

USANDO INTELIGENCIA ARTIFICIAL GENERATIVA PARA APOYAR A LOS ESTUDIANTES EN MATEMÁTICAS

USING GENERATIVE ARTIFICIAL INTELLIGENCE TO SUPPORT STUDENTS IN MATHEMATICS

Julio César Mendívil Casanova

Tecnológico Nacional de México / IT de Culiacán, México
julio.mc@culiacan.tecnm.mx

María Lucía Barrón Estrada

Tecnológico Nacional de México / IT de Culiacán, México
lucia.be@culiacan.tecnm.mx

Ramón Zatarain Cabada

Tecnológico Nacional de México / IT de Culiacán, México
ramon.zc@culiacan.tecnm.mx

Aldo Uriarte Portillo

Tecnológico Nacional de México / IT de Culiacán, México
aldo.up@culiacan.tecnm.mx

Recepción: 18/marzo/2026

Aceptación: 22/mayo/2026

Resumen

Los estudiantes mexicanos de nivel básico exhiben bajo rendimiento académico en matemáticas y ciencia. Este trabajo presenta una herramienta educativa con un tutor virtual para apoyar a los alumnos en el aprendizaje de Matemáticas. La herramienta educativa llamada TutorIA, utiliza el Aprendizaje Basado en Problemas (ABP) mediante un tutor virtual personalizado desarrollado con un Modelo de Lenguaje Grande. El tutor virtual cumple la función de guiar al estudiante durante el proceso de solución de problemas, además de responder a sus dudas y evaluar sus resultados. El uso de la herramienta contribuye en el aprendizaje de Matemáticas, al ofrecer una alternativa personalizada que se adapta a las necesidades del alumno. TutorIA fue evaluada con alumnos de nivel secundaria, obteniendo resultados favorables tanto para el aprendizaje como para la motivación de los estudiantes al usar la tecnología permitiendo a los estudiantes aprender de manera más autónoma y a su propio ritmo.

Palabras Clave: Aprendizaje Basado en Problemas (ABP), aprendizaje personalizado, inteligencia artificial generativa, tutor virtual.

Abstract

Mexican junior high school students have low performance in mathematics and science. This paper presents an educational tool with a virtual tutor to support students in learning mathematics. The educational tool, called TutorIA, implements Problem-Based Learning (PBL) with a personalized virtual tutor developed with a Large Language Model. The virtual tutor guides the students through the problem-solving process, answers questions, and evaluates their reasoning. The use of the tool contributes to mathematics learning by providing a personalized approach that adapts to the individual students' needs. TutorIA was evaluated with junior high school students, obtaining favorable results both for learning and for student motivation when using the technology, allowing students to learn more autonomously and at their own pace.

Keywords: *Generative artificial intelligence, personalized learning, Problem-Based Learning (PBL), virtual tutor.*

1. Introducción

Actualmente el aprendizaje de las Matemáticas a nivel básico en México es un problema persistente. De acuerdo con los datos obtenidos de la prueba de PISA 2023, realizada por la Organización para la Cooperación y el Desarrollo Económicos (OECD por sus siglas en inglés), los resultados promedio de México en el año 2022 disminuyeron en comparación con 2018 en Matemáticas y Ciencias, los resultados obtenidos muestran una pérdida casi total de las ganancias observadas durante el período 2003-2009, y los puntajes promedio retrocedieron acercándose a los observados en 2003 o 2006. En México, el 34% de los estudiantes alcanzó al menos el nivel 2 de competencia en Matemáticas, significativamente menos que el promedio en los países de la OECD (promedio: 69%). Esto indica que la mayoría de los estudiantes mexicanos no son capaces de representar matemáticamente situaciones simples sin indicaciones directas [OECD, 2023].

Debido a esta necesidad de mejorar el aprendizaje de las Matemáticas y demás materias, en México se han hecho cambios en el sistema educativo y se ha presentado la propuesta de la Nueva Escuela Mexicana, la cual impulsa el uso de nuevas metodologías, una de éstas es el aprendizaje basado en problemas (ABP) [Mendieta, 2021]. Este enfoque coloca al estudiante en el centro del proceso de aprendizaje, dándole la responsabilidad de construir su propio conocimiento a través de la resolución de problemas [Secretaría de Educación Pública, 2022a].

El ABP requiere la presencia de un tutor que oriente al estudiante, proporcionando pistas y sugerencias sin darle directamente la respuesta. Este tutor es clave para que el estudiante no se sienta perdido en el proceso de aprendizaje. Sin embargo, la necesidad de retroalimentación personalizada se vuelve una limitación importante en aulas con muchos alumnos, ya que un solo maestro no puede brindar atención individualizada a cada estudiante.

El uso de la Inteligencia Artificial (IA) es una solución a la necesidad de un tutor que sirva de guía para el estudiante, tecnologías como los Modelos de Lenguaje Grandes (LLM por sus siglas en inglés), que es un sistema de inteligencia artificial capaz de procesar y generar lenguaje natural de manera avanzada, han demostrado ser efectivos en tareas como responder preguntas, generar explicaciones detalladas y proporcionar asistencia personalizada.

