

Sistema de Recomendación de Cursos en Línea basado en el Perfil de Competencias TIC del Docente

Online course recommendation system based on the teacher's ICT competence profile

Jorge Maldonado-Mahauad¹ <https://orcid.org/0000-0003-1953-390X>, Diego Moscoso Lozano¹ <https://orcid.org/0009-0008-0073-5687>, Javier Pacheco¹ <https://orcid.org/0009-0004-3990-5277>

¹Universidad de Cuenca, Cuenca, Ecuador

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



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

Enviado: 2024/07/02

Aceptado: 2024/10/03

Publicado: 2024/10/15

Resumen

La transformación digital en la educación ha generado la necesidad de capacitar a los docentes y mejorar sus competencias digitales. Sin embargo, los procesos de formación no suelen ser efectivos dada la falta de personalización en las propuestas y la variedad en los perfiles de los docentes. Dado lo anterior, este artículo aborda el desarrollo de una herramienta web diseñada para diagnosticar las capacidades digitales de los profesores y proporcionar recomendaciones personalizadas de MOOCs para apoyar el desarrollo de sus competencias. Para esto se propone un sistema que utiliza técnicas de Procesamiento de Lenguaje Natural (PLN) y un modelo de recomendación basado en contenido. Los datos de los cursos se obtienen mediante web scraping y se procesan con técnicas de PLN, utilizando un modelo de lenguaje avanzado (BERT) para generar representaciones vectoriales de los cursos y las competencias. El sistema web permite a los docentes autoevaluar sus competencias, visualizar los resultados y recibir recomendaciones personalizadas de cursos MOOC. La implementación de este sistema ha mostrado ser eficaz en la identificación de las áreas de mejora de los docentes y en la sugerencia de cursos relevantes para su desarrollo profesional. Los resultados sugieren que el uso de este sistema puede ser una solución efectiva para la capacitación en competencias digitales de los docentes.

Palabras clave: competencias, inteligencia artificial, pentágono TIC.

Sumario: Introducción, Metodología, Evaluación del Sistema, Discusión y Conclusiones.

Como citar: Maldonado-Mahauad, J., Moscoso, D. & Pacheco, J. (2024). Sistema de Recomendación de Cursos en Línea basado en el Perfil de Competencias TIC del Docente. *Revista Tecnológica - Espol*, 36(2), 196-214. Recuperado a partir de <https://rte.espol.edu.ec/index.php/tecnologica/article/view/1201>

Abstract

Digital transformation in education has created the need to train teachers and improve their digital competencies. However, training processes are often ineffective due to the lack of personalization in the proposals and the variety in teacher profiles. Given this, this article addresses the development of a web tool designed to diagnose teachers' digital competencies and provide personalized recommendations for MOOCs to support the development of their skills. To achieve this, a system that uses Natural Language Processing (NLP) techniques and a content-based recommendation model is proposed. Course data is obtained through web scraping and processed with NLP techniques, using an advanced language model (BERT) to generate vector representations of the courses and competencies. The web system allows teachers to self-assess their competencies, visualize the results, and receive personalized recommendations for MOOCs. The implementation of this system has proven effective in identifying teachers' areas for improvement and suggesting relevant courses for their professional development. The results suggest that the use of this system can be an effective solution for training teachers in digital competencies.

Keywords: pedagogical tools, artificial intelligence, digital competencies, educational technology.

Introducción

La era digital ha provocado una transformación significativa en diversos aspectos de la sociedad, con la educación como uno de los sectores más afectados. Este cambio profundo ha impulsado una revisión de las metodologías de enseñanza y aprendizaje, situando a las Tecnologías de la Información y la Comunicación (TIC) en el núcleo de la interacción educativa (Figueredo, 2013). La incorporación de las TIC en la educación no solo favorece un enfoque más interactivo y dinámico en los procesos educativos, sino que también exige una evolución en el papel del docente. En este contexto, los educadores se enfrentan al reto de adaptarse a estos nuevos entornos digitales, lo que requiere el desarrollo y perfeccionamiento constante de sus competencias digitales y relacionadas con las TIC.

No obstante, las instituciones educativas enfrentan serios desafíos en el proceso de capacitar eficazmente a los docentes en el uso de las TIC. Estos retos incluyen desde la carencia de recursos tecnológicos adecuados, como computadoras y acceso a Internet, hasta limitaciones en la formación y desarrollo profesional de los docentes en el ámbito tecnológico (Ayabaca et al., 2019). Además, uno de los obstáculos más complejos es la falta de una referencia estandarizada que permita evaluar con precisión el nivel inicial de competencia TIC de los docentes. Esta ausencia de criterios uniformes complica la identificación de necesidades específicas de formación, lo que dificulta la personalización de los programas de capacitación para mejorar las habilidades digitales. La disparidad en los niveles de competencia digital entre docentes de diferentes regiones, instituciones e incluso dentro de la misma institución subraya la importancia de enfoques adaptativos en la capacitación docente.

Para superar estos desafíos, se han desarrollado diversos modelos adaptados a los distintos contextos regionales, con el objetivo de establecer una base para evaluar y mejorar las competencias digitales de los docentes. Entre estos, destacan los “Estándares ISTE para educadores” publicados por la Sociedad Internacional de Tecnología Educativa de Estados Unidos (ISTE) en 2008, que proporcionan directrices para la integración de las TIC en la práctica docente (Trust, 2018). En Europa, el Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado (INTEF) ha propuesto el “Marco Común de Competencia Digital Docente”, alineado con el Marco Europeo de Competencia Digital para Ciudadanos y

Educadores, el cual sirve como una herramienta para identificar las necesidades de formación en el ámbito digital (INTEF, 2023; Redecker, 2022). En América Latina, el Ministerio de Educación de Colombia (MEN, 2013) presenta un modelo enfocado en cinco competencias TIC fundamentales.

Los modelos antes mencionados brindan un marco de referencia para establecer una base que permita evaluar las competencias digitales de los docentes. Sin embargo, el verdadero valor de estas evaluaciones radica en utilizar sus resultados para apoyar la formación docente. Si bien es importante medir las competencias TIC, es aún más crucial actuar sobre esa información. No obstante, organizar capacitaciones que lleguen a grandes grupos de docentes y que aborden las competencias y el nivel requerido para cada uno según su perfil evaluado es costoso y complicado para las instituciones educativas. Esto evidencia la necesidad de una herramienta tecnológica que permita ayudar a la formación docente de manera masiva, automática y personalizada. Para asegurar que este sea un proceso automatizado que no requiera mayores esfuerzos organizativos, se pueden utilizar técnicas de Procesamiento de Lenguaje Natural (PLN) para analizar contenidos en la web y determinar su pertinencia con respecto a los objetivos de formación planteados.

Este trabajo propone el desarrollo de una herramienta web diseñada para evaluar y presentar el perfil de competencias digitales de los docentes, y utilizar esos datos para recomendar automáticamente cursos MOOC que ayuden a mejorar cada competencia evaluada. Para lograr esto, se define un marco de referencia para la evaluación de competencias digitales, se recolectan y preprocesan datos de cursos utilizando técnicas de PLN para identificar su relevancia en relación con cada competencia, y se generan representaciones vectoriales tanto de los cursos como de las descripciones de las competencias. A partir de estas representaciones, se calcula la similitud entre ellas, se implementa un modelo de clasificación personalizado y se optimizan sus hiperparámetros. Los cursos se clasifican según su nivel cognitivo para asegurar recomendaciones personalizadas. Finalmente, se implementa un sistema web que realiza búsquedas automáticas y periódicas de cursos, actualizando continuamente las recomendaciones, y permitiendo a los docentes evaluar sus competencias y acceder a sugerencias personalizadas.

