



ISSN 1992-6510
e-ISSN 2520-9299



REALIDAD Y REFLEXIÓN ES UNA PUBLICACIÓN PERIÓDICA DE CARÁCTER SEMESTRAL DE LA UNIVERSIDAD FRANCISCO GAVIDIA
AÑO 25, n.º 62, JULIO-DICIEMBRE 2025. SAN SALVADOR, EL SALVADOR, CENTROAMÉRICA

REALITY AND REFLECTION IS A BIENNIAL PERIODICAL PUBLICATION OF THE FRANCISCO GAVIDIA UNIVERSITY
YEAR 25, n.º 62, JULY-DECEMBER 2025. SAN SALVADOR, EL SALVADOR, CENTRAL AMERICA

UIDE Classifier: inteligencia artificial para la clasificación automática de la retroalimentación estudiantil en la educación superior¹

UIDE Classifier: Artificial Intelligence for the Automatic Classification of Student Feedback in Higher Education

Juan Pablo Bohórquez Erazo

Licenciatura en Administración de Empresas Turísticas, Universidad de Especialidades Turísticas, Ecuador
Máster propio en Dirección de Marketing y Comunicación Política, TECH Universidad Tecnológica, México
Magíster en Inteligencia de Negocios y Comportamiento del Consumidor, Universidad Internacional del Ecuador, Ecuador
Gestor de Apoyo Estudiantil, Universidad Internacional del Ecuador, Ecuador

jbohorquez@uide.edu.ec

jpb9677@hotmail.com

<https://orcid.org/0009-0009-8345-3681>

Fecha de recepción: 11 de julio de 2025

Fecha de aprobación: 06 de octubre de 2025

DOI:



¹ El presente documento fue elaborado por el autor como resultado de una investigación original. Durante su redacción se emplearon herramientas de inteligencia artificial generativa únicamente con fines auxiliares, como la verificación lingüística y la revisión preliminar de estilo. Estas herramientas no sustituyeron en ningún momento la autoría intelectual del contenido ni participaron en la formulación de ideas, argumentos o análisis presentados.

RESUMEN

Este estudio diseña, implementa y evalúa UIDE Classifier, una solución basada en modelos generativos para la clasificación automática de comentarios abiertos de estudiantes en educación superior. Se procesaron 3,328 comentarios provenientes de encuestas NPS aplicadas en la Universidad Internacional del Ecuador (modalidades presencial y en línea; periodos 2024-2 y 2025-1). El enfoque es cualitativo aplicado, con diseño no experimental y corte transversal. La herramienta utiliza ingeniería de *prompts*, una ontología temática institucional cerrada y un archivo de conocimiento con sinónimos y expresiones locales para guiar la clasificación multicategoría e identificar comentarios no informativos. La validación se efectuó frente a un estándar humano (ocho evaluadores) mediante indicadores de desempeño y tiempos de procesamiento. Los resultados muestran una precisión promedio del 96 % y una cobertura temática del 94 %. Operativamente, el tiempo de análisis se redujo en un 83 % (de tres horas humanas a 0.5 horas con inteligencia artificial por cada 1,000 comentarios). Se documentan aciertos y errores recurrentes —como confusiones entre categorías semánticamente cercanas o la inferencia de modalidad en ausencia de contexto—, así como una alta tolerancia del modelo a errores ortográficos y lenguaje informal. Finalmente, se discuten las implicaciones prácticas para la mejora de la gestión de la experiencia estudiantil y se ofrecen recomendaciones para la adopción responsable de inteligencia artificial en entornos universitarios, con énfasis en la trazabilidad, la revisión humana y la gobernanza de datos. El principal aporte consiste en demostrar la viabilidad y transferibilidad de una solución contextualizada al ámbito latinoamericano, que aprovecha la inteligencia artificial generativa para escalar la sistematización de la retroalimentación estudiantil con calidad y eficiencia.

Palabras clave: inteligencia artificial, educación superior, retroalimentación estudiantil, clasificación automatizada, sistematización de información.