La IA se considera una disciplina de la informática que tiene por objetivo permitir que las computadoras y sistemas simulen la inteligencia humana, así como la capacidad para desempeñar tareas y resolución de problemas. Por otra parte, la IA puede definirse como “el estudio de agentes que reciben percepciones del entorno y realizan acciones” [Russell, 2021] con el objetivo de maximizar sus posibilidades de éxito. Esta definición resalta la capacidad de las máquinas para imitar funciones cognitivas humanas a través de algoritmos.

Por otra parte, en 2017 se presentó el modelo *Transformer* [Vaswani, 2017], revolucionando el procesamiento de lenguaje natural con un mecanismo de atención que permite enfocarse de manera selectiva en diferentes partes de la entrada de texto, mejorando la eficiencia y la capacidad de capturar relaciones contextuales complejas en el texto. El mecanismo de atención cambio la forma de

entrenar los modelos y dio lugar a la creación de los Modelos de Lenguaje Grandes (LLM por sus siglas en inglés) por parte de las grandes compañías, entre los LLM más populares se encuentran BERT, GPT, LLaMa, Mistral, y Gemini.

El modelo GPT desarrollado por OpenAI [OpenAI, 2023], se introdujo en 2018 con la ventaja de haber utilizado entrenamiento no supervisado y con un ajuste fino para las tareas específicas. Su arquitectura está basada en el uso de decodificación del modelo *Transformer*. GPT ha tenido una evolución bastante rápida desde el lanzamiento de GPT-2 en 2019 donde sorprendió con su capacidad para generar texto coherente. En 2020 se presentó GPT-3 con 175 millones de parámetros, el cual tenía una gran capacidad para realizar tareas con breves instrucciones. En 2023 el GPT-4 mejoró la comprensión y el razonamiento lógico, ampliando su dominio a diferentes idiomas y tareas especializadas.

Por otra parte, en 2023, Meta presentó la primera versión de Llama (Large Language Model Meta AI), un modelo de lenguaje de código abierto que incluía varias versiones disponibles, que iban desde siete mil millones a sesenta y cinco mil millones de parámetros. Estos modelos fueron entrenados con únicamente datos públicos, demostrando que se podían entrenar modelos que obtuvieran grandes resultados sin la necesidad de datos probados. En 2023 surgió Llama 2, desarrollado en colaboración con Microsoft y también fueron liberados de manera accesible permitiendo a la investigación académica y la industria privada experimentar con modelos de alto rendimiento, eliminando las barreras que antes solo las grandes empresas podían superar [Meta AI, 2024].

El uso de la inteligencia artificial (IA) en la educación tiene el potencial de convertirse en una gran herramienta para facilitar el acceso, equidad y calidad del aprendizaje. El fácil acceso a esta nueva tecnología tiene la capacidad de reducir las barreras al conocimiento y mejorar los métodos pedagógicos aplicados actualmente.

Los *chatbots* de inteligencia artificial son aplicaciones o interfaces que pueden mantener conversaciones similares a los humanos mediante el procesamiento de lenguaje natural. Estos aprovechan los Large Language Models (LLM) para generar respuestas de texto, son capaces de responder preguntas, brindar orientación, resolver dudas, ejecutar tareas específicas, entre otras cosas.

El Aprendizaje Basado en Problemas (ABP) es una metodología que ha ganado popularidad en los últimos años debido a sus beneficios en la educación. Su enfoque coloca al estudiante en el centro del proceso de aprendizaje, dándole la responsabilidad de construir su propio conocimiento a través de la resolución activa de problemas. Esta busca fomentar el desarrollo de habilidades cognitivas, metacognitivas y sociales a través de la resolución de problemas reales.

El objetivo de este trabajo fue desarrollar una aplicación educativa que implemente la metodología de aprendizaje basado en problemas, usando un tutor virtual basado en inteligencia artificial generativa, con el fin de guiar a los estudiantes como lo haría un maestro. Esta aplicación está enfocada en temas de matemáticas básicas de nivel secundaria con base en el programa educativo oficial vigente en México.

2. Métodos

A partir de la identificación del problema, se planteó la necesidad de innovar en la metodología de aprendizaje. Se propuso adoptar la tecnología como apoyo al implementar el ABP a través de un tutor virtual capaz de proporcionar ayuda personalizada a los estudiantes. El tutor virtual se implementa usando inteligencia artificial generativa con el objetivo de guiar al estudiante en su proceso de aprendizaje mediante la resolución de problemas, proporcionando ayuda y explicaciones para que el estudiante logre una comprensión más profunda del tema y desarrolle un pensamiento crítico.

Con la finalidad de garantizar la disponibilidad 24/7, se optó por desarrollar una aplicación web para apoyar a los estudiantes de educación básica en el aprendizaje de las Matemáticas, específicamente en el tema del Teorema de Pitágoras. La herramienta está diseñada para guiar al estudiante a través del proceso de aprendizaje de manera personalizada. En las siguientes secciones, se describen las etapas del desarrollo de software.

Análisis de requisitos

Para definir los requisitos del sistema se llevó a cabo una investigación de los temas que forman parte de la asignatura de Matemáticas en México en el nivel de

educación básica (Secundaria) [Secretaría de Educación Pública, 2022b] abarcando los grados 1º, 2º y 3º. Además, se realizaron entrevistas con docentes de diversas instituciones educativas, tanto públicas como privadas, que imparten la asignatura de Matemáticas. A partir de la información recopilada en este proceso, se establecieron los requerimientos funcionales y de calidad del sistema.

