

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE INGENIERÍA DE SISTEMAS

**DESARROLLO DE UN SIMULADOR DE CLASES
PERSONALIZADAS CON IA GENERATIVA PARA EL
APRENDIZAJE UNIVERSITARIO**

IMPLEMENTACIÓN DEL SISTEMA DE EVALUACIÓN ADAPTATIVA

**TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE
INGENIERO DE SOFTWARE o EN CIENCIAS DE LA COMPUTACIÓN**

CARLOS ANDRÉS CÓRDOVA ACARO
carlos.cordova02@epn.edu.ec

DIRECTOR: ENRIQUE ANDRÉS LARCO AMPUDIA, PhD.
andres.larco@epn.edu.ec

DMQ, enero 2026

CERTIFICACIONES

Yo, **Carlos Andrés Córdova Acaro**, declaro que el trabajo de integración curricular aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y que he consultado las referencias bibliográficas que se incluyen en este documento.

Asimismo, declaro que he utilizado herramientas de inteligencia artificial (ChatGPT, Gemini y Claude) únicamente como apoyo en la generación de tablas y como guía para la investigación de librerías en Python, sin atribuirles autoría, y que todo el contenido derivado ha sido revisado, validado y es de mi exclusiva responsabilidad.

CARLOS ANDRÉS CÓRDOVA ACARO

Certifico que el presente trabajo de integración curricular fue desarrollado por **Carlos Andrés Córdova Acaro**, bajo mi supervisión.

ENRIQUE ANDRÉS LARCO AMPUDIA, PhD.

Director

DECLARACIÓN DE AUTORÍA

A través de la presente declaración, afirmo que el trabajo de integración curricular aquí descrito, así como el producto resultante del mismo, es de carácter público y estará a disposición de la comunidad académica a través del repositorio institucional de la Escuela Politécnica Nacional.

No obstante, la titularidad de los derechos patrimoniales corresponde al autor del presente trabajo, de conformidad con las disposiciones establecidas por el órgano competente en materia de propiedad intelectual, la normativa interna institucional y la legislación vigente.

Carlos Andrés Córdova Acaro

Enrique Andrés Larco Ampudia, PhD.

DEDICATORIA

Dedico este trabajo a mi familia, cuyo apoyo constante, esfuerzo y confianza han sido fundamentales a lo largo de mi formación académica y personal. Su acompañamiento incondicional ha sido un pilar esencial para alcanzar este logro.

AGRADECIMIENTOS

Agradezco a la Escuela Politécnica Nacional por la formación académica brindada a lo largo de mi carrera universitaria. De manera especial, expreso mi gratitud al director de este trabajo, PhD. Enrique Andrés Larco Ampudia, por su orientación, observaciones y acompañamiento durante el desarrollo del presente proyecto.

Asimismo, agradezco a los docentes, compañeros y familiares que, de distintas maneras, contribuyeron con su apoyo y motivación a la culminación de esta etapa académica.

ÍNDICE DE CONTENIDO

ÍNDICE GENERAL

CERTIFICACIONES	I
DECLARACIÓN DE AUTORÍA	II
DEDICATORIA	III
AGRADECIMIENTOS	IV
ÍNDICE DE CONTENIDO	V
RESUMEN	VIII
ABSTRACT	IX
1 DESCRIPCIÓN DEL COMPONENTE	1
1.1 Descripción general del componente	1
1.2 Objetivo general	1
1.3 Objetivos específicos	1
1.3.1 Implementar un sistema de medición dual del perfil del estudiante mediante señales complementarias	1
1.3.2 Desarrollar una política de selección adaptativa basada en IRT para contextos de diagnóstico y certificación	2
1.3.3 Implementar un sistema de rastreo de conocimiento basado en BKT/KT para orientar la práctica guiada	2
1.3.4 Diseñar una arquitectura fundamentada en el patrón MAPE-K para sistemas auto-adaptativos	2
1.3.5 Establecer un contrato explícito de integración entre componentes que garantice trazabilidad	2
1.3.6 Implementar salvaguardas de validez, equidad y mecanismos de monitoreo continuo del sistema	3
1.4 Alcance del componente	3
1.5 Marco teórico	4
1.5.1 Aprendizaje adaptativo	4
1.5.2 Teoría de Respuesta al Ítem (IRT)	4
1.5.3 Sistemas de Tutoría Inteligente (ITS)	6

1.5.4	Algoritmos de selección adaptativa	7
1.5.5	Métricas de desempeño en sistemas adaptativos	8
2	METODOLOGÍA	10
2.1	Enfoque y diseño de la investigación	10
2.2	Clasificación de la investigación	10
2.3	Arquitectura experimental y estrategia de simulación	11
2.4	RESULTADOS, CONCLUSIONES Y RECOMENDACIONES	13
2.4.1	Resultados	13
2.4.2	Conclusiones	13
2.4.3	Recomendaciones	13
	REFERENCIAS BIBLIOGRÁFICAS	14
3	ANEXOS	17
	ANEXO I: Aplicación Móvil	17
	ANEXO II: Aplicación Web	17
	ANEXO III: Página Web	17

