

Metodología para la Identificación de Buenas Prácticas Docentes a Partir de Comentarios Estudiantiles Utilizando Análisis de Sentimientos y Modelado Temático

Methodology for Identifying Good Teaching Practices Through Student Comments Using Sentiment Analysis and Structural Topic Modeling

Diego Moscoso Lozano¹ <https://orcid.org/0009-0008-0073-5687>,
Jorge Maldonado-Mahauad¹ <https://orcid.org/0000-0003-1953-390X>

¹Universidad de Cuenca, Cuenca, Ecuador

diego.moscosol@ucuenca.edu.ec, jorge.maldonado@ucuenca.edu.ec



Esta obra está bajo una licencia internacional
Creative Commons Atribución-NoComercial 4.0

Enviado: 2025/06/28

Aceptado: 2025/08/21

Publicado: 2025/10/15

Resumen

El presente estudio propone una metodología empírica para identificar las dimensiones más valoradas de la práctica docente en educación superior, mediante un enfoque mixto que combina análisis cualitativo y cuantitativo. Se analizaron más de 89000 comentarios estudiantiles recopilados durante las evaluaciones docentes en la Universidad de Cuenca (periodo 2024–2025). Utilizando técnicas avanzadas de procesamiento del lenguaje natural, específicamente análisis de sentimiento y modelado estructural de temas (Structural Topic Modeling, STM), se procesó un amplio corpus textual. El análisis se centró en los comentarios positivos dirigidos a docentes destacados según una evaluación integral previa, con el objetivo de detectar patrones recurrentes en la percepción estudiantil. Como resultado, se identificaron cuatro dimensiones clave de valoración estudiantil: conocimiento disciplinar, gestión comunicativa, habilidades socioemocionales y metodología de enseñanza. Estas dimensiones reflejan fortalezas docentes percibidas por los estudiantes y presentan variabilidad significativa según contextos disciplinares y niveles académicos. Por ejemplo, mientras el conocimiento disciplinar es prioritario en áreas técnicas, las habilidades socioemocionales predominan en ciencias sociales y salud. Estos hallazgos destacan que las valoraciones estudiantiles constituyen una fuente de información cualitativa esencial, que supera las limitaciones de las métricas numéricas tradicionales al ofrecer

Sumario: Introducción, Metodología, Resultados y Discusión y Conclusiones.

Cómo citar: Moscoso, D. & Maldonado-Mahauad, J. (2025). Metodología para la Identificación de Buenas Prácticas Docentes a Partir de Comentarios Estudiantiles Utilizando Análisis de Sentimientos y Modelado Temático. *Revista Tecnológica - Espol*, 37(E1), 146-161.
<https://doi.org/10.37815/rte.v37nE1.1353>

un análisis más profundo y contextualizado. Finalmente, la metodología propuesta aporta a las instituciones de educación superior una herramienta práctica para identificar fortalezas específicas del profesorado y para diseñar programas formativos personalizados basados en evidencia, que facilitan una mejora continua y contextualizada de la calidad educativa.

Palabras clave: evaluación docente, análisis de comentarios, modelado temático estructural, educación superior.

Abstract

This study proposes an empirical methodology to identify the most valued dimensions of teaching practices in higher education through a mixed-method approach that integrates qualitative and quantitative analyses. Over 89,000 student comments collected during teacher evaluations at the University of Cuenca (2024–2025) were analyzed. Advanced natural language processing techniques, specifically sentiment analysis and Structural Topic Modeling (STM), were used to process an extensive textual corpus. The analysis focused on positive comments aimed at teachers recognized through a prior comprehensive evaluation, aiming to uncover recurring patterns in student perceptions. Four key dimensions of student evaluation emerged from the analysis: disciplinary knowledge, communication management, socio-emotional skills, and teaching methodology. These dimensions reflect perceived teaching strengths and exhibit significant variability across disciplinary contexts and academic levels. For example, disciplinary knowledge is prioritized in technical areas, while socio-emotional skills are more valued in social sciences and health disciplines. These findings underscore that student evaluations are an essential source of qualitative information, overcoming limitations of traditional numeric metrics by offering deeper, contextual insights. Ultimately, the proposed methodology provides higher education institutions with a practical tool to identify specific faculty strengths and design evidence-based, personalized professional development programs, thus fostering continuous and contextualized improvements in educational quality.

Keywords: teaching evaluation, comment analysis, structural topic modeling, higher education.

Introducción

La evaluación docente constituye un pilar fundamental en la búsqueda de la excelencia académica y la mejora continua de la calidad educativa en las instituciones de educación superior. A través de ella, las universidades rinden cuentas sobre el desempeño de su personal académico y además obtienen información crucial para fomentar el desarrollo profesional y la innovación pedagógica (Spooren et al., 2013). En este sentido, la perspectiva del estudiante se ha consolidado como una fuente invaluable de información para estas evaluaciones, ofreciendo una visión única sobre el proceso de enseñanza-aprendizaje y el impacto directo del desempeño docente. Así, estudios previos han demostrado cómo las prácticas docentes universitarias impactan directamente el interés, el compromiso y el rendimiento académico de los estudiantes (Álvarez-Álvarez y Falcon, 2023).

Tradicionalmente, la evaluación del desempeño docente se ha apoyado en gran medida en métricas cuantitativas, como promedios de calificación obtenidos de instrumentos de evaluación como los cuestionarios. Si bien las valoraciones numéricas ofrecen una visión general, suelen resultar insuficientes para capturar la complejidad de las prácticas pedagógicas efectivas en la enseñanza y el aprendizaje. De hecho, la mayoría de los estudios anteriores en este campo se han centrado en estas valoraciones numéricas, ignorando el valioso contenido cualitativo de los

comentarios abiertos de los estudiantes (Sun y Yan, 2023). Las evaluaciones basadas en escalas Likert limitan la capacidad de los estudiantes para expresar libremente sus percepciones sobre el desempeño docente (Gencoglu et al., 2023). Gran parte de la información que los estudiantes comparten en sus comentarios se pierde cuando todo se resume en promedios numéricos, lo que impide entender con claridad lo que los estudiantes realmente valoran y consideran como buenas prácticas docentes. Esta limitación muestra la importancia de agregar datos cualitativos para aprovechar los comentarios abiertos de los estudiantes en la evaluación para tener una visión más transparente y completa del trabajo que realiza el docente en el aula de clase (Hu et al., 2022).