El artículo se encuentra estructurado de la siguiente manera: en la sección 2 se describe la metodología utilizada para el desarrollo del sistema de recomendación. La sección 3 presenta los resultados de la evaluación del sistema y finalmente la sección 4 presenta las principales conclusiones de este trabajo.

Metodología

En la Figura 1 se detalla la metodología empleada para desarrollar el sistema de recomendación de cursos utilizando un enfoque de filtrado basado en contenido. El enfoque de filtrado basado en contenido elige los elementos a recomendar basándose en la correlación entre el contenido de esos elementos y las preferencias o necesidades previamente identificadas del usuario (Aggarwal, 2016). En este caso, no se requiere información de otros docentes ni de su interacción con otros cursos; lo único necesario es la correlación entre el contenido de los cursos y las descripciones de las competencias a evaluar, que representan las necesidades identificadas. Esta metodología se divide en tres etapas o secciones principales que son:

- Etapa 1: Preparación de los datos.
- Etapa 2: Implementación del Modelo de Recomendación.
- Etapa 3: Implementación del Sistema.

Figura 1

Metodología aplicada para el desarrollo del sistema de recomendación



Fase 1: Preparación de los datos

Definición del Marco de Referencia

El primer paso en esta primera etapa es definir el marco de referencia para la evaluación de las competencias digitales. Después de revisar varios marcos (ver sección 1), se optó por utilizar el Pentágono de Competencias TIC para el Desarrollo Profesional Docente del Ministerio de Educación Nacional (MEN) de Colombia (MEN, 2013). El Pentágono de Competencias ha sido específicamente desarrollado teniendo en cuenta el contexto latinoamericano, lo que lo hace más relevante y aplicable a las realidades y necesidades específicas de nuestra región. Este modelo se centra en cinco competencias TIC principales que cubren una amplia gama de habilidades necesarias para la práctica docente efectiva en el ámbito digital, asegurando una formación completa y equilibrada para los docentes. Además, el modelo ha sido reconocido y validado por instituciones educativas de la región, garantizando su fiabilidad y efectividad en la mejora de las competencias digitales de los docentes. La implementación de este modelo ha mostrado resultados positivos en la identificación de áreas de mejora y en la capacitación efectiva de los docentes en competencias digitales, lo que respalda su elección sobre otros marcos que pueden no estar tan ajustados a las necesidades locales. Este marco sugiere un camino y guía para que los docentes utilicen las TIC de forma estructurada, a través de la evaluación y el desarrollo de cinco competencias que se describen en la Tabla 1.

Tabla 1

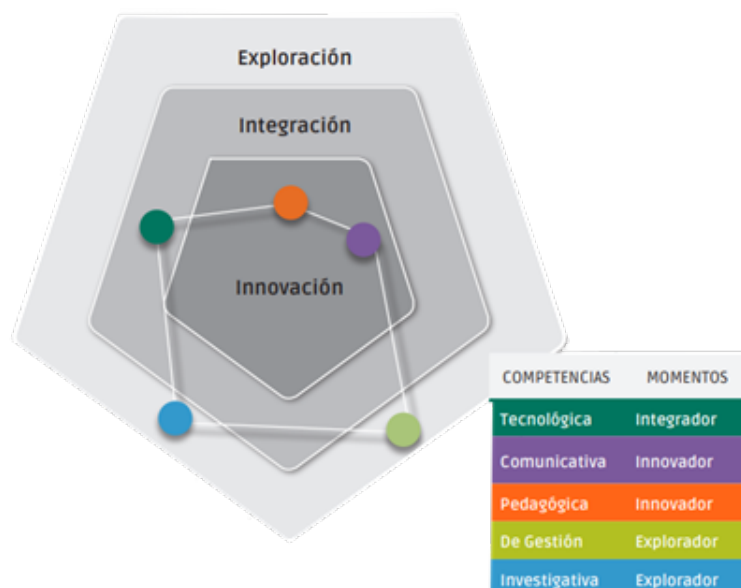
Definición de las competencias digitales según el marco del MEN de Colombia

COMPETENCIA	DESCRIPCIÓN
Tecnológica	Comprensión de las herramientas tecnológicas, desarrollo de contenidos y consideración ética en la utilización de las TIC.
Comunicativa	Capacidades de los docentes para transmitir y compartir información en entornos mediados por las TIC.
Pedagógica	Aprecia el aprendizaje, la resolución de problemas educativos y el diseño de estrategias didácticas apoyadas por las TIC.
De gestión	Aplicación de las TIC en los procesos educativos, las políticas y el crecimiento profesional mediante el uso de TIC.
Investigativa	Comprueba que las actividades de investigación fomenten la reflexión sobre el uso de las TIC, las fuentes de información, el acceso a información a través de TIC y la creación de nuevos conocimientos.

Estas competencias se desarrollan y manifiestan en diferentes niveles o etapas de complejidad. En la primera etapa, conocida como momento de exploración, se facilita el acceso a un conjunto de conocimientos básicos que sirven como fundamento para avanzar hacia niveles más profundos de comprensión. Posteriormente, en la segunda etapa o momento de integración, se promueve la aplicación de los conocimientos adquiridos para resolver problemas en diversos contextos. Finalmente, en la tercera etapa o momento de innovación, se destaca la creatividad, permitiendo ir más allá de lo aprendido y concebir nuevas formas de acción o explicación. En la Figura 2 se muestra un ejemplo del resultado de evaluar las competencias digitales según este modelo.

Figura 2

Resultado de la evaluación de competencias según el modelo del MEN de Colombia



Obtención de los datos de cursos

Una de las principales motivaciones para este trabajo es la necesidad de apoyar la formación docente con contenidos actualizados, al mismo tiempo que se minimizan los esfuerzos de la universidad en términos de tiempo y presupuesto. Una manera de automatizar el proceso de recolección de datos de cursos es emplear una técnica utilizada para extraer datos de plataformas web accesibles a través de Internet denominada web scraping. Esto implica el uso de scripts que recorren páginas web y recopilan información específica de manera automatizada (Patel, 2020). Por tal motivo, la recolección de datos se llevó a cabo utilizando herramientas como Python y bibliotecas como BeautifulSoup y Requests, respetando los términos de servicio de los sitios web objetivo. Las plataformas de cursos MOOC incluyeron sitios como Coursera y Udemy. El proceso consistió en extraer información relevante de cada curso, incluyendo títulos, descripciones y enlaces, asegurando la limpieza y estructuración adecuada de los datos recolectados para su posterior procesamiento. Al tratarse de un proceso calculado se utilizaron términos de búsqueda específicos relacionados con las competencias digitales evaluadas. Estos términos fueron seleccionados cuidadosamente para abarcar una amplia gama de cursos pertinentes y asegurar que los resultados fueran relevantes.