ABSTRACT

This study designs, implements, and evaluates UIDE Classifier, a solution based on generative models for the automatic classification of open-ended student comments in higher education. A total of 3,328 comments were processed from NPS surveys administered at the International University of Ecuador (both on-site and online modalities; academic terms 2024-2 and 2025-1). The research follows an applied qualitative approach with a non-experimental, cross-sectional design. The tool employs prompt engineering, a closed institutional thematic ontology, and a knowledge file containing synonyms and local expressions to guide multi-category classification and detect non-informative comments. Validation was conducted against a human standard (eight evaluators) using performance indicators and processing time metrics. The results show an average accuracy of 96% and thematic coverage of 94%. Operationally, the analysis time was reduced by 83% (from three human hours to 0.5 hours using artificial intelligence per 1,000 comments). The study documents recurrent successes and errors—such as confusions between semantically related categories or modality inference in the absence of context—as well as the model's high tolerance for spelling errors and informal

language. Finally, the paper discusses practical implications for improving student experience management and provides recommendations for the responsible adoption of artificial intelligence in university settings, emphasizing traceability, human review, and data governance. The main contribution lies in demonstrating the feasibility and transferability of a contextualized solution for the Latin American context, leveraging generative AI to scale the systematization of student feedback with quality and efficiency.

Keywords: *artificial intelligence, higher education, student feedback, automated classification, information systematization.*

Introducción

La inteligencia artificial generativa (IAG) está transformando progresivamente el análisis de datos cualitativos en la educación superior. En contextos donde la voz del estudiante constituye un insumo estratégico para los procesos de mejora continua, el procesamiento automatizado de comentarios abiertos representa una alternativa innovadora y eficiente, capaz de transformar los mecanismos tradicionales de retroalimentación. En la Universidad Internacional del Ecuador (UIDE) surgió la necesidad de analizar grandes volúmenes de información recolectada mediante encuestas *Net Promoter Score* (NPS), un sistema ampliamente utilizado para medir la satisfacción estudiantil a través de valoraciones y comentarios textuales.

Este artículo presenta el desarrollo de UIDE Classifier, una herramienta personalizada basada en GPT-4, tecnología derivada del modelo conversacional ChatGPT (OpenAI, 2023), orientada a clasificar comentarios según categorías temáticas institucionales predefinidas, optimizando así la gestión de la retroalimentación académica y fortaleciendo la toma de decisiones sustentadas en evidencia. Su diseño parte del reconocimiento de que, en los entornos educativos contemporáneos, la comprensión de la experiencia estudiantil debe apoyarse no solo en métricas cuantitativas, sino también en el análisis semántico profundo del lenguaje natural.

En la literatura especializada se ha señalado que la inteligencia artificial (IA) constituye una herramienta transformadora en la educación superior, no únicamente por su capacidad para personalizar el aprendizaje, sino también por su potencial para reinterpretar procesos administrativos, evaluativos y de retroalimentación (Holmes *et al.*, 2019). Estos autores destacan que el uso de sistemas inteligentes puede reducir las barreras tradicionales asociadas al análisis manual de datos cualitativos, mejorando la agilidad institucional y permitiendo intervenciones más oportunas y centradas en las necesidades del estudiante.

Más recientemente, Zhai (2023) ha indicado que la IAG, al combinar modelos de lenguaje avanzados con capacidades de razonamiento contextual, posibilita abordar tareas complejas como la clasificación temática, la detección de emociones y la inferencia de significado en textos abiertos. Esto ha originado una nueva línea de investigación que explora cómo la IA puede actuar como mediadora entre el lenguaje informal y los marcos analíticos estructurados, superando incluso los sesgos o imprecisiones gramaticales típicas del discurso estudiantil.

Asimismo, autores como Lan *et al.* (2024) y Lee y Moore (2024) han documentado el potencial del procesamiento de lenguaje natural y de la IAG en entornos educativos, especialmente en la automatización del análisis de retroalimentación, la predicción de comportamientos y la generación de informes adaptativos para docentes y administradores. Estas investigaciones subrayan que el uso de IA puede favorecer la construcción de sistemas educativos más adaptativos, personalizados y basados en datos, capaces de responder de manera proactiva a los desafíos del entorno digital.

Diversos estudios han abordado también el uso del NPS como herramienta de medición de la satisfacción en la educación superior. Por ejemplo, Chicaña Huanca (2018) comparó la percepción de calidad en programas de posgrado presenciales y semipresenciales a partir del NPS, destacando su utilidad para captar la voz del estudiante tanto en términos cuantitativos como cualitativos. Sin embargo, el análisis manual de los comentarios abiertos continúa siendo un desafío para muchas instituciones, debido a su volumen, ambigüedad, subjetividad y a la variabilidad del lenguaje empleado por los estudiantes.