En respuesta a esta limitación, el presente estudio propone una metodología empírica que integra enfoques cuantitativos y cualitativos para el análisis de la evaluación docente. Para esto, se utilizan herramientas de procesamiento de lenguaje natural (PLN), como el análisis de sentimiento y el modelado temático, aplicadas a un corpus de comentarios cualitativos emitidos por estudiantes. El análisis de sentimiento permite clasificar la polaridad de los comentarios para captar la intensidad emocional (Muhammad et al., 2016; Wankmüller y Heumann, 2021) e identificar específicamente las valoraciones positivas. Por su parte, el Structural Topic Model (STM) se emplea para identificar patrones temáticos latentes en los comentarios positivos de los estudiantes (Roberts et al., 2014), con la ventaja de poder explorar cómo varía la prevalencia de estos temas según características contextuales, como la facultad y el nivel académico. Si bien la mayoría de los estudios previos han aplicado técnicas de análisis de sentimiento o topic modeling de forma aislada, la integración de ambos enfoques en esta investigación representa una contribución significativa al proporcionar una comprensión más completa de la percepción estudiantil (Hayat et al., 2024).

Como resultado de aplicar esta metodología, se logró identificar las dimensiones más valoradas de la práctica docente por los estudiantes de la Universidad de Cuenca y entender cómo esas valoraciones cambian según la facultad y el nivel académico, a partir del análisis de los comentarios positivos emitidos hacia docentes con mejor desempeño. Este enfoque busca ofrecer una comprensión más matizada de las expectativas estudiantiles, reconociendo que las percepciones de los estudiantes pueden variar significativamente según el contexto disciplinar (Marsh, 2007) y el momento formativo (Chen y Hoshower, 2003). La metodología propuesta es replicable y puede servir como una herramienta útil para que las universidades diseñen estrategias de formación docente más efectivas y ajustadas a su contexto. Poder identificar fortalezas y planificar capacitaciones con base en los comentarios de los estudiantes abre nuevas posibilidades para la mejora continua de la calidad de la docencia.

El presente artículo se estructura de la siguiente manera: en la sección 2 se detalla la metodología empírica, que se compone de 3 etapas: (1) preparación y segmentación de los datos; (2) modelado temático; y (3) análisis de prevalencia temática. La sección 3 presenta los principales resultados y la discusión. Finalmente, la sección 5 presenta las principales conclusiones del estudio y sus implicaciones para la evaluación y el desarrollo docente en la educación superior.

Metodología

Esta investigación adopta un enfoque mixto que combina métodos cuantitativos (segmentación y modelado temático) y cualitativos (interpretación de los comentarios de los estudiantes), dentro de un diseño exploratorio que busca descubrir patrones temáticos en los comentarios positivos emitidos por estudiantes hacia docentes con desempeño destacado. Para esto, se desarrolló una metodología empírica de tres etapas: (1) la preparación y segmentación

de los datos, (2) el modelado temático y (3) el análisis de prevalencia temática e interpretación contextual. En la Figura 1 se presenta visualmente este proceso metodológico.



Etapas de la metodología empírica

Etapas de la metodología empírica

Recolección de datos

Los datos utilizados para este trabajo corresponden al proceso de evaluación docente realizado en las 12 facultades de la Universidad de Cuenca, específicamente los datos del componente de heteroevaluación (evaluación que realizan los estudiantes a sus docentes). El conjunto de datos utilizado pertenece al periodo lectivo septiembre 2024-febrero 2025 y está conformado por 1891 registros de evaluación; cada uno de estos corresponde a una asignatura evaluada por los estudiantes. Esto implica que un mismo docente puede estar representado en más de un registro en el caso de que imparta distintas asignaturas. Esta decisión metodológica permite considerar el desempeño docente de forma contextualizada por asignatura, ya que un docente puede mostrar fortalezas en determinadas materias o niveles, pero no necesariamente en todas.

Tabla 1
Resumen de las variables utilizadas en el estudio

VARIABLE	TIPO DE DATO	DESCRIPCIÓN
docente_id	Cualitativa nominal	Identificador anónimo único asignado a cada docente evaluado en una o más asignaturas.
facultad	Cualitativa nominal	Facultad a la que pertenece la asignatura evaluada.
carrera	Cualitativa nominal	Carrera en la que se imparte la asignatura evaluada.
asignatura	Cualitativa nominal	Nombre de la materia evaluada por los estudiantes. Un docente puede tener varias asignaturas.
nivel	Cualitativa ordinal	Nivel académico correspondiente al ciclo en el que se dicta la asignatura, codificado del 1 al 10 (de primer a décimo ciclo).
numero_evaluadores	Cuantitativa discreta	Total de estudiantes que completaron la evaluación docente para un registro determinado.
nota_heteroevaluacion	Cuantitativa continua	Promedio obtenido por el docente en la evaluación estudiantil (sobre 40 puntos).
comentarios_destaca	Cualitativa textual	Texto libre escrito por los estudiantes en una pregunta abierta del instrumento de evaluación. Contiene los comentarios por asignatura sobre lo que destaca del docente.
comentarios_mejorar	Cualitativa textual	Texto libre escrito por los estudiantes en una pregunta abierta del instrumento de evaluación. Contiene los comentarios por asignatura sobre lo que el docente debería mejorar.

Dentro de este conjunto de datos, se recopilaron 89281 comentarios abiertos emitidos por los estudiantes en respuesta al apartado cualitativo del instrumento de evaluación aplicado. Estos comentarios conforman el corpus sobre el cual se aplicaron las técnicas de análisis de sentimiento y modelado temático. Para este estudio se consideraron variables de tipo cualitativo y cuantitativo, que se describen en la Tabla 1.

Todos los datos fueron anonimizados antes del análisis y provienen del proceso regular de evaluación docente de la Universidad de Cuenca. No se utilizó información personal de los docentes ni de los estudiantes, por lo que no fue necesario obtener consentimiento adicional para su uso con fines de este trabajo.

Preprocesamiento básico del dataset

Se realizó un proceso de limpieza y estructuración del dataset con el fin de garantizar la integridad y coherencia de los datos. Se filtraron los registros que no contenían comentarios en ninguna de las dos categorías cualitativas (“lo que destaca del docente” y “lo que el docente debería mejorar”), conservando solamente los registros con contenido suficiente para el análisis.

En cuanto al tratamiento de los comentarios, estos fueron obtenidos originalmente como cadenas concatenadas separadas por el carácter “|”. Cada comentario fue desagregado y etiquetado con los prefijos “Destaca:” o “Debe mejorar:”, según la sección del formulario de donde proviene. Finalmente, se integraron en una sola columna ambos tipos de comentarios que se utilizaron para el siguiente paso del análisis de sentimientos. La incorporación de estos prefijos responde a la necesidad de mantener la intención de los comentarios emitidos por los estudiantes, ya que es frecuente que se utilicen términos idénticos con sentidos opuestos según la sección donde se respondan. Esta decisión se basa en el enfoque de análisis de sentimiento contextual planteado por Muhammad et al. (2016), donde se demuestra que identificar correctamente la polaridad de un texto es más preciso cuando se incluye información adicional que ayude a entender el contexto en el que se escribió.