Preprocesamiento y filtrado inicial de los cursos

Para preprocesar los textos de los cursos, se aplicó una serie de pasos de procesamiento de lenguaje natural (PLN) tanto al título como a la descripción. Primero, se llevó a cabo la tokenización, un proceso que divide el texto en unidades más pequeñas llamadas tokens, que

pueden incluir palabras, partes de palabras, caracteres, números, signos de puntuación o símbolos (Lee, 2023). Luego, se aplicó la lematización, que convierte cada palabra a su forma base llamada lema de la palabra (Khyani & B S, 2021) (por ejemplo, "corriendo" se convierte en "correr"). Adicionalmente, se eliminaron las stop words, que son palabras comunes y de poca relevancia (como "y", "el", "de") que no aportan significado al análisis (Kedia & Rasu, 2020). Estos pasos permitieron obtener un conjunto de palabras representativo para cada curso, facilitando su posterior análisis y filtrado.

Durante la obtención automatizada de los cursos, se recolectaron varios cursos que no correspondían a las competencias debido a la naturaleza de los motores de búsqueda en las páginas de cursos, que en ocasiones devuelven resultados irrelevantes. Para abordar este problema, se implementó un proceso de filtrado inicial utilizando términos clave obtenidos de las descripciones de las cinco competencias evaluadas. Para identificar estos términos clave, se utilizó el método TF-IDF (Term Frequency-Inverse Document Frequency), que es ampliamente utilizado en el análisis de textos para resaltar palabras significativas en grandes conjuntos de datos (Leskovec, Rajaraman, & Ullman, 2014). Este método asegura que los términos seleccionados sean verdaderamente representativos de las competencias.

A partir de las descripciones de las competencias, se generó una lista de palabras clave para cada una de ellas. Estas palabras clave se compararon con el conjunto de palabras obtenido de cada curso durante el preprocesamiento. Los cursos se filtraron en función de las coincidencias entre las palabras clave correspondientes a la competencia y el conjunto de palabras de cada curso. Aquellos cursos que no contenían ninguna de las palabras relevantes fueron descartados, asegurando que solo los cursos pertinentes a las competencias evaluadas fueran considerados en las etapas posteriores del análisis. Esta estrategia de filtrado inicial no solo mejora la calidad de los datos, sino que también optimiza los recursos al eliminar cursos irrelevantes desde el inicio.

Etiquetado manual

Se creó un conjunto de datos, extrayendo los 100 cursos con las coincidencias más altas para cada competencia con el fin de facilitar el proceso de etiquetado manual, garantizando que el conjunto de datos fuera manejable y representativo. El objetivo del etiquetado manual es evaluar la pertinencia de cada curso en relación con la competencia para la cual fue seleccionado. Los especialistas revisaron los títulos y descripciones de los cursos y los clasificaron manualmente como "pertinentes" o "no pertinentes". El resultado de este procedimiento fue la creación de un conjunto de datos etiquetados denominado gold estándar, que representa el resultado que se espera del modelo y que sirvió como base para el entrenamiento y evaluación del modelo de recomendación.

Fase 2: Modelo de recomendación

En esta etapa se implementó el modelo de recomendación, para lo cual se utilizó un modelo de lenguaje para convertir tanto las definiciones de las competencias como los títulos y descripciones de los cursos en vectores numéricos denominados embeddings y posteriormente calcular la similitud entre estos embeddings. Usando estos valores de similitud entre los cursos y su respectiva competencia, se implementó un modelo de clasificación que decide la pertinencia del curso. Finalmente, para asegurar la personalización del curso, se implementó una clasificación según el nivel cognitivo de los cursos. En la Figura 3 se explica de manera gráfica el proceso de implementación del modelo de clasificación.

Generación de Embeddings

Para representar mediante embeddings los títulos y descripciones de los cursos, así como las descripciones de las competencias, se utiliza BERT (Bidirectional Encoder Representations from Transformers), el cual es un modelo de lenguaje avanzado que destaca por poder generar representaciones contextuales de las palabras, ya que considera el contexto en el que aparece cada palabra dentro del texto (Devlin et al., 2019). Este modelo tiene un tamaño comparable al de BERT-Base y fue entrenado utilizando la técnica de Whole Word Masking (Cañete et al., 2023).

Para este proceso, no es necesario aplicar ningún tipo de preprocesamiento de textos, ya que puede alterar el contexto, lo cual no es recomendable para ser procesado por BERT. En primer lugar, se utiliza la tokenización específica de BERT, que divide el texto en tokens y añade tokens especiales necesarios para el modelo. Finalmente, para obtener los embeddings, estos tokens se pasan a través del modelo BERT pre entrenado (dccuchile/bert-base-spanish-wwm-uncased), generando un vector de alta dimensión para cada token. Estos vectores capturan información semántica rica sobre las palabras y su contexto en el texto y se promedian para obtener una representación final del texto. Se ha optado por este método de agregación debido a su capacidad para obtener una representación general del texto completo.

Cálculo de similitudes

Evaluar la similitud entre dos textos requiere convertir el texto en un conjunto de números y aplicar operaciones matemáticas para calcular la distancia entre dos vectores, lo que indica su relación. Distancias pequeñas sugieren una alta relación, mientras que distancias grandes indican una baja relación. La similitud coseno es una medida utilizada para evaluar la similitud entre dos vectores en un espacio multidimensional.

Se realizan dos cálculos de similitud por curso: (1) para el título; (2) para la descripción del curso. La función `cosine_similarity` de Sklearn se utiliza para calcular la similitud del coseno. Esta función toma dos vectores como entrada y devuelve un valor de similitud. Este valor puede variar entre -1 y 1, donde los valores cercanos a 1 indican que los vectores son similares, mientras que los valores cercanos a -1 indican que son diferentes. Los resultados obtenidos son la similitud entre el título del curso y la definición de la competencia, así como la similitud entre la descripción del curso y la definición de la competencia.

Implementación del modelo de clasificación

Se implementa un modelo de clasificación personalizado, el cual utiliza los valores de la similitud del coseno calculados para los títulos y descripciones de los cursos en relación con las competencias. Para combinar estos dos valores de similitud en una única puntuación, se calcula un promedio ponderado. El peso del título se denota como w , y el peso de la descripción se calcula como $1-w$.