Requisitos funcionales

Para asegurar el correcto funcionamiento de la aplicación, es necesario identificar con precisión la funcionalidad del sistema, esto garantiza que el usuario alcance el objetivo final: aprender conceptos matemáticos, así como resolver problemas del tema. Los requisitos funcionales describen las características y operaciones que el sistema debe realizar para cumplir el propósito de la herramienta educativa. En la Tabla 1 se presentan algunos requisitos funcionales definidos para TutorIA.

Tabla 1 Requisitos funcionales.

RF	Descripción	Prioridad
1	El sistema debe permitir la autenticación de los estudiantes.	Alta
2	El sistema debe permitir al estudiante elegir el tema que quiere estudiar.	Alta
3	El sistema implementa ABP para presentar problemas matemáticos.	Alta
4	El sistema debe contener variedad de ejercicios matemáticos para resolver.	Media
5	El sistema debe presentar los ejercicios con una imagen y su descripción.	Media
6	La IA debe responder como un tutor, guiando al estudiante sin dar la solución.	Alta
7	El sistema debe registrar el avance del estudiante.	Alta
8	El sistema debe permitir interacción a través de voz.	Baja
9	El sistema debe permitir al estudiante escuchar la recomendación del Tutor.	Baja

Fuente: elaboración propia.

Requisitos de calidad

Los requisitos de calidad (RC) garantizan que el sistema cumpla ciertas características indispensables para el correcto funcionamiento, para esto existen varios tipos de requisitos de calidad, para identificarlos se utilizó la norma ISO 25000 [International Organization for Standardization, 2014] la cual establece categorías de los requisitos de calidad como se muestra en la Tabla 2. La Tabla 3 despliega los requisitos de calidad del sistema, organizados por categoría. Estos requisitos aseguran el correcto funcionamiento del sistema y sirven como referente para la toma de en el diseño y selección de sus componentes.

Tabla 2 Tipos de requisitos de calidad.

Id	Tipos de RC	Descripción
01	Eficiencia	Evalúa el rendimiento del sistema en términos de uso de recursos (CPU, memoria, almacenamiento, etc.) y el tiempo de respuesta.
02	Compatibilidad	Garantiza que el software pueda ejecutarse correctamente en diferentes entornos.
03	Usabilidad	Evalúa si el sistema es intuitivo y fácil de usar, proporciona una experiencia de usuario agradable y eficiente.
04	Fiabilidad	Determina la capacidad del sistema para mantener el funcionamiento estable y libre de fallos, garantizando disponibilidad.
05	Seguridad	Garantiza la confidencialidad, integridad y disponibilidad de la información.
06	Mantenibilidad	Facilidad con la que se modifica el software.

Fuente: elaboración propia.

Tabla 3 Requisitos de calidad del sistema.

RC	Descripción	Id
1	El sistema genera retroalimentación en menos de dos segundos.	01
2	La aplicación debe ejecutarse en cualquier navegador web.	02
3	La aplicación debe ser intuitiva para los estudiantes.	03
4	El sistema debe estar disponible las 24 horas, todos los días.	04
5	La información del usuario debe estar cifrada.	05
6	El sistema permite el ingreso solo a usuarios autenticados y registrados.	05
7	El sistema debe documentarse para mantenimiento futuro.	06

Fuente: elaboración propia.

Actores

En esta sección se describen elementos externos que interactúan con la aplicación, identificados a partir de los requisitos funcionales. En la Tabla 4 se muestran los actores, su descripción y objetivo que cada uno tiene en la aplicación.

Tabla 4 Actores del sistema.

Actor	Descripción	Objetivo
Estudiante	Actor principal, Interactúa con la plataforma para aprender matemáticas y solucionar problemas. Realiza consultas a la inteligencia artificial.	Resolver problemas matemáticos. Fortalecer el pensamiento crítico, hacer un seguimiento de su aprendizaje.
Tutor Virtual	Basado en un LLM, actúa como un agente pedagógico virtual que guía del estudiante, analizando las respuestas y brindarle un apoyo personalizado. Funciona a través de la API de Groq.	Proporciona retroalimentación al estudiante, da pistas y analiza las ideas del estudiante para para potenciar el aprendizaje.

Fuente: elaboración propia.

Arquetipos

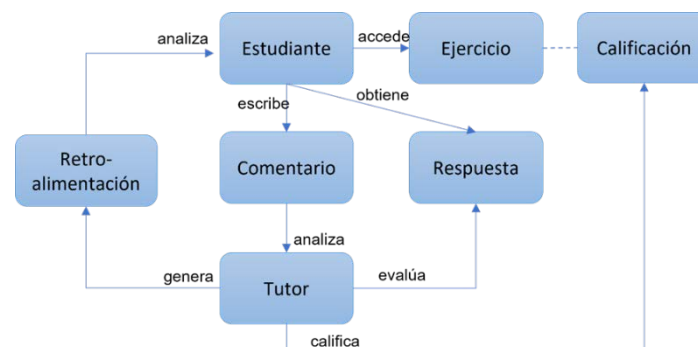
En este apartado se presentan los arquetipos utilizados para representar los elementos internos del sistema mediante abstracciones que representan la

interacción entre ellos. Los arquetipos sirven para definir los componentes clave del sistema, facilitar el diseño de interfaces y priorizar su funcionalidad dentro de la aplicación. La Tabla 5 describe los arquetipos definidos para TutorIA y la Figura 1 muestra la relación entre los arquetipos.