Cálculo del indicador de desempeño compuesto

En el contexto de la evaluación docente, suele ser insuficiente considerar solamente las calificaciones numéricas que resultan de las valoraciones estudiantiles para capturar el desempeño docente. Estas valoraciones de los estudiantes comprenden dimensiones tanto cuantitativas como cualitativas, que deben integrarse para lograr una visión más completa y significativa (Hu et al., 2022). En este sentido, como parte de la metodología, se planteó la construcción de un índice de desempeño compuesto, que combine la calificación promedio otorgada por los estudiantes, la polaridad semántica de los comentarios abiertos y el número de evaluadores por asignatura. Este enfoque responde a la necesidad de interpretar la evaluación más allá de indicadores unidimensionales, permitiendo capturar aspectos no visibles en los promedios numéricos, como la percepción emocional del estudiante o la participación en el proceso evaluativo.

El propósito de este índice es identificar a los docentes que presentan un desempeño sobresaliente desde la percepción estudiantil. Esta identificación permite seleccionar un grupo de referencia sobre el cual aplicar el análisis para reconocer patrones asociados a buenas prácticas docentes que son valoradas por los estudiantes. Este puntaje de desempeño compuesto considera tres variables fundamentales:

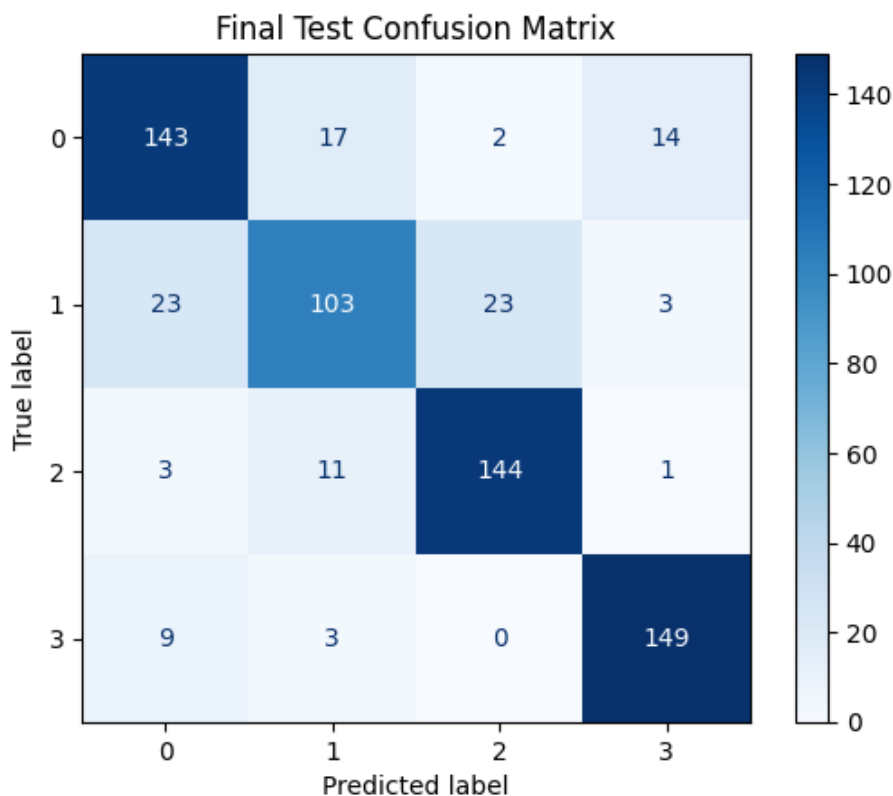
- Nota de heteroevaluación (N_c): Corresponde a la calificación promedio otorgada por los estudiantes a cada docente en una asignatura específica. Dado que la nota

original se encuentra en una escala de 0 a 40, se transformó a una escala de 0 a 100 mediante una regla de tres simple.

- **Calificación de sentimiento (Sc):** Se obtuvo mediante un modelo de tipo BERT (RoBERTuito) previamente ajustado (*fine-tuned*) con un corpus de comentarios etiquetados con los prefijos “Destaca:” y “Debe mejorar:”. La elección de este modelo se justifica por su capacidad para capturar la semántica de las palabras en su contexto, lo que es crucial para el análisis de sentimientos en textos cortos como los comentarios de estudiantes (Prottasha et al., 2022). El modelo clasifica cada comentario en una de cuatro categorías: positivo, neutral, negativo o alerta, y asigna ponderaciones específicas a cada una. Para entrenar y validar el rendimiento del modelo de clasificación utilizado para la obtención del S_c, se empleó un conjunto de 3244 comentarios etiquetados manualmente. A fin de garantizar la robustez del modelo, se aplicó una estrategia de validación cruzada con k-fold (k=5), dividiendo aleatoriamente los datos en cinco subconjuntos para entrenamiento y prueba. El modelo alcanzó un *accuracy* del 85 % y un F1-score del 83 %, lo cual evidencia un buen desempeño en la clasificación de sentimientos en textos cortos. En la Figura 2 se presenta la matriz de confusión del modelo, que permite visualizar el nivel de acierto en cada una de las categorías: positivo, neutral, negativo y alerta. A partir de esta clasificación se construyó un score ponderado de sentimiento, que también fue transformado a una escala de 0 a 100. Esta transformación de una variable textual en una variable continua es especialmente útil para captar la intensidad emocional expresada en textos cortos, como los comentarios que dejan los estudiantes en las evaluaciones docentes (Wankmüller et al., 2021).

Figura 2

Matriz de confusión del modelo RoBERTuito para clasificación de sentimientos



- Número de evaluadores (E_n): como medida indirecta del grado de consenso, esta variable fue transformada mediante logaritmo base 10 y posteriormente normalizada a una escala de 0 a 100 utilizando la técnica de *min-max scaling*, con el fin de reducir la asimetría propia de su distribución, de forma que su influencia en el índice fuera comparable con las demás dimensiones.

La integración de estas tres variables se resume en la Figura 3, donde se presenta gráficamente la estructura del puntaje de desempeño compuesto ($PD_{DOCENTE}$) y sus componentes. La formulación matemática del puntaje es la siguiente:

