



Evaluación de un chatbot basado en aprendizaje supervisado: impacto en la satisfacción y propuestas de mejora

Evaluation of a Supervised Learning-Based Chatbot: Impact on Satisfaction and Improvement Proposals

  Nuria Segovia-García (N.S.-G). Centro de Investigación de Asturias (España)

  Alfredo Guzmán Rincón (A.G.R.). Centro de Investigación de Asturias (España)

RESUMEN

En un contexto donde la rapidez en la resolución de dudas es esencial, las instituciones de educación superior deben reformular sus modelos de atención estudiantil. Tecnologías como los chatbots, basados en procesamiento de lenguaje natural, ofrecen interacciones ágiles y precisas, optimizando los recursos universitarios. Este estudio describe el diseño, implementación y evaluación de un chatbot entrenado mediante aprendizaje supervisado, utilizando preguntas frecuentes de estudiantes reales. Se empleó la metodología Scrum para iterar y ajustar la herramienta en función del feedback recibido, monitoreándola durante un año para mejorar su efectividad. Los resultados evidencian que el chatbot reduce significativamente los tiempos de espera y es percibido como útil para consultas básicas. Sin embargo, se identificaron limitaciones en su fluidez para manejar interacciones más complejas, destacando la necesidad de modelos avanzados de lenguaje que permitan una experiencia más natural y adaptativa. El estudio concluye que, aunque el chatbot representa una herramienta valiosa para mejorar la atención estudiantil, alcanzar un nivel óptimo de satisfacción requiere integrar procesos de ajuste continuo, como el fine-tuning, y desarrollar funcionalidades más especializadas. Estas mejoras podrían transformar el uso de chatbots en una solución efectiva y escalable dentro del ámbito educativo.

ABSTRACT

In a context where the speed of resolving queries is essential, higher education institutions must rethink their student support models. Technologies such as chatbots, based on natural language processing, provide agile and precise interactions, optimizing university resources. This study outlines the design, implementation, and evaluation of a chatbot trained through supervised learning, using real student FAQs. The Scrum methodology was employed to iterate and adjust the tool based on feedback, with monitoring over a year to enhance its effectiveness. The results show that the chatbot significantly reduces response times and is perceived as useful for basic queries. However, limitations were identified in its ability to handle more complex interactions, emphasizing the need for advanced language models to enable a more natural and adaptive user experience. The study concludes that, while the chatbot is a valuable tool for improving student support, achieving optimal satisfaction requires integrating continuous adjustment processes, such as fine-tuning, and developing more specialized functionalities. These enhancements could transform the use of chatbots into an effective and scalable solution within the educational sector.

PALABRAS CLAVE - KEYWORDS

Inteligencia artificial; chatbots, aprendizaje supervisado, procesamiento de lenguaje natural, satisfacción del usuario

Artificial Intelligence, chatbots, supervised learning, natural language processing, user satisfaction



1. INTRODUCCIÓN

El sector educativo está experimentando una era de transformación profunda que redefine de forma global todos los procesos vinculados a este ámbito (García-Peñalvo, 2023; Segovia-García, 2023). El impulso de tecnologías como la inteligencia artificial (IA), asistentes virtuales, o plataformas de aprendizaje adaptativo junto con la incorporación de otras como el Procesamiento del Lenguaje Natural (PLN), el Aprendizaje Automático (AA) y los Modelos de Lenguaje de Gran Escala (LLMs) está revolucionando la manera en cómo se comunican los sistemas educativos, permitiendo una interacción más dinámica, flexible y personalizada que trasciende las barreras que imponen espacio y el tiempo.

En este sentido, una de las herramientas que ejemplifica esta transformación son los chatbots impulsados por IA que son capaces de comprender y responder al lenguaje humano con respuestas en texto o audio, ofreciendo una solución inmediata a las dudas y necesidades de los estudiantes. Esta innovación no sólo acelera el proceso de aprendizaje, sino que también eleva la calidad de la experiencia educativa, brindando un soporte ágil y adaptado a las inquietudes individuales (Huang et al., 2021; Zhang et al., 2023).