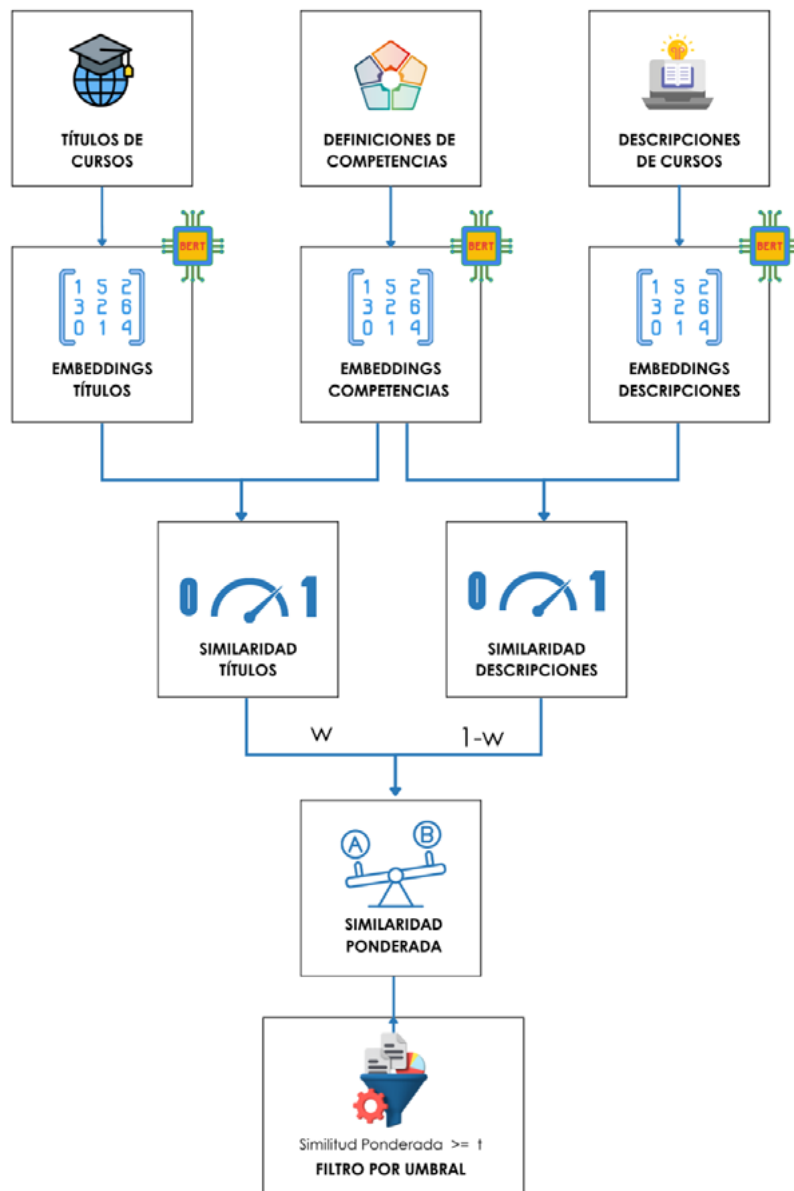
Basándose en la similitud ponderada, el modelo clasifica un curso como pertinente o no pertinente para una competencia. Para tomar esta decisión, se define un umbral denotado como t , que determina el punto de corte para considerar un curso como pertinente. Si la similitud ponderada es mayor o igual al umbral, el curso se clasifica como pertinente.

Optimización y evaluación del modelo

En esta sección se entrenará el modelo con los datos etiquetados para definir los hiperparámetros óptimos para este modelo de clasificación. Al realizar un análisis de los datos etiquetados se pudo observar que las etiquetas están muy desbalanceadas, por lo que se aplicó la técnica de submuestreo al conjunto de datos etiquetados para evitar el sobreajuste y asegurar

un entrenamiento equilibrado del modelo. Esta técnica se utiliza para reducir la cantidad de instancias de la clase mayoritaria, lo que ayuda a equilibrar los conjuntos de datos y mejorar la precisión de los modelos predictivos (Ali et al., 2015). En este caso, la clase mayoritaria son los cursos etiquetados como no pertinentes y la clase minoritaria son los etiquetados como cursos pertinentes. Con este submuestreo, se asegura que el modelo no esté sesgado hacia la clase mayoritaria, es decir, que el modelo no sea propenso a clasificar siempre como no pertinentes a los cursos.

Figura 3
Proceso de implementación del modelo de recomendación



Con el objetivo de buscar los hiperparámetros que brinden un mejor rendimiento para el modelo, se aplicó Grid Search en combinación con validación cruzada. Esta combinación de técnicas es ampliamente utilizada debido a su capacidad para buscar los hiperparámetros óptimos de manera rigurosa y sistemática. Según un artículo en Scikit-learn documentation, este enfoque asegura que los modelos sean evaluados exhaustivamente, utilizando solo una parte del conjunto de datos para la selección de hiperparámetros y otra parte para la evaluación final, lo que reduce el riesgo de sobreajuste (Scikit-learn, 2018).

En la Tabla 2 se observa la descripción de los hiperparámetros a optimizar para este modelo, definiendo los rangos que pueden tener y el incremento que va a ser utilizado para la búsqueda exhaustiva. Para este proceso se definió un conjunto de todas las posibles combinaciones de w y t , para posteriormente realizar una validación cruzada k -fold. En este caso se utilizó un $k=5$; esto implica dividir el conjunto de datos en cinco pliegues para entrenar al modelo con cuatro pliegues y probar con el pliegue restante, reiterando este proceso cinco veces para cada combinación de hiperparámetros.

Para cada combinación de hiperparámetros se obtuvieron las métricas que indican el rendimiento del modelo. Los resultados de estas métricas en las iteraciones de la validación cruzada se promedian para obtener las métricas representativas para cada combinación de hiperparámetros, lo que permite identificar la configuración óptima que maximiza el rendimiento del modelo. Existen diversas métricas para evaluar un modelo y depende de la naturaleza del problema abordado cual es la métrica que se requiere maximizar. En el contexto de esta clasificación de cursos se optó por maximizar la métrica de precisión, ya que en este caso es crucial minimizar el número de falsos positivos, es decir, evitar que se recomienden cursos que no sean relevantes para las competencias definidas. Esta métrica mide la proporción de verdaderos positivos entre todos los clasificados como positivos, y puede ser especialmente útil cuando el objetivo es asegurar que los elementos etiquetados positivamente sean correctos (Saito & Rehmsmeier, 2015). Esta métrica es particularmente relevante en conjuntos de datos desequilibrados, como es el caso en este problema, donde la cantidad de cursos pertinentes es mucho menor que la cantidad de cursos no pertinentes.

Tabla 2
Hiperparámetros del modelo

HIPERPARÁMETRO	DESCRIPCIÓN	RANGO	INCREMENTO
Peso del título	Determina la importancia relativa de la similitud del título en comparación con la similitud de la descripción.	[0,1]	0.1
Umbral	Define el valor mínimo de la similitud ponderada para que un curso sea considerado pertinente.	[0,1]	0.1

La Tabla 3 presenta los mejores resultados obtenidos durante la primera pasada del Grid Search, una técnica utilizada para optimizar la selección de hiperparámetros en el modelo. En este contexto, los hiperparámetros más efectivos fueron un peso del título (w) de 0.2 y un umbral (t) de 0.8. Esta combinación proporcionó una precisión de 0.596, lo que indica que, de las recomendaciones generadas, una proporción considerable fue correcta. Sin embargo, esta mejora en la precisión se obtuvo a costa de una disminución en el recall, que fue de 0.381.

El recall es crucial para asegurar que el modelo identifique la mayoría de los cursos relevantes, y una disminución en esta métrica implica que hay más cursos relevantes que no fueron recomendados (falsos negativos). No obstante, en este escenario, se decidió priorizar la precisión sobre el recall. La razón detrás de esta decisión es que se busca asegurar que los cursos que se recomiendan sean altamente pertinentes para los usuarios, incluso si eso significa no recomendar todos los cursos que podrían ser relevantes. Este enfoque es aceptable, ya que es preferible evitar la inclusión de cursos no pertinentes en las recomendaciones, lo cual podría disminuir la confianza en el sistema. Al asegurar la alta calidad de las recomendaciones, se mantiene la relevancia y efectividad del sistema, alineándose con el objetivo de proporcionar a los docentes cursos que realmente contribuyan a mejorar sus competencias.

Además, el análisis de los resultados sugiere que, aunque existe un trade-off entre precisión y recall, la estrategia seleccionada optimiza el rendimiento del modelo de acuerdo con las prioridades del problema en cuestión, donde la precisión es más crítica para el éxito del sistema de recomendación. Sin embargo, es importante tener en cuenta que futuras iteraciones del modelo podrían explorar un balance diferente o buscar formas de mejorar tanto la precisión como el recall, dependiendo de las necesidades específicas del sistema y el feedback de los usuarios.