$$PD_{docente} = 0,4 \cdot N_c + 0,4 \cdot S_c + 0,2 \cdot E_n$$

Figura 3

Componentes y fórmula del puntaje de desempeño compuesto

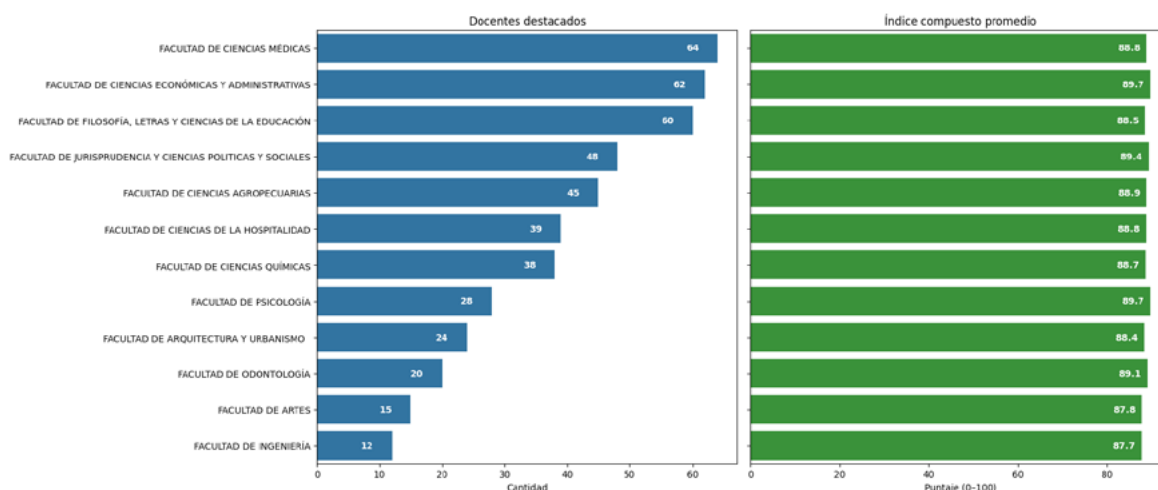
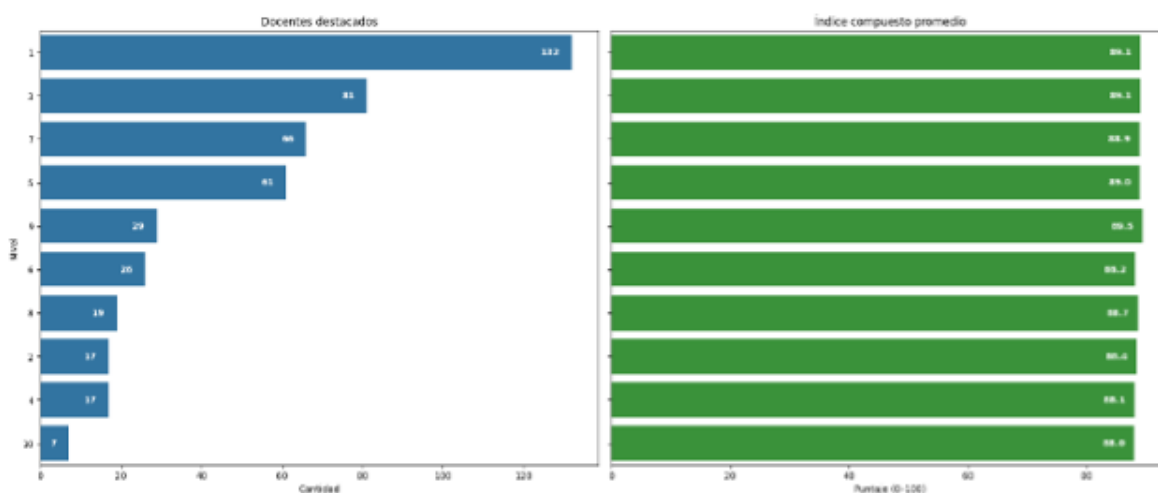


La forma en que se distribuyen los pesos en la fórmula responde a dos criterios principales. Por un lado, se buscó dar la misma importancia a la calificación numérica y al análisis de sentimiento, ya que ambos reflejan dimensiones complementarias del desempeño docente. Por otro lado, se decidió incluir el número de evaluadores con un peso menor, pero significativo, reconociendo que un mayor número de estudiantes aporta mayor confiabilidad a los resultados. Esta decisión está respaldada por estudios que destacan la necesidad de considerar la cantidad de respuestas para asegurar que las valoraciones sean representativas y estables (Spooren et al., 2013).

Segmentación de docentes con desempeño sobresaliente

Una vez calculado el puntaje compuesto de desempeño, se procedió a segmentar a los docentes considerando el cuartil superior de este indicador. De esta manera, se pudo identificar un subconjunto de 455 registros, correspondientes a 334 docentes únicos que presentan un desempeño sobresaliente. La elección del cuartil superior permite trabajar con un subconjunto representativo que concentra a los perfiles mejor valorados desde la percepción estudiantil, y sobre el cual se aplicará el análisis temático posterior.

La distribución de estos registros por facultad se presenta en la Figura 4. En ella se observa, por un lado, la cantidad de docentes destacados en cada unidad académica, y por otro, el promedio del índice de desempeño compuesto para los casos identificados. De forma complementaria, la Figura 5 muestra la misma información desagregada por nivel académico, que resulta útil para contextualizar el análisis posterior sobre patrones temáticos.

Figura 4*Distribución de docentes destacados y promedio del índice de desempeño por facultad***Figura 5***Distribución de docentes destacados y promedio del índice de desempeño por nivel*

Etapas 2: Modelado temático

Análisis de sentimiento y filtrado del corpus

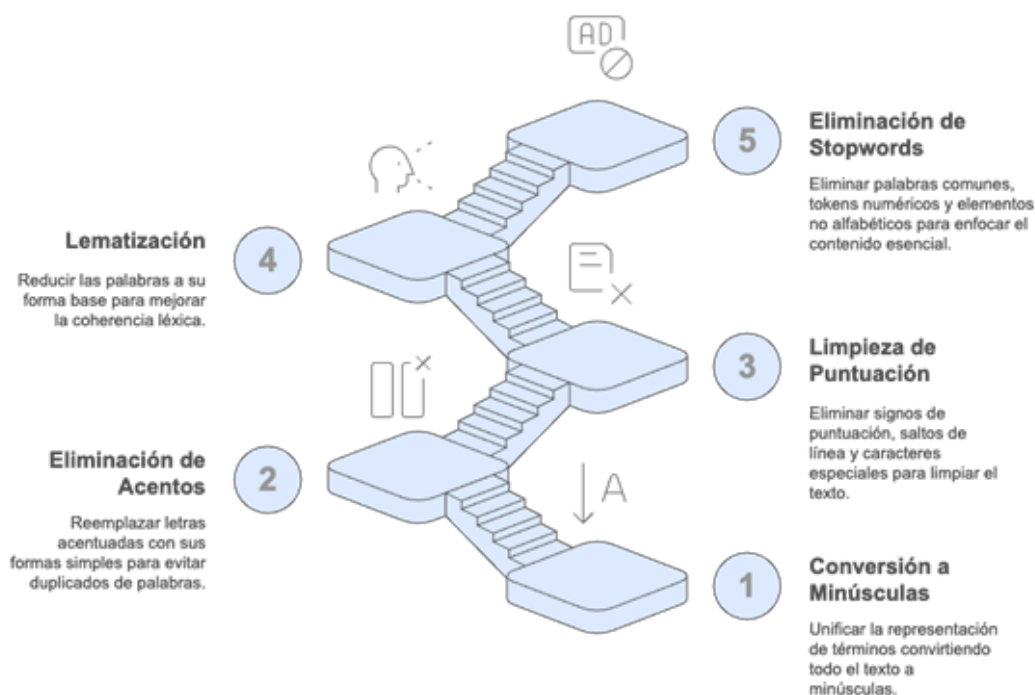
En esta etapa, se definió el subconjunto de comentarios sobre el que se aplicaría el modelado temático. Considerando que el objetivo del estudio es identificar buenas prácticas docentes reconocidas por los estudiantes, se optó por trabajar exclusivamente con los comentarios clasificados como positivos por el modelo de análisis de sentimiento, es decir, aquellos que expresan valoraciones favorables hacia el desempeño del docente.