Dentro del panorama de sistemas de asistencia virtual, uno de los más destacados es ChatGPT (Radford et al., 2019) que permite la interacción a través de chat con entidades virtuales capaces de generar respuestas a las preguntas de los usuarios mediante el análisis de grandes cantidades de datos. Si bien tanto ChatGPT como otros sistemas conversacionales basados en IA comparten el propósito común de brindar respuestas automáticas y asistencia a las consultas de los usuarios, difieren en su enfoque. Mientras algunos chatbots utilizan programaciones específicas y reglas predefinidas para sus respuestas, ChatGPT sobresale por su capacidad para comprender y generar respuestas contextualmente relevantes de manera dinámica (Radford et al., 2019). Este atributo se debe a su entrenamiento con grandes conjuntos de datos, lo que le permite adaptarse de forma más flexible a una variedad más amplia de consultas y contextos de conversación en comparación con muchos otros sistemas conversacionales basados en IA.

En el entorno universitario estas aplicaciones tienen el objetivo de prestar un servicio de atención automatizado que proporcione respuestas instantáneas a una amplia gama de preguntas que los estudiantes pueden tener, desde consultas sobre procesos de inscripción hasta detalles específicos sobre cursos y programas académicos (Labadze et al., 2023). Disponer de herramientas de este tipo es una solución a los problemas administrativos comunes, como la sobrecarga en los departamentos de atención al estudiante, la falta de respuesta oportuna, y la necesidad de automatizar procesos repetitivos. Problemas que afectan a la percepción de la calidad del servicio educativo por parte de los estudiantes, influyendo en su nivel de satisfacción (Belda-Medina & Kokošková, 2023; Segovia-García & Said-Hung, 2021).

La existencia de asistentes virtuales como vía para establecer conversaciones no es nueva, sus predecesores fueron los chatbots basados en reglas como ELIZA, creado por Joseph Weizenbaum en el MIT entre 1964 y 1966 (Sharma et al., 2017). Una herramienta que simulaba una conversación a través de la búsqueda de palabras clave en la entrada del usuario, aplicando reglas predefinidas para generar respuestas. Estos asistentes denominados de Respuesta de Interacción de Texto (ITR) o Dumb Chatbots, se caracterizan por seguir una secuencia de comandos para simular conversaciones a partir de un guion o script y cuyo principal problema

es que debe basarse en preguntas creadas de manera directa y simple, lo que limita su capacidad (Torres Martínez & Cruz Guerrero, 2020).

El avance que está experimentando la IA y las tecnologías como el reconocimiento y la síntesis de voz, han permitido que los chatbots hayan evolucionado hacia modelos cognitivos y conversacionales que pueden personalizarse para satisfacer las necesidades específicas de los estudiantes, ofreciendo soporte adaptativo y retroalimentación personalizada (Balderas et al., 2023; Ifelebuegu, et al., 2023; Segovia-García & Segovia-García, 2024). No obstante, a pesar de los beneficios que esta tecnología puede generar, es importante destacar los retos que, aún existen para ofrecer respuestas adecuadas o dirigir a los estudiantes hacia la información que buscan (Chaves & Gerosa, 2020; Haqimi & Kusuma, 2023). En este sentido, la literatura científica sobre el tema contempla un espacio de mejora importante para que los chatbot recojan realmente el contexto de una conversación, así como sus intenciones, incluyendo estilo, emoción o carácter (Lin, et al. 2023; Pawlik et al., 2022). Investigaciones recientes han destacado varias limitaciones, como la dificultad de estos sistemas para interpretar preguntas complejas o estructuras de oraciones poco convencionales, así como para manejar el lenguaje no estándar o contextual (Belda-Medina & Kokošková, 2023; Caldarini et al., 2022). Estos problemas pueden llevar a respuestas inadecuadas que, lejos de facilitar el diálogo, generan frustración en los estudiantes, comprometiendo su satisfacción y compromiso con su uso (Croes & Antheunis, 2022; Labadze et al., 2023; Stokel-Walker, 2022).

Por otro lado, la interacción con chatbots y otras formas de IA parece influir en la dinámica comunicativa de los usuarios. Estudios como el de Mou y Xu (2017) sugieren que los usuarios adoptan un comportamiento más abierto, amigable y extrovertido frente a seres humanos que ante la IA. Además, aunque estas herramientas continúan evolucionando, el análisis de textos que incluye elementos como el sentimiento que requiere de la complementación con interpretaciones humanas para una mayor precisión (Tabone & De Winter, 2023). Incluso, se ha observado que el uso de chatbots puede incrementar la ansiedad y el temor a evaluaciones o rechazos negativos en los usuarios (Ali et al., 2023), subrayando la necesidad de abordar estos aspectos emocionales en su diseño e implementación.