La IAG constituye una subcategoría emergente dentro del campo más amplio de la IA, caracterizada por su capacidad para producir contenido nuevo a partir de grandes volúmenes de datos. En el ámbito educativo, su aplicación más destacada ha sido a través de modelos de lenguaje como GPT, que permiten generar texto, interpretar lenguaje natural y realizar tareas de clasificación semántica con altos niveles de precisión. A diferencia de los modelos supervisados tradicionales, que requieren grandes cantidades de datos etiquetados, los modelos generativos pueden ajustarse mediante ejemplos contextuales, técnicas de *few-shot learning* o instrucciones específicas, lo que los convierte en herramientas flexibles y adaptables a entornos donde la recolección de datos etiquetados resulta costosa o inviable.

Dentro del campo del procesamiento de lenguaje natural (NLP), la ingeniería de *prompts* se posiciona como una metodología relevante para dirigir el comportamiento de estos modelos. Consiste en diseñar cuidadosamente las instrucciones que recibe el modelo, estructurando entradas que le permitan inferir con mayor precisión la intención del texto. Esta técnica ha resultado particularmente útil en contextos donde se requiere una comprensión profunda del lenguaje informal, como en el caso de los comentarios estudiantiles, que suelen incluir expresiones coloquiales, abreviaturas o errores gramaticales. En este sentido, la ingeniería de *prompts* actúa como un puente entre el lenguaje real de los usuarios y la lógica estructurada necesaria para análisis automatizados precisos.

Pese al creciente interés y la disponibilidad de estas tecnologías, aún existen pocas experiencias documentadas en contextos latinoamericanos que adapten modelos lingüísticos a marcos conceptuales institucionales específicos. Esta brecha resalta la relevancia de propuestas como UIDE Classifier, que no solo aplican tecnologías de frontera, sino que las contextualizan conforme a necesidades reales, lenguajes locales y estructuras organizativas particulares. La presente investigación busca contribuir a este vacío, demostrando que es posible diseñar soluciones de IAG que sean precisas, éticas y adaptables a las realidades de las universidades latinoamericanas.

Método

Enfoque y tipo de estudio

Este estudio se enmarca en un enfoque cualitativo, de tipo aplicado, con diseño no experimental y

corte transversal. Se optó por este enfoque para explorar el comportamiento de un modelo GPT en el análisis automatizado de retroalimentación abierta en un contexto universitario.

El enfoque cualitativo fue elegido por su capacidad para explorar fenómenos complejos relacionados con el lenguaje, la experiencia subjetiva del estudiante y la interpretación contextual de los comentarios abiertos. Este enfoque permite comprender no solo qué se dice, sino cómo se dice, considerando matices, ambigüedades y elementos implícitos. En lugar de reducir los comentarios a indicadores cuantitativos, el análisis cualitativo ofrece una perspectiva más amplia de la percepción estudiantil, sus preferencias, dificultades y formas de interacción con el entorno educativo.

El diseño no experimental, de corte transversal, permitió observar el comportamiento del modelo en un momento específico, sin manipular variables, lo cual resultó adecuado para validar su funcionamiento en condiciones reales. Este tipo de diseño es útil en estudios exploratorios orientados a evaluar la viabilidad de herramientas tecnológicas aplicadas a contextos institucionales concretos.

Población y muestra

La población objetivo incluyó estudiantes de pregrado y posgrado de la UIDE, en modalidades presencial y en línea. La muestra estuvo compuesta por 3,328 comentarios recogidos en los formularios institucionales NPS correspondientes a los periodos 2024-2 y 2025-1. Se utilizó un muestreo no probabilístico por conveniencia, considerando únicamente registros con comentarios abiertos completos. Todos los datos fueron anonimizados para garantizar la confidencialidad.

La inclusión de estudiantes tanto de pregrado como de posgrado respondió a la intención de capturar una diversidad de voces y contextos educativos. Las modalidades presencial y en línea presentan dinámicas diferenciadas en cuanto a interacción, percepción de calidad y uso de plataformas, lo cual enriquece la base de datos y exige al modelo una mayor capacidad inferencial.