RESUMEN

El presente trabajo desarrolla un sistema de evaluación adaptativa orientado al contexto universitario, cuyo objetivo es estimar de forma precisa el nivel de conocimiento del estudiante y personalizar la selección de actividades de aprendizaje. El sistema se fundamenta en un modelo híbrido que integra la Teoría de Respuesta al Ítem (IRT) con técnicas de Knowledge Tracing bayesiano, permitiendo un seguimiento continuo del estado de dominio del estudiante.

La arquitectura propuesta se implementa como un motor adaptativo desacoplado, capaz de interactuar con sistemas externos de generación de contenidos educativos. La validación del sistema se realiza mediante simulación computacional con estudiantes virtuales, analizando métricas de precisión, estabilidad y rendimiento. Los resultados obtenidos evidencian la viabilidad técnica del enfoque propuesto y su potencial aplicación en entornos de aprendizaje personalizados.

ABSTRACT

This work presents the development of an adaptive assessment system designed for higher education environments. The main objective of the system is to accurately estimate students' knowledge levels and to personalize learning activities accordingly. The proposed approach is based on a hybrid model that combines Item Response Theory (IRT) with Bayesian Knowledge Tracing techniques.

The system is implemented as a decoupled adaptive engine capable of interacting with external content generation components. Validation is carried out through computational simulation using virtual students, evaluating accuracy, stability, and performance metrics. The results demonstrate the technical feasibility of the proposed model and its potential applicability in personalized learning systems.

1 DESCRIPCIÓN DEL COMPONENTE

1.1 Descripción general del componente

Este proyecto se compone de cuatro componentes interrelacionados. El Componente B constituye el motor que traduce la evidencia de interacción del estudiante en decisiones pedagógicas accionables. Su función principal es calcular el nivel de habilidad del estudiante, identificar qué conocimientos domina con suficiente confianza y determinar qué actividad o ítem presentar a continuación, incluyendo la dificultad apropiada y el tipo de soporte necesario. Los objetivos del componente son duales: por un lado, medir con precisión el desempeño del estudiante; por otro, favorecer el aprendizaje mediante trayectorias personalizadas, haciendo uso de la retroalimentación constante que mantiene con el Componente A, responsable de generar las actividades y las clases. Este enfoque se alinea con la lógica del aprendizaje adaptativo que predomina en la actualidad, basada en realizar ajustes del proceso de aprendizaje a partir de datos individuales en lugar de seguir rutas fijas predeterminadas [1], [2], [3].

1.2 Objetivo general

Desarrollar un motor adaptativo operativo que traduzca la evidencia de interacción del estudiante en decisiones pedagógicas fundamentadas, mediante la combinación de modelos psicométricos extensivamente validados (IRT y BKT/KT) para medir con precisión el desempeño del estudiante y construir trayectorias de aprendizaje genuinamente personalizadas que se ajusten dinámicamente a partir de datos individuales.

1.3 Objetivos específicos

1.3.1 Implementar un sistema de medición dual del perfil del estudiante mediante señales complementarias

Desarrollar un sistema que recoja eventos de interacción y calcule dos señales complementarias: un nivel continuo de habilidad θ basado en IRT para ordenar ítems informativos y reducir el error estándar de medición, y una probabilidad de dominio por habilidad fundamentada en BKT/KT para guiar la práctica espaciada y el refuerzo cuando el objetivo sea consolidar conocimientos de forma sostenida.

1.3.2 Desarrollar una política de selección adaptativa basada en IRT para contextos de diagnóstico y certificación

Implementar una política de selección que escoja el ítem que maximiza la información en torno al θ estimado, reduciendo rápidamente el error estándar y permitiendo alcanzar precisión localizada donde más importa, incorporando reglas de detención, restricciones de contenido curricular y limitaciones de exposición siguiendo las buenas prácticas establecidas en IRT y CAT.

1.3.3 Implementar un sistema de rastreo de conocimiento basado en BKT/KT para orientar la práctica guiada

Desarrollar mediante BKT/KT un sistema que mantenga una probabilidad de dominio por habilidad y la actualice tras cada interacción, modelando fenómenos como la adivinación y el desliz, de modo que la selección de la actividad subsiguiente se decida según el beneficio esperado: confirmar dominio incipiente, reducir incertidumbre o fomentar el aprendizaje en la zona de desarrollo más productiva.