1.1. Diseño y Entrenamiento de Chatbots Educativos

A pesar de los desafíos inherentes a la implantación de estas tecnologías, estudios recientes han demostrado que un diseño meticuloso y una implementación rigurosa pueden superar los obstáculos presentes en entornos educativos y otros contextos (Lin et al., 2023). Los avances en la tecnología de IA han permitido mejoras significativas en las capacidades de los chatbots, en particular en su destreza para comprender y responder a preguntas complejas con mayor precisión. No obstante, para maximizar el potencial de estas herramientas, es crucial adoptar un enfoque de diseño bien estructurado que incluya la definición clara de objetivos, la comprensión profunda de las características de los usuarios finales y la anticipación de los tipos de consultas que deberán gestionarse.

También es necesario conocer con qué tipo de aprendizaje trabaja la herramienta puesto que existen varias aproximaciones, como: a) aprendizaje supervisado, donde un conjunto de datos etiquetados facilita la predicción de salidas, útil en contextos educativos con preguntas y respuestas predecibles (Torres & Cruz, 2020); b) aprendizaje no supervisado, que utiliza datos

no etiquetados para encontrar patrones subyacentes sin una instrucción explícita sobre la salida; c) aprendizaje por refuerzo, donde el modelo toma decisiones basadas en la experimentación en un entorno y se retroalimenta con recompensas o penalizaciones, adaptándose a las necesidades individuales de los usuarios (Aleedy et al., 2022; Keerthana et al., 2021; Uc-Cetina et al., 2022); d) un modelo híbrido que combine varios enfoques (Rojas, 2020) u otros procesos de entrenamiento adicional como el fine-tuning (o ajuste fino) que basándose en un modelo preentrenado, como ChatGPT que se ajusta con datos específicos para optimizar su desempeño en tareas delimitadas, permitiendo personalizar la salida según necesidades concretas (Latif & Zhai, 2024).

Este estudio tiene como objetivo evaluar el diseño, entrenamiento y desempeño de un chatbot basado en aprendizaje supervisado, implementado durante un año en una institución educativa. Concebida como una herramienta de apoyo para la atención estudiantil, esta solución tecnológica está orientada a resolver dudas frecuentes sobre procesos clave del entorno institucional, mediante respuestas estructuradas y predefinidas que optimizan la experiencia del usuario. En un contexto de rápida evolución tecnológica, donde estas herramientas pueden quedar obsoletas en poco tiempo, se analizan también los desafíos de su implementación, escalabilidad y adaptación. El análisis se centra en identificar los factores que influyen en la satisfacción del usuario, con especial atención a la claridad de las respuestas y la fluidez conversacional, a fin de proponer mejoras que fortalezcan su funcionalidad.

2. MÉTODO

En el contexto de una estrategia institucional orientada a la adopción de soluciones tecnológicas innovadoras, durante 2023 y 2024 se implementó y entrenó un asistente virtual desarrollado bajo un enfoque de aprendizaje supervisado. La iniciativa respondió a la necesidad de optimizar los procesos de atención al estudiante mediante la automatización de consultas frecuentes y específicas, con el fin de mejorar la eficiencia operativa y la experiencia del usuario en el entorno educativo.

A lo largo de un año, el equipo técnico desarrolló flujos conversacionales ajustados a los requerimientos institucionales, utilizando técnicas de aprendizaje supervisado que permitieron al sistema procesar el lenguaje de forma eficiente dentro de escenarios predefinidos. El proyecto se estructuró con base en la metodología Scrum, con el propósito de establecer un marco iterativo e incremental para el diseño e implementación del asistente. Esta metodología facilitó el desarrollo en ciclos cortos y repetitivos (sprints), permitiendo realizar ajustes continuos en función del comportamiento real de los usuarios (Medina et al., 2024; Mekić et al., 2024; Skuridin & Wynn, 2024).

Durante cada sprint el equipo se centró en completar un conjunto de tareas o funcionalidades permitiendo simplificar el abordaje de problemas complejos de una manera eficiente y facilitando la adaptación en tiempo real a los requerimientos a partir de la retroalimentación proporcionada por los usuarios (Pratama & Kristiana, 2023).

La metodología para el desarrollo del chatbot se inició en una primera fase definiendo equipo de trabajo e identificando los roles clave del proyecto y seleccionando la herramienta tecnológica a partir de la evaluación de las necesidades del proyecto, considerando aspectos

como funcionalidades requeridas, integración con sistemas existentes y facilidad de uso. Posteriormente se recopilaban los datos necesarios para proporcionar al chatbot el corpus de conocimiento adecuado garantizar una amplia gama de interacciones durante su entrenamiento. En esta fase se recogió información de diferentes fuentes depurando, corrigiendo y eliminando datos erróneos y estructurando la información en segmentos pertinentes para facilitar respuestas adecuadas y adaptando datos tabulares a formatos textuales.