Tabla 3
Mejores Resultados Grid Search (primera pasada)

PESO DEL TITULO	UMBRAL	PRECISIÓN	RECALL	F1-SCORE	ACCURACY
0.2	0.8	0.596	0.381	0.454	0.789
0.3	0.8	0.496	0.147	0.22	0.777
1	0.7	0.466	0.103	0.168	0.766
0.1	0.8	0.458	0.536	0.485	0.726
0.8	0.7	0.457	0.259	0.325	0.748
0.9	0.7	0.434	0.155	0.227	0.756
0.7	0.7	0.406	0.397	0.399	0.724
0.5	0.8	0.4	0.017	0.033	0.769
0.4	0.8	0.35	0.043	0.076	0.769
0.6	0.7	0.325	0.543	0.404	0.63

Después de los resultados obtenidos en la primera pasada del Grid Search, se decidió realizar una segunda pasada para afinar aún más la precisión del modelo. En esta iteración, se utilizó un rango más específico de valores de los hiperparámetros, centrado alrededor de los valores previamente identificados como prometedores. El objetivo de esta segunda búsqueda era optimizar con mayor precisión los hiperparámetros clave, asegurando que se exploraran las configuraciones más cercanas al óptimo para mejorar el rendimiento del modelo.

Como resultado de esta segunda búsqueda, los valores óptimos identificados fueron un peso del título (w) de 0.215 y un umbral (t) de 0.811. Como se muestra en la Tabla 4, estos hiperparámetros produjeron una precisión de 0.624, una mejora notable respecto a los resultados anteriores. Aunque el recall no mostró una mejora significativa y se mantuvo en valores similares a los de la primera pasada (0.251), la precisión fue la métrica que experimentó el mayor beneficio.

La mejora en la precisión significa que el modelo es ahora más eficaz en la identificación de cursos altamente relevantes para los usuarios, minimizando la inclusión de cursos que no cumplen con los criterios deseados. Este resultado es especialmente relevante para el contexto en el que se aplica, donde la exactitud en las recomendaciones es prioritaria. La ligera reducción en el recall es un costo aceptable en este caso, ya que la estrategia sigue siendo asegurar que los cursos recomendados sean lo más pertinentes posible, aún si esto implica no captar todos los cursos relevantes disponibles.

Además, la Tabla 4 muestra que la precisión lograda con estos hiperparámetros también está acompañada por un F1-Score de 0.349 y una precisión global (accuracy) de 0.796, lo que

indica un buen balance general entre las métricas de rendimiento. Estos resultados sugieren que el ajuste fino realizado en esta segunda pasada ha permitido al modelo alcanzar un nivel más elevado de rendimiento, adecuándose mejor a los objetivos del sistema de recomendación.

Finalmente, es importante destacar que este proceso iterativo de ajuste y optimización es fundamental para mejorar continuamente la calidad de las recomendaciones, y futuras iteraciones podrían continuar afinando los hiperparámetros para adaptarse a cambios en los datos o en las necesidades de los usuarios.

Tabla 4
Mejores Resultados Grid Search (segunda pasada)

PESO DEL TITULO	UMBRAL	PRECISIÓN	RECALL	F1-SCORE	ACCURACY
0.215	0.811	0.624	0.251	0.349	0.796
0.216	0.811	0.618	0.243	0.339	0.794
0.221	0.81	0.618	0.243	0.339	0.794
0.227	0.808	0.611	0.243	0.338	0.792
0.24	0.805	0.611	0.243	0.338	0.792
0.232	0.807	0.611	0.243	0.338	0.792
0.231	0.807	0.611	0.243	0.338	0.792
0.239	0.805	0.611	0.243	0.338	0.792
0.223	0.809	0.611	0.243	0.338	0.792
0.228	0.808	0.611	0.243	0.338	0.792

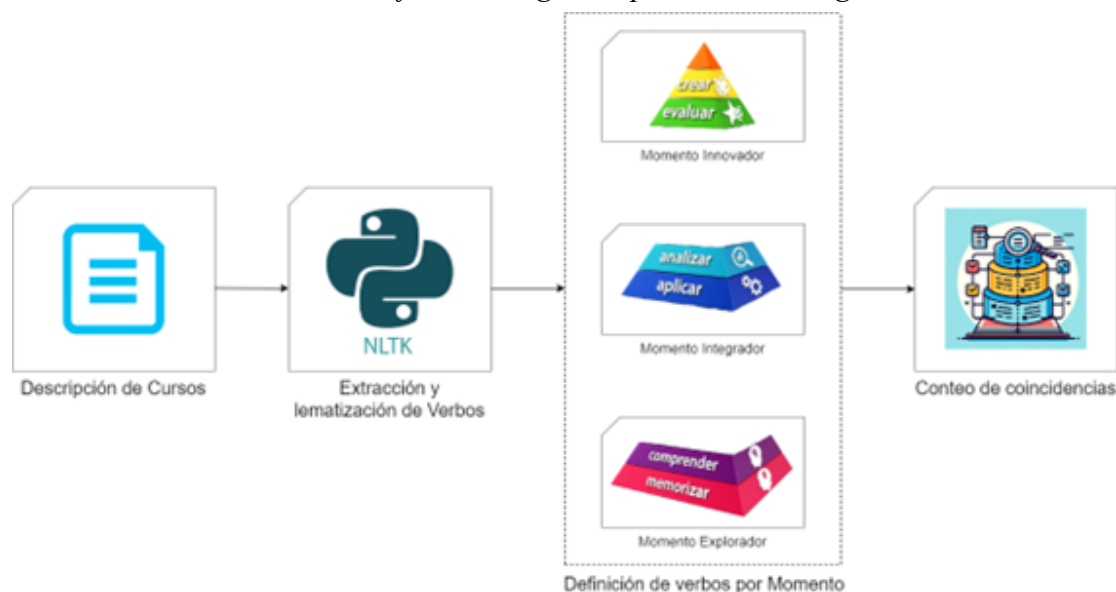
Los resultados de este proceso de optimización y evaluación muestran que el modelo actual proporciona recomendaciones de moderada precisión, lo cual es adecuado para las necesidades actuales del sistema, pero es importante trabajar a futuro en la optimización de este modelo para poder asegurar que se recomienden todos los cursos pertinentes posibles y se minimicen los falsos negativos en futuras iteraciones del sistema.

Clasificación utilizando la clasificación de Bloom

La clasificación de Bloom, creada por Benjamin Bloom en 1956, es un marco para clasificar los objetivos educativos, que ha sido ampliamente utilizado en el ámbito educativo desde su desarrollo. Los educadores emplean esta clasificación o taxonomía para establecer objetivos de aprendizaje, diseñar actividades de enseñanza y evaluación, y garantizar que se aborden diferentes niveles cognitivos. En su artículo "A Revision of Bloom's Taxonomy: An Overview", Krathwohl (2002) enfatiza la relevancia de la taxonomía revisada en la evaluación y planificación educativa. Esta taxonomía inicialmente clasifica el aprendizaje cognitivo en seis niveles: Conocer, Entender, Aplicar, Analizar, Evaluar y Crear. En el contexto de la educación en línea, la taxonomía de Bloom se considera una herramienta útil para clasificar y analizar la profundidad cognitiva de los cursos MOOCs. En este estudio, se adaptó la taxonomía de Bloom para categorizar los cursos en tres etapas de competencia según el marco referencial seleccionado: "Explorador", "Integrador" e "Innovador", representando un incremento progresivo en la complejidad cognitiva. Para realizar esta clasificación, se siguió el proceso que se ilustra en la Figura 4:

Figura 4

Proceso de clasificación según etapas o niveles cognitivos



- Definición de Niveles Cognitivos para los momentos del Pentágono TIC: Con el fin de simplificar la taxonomía de Bloom y ajustarla a nuestro contexto, se combinaron sus seis niveles cognitivos en tres categorías (que corresponden con los momentos), como se muestra en la Tabla 5.