1.3.4 Diseñar una arquitectura fundamentada en el patrón MAPE-K para sistemas auto-adaptativos

Implementar el Componente B como un bucle MAPE-K que monitorice respuestas y patrones de desempeño, analice el perfil del estudiante, planifique la siguiente actividad especificando ítem, dificultad y tipo de apoyo, y ejecute enviando recomendaciones explícitas al Componente A, garantizando trazabilidad, explicabilidad y la posibilidad de integrar organización semántica del contenido mediante grafos o rutas de aprendizaje.

1.3.5 Establecer un contrato explícito de integración entre componentes que garantice trazabilidad

Diseñar la integración $A \leftrightarrow B$ como un contrato de datos simple y explícito donde el Componente A envíe datos del usuario, tema activo e historial de interacciones, y el Componente B responda con el ítem propuesto, dificultad objetivo, tipo de ayuda recomendada y justificación concisa de la decisión, haciendo las decisiones auditables y proporcionando al docente evidencia clara del proceso.

1.3.6 Implementar salvaguardas de validez, equidad y mecanismos de monitoreo continuo del sistema

Desarrollar salvaguardas que resguarden la validez y equidad mediante el reporte de precisión alcanzada, validación del ajuste del modelo, monitorización de exposición equilibrada de temas e ítems, reporte de ganancia pre-post, control de métricas predictivas como AUC o log-loss, y realización de pruebas DIF y análisis de brechas entre perfiles para detectar posibles sesgos.

1.4 Alcance del componente

El alcance del Componente B comprende el desarrollo de un motor adaptativo operativo con las siguientes características funcionales y técnicas:

- I. **Integración de modelos complementarios:** Combinar IRT para diagnóstico o evaluación sumativa con BKT/KT para guiar la práctica continua, aprovechando las fortalezas de ambos enfoques para ofrecer una evaluación integral que mida y fomente el aprendizaje.
- II. **Orquestación de la selección de ítems:** Orquestrar la selección de ítems mediante reglas claras de parada, cobertura curricular y exposición equilibrada que permitan determinar cuándo la evaluación ha alcanzado suficiente precisión, garantizando que todos los temas relevantes sean cubiertos y evitando la sobreutilización de ítems específicos.
- III. **Panel de métricas para validación docente:** Publicar un panel de métricas (precisión diagnóstica, eficiencia, progreso del estudiante, calidad predictiva y equidad) accesible para que los docentes validen el funcionamiento del sistema, facilitando la transparencia y permitiendo intervenciones informadas cuando resulte necesario.
- IV. **Gobernanza de datos y privacidad:** Documentar la gobernanza de datos y privacidad, especificando qué se registra, para qué propósito y cómo se protege la información, asegurando el cumplimiento de estándares éticos y regulatorios en el manejo de datos educativos sensibles.

Todo ello integrado con el Componente A de manera que cada estudiante reciba un reto apropiado, en el momento oportuno, con las explicaciones y los apoyos adecuados a su nivel y necesidades específicas, logrando una experiencia de aprendizaje genuinamente personalizada y fundamentada en evidencia [1], [3], [4], [5], [6].

1.5 Marco teórico

1.5.1 Aprendizaje adaptativo

El aprendizaje adaptativo es una forma de enseñanza que se ajusta a cada estudiante en tiempo real. En vez de proponer la misma ruta para todos, el sistema observa evidencias (aciertos, errores, tiempo de respuesta, interacciones) y decide qué contenido, qué nivel de dificultad y qué apoyo conviene a continuación. Así, la progresión deja de ser lineal y se vuelve personalizada, manteniendo el foco en el dominio gradual de objetivos. Esta idea se formaliza y se sostiene en la literatura reciente sobre evaluación, personalización y uso responsable de IA en educación [1], [3].

Conviene distinguir lo adaptativo de lo simplemente "personalizado". La personalización puede implicar variedad de actividades o estilos, pero no siempre supone que el sistema mida y ajuste continuamente con base en datos. Lo adaptativo, en cambio, depende de un ciclo continuo de diagnóstico-retroalimentación-reajuste, y se apoya en un buen diseño instruccional: objetivos claros, progresiones definidas y evidencias útiles para decidir los siguientes pasos [3].

En la práctica, el ciclo luce así: (1) un breve diagnóstico, (2) selección de recursos y tareas ajustadas al nivel detectado, (3) retroalimentación oportuna, (4) una nueva medición que confirma avances o sugiere refuerzos y (5) ajustes de la ruta. La clave no es aumentar la cantidad de ejercicios, sino ofrecer los adecuados en el momento preciso. Este principio didáctico se alinea con marcos como los de Reigeluth y los primeros principios de Merrill, que recomiendan activar saberes previos, demostrar, aplicar e integrar lo aprendido [3].