Tabla 5 Descripción de Arquetipos.

Arquetipos	Descripción
Estudiante	Actor principal del sistema, interactúa con la plataforma resolviendo los ejercicios diseñados bajo la metodología ABP.
Tutor	Analiza las ideas, comentarios, dudas y observaciones del estudiante. Genera retroalimentación. Evalúa la respuesta y asigna calificación.
Ejercicio	Representación de un problema matemático diseñado bajo la metodología ABP. Proporcionan el escenario de aprendizaje para el estudiante.
Comentario	Son ideas, dudas, comentarios o preguntas que el estudiante escribe al tutor para que lo guíe a la solución del problema.
Respuesta	Es la explicación del proceso usado por el alumno para resolver el problema junto con el resultado obtenido.
Calificación	Valoración cuantitativa y cualitativa asignada a la resolución de un ejercicio, incluye las respuestas, número de intentos para encontrar la solución y el proceso usado por del estudiante al resolver el problema.
Retroalimentación	Sugerencias y comentarios generados por el tutor virtual con el propósito de guiar al estudiante en el proceso de resolver los ejercicios.

Fuente: elaboración propia.



Fuente: elaboración propia.

Figura 1 Diagrama de arquetipos.

Casos de uso

Los casos de uso representan los escenarios en los que uno o más actores interactúan con el sistema. Cada caso de uso describe una secuencia específica de acciones que los actores realizan para alcanzar un objetivo. En la Tabla 6 se

presentan algunos casos de uso que representan la funcionalidad esencial para garantizar el cumplimiento de los objetivos educativos de TutorIA.

Tabla 6 Casos de uso.

Código	Nombre	Descripción
CU – 01	Iniciar sesión	Permite al usuario autenticarse en el sistema mediante credenciales válidas (cuenta de usuario y contraseña).
CU – 02	Resolver ejercicio	El estudiante accede para resolver un ejercicio en el sistema.
CU – 03	Recibir retroalimentación	El tutor virtual analiza los comentarios del estudiante y proporciona retroalimentación al estudiante.
CU – 04	Consultar progreso	Permite al usuario consultar los ejercicios resueltos y número de intentos.
CU – 05	Consultar al tutor	El estudiante se comunica con el tutor virtual cuando tiene dudas o necesita de una aclaración.
CU – 06	Seleccionar ejercicio	El usuario puede seleccionar un tema y el ejercicio que quiere resolver.
CU – 07	Cerrar sesión	Permite al usuario finalizar la sesión activa y salir del sistema.

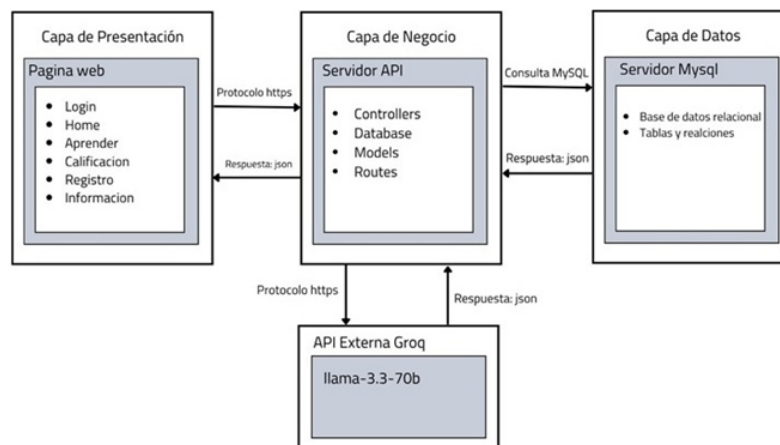
Fuente: elaboración propia.

Arquitectura

Para la arquitectura del sistema se seleccionó el modelo cliente-servidor utilizando una arquitectura de capas (Figura 2), este enfoque permite organizar el sistema en diferentes niveles, capa de presentación, capa de negocios y la capa de datos, distribuyendo estas en los componentes cliente y servidor:

- **Capa de presentación.** La Capa de Presentación es la parte del sistema encargada de interactuar con el usuario. Su principal función es recibir las entradas del usuario, enviarlas a la Capa de Negocio para su procesamiento y mostrar resultados de manera clara y comprensible. Esta capa puede implementarse en un navegador web como una aplicación de escritorio mediante una interfaz gráfica de usuario. La Capa de Presentación se suele desarrollar utilizando diferentes tecnologías.
- **Capa de negocios.** Esta capa es la responsable de gestionar toda la lógica del sistema, en ésta se encuentran las reglas que definen el funcionamiento del sistema. Su función principal consiste en procesar los datos recibidos desde la Capa de Presentación, y, cuando es necesario, realizar consultas a la Capa de datos. También es la capa encargada de la comunicación con sistemas externos.