El muestreo por conveniencia, aunque no permite la generalización estadística, resulta pertinente en estudios cualitativos cuyo propósito es profundizar en los significados más que estimar proporciones. La inclusión exclusiva de comentarios completos garantizó que el material de análisis fuera lo suficientemente informativo para extraer patrones temáticos y emocionales relevantes.

Técnicas de análisis

Se emplearon técnicas de ingeniería de *prompts* para entrenar un modelo GPT personalizado a través de la plataforma OpenAI, utilizando archivos de referencia institucional, etiquetas diseñadas *ad hoc* y afinación mediante instrucciones específicas.

Las técnicas de ingeniería de *prompts* constituyeron el eje metodológico del entrenamiento del modelo. Se diseñaron instrucciones detalladas que simulaban el razonamiento humano al clasificar comentarios, incorporando ejemplos positivos y negativos, manejo de ambigüedades y diferenciación entre categorías semánticamente cercanas.

Asimismo, se implementó lógica condicional dentro de los *prompts* para detectar múltiples temas en un mismo comentario y para inferir la modalidad educativa aun en ausencia de menciones explícitas. Este enfoque permitió entrenar al modelo no solo para ejecutar tareas, sino también para tomar decisiones interpretativas dentro de límites previamente definidos.

El proceso se apoyó en pruebas iterativas, en las cuales las salidas del modelo fueron evaluadas por humanos y ajustadas mediante la reformulación progresiva de las instrucciones. Esto garantizó una mayor alineación entre el comportamiento del modelo y los criterios conceptuales institucionales, fortaleciendo su confiabilidad como herramienta de análisis.

Entrenamiento del modelo

Se utilizó una instancia personalizada de GPT-4 a través de la plataforma OpenAI, configurada específicamente para ejecutar tareas de clasificación semántica bajo condiciones controladas. Este modelo se fundamenta en los principios de alineación mediante instrucciones descritos en *InstructGPT* (OpenAI, 2022), los cuales permiten adaptar su comportamiento a tareas específicas a partir de ejemplos y directrices semánticas. El entrenamiento no implicó un ajuste de pesos del modelo base (*fine-tuning*), sino un proceso de afinación mediante ingeniería de *prompts*, con una guía de clasificación diseñada *ad hoc* y estructurada en seis componentes principales.

Definición del propósito. El objetivo del modelo fue analizar comentarios abiertos provenientes de encuestas NPS institucionales con el fin de identificar: 1) Categorías temáticas relevantes (a partir de una lista cerrada de 66 opciones oficiales); 2) Sentimiento expresado (positivo, negativo o neutral); y 3) Modalidad inferida (presencial u *online*), incluso cuando no fuera declarada explícitamente.

Diseño de ontología institucional. Se definió una estructura semántica cerrada que agrupa temas relacionados con áreas académicas, administrativas, tecnológicas y de infraestructura. Esta ontología resultó crucial para garantizar la consistencia terminológica, reducir ambigüedades y facilitar la posterior agregación de resultados para los análisis institucionales.

Archivo de conocimiento. Se elaboró un conjunto de instrucciones detalladas que contenía sinónimos, expresiones frecuentes, patrones lingüísticos y excepciones asociadas a cada categoría. Este archivo también incluía lineamientos específicos para interpretar comentarios vagos, ambiguos o con

inferencias implícitas. Actuó como una «capa semántica intermedia» entre el lenguaje informal de los estudiantes y el modelo.

Ejemplos anotados. Se construyó un *corpus* con más de 80 comentarios reales previamente clasificados por expertos, que sirvió como guía para orientar el comportamiento del modelo. Estos ejemplos permitieron reforzar criterios de clasificación, contextualizar decisiones inferenciales y reducir errores por sobreajuste al *prompt* principal.

Procesamiento por lotes. Dado que los modelos generativos pueden presentar pérdida de precisión ante entradas extensas o irregulares, se implementó una lógica de análisis por bloques de hasta 50 comentarios por sesión. Esto permitió mantener la estabilidad en el rendimiento, controlar la carga cognitiva del modelo y facilitar la revisión parcial de resultados.