La IA generativa ha aportado un motor útil para redactar explicaciones, proponer ejemplos y crear ejercicios alineados con la ruta de cada estudiante. Sin embargo, la evidencia disponible también advierte que estas herramientas no reemplazan la pedagogía ni la evaluación rigurosa; su valor aumenta cuando operan bajo criterios claros de calidad, ética y supervisión docente [1].

Un ejemplo concreto es PathRAG, que organiza el conocimiento como un grafo (conceptos y relaciones) para trazar caminos pertinentes según el perfil del estudiante. Estudios recientes en contextos universitarios híbridos reportan mejoras en participación, logro de competencias y percepción de inclusión cuando se integran rutas personalizadas con apoyo de IA generativa; aun así, subrayan límites metodológicos y la necesidad de diseños más robustos [2].

1.5.2 Teoría de Respuesta al Ítem (IRT)

La Teoría de Respuesta al Ítem (TRI o IRT) es una forma moderna de entender las pruebas: en lugar de mirar solo el puntaje total, analiza cómo responde una persona a cada ítem y, a partir de ello, estima su nivel en el rasgo que se quiere medir θ . Con esa estimación, es posible seleccionar mejores preguntas, ubicar la dificultad donde más hace falta y conocer cuán precisa es la medición en cada tramo del continuo. Frente a la Teoría Clásica de Tests,

su aporte central es la invariancia': medir con la misma escala, aunque cambien los sujetos o los ítems (dentro de ciertos supuestos) [1], [2].

1.5.2.1 Ideas clave

- **Curva característica del ítem (CCI):** es un gráfico que muestra, para cada nivel de θ , la probabilidad de elegir la opción "clave" del ítem (por ejemplo, responder correctamente o indicar mayor rasgo). Su forma creciente refleja que, a mayor nivel del rasgo, mayor probabilidad de dar la respuesta asociada al rasgo [1], [2].
- **Parámetros a , b y c :** a indica cuánto discrimina el ítem (qué tan bien separa a personas con niveles cercanos de θ), b ubica la dificultad del ítem (el punto de la escala donde el ítem decide), y c modela el azar o pseudo-adivinación en ítems de opción correcta/incorrecta. No todos los modelos usan los tres: Rasch (1PL) usa solo b , 2PL usa a y b , y 3PL usa a , b , c [1], [2].
- **Información del ítem/test:** indica dónde el ítem o el conjunto de ítems mide con mayor precisión. En IRT la precisión no es plana: puede ser excelente en un rango de θ y más baja en otros. Esto permite construir bancos de ítems que cubran la escala con precisión donde más importa [2].

1.5.2.2 Supuestos relevantes

- **Unidimensionalidad:** los ítems de una escala deben reflejar esencialmente un solo rasgo dominante; si influyen varios rasgos a la vez, conviene usar modelos multidimensionales o depurar la escala [2].
- **Independencia local:** si ya sabemos el nivel de θ , las respuestas a ítems distintos no deben depender entre sí. Cuando hay pistas entre ítems o se agrupan demasiado, este supuesto se rompe y la medición pierde calidad [1], [2].

1.5.2.3 Modelos más usados (visión práctica)

- **Ítems dicotómicos (correcto/incorrecto):** 1PL (Rasch), 2PL y 3PL. Rasch asume igual discriminación y sin azar; 2PL permite que la discriminación varíe; 3PL incluye el parámetro de pseudo-adivinación. Elegir el modelo depende del contexto y los datos [1], [2].
- **Ítems politómicos (escalas Likert):** Modelos como Respuesta Graduada (Samejima) o Crédito Parcial. En el Modelo de Respuesta Graduada, cada salto entre categorías tiene un umbral de dificultad, y un único parámetro a de discriminación para el ítem. Esto es muy útil para cuestionarios con varias opciones de respuesta [7].

1.5.2.4 Bancos de ítems y pruebas adaptativas (CAT)

Al estimar θ en tiempo real y conocer la información de cada ítem, es posible elegir la siguiente pregunta que aporte máxima precisión justo alrededor del nivel estimado del estudiante. Así nacen los tests adaptativos: cada persona responde un conjunto distinto de preguntas, pero todos son evaluados en la misma escala. En tu proyecto, esto es clave para que el Componente B seleccione o recomiende ítems con mayor "ganancia informativa" [7], [8].

1.5.2.5 Integración con los Componentes A y B

1. (B) A partir de las respuestas del estudiante, se estima θ con un modelo IRT apropiado (2PL/3PL para ítems dicotómicos; Respuesta Graduada para Likert).
2. (B) se consulta el banco de ítems para identificar cuáles ofrecen más información alrededor del θ actual (o del umbral de dominio).
3. (B→A) se envía al Componente A la dificultad objetivo y, si aplica, los ítems recomendados o las pautas de complejidad.
4. (A) el Componente A genera la siguiente actividad con esa dificultad y pistas adecuadas.
5. (B) tras la actividad, se actualiza θ y se repite el ciclo. Este bucle mantiene rutas personalizadas y medibles [7], [9].