- **Capa de datos.** Es la capa responsable de almacenar y gestionar la información necesaria para el funcionamiento del sistema. Mantiene comunicación con la Capa de Negocio para almacenar, recuperar, actualizar o eliminar información. Para esta aplicación se seleccionó el gestor de base de datos MySQL, debido a su estabilidad y facilidad de uso, gran comunidad de soporte y su alta compatibilidad con aplicaciones web. Permite llevar un control de los datos y es compatible con las tecnologías que se utilizan. Está diseñado para relacionar las tablas entre sí a través de claves, facilitando la consulta de los datos.



Fuente: elaboración propia.

Figura 2 Arquitectura de TutorIA.

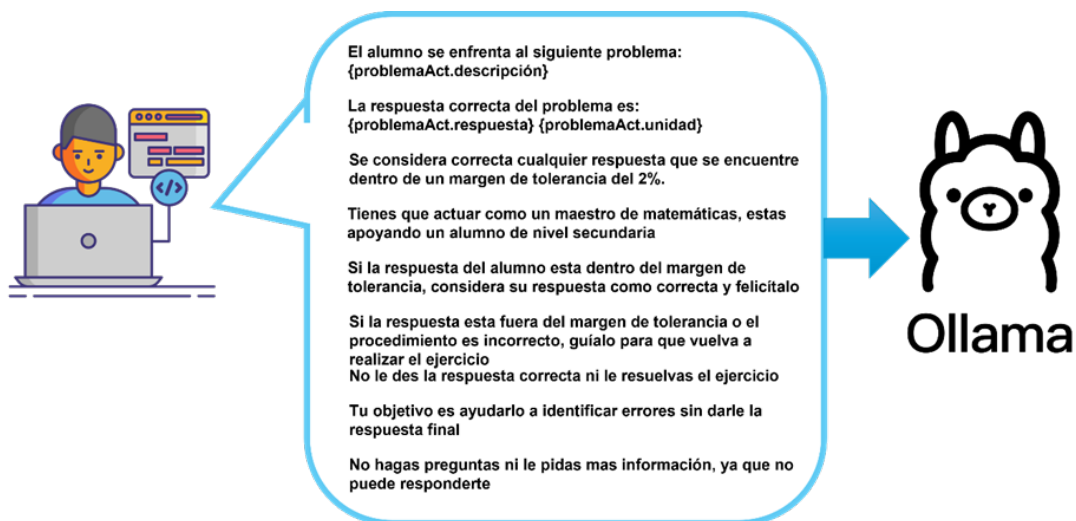
Implementación del tutor virtual

Para la implementación del tutor virtual se utilizó un Modelo de Lenguaje Grande (LLM en inglés Large Language Model). Estos modelos de inteligencia artificial permiten generar respuestas personalizadas y pueden realizar instrucciones complejas, esto se logra por medio del uso de prompts, que son instrucciones en lenguaje natural que guían el comportamiento esperado del modelo de lenguaje. Para este trabajo se implementó el modelo de LLaMA “*llama – 3.3 – 70b*”. Este modelo fue integrado a través de una API, permitiendo obtener respuestas rápidas y sin la necesidad de alojar el modelo en un servidor propio. A continuación, se detalla la información acerca de las decisiones tomadas para la elección del modelo y su implementación.

LLaMA 3.3 – 70B cuenta con setenta mil millones de parámetros, una versión avanzada de la familia de modelos de LLaMA desarrollado por la compañía Meta. Esta versión fue lanzada en abril de 2024, la cual destaca por su rendimiento mejorado en comparación a versiones anteriores. El modelo sobresale en tareas como seguir instrucciones complejas, razonamiento, programación y análisis de información [Meta AI, 2024].

Prompt utilizado para el modelo de lenguaje

El prompt enviado al modelo de lenguaje es fundamental para obtener el comportamiento esperado de la inteligencia artificial, por lo que para este trabajo se diseñó un *prompt* para el rol “System” que incluye las siguientes consideraciones: contextualización, instrucciones pedagógicas, restricciones e instrucciones de objetivo (Figura 3).



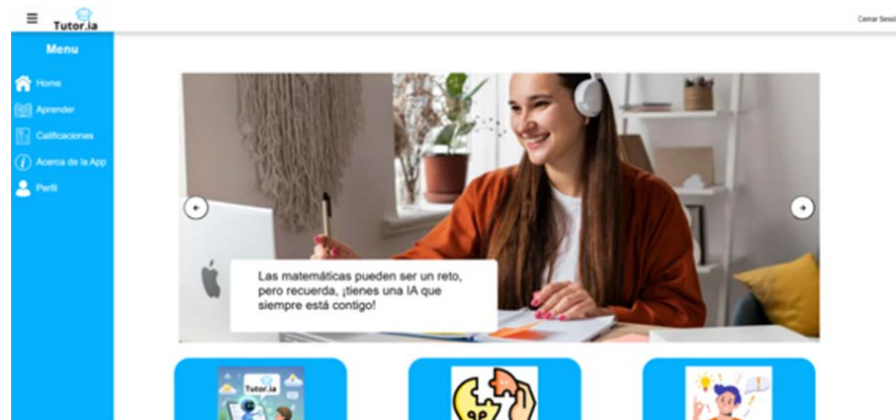
Fuente: elaboración propia.

Figura 3 Proceso del prompt.