Este filtro se aplicó a los registros pertenecientes al cuartil superior del índice de desempeño compuesto, lo que permitió obtener un total de 19240 comentarios positivos que conforman el corpus final de análisis. Esta selección permite enfocar el modelado en expresiones directamente vinculadas con aspectos reconocidos como buenas prácticas por parte de los estudiantes, reduciendo así la presencia de ruido derivado de comentarios neutros o negativos. El conjunto resultante constituye una muestra textual coherente con los objetivos del estudio, sobre la cual se aplicarán los procesos de preprocesamiento lingüístico y modelado temático descritos en los pasos siguientes.

Preprocesamiento lingüístico

Antes de aplicar el modelado temático, se llevó a cabo un proceso de limpieza y normalización de los comentarios positivos, con el objetivo de reducir el ruido textual y asegurar la coherencia léxica en el corpus. Este procesamiento se implementó en Python, utilizando expresiones regulares y la biblioteca spaCy con el modelo en español es_core_news_md. Las transformaciones aplicadas se resumen en la Figura 6.

Figura 6
Etapas del preprocesamiento lingüístico del corpus



Entrenamiento del modelo STM

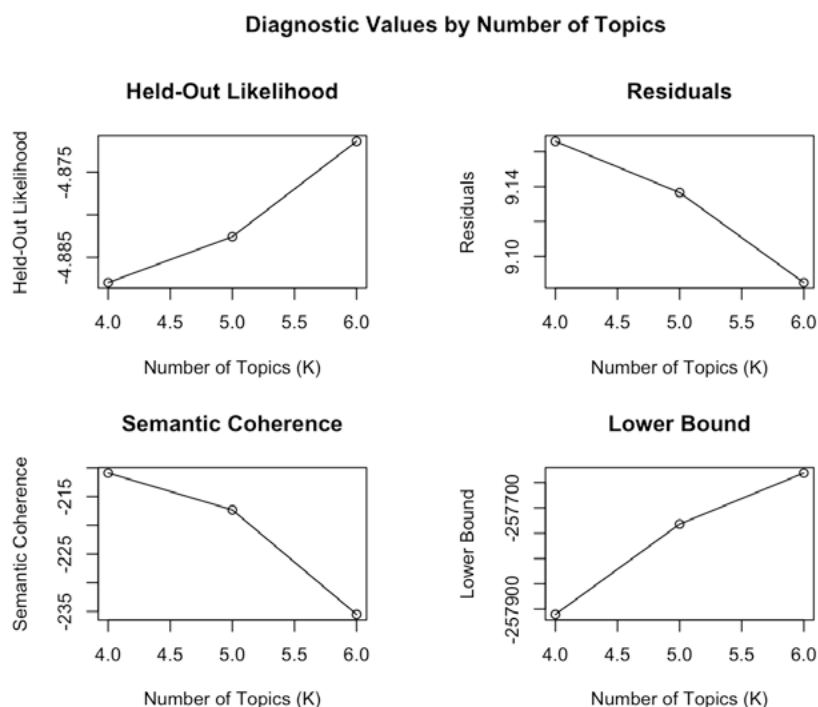
Para identificar los patrones temáticos presentes en los comentarios positivos emitidos por los estudiantes, se utilizó el modelo Structural Topic Model (STM) propuesto por Roberts et al. (2014). Este enfoque permite identificar temas latentes en textos, tomando en cuenta información adicional de los documentos como variables explicativas, lo que facilita explorar cómo varía la prevalencia de los tópicos según características contextuales. A diferencia de otros métodos como LDA (Latent Dirichlet Allocation), el STM permite estimar directamente los efectos de variables externas sobre la distribución temática de cada documento, lo que lo hace especialmente útil para estudios en ciencias sociales y educación.

El modelo se entrenó en R utilizando la librería stm (Roberts et al., 2019), con el corpus previamente procesado y representado mediante los objetos documents, vocab y meta. El preprocesamiento del texto, que incluyó la tokenización y la aplicación de umbrales de frecuencia para palabras raras, se realizó de manera automática por las funciones integradas en este paquete. Este proceso asegura la rigurosidad y la calidad de los datos de entrada al modelo, y concentra el análisis en palabras con suficiente representatividad. La estimación se realizó con el método de inicialización Spectral, recomendado por los autores del modelo para facilitar la convergencia y mejorar la estabilidad de los temas generados. En la fórmula de prevalencia se incluyeron dos covariables: facultad y nivel.

Se evaluaron modelos con 4, 5 y 6 temas (K) con base en métricas cuantitativas como la Held-Out Likelihood, residuos y convergencia. Si bien los resultados (ilustrados en la Figura 7) mostraron diferencias marginales entre las configuraciones, la literatura reciente señala que la selección del número óptimo de tópicos no puede basarse exclusivamente en métricas cuantitativas (Hayat et al., 2024). Un modelo puede mostrar métricas ligeramente superiores, pero su interpretabilidad y relevancia contextual pueden ser bajas. Por ello, se consideró necesario complementar este análisis con una validación cualitativa más detallada, abordada en la siguiente sección.

Figura 7

Métricas de diagnóstico para la selección del número óptimo de temas (K) en el modelo STM



Validación cualitativa

Una vez entrenados los modelos con diferentes cantidades de temas, se procedió a realizar una validación cualitativa para determinar cuál de las configuraciones generadas ofrecía una mayor coherencia semántica y claridad interpretativa, un criterio considerado fundamental en el modelado de temas (Hayat et al., 2024). Esta revisión se centró en analizar las palabras clave asociadas a cada tema, así como ejemplos representativos de comentarios asignados a cada uno.

La validación fue realizada por una experta en educación superior, con experiencia en evaluación docente y análisis cualitativo. Su revisión consideró tres criterios principales: qué tan bien se entendía cada tema a partir de sus palabras clave, si los temas estaban claramente diferenciados entre sí, y si los comentarios asignados a cada tópico realmente reflejaban lo que sugerían esas palabras.

Aunque los modelos con 5 y 6 temas presentaban ventajas cuantitativas, la revisión cualitativa reveló que esos modelos generaban temas con mayor solapamiento, menor claridad semántica y una división forzada de tópicos que dificultaba su interpretación. Por el contrario, la configuración con $K = 4$ ofrecía una mejor coherencia conceptual, diferenciación clara entre

tópicos y relevancia directa respecto al objetivo del estudio. Esta validación sustentó la adopción del modelo con $K = 4$ como base para el posterior proceso de etiquetado e interpretación de los temas.