1.5.3 Sistemas de Tutoría Inteligente (ITS)

Un Sistema de Tutoría Inteligente (ITS) es un software que intenta parecerse a una tutoría humana: observa cómo aprende el estudiante, le ofrece explicaciones y actividades a la medida, y retroalimenta en los momentos clave. La idea no es reemplazar al docente, sino multiplicar su apoyo para que cada persona avance a su propio ritmo y con la ayuda justa. Las revisiones recientes muestran que, bien implementados, los ITS mejoran el rendimiento y la participación, personalizan contenidos y apoyan la autorregulación del aprendizaje [4].

«“<

»”1.5.3.1 Componentes típicos

- **Modelo del estudiante:** mantiene un perfil dinámico con aciertos, errores, tiempos y progreso.
- **Modelo del tutor:** decide qué explicar, qué actividad proponer y qué tipo de pista o intervención aplicar.
- **Modelo de dominio:** representa el conocimiento de la materia (conceptos, habilidades y relaciones).

- **Interfaz:** canaliza ejercicios, retroalimentación y seguimiento.

»"Estas piezas permiten ajustar dificultad, secuencias y apoyos en tiempo real [5].

»"1.5.3.2 Integración con los Componentes A y B

1. »"(B) El ITS observa respuestas, estima el nivel y detecta dificultades relevantes.
2. »"(B→A) Envía al Componente A la dificultad recomendada, objetivos prioritarios y el tipo de intervención.
3. »"(A) El Componente A genera la actividad o clase con esa dificultad y apoyo.
4. »"(B) Tras la actividad, se vuelve a medir y se ajusta la ruta, manteniendo trayectorias personalizadas [4], [5].

»"1.5.4 Algoritmos de selección adaptativa

»"Seleccionar la siguiente actividad no es aleatorio: implica decidir con evidencia qué ítem conviene presentar para medir mejor o para favorecer el aprendizaje. En este proyecto, dicha decisión se ejecuta en el Componente B y retroalimenta al Componente A. A grandes rasgos, existen dos familias principales: selección guiada por IRT (orientada a medición precisa) y selección guiada por BKT/KT (orientada a adquisición gradual de habilidades).

»"1.5.4.1 Selección guiada por IRT (medición precisa)

»"IRT modela la probabilidad de respuesta correcta según θ y los parámetros del ítem. Sobre esa base, la selección adaptativa típica elige el ítem con mayor información alrededor del θ estimado, reduciendo el error estándar con menos preguntas y manteniendo la dificultad donde más informa. En la práctica se inicia con un θ neutro o un breve *warm-up*; luego, tras cada respuesta, se reestima θ y se elige el ítem que maximiza la información bajo restricciones de contenido y control de exposición, con reglas de parada por longitud o precisión [7], [8].

»"Esto aporta a la integración $A \leftrightarrow B$ al permitir que el Componente B devuelva a A un rango de dificultad recomendado y cobertura temática pendiente, evitando sobreestimar o subestimar la exigencia de la siguiente actividad [7], [8].

»"1.5.4.2 Selección guiada por BKT/KT (apoyo al aprendizaje)

»"BKT mantiene, para cada habilidad, una probabilidad de dominio que evoluciona a lo largo del tiempo y modela cuatro parámetros usuales (conocimiento inicial, aprendizaje tras práctica, adivinación y desliz). Con este perfil, el sistema decide el siguiente ejercicio según beneficio esperado: reducir incertidumbre, confirmar dominio o provocar aprendizaje en la zona más productiva. En entornos reales también se integran prerrequisitos, espaciado y

señales de compromiso (tiempos, rachas). La evidencia muestra que BKT es eficaz para personalizar secuencias cuando la meta es progresar en habilidades específicas y no solo medir una vez con precisión [10], [11].

»"Para $A \leftrightarrow B$, BKT permite priorizar habilidad objetivo, tipo de apoyo y el momento oportuno para refuerzo o espaciado, mientras que el Componente B actualiza probabilidades tras cada actividad ejecutada por A [10], [11].

»"1.5.4.3 Estrategia híbrida

»"En etapas tempranas o con bancos pequeños, BKT tiende a ser más práctico porque requiere menos calibración de ítems y entrega señales útiles para guiar práctica. A medida que el banco crece y se busca una estimación fina del nivel, IRT adquiere mayor relevancia al permitir fijar una precisión objetivo y optimizar la selección. Una política viable es usar BKT para guiar práctica continua por habilidades y activar selección IRT en cortes de evaluación (diagnóstico o certificación). En este proyecto, dicha coordinación se expresa como un ciclo MAPE-K [6].