Capacidades extendidas del modelo. 1) Clasificación múltiple por comentario, en caso de detectarse referencias a más de una categoría temática; 2) Inferencia de modalidad educativa a partir de pistas contextuales, como el uso de términos («Zoom», «presencialmente», «en línea»); y 3) Etiquetado de comentarios irrelevantes, ambiguos o no informativos mediante la categoría <SIN CLASIFICACIÓN>.

Este enfoque permitió construir un clasificador alineado con el lenguaje natural de los estudiantes, sensible al contexto institucional y suficientemente robusto para operar con datos reales, diversos y con errores lingüísticos.

Validación del modelo

La validación del modelo se centró en comparar los resultados de una clasificación automática realizada mediante GPT con una clasificación manual de referencia. Para ello, se conformó un equipo de ocho evaluadores humanos independientes, quienes clasificaron una base de 4,093 comentarios según categoría temática y sentimiento. La clasificación humana fue considerada el estándar de oro (*gold standard*), y los resultados del modelo se cotejaron mediante un análisis de concordancia intercodificadores, aplicando indicadores de precisión, tiempo y cobertura temática.

Se empleó una estrategia de validación cruzada doble: por un lado, se compararon los resultados individuales del modelo con las etiquetas humanas; por otro, se evaluó la consistencia interna entre los evaluadores humanos para identificar posibles sesgos o divergencias interpretativas. Cuando surgieron discrepancias relevantes entre el modelo y el juicio humano, se aplicó un proceso de revisión por consenso entre pares, priorizando el análisis semántico del contexto textual.

Este procedimiento permitió no solo medir la precisión global del clasificador, sino también evaluar su desempeño en escenarios complejos, como la identificación de múltiples temas en un mismo

comentario, la inferencia de modalidad educativa implícita y la clasificación de comentarios redactados en lenguaje informal. Además, las discrepancias detectadas durante la validación aportaron insumos valiosos para afinar los *prompts* y enriquecer el archivo de conocimiento, consolidando un ciclo de mejora continua en el entrenamiento del modelo.

Limitaciones metodológicas

En el marco de la transparencia metodológica, es pertinente señalar algunas limitaciones asociadas al diseño de este estudio. En primer lugar, al tratarse de un enfoque cualitativo de tipo aplicado, con un diseño no experimental y corte transversal, los hallazgos no pueden generalizarse a toda la población estudiantil ni establecer relaciones causales entre variables. El propósito principal fue validar el desempeño del modelo en un contexto institucional específico, más que derivar conclusiones extrapolables.

El muestreo no probabilístico por conveniencia, aunque adecuado para estudios exploratorios, representa una limitación en términos de representatividad. Si bien se consideraron diversas modalidades y niveles educativos, no se garantiza que todos los perfiles estudiantiles estén plenamente reflejados. Esto podría afectar la diversidad de expresiones y temáticas presentes en los comentarios analizados.

Asimismo, la interpretación inferencial realizada por el modelo se basa en un archivo de conocimiento institucional y en ejemplos previamente clasificados. Esto implica que, en casos de ambigüedad extrema, errores ortográficos graves o expresiones locales no contempladas, el modelo podría efectuar clasificaciones inexactas o subestimar contenidos relevantes. Aunque se incorporaron mecanismos de curaduría y validación humana, estos escenarios continúan siendo posibles.

Por último, el proceso de validación se efectuó sobre una base de comentarios limitada a dos periodos académicos. No se ha evaluado aún el comportamiento del modelo en ciclos futuros ni su capacidad de adaptación a cambios en el lenguaje estudiantil, por lo que se recomienda ampliar la validación en estudios posteriores con una perspectiva longitudinal.

Resultados

La efectividad del modelo se evaluó mediante la comparación de sus resultados con una clasificación manual de referencia (*gold standard*), utilizando medidas de precisión, cobertura temática y tiempo de procesamiento:

- Precisión promedio del modelo: 96 % en comparación con el *gold standard* humano.
- Reducción del tiempo de procesamiento: de tres horas manuales a 0.5 horas con IA por cada 1,000 comentarios.
- Cobertura temática: identificación correcta de más del 94 % de las categorías relevantes.
- Satisfacción del usuario interno: los equipos de gestión valoraron de forma positiva la utilidad del clasificador.

La Tabla 1 resume los principales indicadores de desempeño cuantitativo obtenidos.