Etiquetado de temas

Una vez seleccionado el modelo STM con cuatro tópicos, se realizó el proceso de interpretación de los resultados. Para cada tema, se revisaron las palabras clave con mayor exclusividad (FREX) y se leyeron comentarios representativos asignados por el modelo. A partir de esta revisión, se asignó un nombre que resume de forma clara la idea principal de cada grupo de comentarios, tal como se sugiere en el enfoque propuesto por Roberts et al. (2014).

En la Tabla 2 se presentan los cuatro temas identificados, junto con sus palabras clave más relevantes y un ejemplo de comentario real para cada uno. Estos temas servirán como base para los análisis posteriores por facultad y nivel.

Tabla 2
Temas identificados mediante STM

TÓPICO	NOMBRE ASIGNADO	PALABRAS CLAVE (FREX)	COMENTARIO REPRESENTATIVO
T1	Gestión Comunicativa	lenguaje, actividad, organización, dinámica, creatividad, impartir, interactivo	“El docente destaca en su expresión corporal y su lenguaje, a la hora de impartir su clase y su manera llamativa de interactuar con el grupo”.
T2	Metodología de Enseñanza	método, contenido, explicación, claridad, teórico, dinámico, metodología	“Destaca la aplicación de metodologías dinámicas y participativas que fomentan el aprendizaje significativo, el pensamiento crítico y la resolución de problemas”.
T3	Conocimiento Disciplinar	amplio, explicar, comprensión, claro, dedicación, excelente, tema, área, retroalimentación	“El docente demuestra un alto nivel de conocimiento en la asignatura y explica los temas de manera clara y organizada”.
T4	Habilidades Socioemocionales	manejo, ambiente, ejemplo, despertar, profesor, momento, forma, fácil	“Comprensiva y paciente, siempre con críticas constructivas que me ayudan a mejorar”.

Etapa 3: Análisis de prevalencia temática

Estimación de efectos marginales

Después de etiquetar los temas obtenidos con el modelo STM, se estimó cómo varía la presencia de cada tópico en función de dos variables contextuales: la facultad y el nivel académico de la asignatura evaluada. Esto permitió analizar si las prácticas valoradas por los estudiantes, que fueron identificadas en el modelo tienden a aparecer con más frecuencia en algunas facultades o en determinados ciclos formativos.

Para este análisis se utilizó la función `estimateEffect()` del paquete `stm` en R, que permite modelar la relación entre los temas generados y las covariables incluidas en la fórmula de prevalencia. En este caso, se trabajó con facultad y nivel, variables que representan el contexto institucional y formativo en el que se imparten las clases. La elección de estas covariables se basa en evidencia que muestra que las percepciones estudiantiles pueden variar según la unidad académica, por diferencias disciplinares (Marsh, 2007), así como por el momento formativo del estudiante, ya que sus expectativas y valoraciones tienden a cambiar conforme avanza en la carrera (Chen y Hoshower, 2003).

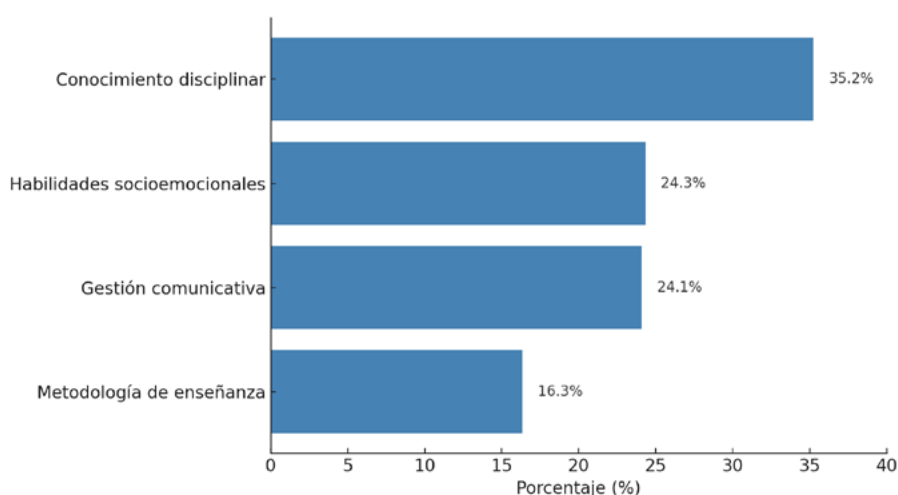
Resultados y Discusión

Esta sección presenta los resultados obtenidos a partir del análisis temático de los comentarios positivos emitidos por los estudiantes a los docentes con mejor desempeño, según

el índice compuesto calculado. El objetivo es identificar patrones discursivos que revelen buenas prácticas docentes valoradas por los estudiantes. A continuación, se describe la distribución general de los temas extraídos por el modelo STM, y luego se analiza su prevalencia según dos variables contextuales: facultad y nivel académico.

El análisis temático permitió identificar cuatro grandes categorías recurrentes en los comentarios positivos de los estudiantes hacia docentes destacados. La Figura 8 presenta la distribución de estos temas según su frecuencia media en el corpus. El tema más representado fue conocimiento disciplinar (35,2 %), seguido de habilidades socioemocionales (24,3 %) y gestión comunicativa (24,1 %). En menor proporción, aunque aún relevante, se identificó la metodología de enseñanza, con un 16,3 % de aparición promedio en los documentos. Esta primera aproximación muestra que los estudiantes tienden a valorar con mayor frecuencia el dominio del contenido por parte del docente, pero también reconocen aspectos actitudinales y comunicativos en sus valoraciones positivas.

Figura 8
Frecuencia de aparición de los temas identificados



La Figura 9 muestra cómo varían los temas identificados en los comentarios positivos según la facultad de la Universidad de Cuenca. En términos generales, el conocimiento disciplinar predomina notablemente y alcanza su máxima presencia en la Facultad de Odontología (38,9 %), seguida por Ciencias de la Hospitalidad (38,4 %) e Ingeniería (37,7 %), indicando una fuerte apreciación estudiantil por el dominio conceptual y técnico en estas áreas. Por otro lado, la facultad con menor valoración en este aspecto es Artes (27,0 %).