»"1.5.5 Métricas de desempeño en sistemas adaptativos

»"Las métricas permiten demostrar que el sistema mide con precisión, promueve aprendizaje y opera de forma eficiente y justa. En un sistema adaptativo se consideran dos planos: calidad de medición (precisión y validez) y calidad de enseñanza (progreso y retención), complementados con métricas predictivas y de equidad [6], [8], [10], [12].

»"1.5.5.1 Precisión y validez de medición (IRT/CAT)

»"En pruebas adaptativas orientadas a medir con exactitud, la precisión se observa en el error estándar $SE(\hat{\theta})$ y en la información del test alrededor del nivel estimado. Un buen algoritmo reduce $SE(\hat{\theta})$ con menos ítems, balanceando precisión y longitud del test. Para comparar estrategias, se puede fijar una longitud media y medir la precisión resultante, o fijar precisión objetivo y medir cuántos ítems se requieren [12].

»"También es relevante el ajuste del modelo: sesgo (diferencia promedio entre estimado y referencia), RMSE y consistencia al recalibrar. En enfoques flexibles se emplean medidas de divergencia o criterios de información para selección de modelo, que afectan la precisión final [12].

»"1.5.5.2 Eficiencia y carga de respuesta

»"La eficiencia se refleja en el número promedio de ítems administrados, su variabilidad entre estudiantes, el tiempo por objetivo alcanzado y el ahorro respecto del banco total. Además, es útil analizar cómo cambia la carga según el nivel del estudiante, ya que algunos algoritmos demandan más ítems en extremos de la escala [12].

»'1.5.5.3 Aprendizaje y progreso

»'Cuando el objetivo es aprender (no solo medir), importan métricas de ganancia pre–post, tasa de dominio por unidad, tiempo o ítems hasta alcanzar un umbral y retención tras un intervalo. En KT/BKT, se reporta la evolución de probabilidad de dominio por habilidad, verificando que refuerzos y apoyos incrementen la probabilidad de forma sostenida [10], [11].

»'1.5.5.4 Calidad predictiva de la política adaptativa

»'Para validar la consistencia del motor, se evalúa su capacidad predictiva sobre respuestas futuras mediante métricas como AUC/ROC, exactitud de clasificación de dominio y *log-loss*. Estas métricas no sustituyen evidencia de aprendizaje, pero ayudan a verificar estabilidad y coherencia del sistema [10], [11].

»'1.5.5.5 Equidad y robustez

»'La equidad se analiza con DIF, brechas de error y precisión entre perfiles, y auditorías de exposición de ítems. La robustez requiere pruebas de sensibilidad a supuestos (unidimensionalidad, independencia local) y validación al recalibrar el banco. Además, la transparencia en el uso de datos y la interpretabilidad de reportes son claves para confianza y adopción [7], [8], [12].

»'1.5.5.6 Reporte para la integración $A \leftrightarrow B$

»'Para cerrar el ciclo $A \leftrightarrow B$, el Componente B debe devolver un resumen compacto: (i) precisión alcanzada (por ejemplo $SE(\hat{\theta})$), (ii) eficiencia (ítems/tiempo vs. objetivo), (iii) progreso por habilidad (probabilidad de dominio y ganancia), (iv) calidad predictiva (AUC/log-loss) y (v) equidad (DIF y exposición balanceada). Con ese panel, el Componente A puede ajustar dificultad, apoyo y espaciado con criterio [6], [10], [12].

»'2 METODOLOGÍA

»'2.1 Enfoque y diseño de la investigación

»'En el desarrollo del Sistema de Evaluación Adaptativa (Componente B) se utilizó un enfoque de investigación cuantitativa en el que la evaluación era objetiva, numérica, reproducible y medible de las variables pedagógicas y computacionales. Este enfoque era lógico ya que se relaciona con el tipo de problema abordado, es decir, optimizar procesos de evaluación y validar un sistema de software fundamentado en modelos matemáticos, estadísticos y probabilísticos que centra la evaluación en niveles profundos del conocimiento del estudiante. El método cuantitativo permite la evaluación de la forma de actuar del motor adaptativo mediante la valoración de indicadores con los que se puede medir, como pueden ser la estimación de la habilidad latente del aprendiz (θ), el ajuste en base a la precisión de las métricas como el error cuadrático medio (RMSE), la fiabilidad de las probabilidades analizada con la métrica de Brier Score; además tomando métricas concretas de la ingeniería de ciencias computacionales como la latencia de la respuesta del sistema, percentiles de tiempo de procesamiento (P50 y P95), y la tasa de peticiones por segundo (RPS). Estas métricas sirven para el diagnóstico en base a criterios cuantificables de la precisión, la eficiencia y la escalabilidad del sistema propuesto. La utilización de esta vertiente metodológica se apoya en la documentación especializada de los sistemas de aprendizaje adaptativo y la evaluación psicométrica, la cual establece que para la obtención de indicadores robustos de aprendizaje deben emplearse modelos estadísticos que permiten inferir variables latentes a partir de la evidencia empírica observable, particularmente en el caso de la Teoría de Respuesta al Ítem (IRT) y en los modelos bayesianos de rastreo de conocimiento^{4, 5, 11, 12}.