Tabla 1

Desempeño del modelo UIDE Classifier en comparación con evaluación humana (gold standard)

Métrica	Resultado
Precisión	96 %
Cobertura temática	94 %
Reducción de tiempo	83 %

Fuente: elaboración propia.

Más allá de las métricas cuantitativas, el proceso de validación permitió identificar ciertos errores recurrentes que ofrecen oportunidades relevantes para el perfeccionamiento del modelo. Entre los casos más frecuentes se observaron los siguientes:

- Clasificaciones erróneas en comentarios vagos o excesivamente breves, en los cuales no se expresaba un contenido temático claro.
- Dificultades ocasionales para distinguir entre categorías semánticamente cercanas, como *docentes* y *metodología de enseñanza*, especialmente en frases ambiguas.
- Inferencias incorrectas de modalidad (presencial u *online*) cuando el comentario carecía de suficientes pistas contextuales.

Estos hallazgos resaltan la importancia de mantener un sistema de curaduría y revisión humana complementaria, en particular durante las etapas iniciales de implementación.

La Tabla 2 muestra el nivel de concordancia entre el juicio humano y el modelo, así como algunas diferencias interpretativas observadas.

Tabla 2

Ejemplos comparativos entre clasificación humana y clasificación automática

Comentario estudiantil	Clasificación humana	Clasificación: UIDE Classifier
«La docente explica muy bien, aunque la plataforma a veces falla»	Docentes, aula virtual	Docentes, aula virtual
«No entendí el objetivo de varias materias, y los horarios me afectan mucho»	Plan de estudios, horarios	Plan de estudios, horarios
«Todo bien, pero a veces no se entienden las clases en Zoom»	Aula virtual	Metodología de enseñanza

Fuente: elaboración propia.

En el último ejemplo se observa una discrepancia: mientras la clasificación humana priorizó el medio

tecnológico (Aula virtual), el modelo identificó un posible problema pedagógico (Metodología de enseñanza), basándose en la expresión «no se entienden las clases», lo cual evidencia un juicio inferencial. Este tipo de diferencias, aunque poco frecuentes, resulta relevante para la mejora continua del sistema.

Asimismo, se identificó que el modelo mostró una alta tolerancia a errores ortográficos, gramaticales y sintácticos, aspecto especialmente valioso en contextos reales donde los comentarios estudiantiles suelen redactarse de forma informal o descuidada. A pesar de estas imperfecciones, el clasificador logró realizar asignaciones correctas en la mayoría de los casos.

La Tabla 3 muestra la capacidad del modelo para procesar adecuadamente entradas con errores ortográficos y redacción informal.

Tabla 3

Ejemplos de desempeño del modelo ante errores ortográficos y de redacción informal

Comentario estudiantil con errores	Clasificación: UIDE Classifier
«excelente profe pero la plata forma a vecs se cae»	Docentes, aula virtual
«no entiendo muxo la materia el orario no me ayuda»	Plan de estudios, horarios
«las clases en zoom a vezce son raraz»	Metodología de enseñanza

Fuente: elaboración propia.

Este comportamiento demuestra la capacidad inferencial del modelo, sustentada en el reconocimiento semántico más allá de la forma textual precisa. De este modo, incluso ante errores tipográficos o expresiones no normativas, el sistema logra identificar correctamente las categorías temáticas pertinentes.

Una de las aportaciones más relevantes del presente estudio es la evidencia empírica que muestra la ventaja del uso de la IAG frente a los métodos tradicionales de análisis cualitativo. En los enfoques manuales, el procesamiento de comentarios abiertos suele requerir la codificación temática por parte de múltiples evaluadores humanos, lo que demanda tiempo, incrementa los costos y puede generar inconsistencias debido a la subjetividad del juicio humano.

En contraste, el modelo UIDE Classifier permitió reducir en un 83 % el tiempo necesario para procesar 1,000 comentarios, manteniendo una precisión del 96 %. Además, mostró una mayor estabilidad en la asignación de categorías múltiples y en el reconocimiento del lenguaje informal. Estos resultados evidencian no solo una mejora considerable en la eficiencia operativa, sino también una mayor coherencia en la toma de decisiones, aspecto relevante para fortalecer los procesos institucionales basados en evidencia y centrados en la experiencia estudiantil.