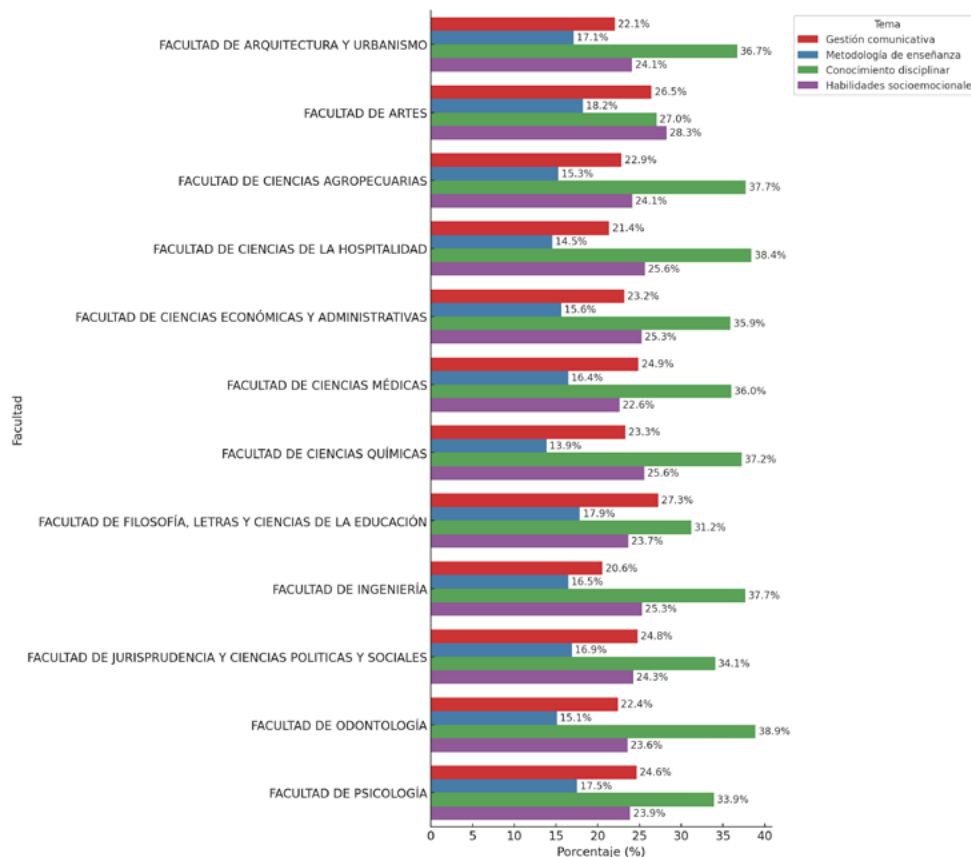
En contraste, las habilidades socioemocionales muestran una presencia significativa en facultades de carácter más humanístico y social, destacando especialmente en Psicología (23,9 %) y Arquitectura y Urbanismo (22,1 %), mientras que su representación más baja aparece en Ciencias Químicas (13,9 %).

Respecto a la gestión comunicativa, esta dimensión destaca particularmente en Artes (26,5 %), Arquitectura y Urbanismo (24,1 %) y Ciencias Médicas (22,6 %), lo que refleja una alta valoración hacia la claridad, interacción y empatía comunicativa por parte de los docentes en dichas áreas. Su representación más limitada se observa en Ciencias Agropecuarias (15,3 %) y Ciencias Económicas y Administrativas (15,6 %).

Finalmente, la dimensión de metodología de enseñanza, aunque presenta una menor prevalencia en general, alcanza un valor máximo en Filosofía, Letras y Ciencias de la Educación (27,3 %), resaltando la importancia que adquiere la estructura y dinámica pedagógica en esta facultad. En contraste, Arquitectura y Urbanismo muestra el mínimo valor para esta dimensión (17,1 %).

Estas diferencias ilustran claramente cómo la valoración estudiantil sobre las prácticas docentes varía según el contexto disciplinar, subrayando la importancia de considerar estas particularidades en la interpretación y aplicación de resultados de evaluación docente.

Figura 9
Frecuencia media estimada de los temas por facultad



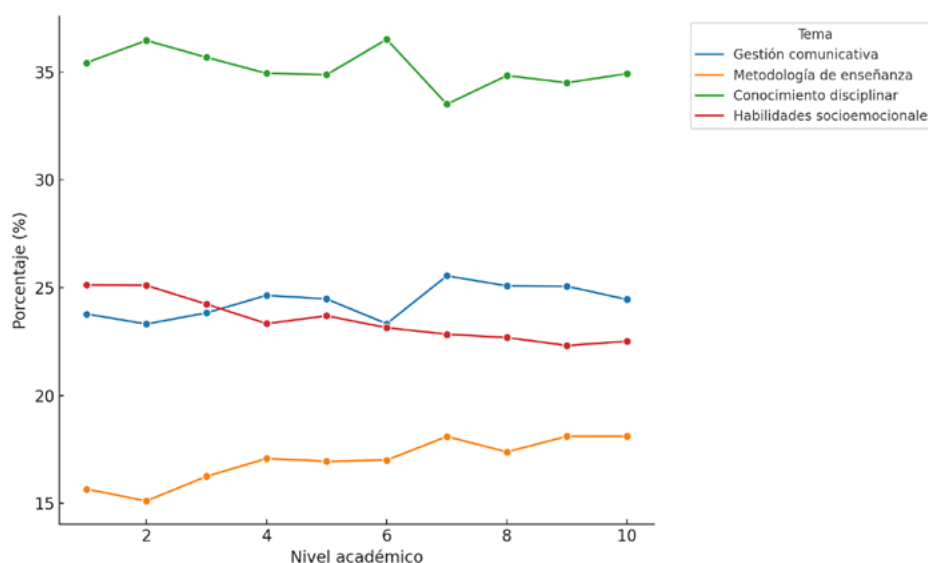
La Figura 10 presenta la evolución porcentual de los temas identificados en los comentarios positivos a lo largo de los diferentes niveles académicos. El conocimiento disciplinar se mantiene consistentemente como el más frecuente en todos los niveles, registrando valores cercanos al 35 %, aunque muestra una leve disminución hacia los niveles superiores. Esto podría indicar que, aunque la solidez conceptual y técnica sigue siendo prioritaria, otros aspectos cobran más relevancia conforme avanzan los estudiantes en su formación.

En contraste, la gestión comunicativa y las habilidades socioemocionales muestran una tendencia ligeramente ascendente a medida que aumentan los niveles académicos. En particular, la gestión comunicativa se consolida claramente por encima del 25 % a partir del nivel 5, destacando la importancia que adquieren las capacidades de interacción y comunicación efectiva del profesorado en etapas académicas avanzadas. Por su parte, las habilidades socioemocionales presentan un leve repunte en los niveles superiores, sugiriendo que los estudiantes valoran crecientemente la empatía y el apoyo emocional conforme avanzan en sus estudios.

Finalmente, la metodología de enseñanza se mantiene relativamente estable, con porcentajes en torno al 15-18 % en todos los ciclos académicos, constituyendo el tema menos frecuente a lo largo del recorrido formativo analizado. Esto podría reflejar que, aunque la metodología didáctica es apreciada, los estudiantes asignan consistentemente mayor peso a otros aspectos relacionados directamente con el contenido y la interacción docente-estudiante.