Tabla 5

Taxonomía de Bloom y categorías del Pentágono TIC

MOMENTOS	NIVEL COGNITIVO DE BLOOM COMBINADO	DETALLE
Explorador	Conocer y Entender	Cursos enfocados en la introducción y comprensión fundamental de conceptos.
Integrador	Aplicar y Analizar	Cursos que fomentan la aplicación de conocimientos y el análisis crítico.
Innovador	Evaluar y Crear	Cursos avanzados que impulsan la evaluación crítica y la generación de nuevos conocimientos o productos.

- Identificación de Verbos Clave para cada momento: Se seleccionó un conjunto de verbos asociados originales de la taxonomía de Bloom. Estos verbos fueron lematizados y sirven como indicadores del nivel cognitivo que se espera que el curso logre.
- Procesamiento de Descripciones de Cursos: Las descripciones de los cursos disponibles en las plataformas fueron procesadas para extraer los verbos clave, utilizando la biblioteca NLTK de Python. Esta herramienta permitió llevar a cabo dos procesos fundamentales: la lematización de verbos y la eliminación de “stop words”. La lematización redujo los verbos a su forma base, lo cual es crucial para un análisis comparativo preciso, mientras que la eliminación de “stop words” ayudó a centrarse en las palabras más relevantes de las descripciones.
- Comparación y Asignación de Momentos: Utilizando los verbos lematizados

extraídos de las descripciones junto con los verbos clave definidos para cada nivel cognitivo, se realizó una comparación detallada. A cada curso se le asignó el nivel correspondiente basado en la mayor frecuencia de verbos asociados. Este enfoque cuantitativo garantiza una clasificación objetiva y sistemática de los cursos según su nivel de profundidad cognitiva.

Fase 3: Implementación del sistema

En esta etapa final se implementó un sistema web que permite a los docentes acceder a realizar una autoevaluación de sus competencias, ver sus resultados y obtener las recomendaciones de cursos para capacitarse en sus competencias digitales. Además, este sistema utiliza técnicas de web scraping para obtener nuevos cursos y utiliza el modelo generado en la etapa 2 para filtrar los cursos a recomendarse y almacenarlos en la base de datos. La arquitectura de este sistema se describe de manera gráfica en la Figura 5.

Figura 5
Arquitectura del sistema de recomendación

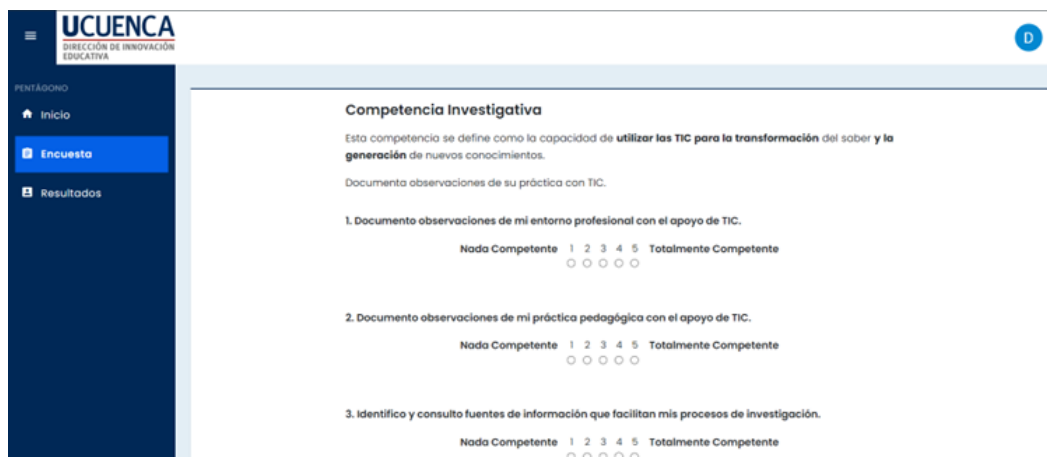


Autoevaluación de capacidades

A partir del instrumento adaptado, se implementó en el sistema una herramienta de evaluación adaptada al contexto universitario para medir las competencias TIC de los docentes. Este instrumento evalúa competencias en áreas tecnológicas, pedagógicas, comunicativas, de gestión e investigativas. Consta de 80 preguntas y utiliza una escala de calificación basada en la escala de Likert (“1- Nada Competente” a “5- Totalmente competente”). Además, el cuestionario recoge información sobre la facultad y la carrera en las que el docente imparte clases, así como su número de identificación personal. La interfaz de usuario para esta sección se muestra en la Figura 6.

Entre las preguntas diseñadas para cada competencia, se incluyen preguntas específicas que describen los diferentes momentos o niveles en los que un docente puede encontrarse. De este modo, para cada competencia, se ubicará al docente en el nivel correspondiente al grupo de preguntas en el que obtuvo la mayor calificación promedio.

Figura 6
Interfaz para completar la autoevaluación



Resultados y recomendaciones

Después de que los docentes completen la autoevaluación, podrán acceder a sus resultados en la sección de resultados, la cual muestra una representación de su nivel de competencia TIC. Para visualizar los resultados, se utilizó un gráfico de radar en forma de pentágono, que consta de tres niveles. Cada punta del pentágono representa una competencia, y cada nivel indica el momento en que se encuentra el docente, siendo el nivel más interno el Momento Explorador, el nivel intermedio el Momento Integrador, y el nivel más externo el Momento Innovador. Además, se ofrece una interpretación detallada de cada competencia y sus respectivos niveles, permitiendo a los docentes comprender los resultados obtenidos en la encuesta (ver Figura 7).

Figura 7
Vista de la interfaz de usuario para visualizar los resultados de la evaluación



La sección de cursos recomendados (ver Figura 8) muestra una lista de todos los cursos que se recomiendan al profesor y que están alineados con su nivel de competencias TIC. Los cursos recomendados están organizados en grupos de 3 por cada competencia. Cada curso es accesible y presenta una imagen, un título, la descripción breve, puntuación, creador y la plataforma de origen.

Figura 8

Vista de la interfaz de usuario para visualizar las recomendaciones de los cursos



Búsqueda automatizada de cursos

El sistema realiza una búsqueda automatizada de nuevos cursos en las plataformas MOOC de forma periódica. Esto se realiza para mantener actualizada la base de datos de cursos para recomendar utilizando técnicas de web scraping y los términos de búsqueda definidos. Para esto se implementó un script que realiza el web scraping de manera periódica y se utilizan técnicas de PLN y el modelo de clasificación implementado para definir la pertinencia de los cursos de manera automatizada.