»'2.2 Clasificación de la investigación

»'Tomando en cuenta el marco metodológico que se ha seguido, esta investigación puede ser clasificada como una investigación tecnológica aplicada, la cual se halla orientada al diseño, a la validación y a la implementación de un artefacto de software funcional y operativo que tiene la finalidad de proporcionar una solución a un problema de práctica educativa en contextos locales, específicamente, la modulación adaptativa de evaluaciones mediante la integración de modelos de Machine Learning en la educación superior ^{7, 10}.

»'2.3 Arquitectura experimental y estrategia de simulación

»'La arquitectura experimental se apoya en la simulación computacional, pues las limitaciones logísticas, éticas y operativas de llevar a cabo un elevado número de pruebas con estudiantes reales en una fase incipiente del desarrollo nos llevaron a adoptar esta opción metodológica. En este contexto, la simulación estocástica y los métodos Monte Carlo conforman una opción metodológica argumentada teóricamente, pero también muy extendida y validada en la literatura para evaluar sistemas adaptativos complejos 3, 9, 10.

»'Este modelo propone la construcción de un entorno de simulación en el que fueron modelados perfiles de estudiantes virtuales con ciertos parámetros psicométricos controlados, entre los que podemos encontrar el nivel de habilidad inicial (θ), la consistencia de respuesta ante ítems de dificultad variable y el ratio de aprendizaje. Este entorno permite generar un elevado número de interacciones simuladas entre el sistema y perfiles de estudiantes heterogéneos, lo que permite estudiar la convergencia del algoritmo adaptativo, su comportamiento bajo distintas condiciones operativas y evaluar la robustez frente a situaciones adversas o excepcionales. De igual manera, la simulación computacional favoreció la validación operativa del sistema en situaciones de baja probabilidad de ocurrencia o difícilmente reproducibles en contextos reales de aplicación, tales como patrones de respuestas erráticas por parte de los estudiantes, ejecución de múltiples sesiones de evaluación de manera concurrente, y escenarios de escasez de ítems calibrados en el banco de preguntas. Todo ello contribuyó a proporcionar consistencia empírica a la evaluación de la tolerancia a fallos y la robustez estructural del motor adaptativo. Finalmente, el diseño metodológico propuesto permite establecer que esta fase corresponde a una validación algorítmica y técnica del sistema en condiciones controladas. La arquitectura del estudio contempla la ejecución de pruebas con usuarios reales para una fase posterior del proceso de investigación, una fase que estará orientada al análisis del impacto pedagógico efectivo del sistema y al estudio de factores cualitativos emergentes en contextos auténticos de aprendizaje.

»'En el siguiente apartado se presenta la relación detallada de las variables e indicadores de validación en la Tabla 2.3, la cual recoge la síntesis estructurada de los criterios cuantitativos que se han establecido para la evaluación sistemática del desempeño del motor adaptativo.

»"Variable operacio- nalizada	»"Descripción conceptual	»"Función en el proceso de validación
»"θ (parámetro de habi- lidad estimado)	»"Inferencia del nivel latente de dominio cognitivo del estudiante	»"Evaluar la precisión diagnósti- ca del modelo psicométrico
»"RMSE / MAE	»"Cuantificación del error entre el parámetro real de habilidad y su estimación computacional	»"Medir exactitud predictiva del sistema implementado
»"Brier Score	»"Función de pérdida cuadrá- tica aplicada a probabilidades predichas	»"Evaluar calidad y calibración de las predicciones probabilísti- cas
»"Latencia (ms)	»"Duración temporal del proce- samiento de peticiones del siste- ma	»"Analizar rendimiento compu- tacional y eficiencia algorítmica
»"RPS	»"Volumen de peticiones proce- sadas exitosamente por unidad temporal	»"Evaluar escalabilidad horizon- tal y capacidad de concurrencia
»"P50 / P95	»"Percentiles de la distribución de tiempos de respuesta	»"Detectar degradación del ren- dimiento bajo condiciones de carga elevada

Cuadro 2.1: Criterios cuantitativos utilizados para la validación del motor adaptativo

»'''2.4 RESULTADOS, CONCLUSIONES Y RECOMENDACIONES

»'''2.4.1 Resultados

»'''2.4.2 Conclusiones

»'''2.4.3 Recomendaciones

REFERENCIAS BIBLIOGRÁFICAS

»'''