Interfaz gráfica

Para el desarrollo de la interfaz gráfica se estableció como requisito principal la facilidad de uso, considerando que el sistema está dirigido a estudiantes de secundaria. En este sentido, la interfaz debe ser clara, intuitiva y accesible, de manera que favorezca la interacción con el sistema sin requerir conocimientos

técnicos previos. Además, se busca que la interfaz sea motivadora, incorporando elementos gráficos que alienten al estudiante a seguir utilizándola (Figura 4). Para lograr esto se emplearon distintas tecnologías de programación como son: React [Meta, 2023], CSS y HTML.



Fuente: elaboración propia.

Figura 4 Pantalla de ingreso.

Pruebas

En este apartado, se abordan las intervenciones realizadas con los estudiantes. Se optó por utilizar un diseño cuasi experimental, con el propósito de identificar una posible relación de causa-efecto en el uso de TutorIA. Para esto, se diseñaron los instrumentos de medición orientados a evaluar si tanto la ganancia de aprendizaje en los estudiantes como la aceptación de la herramienta. Este tipo de diseño permitió realizar una intervención controlada, sin requerir la asignación aleatoriamente a los participantes.

La intervención se llevó a cabo en una institución privada de Culiacán, Sinaloa con una muestra de 25 *participantes* ($N = 25$) de segundo y tercer grado de secundaria. La población estuvo conformada por 11 *estudiantes* de segundo grado y 14 *estudiantes* de tercer grado, con edades comprendidas entre 13 y 15 años.

El procedimiento experimental se llevó a cabo en dos sesiones. En la primera sesión, se presentó el tema y se aplicó el instrumento de medición Pretest. En la segunda sesión, los estudiantes tuvieron 50 *minutos* de tiempo total para uso de la herramienta de manera libre, por lo que cada estudiante avanzó a su ritmo

resolviendo los ejercicios. Con el objetivo de analizar el uso que se le dio a la herramienta, se registraron los ejercicios resueltos por cada estudiante, cantidad de intentos y las interacciones que tuvieron con el tutor virtual. Posteriormente, se procedió a dar respuesta al instrumento posttest. Finalmente, los estudiantes dieron respuesta a la encuesta de aceptación de la herramienta. La Figura 5 ilustra el proceso de la intervención con estudiantes en el uso de TutorIA.



Fuente: elaboración propia.

Figura 5 Proceso de la intervención

Instrumentos de medición

Con el propósito de evaluar el impacto de la herramienta de aprendizaje TutorIA sobre los estudiantes en el aprendizaje del Teorema de Pitágoras, se diseñaron tres instrumentos de medición: pretest, posttest y encuesta de aceptación.

El instrumento pretest es una prueba que se responde antes de la intervención, su finalidad es medir el conocimiento previo del estudiante sobre el tema; mientras que el posttest se responde después del uso de la herramienta para evaluar los conocimientos adquiridos. Ambos instrumentos con el mismo grado de complejidad constan de ocho preguntas, cinco de opción múltiple y tres preguntas prácticas relacionados con el Teorema de Pitágoras. Cada pregunta tiene cuatro respuestas donde solamente una opción es correcta. El tercer instrumento corresponde a la encuesta de aceptación, integrado por diez reactivos orientados a valorar la utilidad percibida, la motivación, la autonomía, facilidad de uso y la confianza en la inteligencia artificial. Este instrumento busca medir el interés que tienen los estudiantes por el uso de tecnologías basadas en IA en contextos educativos.

3. Resultados

En este apartado se presentan los resultados derivados de la intervención realizada con los estudiantes mediante el uso de la herramienta. Además, se hace una interpretación de los datos con el propósito de obtener conclusiones significativas respecto al impacto de la aplicación en el aprendizaje y la aceptación de esta. A continuación, se analizan los resultados obtenidos en las pruebas pretest, posttest y la encuesta de aceptación.

Resultados pretest

En este apartado se presentan, en primer lugar, los resultados del grupo de segundo grado, seguido por los de tercer grado y finalmente, de la población total. El grupo de segundo grado ($N = 11$), sin conocimientos previos en el tema, reflejó en un desempeño bajo y acorde a lo esperado ($M = 2.38$, $SD = 1.35$). Este resultado confirma que, al ser un tema no abordado previamente en el aula, los estudiantes presentaron un nivel inicial limitado en el tema. Por otra parte, el grupo de tercer grado ($N = 14$), quienes ya habían visto el tema semanas previas, obtuvo un puntaje mayor ($M = 3.93$, $SD = 1.88$).

En la población total, el puntaje obtenido fue bajo ($M = 3.25$, $SD = 1.87$), lo que evidencia que, en conjunto, los estudiantes partían de una base conceptual insuficiente sobre el Teorema de Pitágoras.

Uso de la herramienta

Durante el uso de la herramienta, el grupo de segundo grado resolvió en promedio 3.36 ejercicios ($M = 3.36$, $SD = 1.02$) en 5.45 intentos ($M = 5.45$, $SD = 2.84$) reflejando un desempeño inicial limitado por la falta de conocimientos previos. El grupo de tercer grado mostró un rendimiento ligeramente superior, con un promedio de 3.62 ejercicios resueltos ($M = 3.62$, $SD = 1.50$) y 5.11 intentos ($M = 5.11$, $SD = 2.84$), lo que evidencia una mayor eficiencia atribuible a su experiencia previa. En la muestra total, el promedio fue de 3.47 ejercicios resueltos ($M = 3.47$, $SD = 1.21$) y 5.30 intentos ($M = 5.30$, $SD = 2.55$). Estos resultados sugieren que, aún con un conocimiento inicial limitado, los estudiantes lograron avanzar en la

comprensión del tema, mientras que quienes tenían preparación previa resolvieron los ejercicios ligeramente mejor.