Evaluación del Sistema

En esta sección se presenta la evaluación del sistema de recomendación, para esto se utilizó dos criterios clave: (1) usabilidad, y (2) calidad de la recomendación. Para evaluar (1) la usabilidad, se utilizó el cuestionario System Usability Scale (SUS) (Brooke, 1995), que consta de 10 preguntas con cinco opciones de respuesta. Para evaluar (2) la calidad de la recomendación, se diseñó un cuestionario específico de 3 preguntas que permite analizar el grado de satisfacción de los docentes con respecto a los cursos sugeridos. La lista de preguntas se muestra en la Tabla 6. La evaluación fue realizada por una muestra de $n=140$ docentes de la Universidad de Cuenca utilizando Google Forms.

Tabla 6

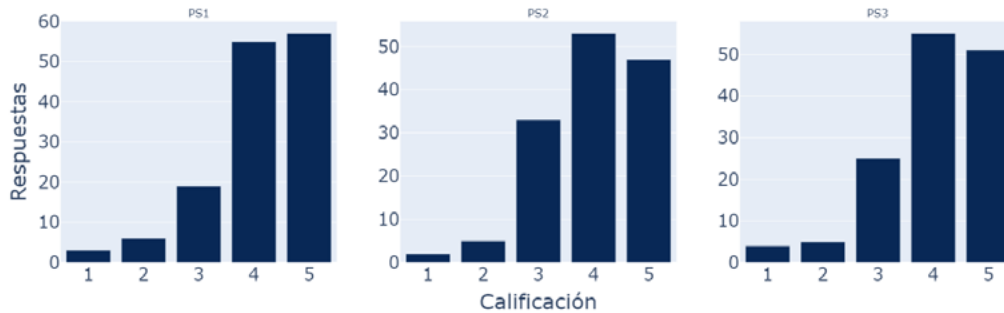
Preguntas para medir el grado de satisfacción de los docentes (calidad de la recomendación)

IDENTIFICADOR	PREGUNTA	MÉTRICA MÍNIMA	MÉTRICA MÁXIMA
PS1	¿Cómo calificaría la relevancia de los cursos recomendados?	Nada relevante.	Muy relevante.
PS2	En términos de ajuste a sus intereses y necesidades, ¿cómo calificaría las recomendaciones?	Nada útil.	Muy útil.
PS3	¿Considera que los cursos recomendados están alineados a las cinco competencias TIC (tecnológica, comunicativa, de gestión, investigativa y pedagógica)?	No alineadas en absoluto.	Totalmente alineadas.

En cuanto al cuestionario de satisfacción sobre la calidad de la recomendación de los cursos, la tendencia general mostrada en la Figura 9 revela que las respuestas tienden a concentrarse en las calificaciones más altas (4 y 5).

Figura 9

Agrupación de las respuestas por cada pregunta del cuestionario de satisfacción



En la Tabla 7, se puede observar que todos los promedios para cada una de las preguntas superan el valor de 3.5 (media), lo que sugiere una percepción generalmente positiva por parte de los docentes sobre los cursos recomendados. Los valores de la desviación estándar reflejan una ligera variabilidad en las respuestas para cada pregunta, lo que indica que, aunque la percepción es favorable, hay algunas diferencias en cómo los docentes valoran ciertos aspectos. En general, los resultados muestran que los cursos recomendados son vistos como relevantes, bien ajustados a sus intereses y necesidades, y alineados con las competencias TIC. En relación con el cuestionario SUS, el puntaje medio de usabilidad registrado es de aproximadamente 62.27, con una mediana de 60. Según la Figura 10, que proporciona una interpretación de los resultados del cuestionario, la herramienta recibe una calificación de 'D' y un adjetivo descriptor de 'OK'. En términos de aceptabilidad, se clasifica como 'Marginal'. La desviación estándar, que se sitúa en torno a 16.96, indica una variabilidad moderada en los puntajes de usabilidad entre los docentes.

Tabla 7

Resultados del cuestionario de satisfacción

IDENTIFICADOR	MEDIA	DESVIACIÓN ESTÁNDAR
PS1	4.12	0.95
PS2	3.98	0.92
PS3	4.03	0.97

Figura 10

Formas de interpretar el puntaje SUS

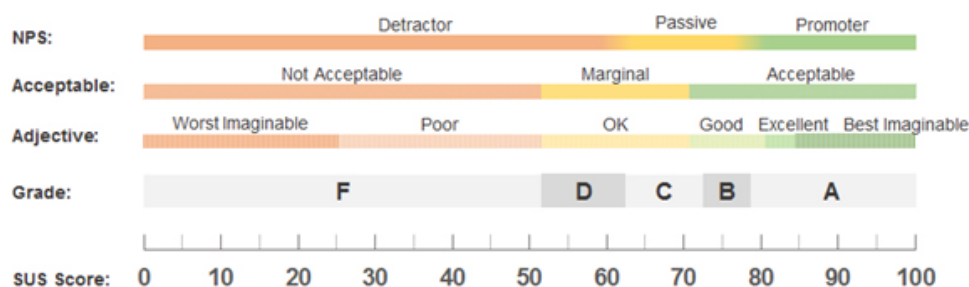
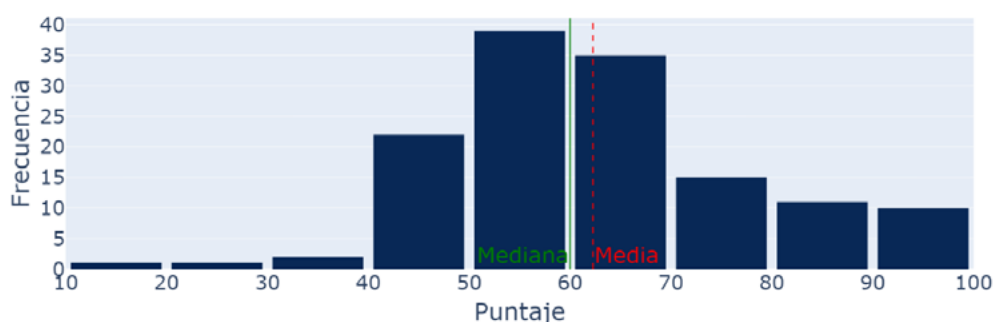


Figura 11

Agrupación de las respuestas por cada pregunta del cuestionario de satisfacción



Como se muestra en la Figura 11, los resultados sugieren que los docentes consideran que el sistema tiene una usabilidad moderada o promedio. Aunque el puntaje no es bajo, indica que existe un margen significativo para mejorar la usabilidad del sistema. Esto implica que la herramienta podría beneficiarse de una revisión para abordar los aspectos que han conducido a valoraciones más bajas. En resumen, mientras los docentes están satisfechos con los cursos recomendados, encuentran dificultades en el uso de la herramienta, lo que subraya la necesidad de mejorar ciertos aspectos de su usabilidad para optimizar la experiencia del usuario.

Discusión y Conclusiones

El campo de las Tecnologías de la Información y la Comunicación (TIC) está en constante evolución, lo que impone la necesidad de una revisión continua de las competencias docentes en este ámbito. A lo largo de este estudio, la implementación de un sistema de recomendación de cursos personalizados ha demostrado ser una herramienta valiosa para apoyar el desarrollo profesional continuo de los docentes, contribuyendo así a la mejora de la calidad educativa en las instituciones.

Los resultados obtenidos sugieren que, si bien los cursos recomendados son percibidos como relevantes y bien alineados con las necesidades y competencias TIC de los docentes, la usabilidad del sistema presenta ciertos desafíos. El análisis de la usabilidad muestra que, aunque la herramienta cumple su función, existe un margen considerable para mejorar la experiencia del usuario, lo cual es esencial para maximizar la efectividad y la adopción de la plataforma.