Una vez completado el desarrollo iterativo del chatbot mediante los sprints, fue esencial realizar una evaluación exhaustiva de su desempeño para garantizar que cumpliera con los objetivos pedagógicos y técnicos definidos. Para ello se recopilaban y analizaban los resultados obtenidos durante cada sprint, aplicando métricas clave como número de interacciones, eficiencia en la interacción y la satisfacción del usuario. Y por otro lado se socializó un instrumento de evaluación diseñado específicamente para medir la satisfacción de los estudiantes con el Chatbot permitiendo identificar oportunidades de mejora basadas en datos concretos (Tabla 1).

Tabla 1

Instrumento de evaluación satisfacción

Ítem	Pregunta	Tipo de Respuesta	Soporte teórico
1	¿Conoces a Sofía, el chatbot implementado en Uniasturias para resolver dudas?	Sí / No	Balderas et al., 2023
2	¿Qué nivel de satisfacción tienes con tu experiencia general al usar el chatbot Sofía?		Nguyen et al., 2021; Segovia-García, N., y Segovia-García, L. 2024
3	¿Recomendarías el uso de Sofía a otros estudiantes?		Nguyen et al., 2021
4	¿La experiencia con Sofía ha cumplido tus expectativas?		Aleedy, M et al., 2022; Belda-Medina y Kokošková, 2023
5	¿Consideras que Sofía es eficiente en la resolución de tus dudas?	Escala Likert de 5 niveles	Balderas et al., 2023; Belda-Medina y Kokošková, 2023
6	¿Percibes que Sofía te ahorra tiempo en la búsqueda de respuestas?		Balderas et al., 2023; Belda-Medina y Kokošková, 2023
7	¿Encuentras útil el uso de Sofía para consultas académicas?		Aleedy, M et al., 2022; Balderas et al., 2023; Belda-Medina y Kokošková, 2023
8	¿Las respuestas de Sofía te parecen claras y comprensibles?		Aleedy, M et al., 2022;Balderas et al., 2023
9	¿Consideras que las respuestas de Sofía son precisas y relevantes?		Belda-Medina y Kokošková, 2023;

Ítem	Pregunta	Tipo de Respuesta	Soporte teórico
10	¿Sofía ofrece opciones suficientes para continuar con tus consultas?		Isa et al., 2024
11	¿Te resulta fácil utilizar Sofía?		Aleedy, M et al., 2022; Balderas et al., 2023
12	¿Sofía entiende tus preguntas cuando usas un lenguaje informal?		Aleedy, M et al., 2022; Belda-Medina y Kokošková, 2023
13	¿La respuesta de Sofía a tus preguntas es rápida?		Balderas et al., 2023; Belda-Medina y Kokošková, 2023
14	¿Tienes algún comentario para mejorar la atención de Sofía?	Respuesta abierta	

A partir de la información recopilada se aplicó un análisis de componentes principales (PCA) para identificar los factores clave que influyen en la percepción de los usuarios y un análisis de contenido para las respuestas abiertas.

3. RESULTADOS

En este apartado se presentan los resultados obtenidos a lo largo de las distintas etapas del proyecto. Cada fase se ha abordado meticulosamente para garantizar el desarrollo efectivo y la implementación del chatbot de atención al estudiante, acorde con las necesidades identificadas.

3.1. Backlog del Producto

La primera fase del proceso fue clave para alinear los objetivos del proyecto con las necesidades de los usuarios. En esta etapa, se estableció un equipo multidisciplinario compuesto por figuras clave, como el Vicerrector Académico y los responsables de Servicio al Estudiante, Bienestar, Registro y Control y Tecnología Educativa. Este enfoque colaborativo aseguró que el diseño del asistente virtual integrara de manera articulada las perspectivas académicas, de apoyo estudiantil y tecnológicas

Se designó al responsable de Bienestar universitario como Product Owner y junto con el resto de stakeholders se trabajó en la creación y priorización del Product Backlog, identificando y organizando las funcionalidades clave necesarias para el chatbot. Esta fase no solo permitió definir y priorizar las tareas esenciales del backlog, sino también estableció una visión clara y compartida del proyecto, que guía al equipo de desarrollo en los siguientes sprints (Tabla 2):