»'''BIBLIOGRAFÍA

- [1] »'''M. Zapata Ros, «IA generativa y ChatGPT en Educación: Un reto para la evaluación y ¿una nueva pedagogía?» *Revista Paraguaya de Educación a Distancia*, vol. 5, n.º 1, págs. 12-44, 2024. DOI: 10.56152/reped2024-vol5num1-art2
- [2] »'''R. Juárez Cádiz, «PathRAG application in adaptive learning with generative AI for inclusive and sustainable education,» *RIED-Revista Iberoamericana de Educación a Distancia*, vol. 29, n.º 1, 2026. DOI: 10.5944/ried.45378
- [3] »'''G. C. Tenorio-Sepúlveda, A. Soberanes-Martín y M. Martínez-Reyes, «Diseño instruccional con aprendizaje adaptativo de un curso en línea: Redacción de protocolos de investigación,» *Revista de Gestión Universitaria*, vol. 2, n.º 3, págs. 9-16, mar. de 2018.
- [4] »'''N. Carbonell Bernal y M. Á. Hernández Prados, «Impacto de los Sistemas de Tutoría Inteligente. Una revisión sistemática,» *EDUTEC. Revista Electrónica de Tecnología Educativa*, n.º 89, págs. 121-132, sep. de 2024. DOI: 10.21556/edutec.2024.89.3025
- [5] »'''M. H. Rodríguez Chávez, «Sistemas de tutoría inteligente y su aplicación en la educación superior,» *RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, vol. 12, n.º 22, e175, 2021. DOI: 10.23913/ride.v11i22.848
- [6] »'''O. Gheibi, D. Weyns y F. Quin, «Applying Machine Learning in Self-adaptive Systems: A Systematic Literature Review,» *ACM Transactions on Autonomous and Adaptive Systems*, vol. 15, n.º 3, Article 9, 2021. DOI: 10.1145/3469440
- [7] »'''M. D. Hidalgo-Montesinos y B. F. French, «Una introducción didáctica a la Teoría de Respuesta al Ítem para comprender la construcción de escalas,» *Revista de Psicología Clínica con Niños y Adolescentes*, vol. 3, n.º 2, págs. 13-21, jul. de 2016.
- [8] »'''H. F. Attorresi, G. S. Lozzia, F. J. P. Abal, M. S. Galibert y M. E. Aguerri, «Teoría de Respuesta al Ítem. Conceptos básicos y aplicaciones para la medición de constructos psicológicos,» *Revista Argentina de Clínica Psicológica*, vol. 18, n.º 2, págs. 179-188, ago. de 2009.
- [9] »'''F. J. P. Abal, G. S. Lozzia, M. E. Aguerri, M. S. Galibert y H. F. Attorresi, «La escasa aplicación de la teoría de respuesta al ítem en tests de ejecución típica,» *Revista Colombiana de Psicología*, vol. 19, n.º 1, págs. 111-122, 2010.
- [10] »'''Y. Hicke, *Knowledge Tracing Challenge: Optimal Activity Sequencing for Students*, arXiv:2311.14707v1, 2023.

- [11] »"S. Xu, M. Sun, W. Fang, K. Chen, H. Luo y P. X. W. Zou, «A Bayesian-based knowledge tracing model for improving safety training outcomes in construction: An adaptive learning framework,» *Developments in the Built Environment*, vol. 13, pág. 100 111, 2023.
- [12] »"A. Psychogiopoulos, N. Smits y L. A. van der Ark, «Estimating the Joint Item-Score Density Using an Unrestricted Latent Class Model,» *Journal of Computerized Adaptive Testing*, vol. 12, n.º 3, págs. 136-151, jul. de 2025. DOI: 10.7333/2507-1203136
- [13] »"L. Vargas Peña, «Estudio comparativo de modelos adaptativos y análisis de su uso en un ambiente de m-Learning implementando la técnica adaptativa de la Teoría de Respuesta al Ítem (IRT),» Tesis de mtría., Universidad Rey Juan Carlos, Madrid, España, 2012.
- [14] »"E. Hernández-Salazar y C. A. Beltrán, «SCRUM, un enfoque práctico de metodología ágil para la ingeniería de software,» *Revista Tecnología, Investigación y Academia (TIA)*, vol. 8, n.º 2, págs. 61-73, 2020.
- [15] »"K. Schwaber y J. Sutherland, *La Guía Scrum: La guía definitiva de Scrum Las reglas del juego*, Versión 2020. Licencia Creative Commons Attribution Share-Alike 4.0, 2020. dirección: <https://scrumguides.org>
- [16] »"J. A. Sarango y M. Villamar, «Aplicación de metodologías de desarrollo de software en proyectos de ingeniería,» Escuela Superior Politécnica del Litoral, Ecuador, 2022.

»'''3 ANEXOS

»'''ANEXO I: Aplicación Móvil

»'''ANEXO II: Aplicación Web

»'''ANEXO III: Página Web