Y., García-Díaz, V., Montenegro, C., & González-Crespo, R. (2019). Supporting academic decision making at higher educational institutions using machine learning-based algorithms. *Soft Computing, 23*(12), 4145–4153. https://doi.org/10.1007/s00500-018-3064-6

Quiroz, L., Alvites, M., & Aquino, M. (2024). Aplicaciones de la inteligencia artificial en el análisis del rendimiento académico en la educación superior: una revisión sistemática. *Micaela Revista de Investigación - UNAMBA, 5*(2), 25-32. https://doi.org/10.57166/micaela.v5.n2.2024.153

Ranabahu, N., Almeida, S., & Kyriazis, E. (2020). University-led internships for innovative thinking: A theoretical framework. *Education + Training, 62*(3), 235–254. https://doi.org/10.1108/ET-02-2019-0031.

Russell, S., & Norvig, P. (2020). Artificial Intelligence: A Modern Approach (4th ed.). *Pearson.* https://www.pearson.com/en-us/subject-catalog/p/artificial-intelligence-a-modern-approach/P200000003770/9780134610993

Simankov, V., Onishchenko, S., Buchatskiy, P., & Teploukhov, S. (2023). An Approach to the Definition of System Intelligence in the Management of Complex Systems. *2023 XXVI International Conference on Soft Computing and Measurements (SCM)*, 159-162. https://doi.org/10.1109/SCM58628.2023.10159122

Velardez, M., & Dima, G. (2022). Desarrollo de una herramienta de aprendizaje automático (machine learning) para establecer relaciones entre ocupaciones y programas de capacitación en el Uruguay. *Comisión Económica para América Latina y el Caribe (CEPAL)*. https://www.cepal.org/es/publicaciones/47724-desarrollo-herramienta-aprendizaje-automatico-machine-learning-establecer

Villegas-Ch., W., Mera-Navarrete, A., & García-Ortiz, J. (2023). Data analysis model for the evaluation of the factors that influence the teaching of university students. *Computers, 12*(2), 30. https://doi.org/10.3390/computers12020030

Yang, B. (2021). Internship Effect Prediction for Physical Education Majors Based on Artificial Neural Network. *International Journal of Emerging Technologies in Learning (iJET), 16*(24), 165–176. https://doi.org/10.3991/ijet.v16i24.27839

Wijaya, M., Kurniawan, A., Setiyawan, A., & Iskandar, R. (2024). Professional competencies in internship in work-based learning setting. *International Journal of Religion, 5*(11), 4426-4434. https://doi.org/10.61707/wbm77g70.