En términos generales, este sistema no solo facilita la actualización continua de las competencias docentes en un entorno educativo dinámico, sino que también promueve una integración más efectiva de las TIC en la enseñanza. Esto tiene un impacto directo en la calidad de la educación y enriquece la experiencia de aprendizaje de los estudiantes, preparándolos mejor para un mundo cada vez más digital. No obstante, para que el sistema alcance su pleno potencial, es fundamental continuar mejorando tanto la calidad de las recomendaciones como la usabilidad del sistema. Estas mejoras son necesarias para garantizar que la herramienta siga siendo relevante y útil en un contexto educativo en constante cambio.

El uso de técnicas de filtrado basado en contenido y procesamiento de lenguaje natural (PLN) ha demostrado ser efectivo para categorizar cursos y alinear las recomendaciones con las necesidades específicas de los docentes. No obstante, la precisión en la interpretación del lenguaje y la clasificación de contenidos sigue siendo un área de mejora. La implementación

de un modelo híbrido, que integra la Taxonomía de Bloom para identificar el nivel cognitivo de los cursos, ha permitido una mayor personalización, pero también ha subrayado la complejidad de automatizar estas recomendaciones de manera precisa.

Como lecciones aprendidas se pueden indicar las siguientes: (1) Respecto de la Evaluación de Competencias TIC: La diversidad en las habilidades y conocimientos TIC entre los docentes destacó la necesidad de herramientas de evaluación más precisas y objetivas. La subjetividad en las autoevaluaciones puede afectar la exactitud de las recomendaciones, sugiriendo la necesidad de métodos más estandarizados y menos dependientes de la percepción individual. (2) Respecto del Desarrollo Tecnológico: La colaboración con expertos en diseño y desarrollo web fue crucial para crear una herramienta que no solo fuera funcional, sino también alineada con la identidad institucional. La flexibilidad y escalabilidad del sistema permitieron ajustes continuos y mejoras basadas en la retroalimentación de los usuarios. (3) Respecto de la Retroalimentación y Adaptación: La retroalimentación de los docentes indicó que, aunque la mayoría estaba satisfecha con los cursos recomendados, la usabilidad de la interfaz necesitaba mejoras.

Lo anterior resalta la importancia de considerar tanto el contenido como la experiencia del usuario en el diseño de sistemas educativos. Sin embargo, este desarrollo logró exitosamente diseñar e implementar un sistema de recomendación de cursos en línea para mejorar las competencias TIC de los docentes de la Universidad de Cuenca. Esta herramienta no solo facilita el acceso a recursos educativos personalizados, sino que también ayuda significativamente a la adaptación de la educación superior a las demandas digitales actuales. La integración de la Taxonomía de Bloom y técnicas de procesamiento de lenguaje natural (PLN) permitió realizar recomendaciones más precisas y relevantes, aunque todavía hay espacio para futuras mejoras.

A pesar del éxito, el desarrollo enfrentó varias limitaciones. La precisión en la medición de competencias se vio afectada por la subjetividad de la autoevaluación y la complejidad de las competencias TIC, lo que subraya la necesidad de herramientas de evaluación más objetivas. Además, la constante evolución de las tecnologías y metodologías pedagógicas requiere que el sistema sea dinámico y adaptable. Los desafíos del PLN, como la interpretación correcta de términos y contextos, también presentaron dificultades, especialmente en la clasificación precisa del contenido educativo. En resumen, el desarrollo ha establecido una base sólida para futuras investigaciones y desarrollos, ofreciendo valiosas perspectivas para la mejora continua de las herramientas educativas digitales.

Para trabajos futuros, se recomienda evaluar la eficacia del uso de la Taxonomía de Bloom para categorizar los cursos según su nivel cognitivo. Además, se sugiere explorar posibles soluciones basadas en técnicas de machine learning para mejorar la personalización de los cursos. Esto podría incluir el uso de modelos más avanzados de aprendizaje automático para analizar patrones en las preferencias y necesidades de los docentes, con el fin de ofrecer recomendaciones de cursos aún más precisas y adaptadas.

Agradecimientos

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

Referencias

- Ali, Aida & Shamsuddin, Siti Mariyam & Ralescu, Anca. (2015). Classification with class imbalance problem: A review. *International Journal of Advance Soft Computing Applications*, 7, 176-204.
- Aggarwal, C. C. (2016). *Recommender Systems*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-29659-3>
- Ayabaca, D. M. G., Alba, J. A. J., & Guamán, E. E. E. (2019). Implementación de las TIC en el ámbito educativo ecuatoriano. *Sociedad & Tecnología*, 2(2), Article 2. <https://doi.org/10.51247/st.v2i2.49>
- Brooke, J. (1995). SUS: A quick and dirty usability scale. *Usability Eval. Ind.*, 189.
- Cañete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H., & Pérez, J. (2023). Spanish Pre-trained BERT Model and Evaluation Data (arXiv:2308.02976). *arXiv*. <https://doi.org/10.48550/arXiv.2308.02976>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv*. <http://arxiv.org/abs/1810.04805>
- Figueredo, O. R. B. (2013). Tecnologías emergentes en la educación: Una experiencia de formación de docentes que fomenta el diseño de ambientes de aprendizaje. *Educação & Sociedade*, 34(123), 531–548.
- INTEF (Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado). (2023). *Marco de Referencia de la Competencia Digital Docente (MRCDD)*. Secretaría General Técnica del Ministerio de Educación y Formación Profesional de España.
- Kedia, A., & Rasu, M. (2020). *Hands-On Python Natural Language Processing: Explore tools and techniques to analyze and process text with a view to building real-world NLP applications*. Packt Publishing Ltd.
- Khyani, D., & B S, S. (2021). An Interpretation of Lemmatization and Stemming in Natural Language Processing. *Shanghai Ligong Daxue Xuebao/Journal of University of Shanghai for Science and Technology*, 22, 350–357.
- Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, 41(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2
- Lee, R. S. T. (2023). *Natural Language Processing: A Textbook with Python Implementation*. Springer Nature.
- Leskovec, J., Rajaraman, A., & Ullman, J. D. (2014). *Mining of Massive Datasets*. Cambridge University Press.
- MEN (Ministerio de Educación Nacional de Colombia). (2013). *Competencias TIC para el Desarrollo Profesional Docente*. Oficina de Innovación Educativa con Uso de Nuevas Tecnologías.
- Scikit-learn. (2018). Parameter estimation using grid search with cross-validation. *Scikit-learn*. Retrieved May 27, 2024, from https://scikit-learn.org/0.18/auto_examples/model_selection/grid_search_digits.html
- Patel, J. M. (2020). *Getting Structured Data from the Internet: Running Web Crawlers/Scrapers on a Big Data Production Scale*. Apress.
- Redecker, C. (2020). *Marco Europeo para la Competencia Digital de los Educadores: DigCompEdu* (Fundación Universia & Ministerio de Educación y Formación Profesional, Trans.). Secretaría General Técnica del Ministerio de Educación y Formación Profesional de España.
- Saito, T., & Rehmsmeier, M. (2015). The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLOS ONE*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- Trust, T. (2018). 2017 ISTE Standards for Educators: From Teaching With Technology to Using Technology to Empower Learners. *Journal of Digital Learning in Teacher Education*, 34(1), 1–3. <https://doi.org/10.1080/21532974.2017.1398980>