Discusión

El uso de modelos generativos para el análisis de datos cualitativos representa una oportunidad estratégica para las instituciones de educación superior, en especial en contextos donde la voz del estudiante constituye un insumo relevante para la evaluación institucional, la mejora continua y el diseño de políticas académicas. UIDE Classifier no solo agiliza el procesamiento de grandes volúmenes de retroalimentación, sino que también permite identificar patrones emergentes, señales de alerta temprana y focos de intervención institucional con mayor rapidez y consistencia que los métodos tradicionales.

Al tratarse de un modelo entrenado específicamente con datos, estructuras y marcos semánticos de la Universidad Internacional del Ecuador, su aplicabilidad práctica se ve considerablemente fortalecida, al ofrecer respuestas adaptadas a las necesidades y realidades del entorno universitario. Esta capacidad de adaptación contextual diferencia a la herramienta de soluciones genéricas y facilita su integración operativa en los ciclos institucionales de mejora continua.

Durante la fase de validación se identificaron márgenes de error e inconsistencias, sobre todo en comentarios breves, vagos o ambivalentes, donde el modelo tendió a realizar inferencias que no siempre coincidieron con la interpretación humana. Estos hallazgos resaltan la importancia de mantener mecanismos de supervisión humana en las etapas críticas del proceso, especialmente durante la implementación inicial, así como de continuar alimentando al modelo con nuevos ejemplos representativos del lenguaje estudiantil real, caracterizado por su tono informal, emocional y heterogéneo.

Estos resultados se alinean con estudios previos que señalan cómo la IAG puede desempeñar un papel relevante en los procesos de retroalimentación educativa, al permitir una comprensión más profunda de las experiencias estudiantiles y facilitar la toma de decisiones pedagógicas más informadas y sensibles al contexto (Luckin *et al.*, 2016). En particular, la capacidad del modelo para detectar temas latentes o emocionales a partir de expresiones coloquiales amplía el espectro del análisis cualitativo institucional, hasta ahora limitado por las capacidades humanas disponibles.

No obstante, la implementación de sistemas de IA en procesos institucionales conlleva consideraciones éticas significativas. Uno de los principales desafíos identificados fue garantizar la trazabilidad de las decisiones automatizadas, especialmente en los casos en que el modelo infirió modalidad educativa o sentimientos sin evidencia explícita en el texto. Esto plantea interrogantes sobre los límites de la interpretación algorítmica y la necesidad de establecer protocolos que regulen su uso.

Asimismo, el manejo de comentarios con errores ortográficos, ambigüedades sintácticas o expresiones poco convencionales obligó a definir criterios claros de interpretación contextual, evitando sobre

interpretaciones que pudieran distorsionar la voz genuina del estudiante. Estos desafíos destacan la importancia de establecer una arquitectura de gobernanza de la IA que combine transparencia operativa, revisión humana y adaptabilidad continua.

Por otra parte, el respeto por la privacidad de los datos, la auditabilidad del sistema y la posibilidad de retroalimentar y mejorar los algoritmos sin comprometer la integridad del discurso estudiantil se constituyen en principios esenciales para una adopción ética. En los contextos universitarios latinoamericanos, donde la confianza institucional es un recurso estratégico, estos elementos adquieren una relevancia aún mayor.

En conjunto, la experiencia con UIDE Classifier demuestra que es posible combinar eficacia tecnológica con sensibilidad institucional, siempre que se mantenga un equilibrio entre automatización y supervisión crítica, y se asegure que la inteligencia artificial funcione como una herramienta al servicio de las personas, y no como sustituto del juicio humano.

Síntesis y proyecciones

La implementación del modelo UIDE Classifier evidencia que la IAG puede aplicarse de forma ética, efectiva y contextualizada en los procesos de análisis cualitativo dentro de la educación superior. A partir de un diseño basado en ingeniería de *prompts*, entrenamiento semántico y validación experta, fue posible automatizar la clasificación temática y emocional de los comentarios estudiantiles con altos niveles de precisión y una notable reducción del tiempo de procesamiento.

Los resultados del estudio confirman el potencial de este tipo de herramientas para fortalecer los mecanismos institucionales de escucha activa, seguimiento de la experiencia estudiantil y toma de decisiones sustentadas en evidencia. Asimismo, se comprobó que es viable desarrollar modelos que respeten el lenguaje natural de los estudiantes y se adapten a los marcos conceptuales de cada institución.