Resultados posttest

Para evaluar el conocimiento adquirido se aplicó el posttest. El grupo de segundo grado mostró una mejora significativa ($M = 5.57$, $SD = 1.62$), lo que indica que la aplicación tuvo un impacto positivo en el puntaje de los estudiantes sin conocimientos previos. El grupo de tercero también mejoró su desempeño ($M = 5.89$, $SD = 2.79$), aunque en menor medida, lo que sugiere una consolidación de conocimientos previos. Finalmente, al considerar la población total, el promedio general fue favorable ($M = 5.75$, $SD = 2.31$), reflejando que el uso de TutorIA contribuyó tanto a la adquisición como el refuerzo de conocimientos en el tema.

Resultados de la encuesta de aceptación

Para evaluar la aceptación y motivación de los estudiantes se aplicó una encuesta basada en la escala de Likert (escala 1 – 5). El promedio general obtenido ($M = 3.63$) indica una percepción positiva del uso de la herramienta. Los estudiantes valoraron la ayuda del tutor virtual para comprender el tema, la claridad en las explicaciones, indicaron confianza en la inteligencia artificial como apoyo de aprendizaje. No obstante, el promedio también refleja que existen áreas de mejora en el diseño de la herramienta y en la integración de la IA. Además, también refleja que este tipo de herramientas tienen potencial significativo en el ámbito educativo. Un ejemplo de un reactivo de esta encuesta es: “Las explicaciones del tutor virtual fueron claras, fáciles de entender, y útiles para mi aprendizaje”.

Comparación de resultados entre Pretest y Posttest

Para evaluar el impacto de la herramienta se compararon los resultados entre el pretest y el posttest mediante análisis estadístico. Dado que los datos no cumplieron con una distribución normal, se optó por aplicar pruebas no paramétricas. Se aplicó la prueba de los rangos con signo de Wilcoxon para determinar diferencias significativas entre los datos del pretest y posttest. En el caso del grupo de segundo

grado, el puntaje promedio pasó de 2.38 a 5.57 mostrando una diferencia estadísticamente significativa ($Z = -2.675$, $p = 0.007$), lo que confirmó el impacto positivo de la herramienta en estudiantes sin conocimientos previos. De manera similar, el grupo de tercer grado, el promedio aumentó de 3.93 a 5.89, donde existe una diferencia estadísticamente significativa ($Z = -2.218$, $p = 0.027$). En ambos casos, la prueba de Wilcoxon permitió evidenciar mejoras en la ganancia de aprendizaje tras la intervención con TutorIA. Para evaluar la diferencia global entre pretest y posttest sobre la población total, se aplicó la prueba de U de Mann-Whitney. Los resultados mostraron un rango promedio de 17.52 en el pretest y 33.48 en el posttest, con valores estadísticamente significativos ($U = 113.0$, $Z = -3.937$, $p < 0.001$). Esto permite confirmar una diferencia estadísticamente significativa en el desempeño general de los estudiantes. Este hallazgo respalda que TutorIA produjo un incremento en la ganancia de aprendizaje, tanto en estudiantes con conocimientos previos como en aquellos sin experiencia en el tema.

4. Discusión

Los resultados de la intervención en rangos generales tuvieron un impacto positivo en la ganancia de aprendizaje del Teorema de Pitágoras. La intervención, diseñada en dos sesiones, permitió identificar factores relevantes que pudieron influir en el desempeño de los estudiantes. Durante la aplicación del Pretest, se observó una falta de interés o baja disposición inicial por parte de los estudiantes, posiblemente por tratarse de un examen tradicional. Los estudiantes mostraron poca disposición para resolver los ejercicios de la prueba, lo cual pudo haber influido en los resultados. Sin lugar a duda la motivación de los estudiantes fue un factor muy importante que se hizo notar desde las primeras etapas de la intervención. Al inicio, durante el uso de la herramienta, los estudiantes requirieron de un tiempo de adaptación para comprender el funcionamiento, por lo que poco a poco se pudo percibir un aumento en el interés de los estudiantes.

El análisis estadístico de los datos y la observación directa durante la intervención confirmaron que sí hubo un avance significativo en el aprendizaje, siendo la etapa del uso de la aplicación donde más se notó la motivación y el entusiasmo por parte

de los estudiantes. Aquellos alumnos con mayor entusiasmo lograron resolver un mayor número de ejercicios, lo que favoreció el aprovechamiento de la herramienta. Aunque solo una minoría mostró resistencia al uso de la aplicación, su participación limitada impactó ligeramente en los promedios de las pruebas. Algunos de estos estudiantes decidieron hacer un uso mínimo de la herramienta en el tiempo establecido, lo cual se vio reflejado en los resultados de la prueba Posttest.