Tabla 2

Instrumento de evaluación satisfacción

Sprint	Tareas	Prioridad	Horas Estimadas
Análisis y Planificación estratégica	Definir el propósito del chatbot, requisitos técnicos y operativos, preguntas frecuentes (FAQ) y flujo conversacional básico.	Alta	35
Diseño de Interfaz de Usuario (UI) y Prototipo	Diseñar interfaz inicial, crear prototipos y validarlos con stakeholders, adaptando el diseño al perfil del usuario.	Media	20
Desarrollo Inicial del Chatbot	Programar respuestas a preguntas comunes, integrar PLN básico, realizar pruebas internas y ajustar lenguaje informal.	Media	50
Implementación Avanzada y Prueba Piloto	Entrenar con corpus ampliado, integrar PLN avanzado, realizar pruebas piloto con usuarios reales y evaluar funcionamiento.	Alta	80
Optimización y Despliegue	Ajustar respuestas según retroalimentación, optimizar tiempos, gestionar consultas ambiguas y desplegar el chatbot.	Alta	45
Mantenimiento y Mejora Continua	Aplicar encuestas periódicas, ampliar funcionalidades según retroalimentación y monitorear métricas clave en tiempo real.	Alta	Continuo / (1 h. diaria)

3.2. Planificación y Diseño (sprint 1 y 2)

En el primer sprint se acuerda el objetivo del chatbot acotándolo como una herramienta de apoyo diseñada para mejorar el flujo de comunicación entre los estudiantes y la institución. Su propósito principal es ofrecer respuestas rápidas y eficientes a preguntas frecuentes relacionadas con procesos administrativos, académicos y financieros, mejorando así la accesibilidad y la experiencia de los usuarios en estos ámbitos clave.

En esta fase, se llevó a cabo una evaluación funcional de diversas herramientas de desarrollo, teniendo en cuenta sus capacidades de PLN, facilidad de integración y costos asociados. Tras este análisis, se optó por implementar una solución comercial, que destacó por ofrecer un equilibrio adecuado entre funcionalidad avanzada, flexibilidad en su integración y una relación costo-beneficio favorable.

La definición de las preguntas frecuentes que deben nutrir al corpus de conocimiento de la herramienta se elaboró a partir de un análisis exhaustivo de las consultas más comunes realizadas por los estudiantes en el Centro de Ayuda de la institución. Este análisis fue complementado con la revisión de los requerimientos más solicitados durante el último año académico, lo que permitió identificar las áreas de mayor interés para los usuarios entre las

que destacan las consultas técnicas referentes al acceso a la plataforma de estudio (credenciales, reinicio de contraseña, fallos en el acceso), financieras (medios de pago), académicos (relacionadas con la metodología de estudio, fechas de entrega, etc.) entre otras.

En el diseño del flujo conversacional se incluyó la creación de diagramas de interacción detallados sobre posibles escenarios de consulta y sus respectivas respuestas para desarrollar rutas dinámicas capaces de gestionar consultas complejas adaptándose a las necesidades del usuario, así como mensajes predefinidos diseñados para responder de manera eficiente a preguntas frecuentes.

Por otro lado, en este diseño se tuvo en cuenta la caracterización del perfil de alumno y sus necesidades específicas. Para ello se analizaron los datos sociodemográficos de los 7.760 estudiantes activos entre marzo y junio de 2023 advirtiendo que la edad promedio de los estudiantes no superaba los 35 años y que el 65.8% eran mujeres. Además, se identificó que el 99% de los estudiantes eran de Colombia, con una alta concentración en ciudades principales como Bogotá. Con base a este perfil de usuarios, el equipo del proyecto determinó que el chatbot debería emplear un lenguaje claro, accesible y preciso, evitando tecnicismos innecesarios. Asimismo, se adoptó un tono formal pero cercano en la interfaz, asegurando que la comunicación fuera efectiva y acorde a las expectativas de los estudiantes.

A nivel gráfico se diseñó una interfaz de la aplicación intuitiva integrándola en la página de acceso al portal del estudiante y en la sección de contacto que complementa la estrategia de comunicación (figura 1).