Desde una perspectiva de innovación educativa, UIDE Classifier representa una solución escalable, replicable y de bajo costo relativo, con un alto valor estratégico para las universidades latinoamericanas que enfrentan desafíos similares en la gestión de datos cualitativos.

A partir de la experiencia desarrollada con UIDE Classifier, se proponen las siguientes recomendaciones para aquellas instituciones que deseen replicar este modelo en sus propios contextos:

- Adaptar la ontología temática a las categorías institucionales vigentes, asegurando un marco cerrado y contextualizado que refleje las prioridades académicas, administrativas y tecnológicas de la institución.
- Elaborar un archivo de conocimiento con expresiones clave y sinónimos comunes utilizados por los

estudiantes, especialmente en lenguaje informal. Este insumo resulta esencial para garantizar una clasificación semántica precisa.

- Implementar una fase de validación cruzada con evaluadores humanos que permita detectar errores frecuentes, sesgos y discrepancias relevantes. La revisión por consenso constituye un paso crucial para afinar los criterios de clasificación.
- Utilizar herramientas de IA en modo asistido, manteniendo siempre mecanismos de revisión humana en las etapas críticas del proceso. Esta práctica asegura un uso responsable, ético y adaptativo del sistema.
- Documentar exhaustivamente todo el proceso, desde la recolección de datos hasta la afinación de *prompts* y la definición de criterios, con el fin de facilitar futuras auditorías, transferencias o mejoras del modelo.

En conjunto, este estudio aporta evidencia metodológica y aplicada sobre cómo adaptar modelos generativos a marcos institucionales cerrados en América Latina, abriendo nuevas rutas para la gestión automatizada de la retroalimentación estudiantil. Su desarrollo demuestra que es posible implementar soluciones de IA que respeten la especificidad semántica de cada institución, integrándose a procesos de mejora continua sin comprometer la voz auténtica del estudiante.

Asimismo, la propuesta establece un precedente replicable para universidades que enfrentan desafíos similares de volumen, ambigüedad y variabilidad en los datos cualitativos. Al combinar accesibilidad tecnológica con efectividad operativa, UIDE Classifier se posiciona como una iniciativa pionera para transformar el análisis de datos en contextos educativos de la región, sentando las bases para futuras aplicaciones en ámbitos como alertas tempranas, visualización de tendencias o evaluación institucional integral.

Referencias

- Chicaña Huanca, B. D. (2018). *Estudio comparativo sobre el índice de satisfacción de los servicios educativos, empleando el Net Promoter Score (NPS®), de las maestrías presenciales y semipresenciales de la Escuela de Postgrado de la UCSM. Arequipa-2016* [Tesis de maestría, Universidad Católica de Santa María] Repositorio de tesis UCSM. <https://repositorio.ucsm.edu.pe/server/api/core/bitstreams/3c88da56-f1d3-4d2f-b8a9-4d7a46b8e4a2/content>
- Holmes, W., Bialik, M. y Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign. <https://curriculumredesign.org/wp-content/uploads/AI-in-Education-Promises-and-Implications.pdf>
- Lan, Y., Li, X., Du, H., Lu, X., Gao, M., Qian, W. y Zhou, A. (2024). Survey of Natural Language Processing for Education: Taxonomy, Systematic Review, and Future Trends. *arXiv*. <https://arxiv.org/abs/2401.07518>
- Lee, S. S. y Moore, R. L. (2024). Harnessing Generative AI (GenAI) for Automated Feedback in

Higher Education: A Systematic Review. *Online Learning*, 28(3).

<https://doi.org/10.24059/olj.v28i3.4593>

Luckin, R., Holmes, W., Griffiths, M. y Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education. https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/about-pearson/innovation/open-ideas/Intelligence_Unleashed_Publication.pdf

OpenAI. (2022). *InstructGPT: Aligning language models to follow instructions*.

<https://openai.com/research/instruction-following>

OpenAI. (2023). *ChatGPT: Optimizing language models for dialogue*. <https://openai.com/blog/chatgpt>

Zhai, X. (2023). Generative AI as a transformative technology for education: Evidence, opportunities, and challenges. *Educational Technology Research and Development*, 71, pp. 1149–1153.

<https://doi.org/10.1007/s11423-023-10168-4>