En general, estos hallazgos indican que la motivación de los estudiantes juega un rol fundamental en el proceso de aprendizaje asistido por tecnologías que incorporan inteligencia artificial, y que la herramienta además de facilitar la comprensión del Teorema de Pitágoras favoreció a aumentar la disposición de los estudiantes a la solución de problemas matemáticos.

5. Conclusiones

El uso de un software tutor que combina la metodología de aprendizaje basado en problemas con un modelo de inteligencia artificial generativa demostró favorecer tanto el aprendizaje de las Matemáticas como el desarrollo habilidades del pensamiento crítico y resolución de problemas.

Los resultados obtenidos fueron prometedores. El análisis mediante la prueba de rangos con signo de Wilcoxon reveló diferencias estadísticamente significativas entre los puntajes del Pretest ($M = 3.25$) y del Posttest ($M = 5.75$), con un valor de significancia de $p = 0.001$, lo que evidencia una mejora significativa en la ganancia de aprendizaje. Por otra parte, los resultados de la prueba U de Mann-Whitney confirmaron diferencias significativas entra las calificaciones de Pretest y Posttest ($U = 113.0$, $Z = -3.937$, $p < 0.00$), lo que permitió confirmar que los estudiantes lograron un aprendizaje significativo tras el uso de la herramienta.

Estos hallazgos confirman la validez del uso de LLM para el aprendizaje con la metodología ABP [Maalek, 2024] y son consistentes con los resultados de otros estudios que destacan la eficiencia de tutores virtuales en mejorar tanto el rendimiento académico como el factor motivacional de los estudiantes [Ateş, 2025]. Los resultados sugieren un avance en el desempeño y una percepción favorable del uso de TutorIA. Sin embargo, conviene interpretarlos con mesura. La evaluación se

realizó con una muestra acotada ($N = 25$) en una sola institución, durante un periodo breve y enfocada únicamente en el Teorema de Pitágoras, por lo que no es posible generalizar los hallazgos a otros contextos, niveles o contenidos sin evidencia adicional. Además, al tratarse de un diseño cuasi experimental, es posible que variables no controladas, tales como diferencias previas entre grupos, motivación o efecto de novedad, que hayan podido influir en los resultados.

Como trabajo futuro, se puede contemplar ampliar la muestra con mayor número de estudiantes e incluso, con diferentes contextos sociales y educativos. Asimismo, incorporar otros temas de educación básica, a otros temas e incluso a una unidad completa de un curso con el fin de ampliar los datos por recopilar y realizar un análisis estadístico más amplio, completo y robusto que permita afirmar el impacto de este tipo de herramientas con énfasis en el ámbito educativo.

6. Referencias y Bibliografía

- [1] Ateş, H., (2025). Integrating Augmented Reality into Intelligent Tutoring Systems to Enhance Science Education Outcomes. *Education and Information Technologies*, 30(4), pp. 4435-4470. <https://doi.org/10.1007/s10639-024-12970-y>.
- [2] International Organization for Standardization, (2014). *Systems and Software Engineering - Systems and Software Quality Requirements and Evaluation (SQuaRE) - Guide to SQuaRE (ISO/IEC 25000:2014)*. ISO. Geneva, Suiza. <https://www.iso.org/standard/64764.html>.
- [3] Maalek, R., (2024). Integrating Generative Artificial Intelligence and Problem-Based Learning into the Digitization in Construction Curriculum. *Buildings*, 14(11), 3642. <https://doi.org/10.3390/buildings14113642>.
- [4] Mendieta, J. B., El Aprendizaje Basado en Problemas para Mejorar el Pensamiento Crítico: Revisión Sistemática. *INNOVA Research Journal*, 6(2), pp. 77-89. Mayo, 2021.
- [5] Meta (2023). *React – A Javascript Library for Building User Interfaces*. React, California, USA. <https://react.dev>.

- [6] Meta AI., (2024). Llama: Open Foundation Models. Llama.com, California.
<https://www.llama.com/>.
- [7] OpenAI (2023). GPT-4 Technical Report. openai.com, San Francisco,
disponible en: <https://openai.com/research/gpt-4>.
- [8] Organization for Economic Co-operation and Development, (2023). PISA
2022 Results (Volume I): The State of Learning and Equity in Education.
OECD Publishing, París, Francia.
- [9] Russell, S. J., Norvig, P. Artificial Intelligence: A Modern Approach, Global
Edition (4th ed.). Pearson. Abril 2021.
- [10] Secretaría de Educación Pública, (2022a). Avance del Contenido para el
Libro del Docente. Primer grado. SEP, Ciudad de México.
- [11] Secretaría de Educación Pública (2022b). Plan de Estudio para la Educación
Preescolar, Primaria y Secundaria 2022. Dirección General de Desarrollo
Curricular, Ciudad de México, México.
[https://educacionbasica.sep.gob.mx/materiales-de-apoyo-a-la-apropiacion-
del-plan-y-programas-de-estudio-2022/](https://educacionbasica.sep.gob.mx/materiales-de-apoyo-a-la-apropiacion-del-plan-y-programas-de-estudio-2022/).
- [12] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.,
Kaiser, L., & Polosukhin, I., Attention is All you Need. Advances in Neural
Information Processing Systems, Curran Associates, Inc., vol. 30, Long
Beach. December 2017.