Estos resultados sugieren que, conforme avanzan en sus estudios, las prioridades de los estudiantes experimentan cambios sutiles pero relevantes, destacando progresivamente aspectos relacionales y comunicativos del proceso formativo, además del continuo énfasis en la claridad conceptual.

Figura 10
Proporción media de los temas por nivel



Al observar la distribución de los temas por área académica, se evidencian diferencias asociadas al enfoque de formación. En carreras técnicas y aplicadas, predominan valoraciones relacionadas con el conocimiento disciplinar, mientras que en áreas de salud, educación y ciencias sociales emergen con mayor frecuencia comentarios sobre habilidades comunicativas y relacionales. En cuanto al nivel académico, se observó que la proporción de comentarios asociados a Gestión comunicativa tiende a incrementarse levemente en los ciclos superiores, mientras que las Habilidades socioemocionales presentan una ligera disminución. Esto sugiere que, conforme avanzan en su formación, los estudiantes podrían valorar más la capacidad del docente para sostener una comunicación clara y estructurada que aspectos de tipo empático o relacional.

Conclusiones

El estudio ha demostrado con éxito la eficacia de la metodología empírica, al combinar análisis cualitativos y cuantitativos mediante técnicas avanzadas de procesamiento de lenguaje natural para comprender las percepciones estudiantiles sobre las prácticas docentes. Las cuatro dimensiones identificadas —conocimiento disciplinar, gestión comunicativa, habilidades socioemocionales y metodología de enseñanza— destacan aspectos críticos de la docencia valorados por los estudiantes. Particularmente, el predominio del conocimiento disciplinar en la mayoría de las facultades subraya la importancia continua de un sólido dominio del contenido,

mientras que las variaciones en el énfasis sobre aspectos socioemocionales y comunicativos reflejan expectativas diferenciadas según el contexto disciplinar y la progresión académica.

A pesar de la solidez de la metodología utilizada, este estudio presenta ciertas limitaciones. Primero, el análisis se restringió a comentarios positivos de estudiantes que evaluaron a docentes altamente calificados, lo que podría dejar fuera percepciones valiosas presentes en comentarios neutros o negativos. En segundo lugar, la generalización contextual podría verse limitada dado que los hallazgos corresponden a una única institución y a un periodo temporal específico (2024–2025).

Investigaciones futuras deberían ampliar el análisis para incluir comentarios con distintos niveles de sentimiento, permitiendo una comprensión más integral tanto de fortalezas como de áreas de mejora. Además, estudios comparativos entre diversas instituciones y a lo largo de periodos más extendidos enriquecerían la generalizabilidad y robustez de los resultados obtenidos. Finalmente, explorar análisis longitudinales podría revelar tendencias a largo plazo, brindando perspectivas valiosas sobre la evolución de las percepciones estudiantiles y la efectividad institucional de las iniciativas de desarrollo profesional docente.

Reconocimientos y Declaraciones

Este artículo recibió el apoyo de la Vicerrectoría Académica de la Universidad de Cuenca y la Dirección de Innovación Educativa.

Este trabajo forma parte de una investigación realizada por la Dirección de Innovación Educativa de la Universidad de Cuenca.

Los autores desean expresar su agradecimiento a la Universidad de Cuenca, al Vicerrectorado Académico y a la Dirección de Innovación Educativa por el apoyo institucional brindado para el desarrollo de la presente investigación.

En cuanto a la contribución de los autores, siguiendo la taxonomía CRediT se declara que la participación ha sido equitativa en todas las fases del trabajo.

Asimismo, los autores manifiestan que, en la elaboración de este artículo, se han utilizado herramientas de inteligencia artificial como apoyo al proceso de investigación y redacción. En particular:

- Consensus, para la búsqueda y análisis de literatura científica.
- Notebook LM, para la revisión, organización y síntesis de la información recolectada.
- ChatGPT-4o, para la redacción de las secciones de Introducción y Metodología.

Referencias

- Álvarez-Álvarez, C., Falcon, S. Students' preferences with university teaching practices: analysis of testimonials with artificial intelligence. *Education Tech Research Dev* 71, 1709–1724 (2023). <https://doi.org/10.1007/s11423-023-10239-8>
- Chen, Y., & Hoshower, L. B. (2003). Student evaluation of teaching effectiveness: An assessment of student perception and motivation. *Assessment & Evaluation in Higher Education*, 28(1), 71–88. <https://doi.org/10.1080/02602930301683>

- Gencoglu, B., Helms-Lorenz, M., Maulana, R., Jansen, E. P. W. A., & Gencoglu, O. (2023). Machine and expert judgments of student perceptions of teaching behavior in secondary education: Added value of topic modeling with big data. *Computers & Education*, 193, 104682. <https://doi.org/10.1016/j.compedu.2022.104682>
- Hayat, F., Shatnawi, S., & Haig, E. (2024). Comparative analysis of topic modelling approaches on student feedback. En *Proceedings of the 16th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (KDIR2024)* (pp. 226-233). SciTePress. <https://doi.org/10.5220/0012890400003838>
- Hu, Y., Zhang, S., Sathy, V., Panter, A., & Bansal, M. (2022). SETSum: Summarization and Visualization of Student Evaluations of Teaching. En *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: System Demonstrations* (pp. 71–89). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2022.naacl-demo.9>
- Marsh, H. W. (2007). Do university teachers become more effective with experience? A multilevel growth model of students' evaluations of teaching over 13 years. *Journal of Educational Psychology*, 99(4), 775–790. <https://doi.org/10.1037/0022-0663.99.4.775>
- Muhammad, A., Wiratunga, N., & Lothian, R. (2016). Contextual sentiment analysis for social media genres. *Knowledge-Based Systems*, 108, 92–101. <https://doi.org/10.1016/j.knosys.2016.05.032>
- Prottasha, N. J., As Sami, A., Kowsher, M., Murad, S. A., Bairagi, A. K., Masud, M., & Baz, M. (2022). Transfer learning for sentiment analysis using BERT based supervised fine-tuning. *Sensors*, 22(11), 4157. <https://doi.org/10.3390/s22114157>
- Roberts, M. E., Stewart, B. M., & Airoldi, E. M. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58(4), 1064–1082. <https://doi.org/10.1111/ajps.12103>
- Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). stm: An R package for structural topic models. *Journal of Statistical Software*, 91(2), 1–40. <https://doi.org/10.18637/jss.v091.i02>
- Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, 83(4), 598–642. <https://doi.org/10.3102/0034654313496870>
- Sun, J., & Yan, L. (2023). Using topic modeling to understand comments in student evaluations of teaching. *Discover Education*, 2(25). <https://doi.org/10.1007/s44217-023-00051-0>
- Wankmüller, S., & Heumann, C. (2021). How to estimate continuous sentiments from texts using binary training data. En *Proceedings of the Conference on Natural Language Processing (KONVENS 2021)* (pp. 131–139). German Society for Computational Linguistics & Language Technology. <https://aclanthology.org/2021.konvens-1.16/>