Figura 1

Interface del Chatbot



3.3. Desarrollo y pruebas funcionales (sprint 3 y 4)

En esta fase del proyecto el objetivo fue llevar a cabo la implementación técnica del chatbot asegurando su capacidad para responder automáticamente a las preguntas frecuentes (FAQ) más comunes. Se configuraron y optimizaron las capacidades de PLN incluidas en la

herramienta para comprender y responder de manera eficaz tanto a consultas formuladas de manera precisa como a aquellas con lenguaje informal o variaciones lingüísticas y se entrenaron 30 preguntas seleccionadas cuidadosamente a partir de las consultas más recurrentes relativas a:

- Procesos académicos (Pruebas Saber Pro y TyT, Certificado de estudios, notas y/o temáticos, Opciones de grado, electivas, entre otras)
- Procesos administrativos (procesos de grado y admisión, reingreso, retiro, aplazo, ampliaciones de plazo, etc.)
- Cuestiones financieras (Comunicación apoyo financiero, medios de pago, devoluciones, etc.)
- Tecnológico (activación en plataforma, recuperación de credenciales, etc.)
- Otras (comunicación con mentor, carné estudiantil, etc.).

Adicional se implementaron 118 preguntas predefinidas que permitieron al chatbot interactuar de manera más natural y humana con los usuarios. Estas preguntas no están relacionadas directamente con las funcionalidades específicas del chatbot o los procesos técnicos, sino que buscan dar personalidad al asistente virtual, haciendo que la experiencia del usuario sea más cercana, amigable y auténtica.

Después de la configuración de las respuestas automáticas y el ajuste del sistema de PLN, se realizaron pruebas internas para evaluar su rendimiento e identificar posibles problemas técnicos como respuestas incorrectas o dificultades para procesar preguntas más complejas. Con base en los resultados de las pruebas, se implementaron mejoras adicionales en la configuración de la herramienta: a) integración de botones interactivos que ofrecen a los usuarios opciones predefinidas para continuar el diálogo sin necesidad de escribir una respuesta reduciendo la posibilidad de que el usuario formule preguntas ambiguas o mal estructuradas; b) mensajes y evasivas diseñadas para gestionar diferentes tipos de respuestas del chatbot, desde mensajes generales hasta formas más sofisticadas de manejar situaciones inciertas o difíciles de responder.

Posteriormente, el chatbot fue entrenado con un corpus de datos ampliado, y entre enero y mayo de 2024 se realizó una prueba piloto para evaluar su rendimiento en contexto real. Durante este periodo, se observó un incremento en las interacciones, pasando de 624 en enero a 3.444 en mayo, atribuible a campañas institucionales que fomentaron su uso (figura 2). La eficiencia del chatbot fue alta, respondiendo satisfactoriamente a más del 88,02% de las consultas, con un margen de interacciones negativas inferior al 11%, atribuible a consultas complejas.

Figura 2

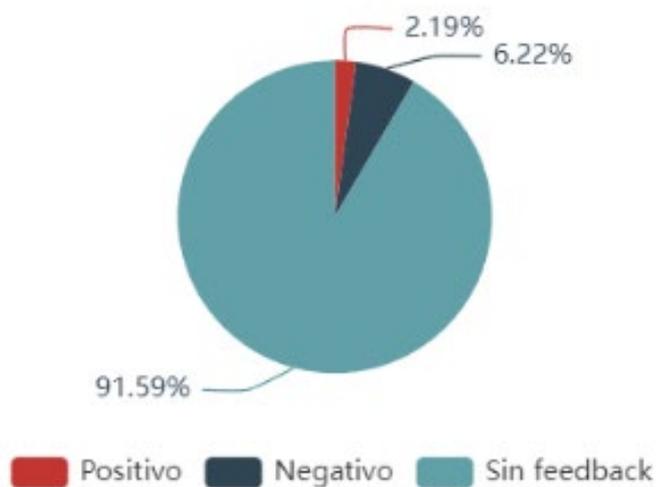
Incremento interacciones chatbot. Información extraída de plataforma AIVO



Sin embargo, las estadísticas de la herramienta delatan que, aunque el sistema funciona, la falta de retroalimentación es el principal desafío, ya que un porcentaje muy alto de interacciones no ofrecen información evaluativa. Esto implica que, para identificar oportunidades de mejora o reforzar las respuestas, el equipo debe analizar las interacciones registradas en el sistema.

Figura 3

Feedback emitido. Información extraída de plataforma AIVO



En cuanto a satisfacción por categorías, se destacan áreas clave que requieren atención específica para mejorar la experiencia del usuario:

Tabla 3

Satisfacción respuesta herramienta

Categoría	Promedio de Estrellas	% de 1 Estrella	% de 5 Estrellas
Activación de plataforma	2.73	44.44%	29.63%
Comunicación apoyo financiero	2.19	63.29%	18.99%
Contraseña	3.11	33.57%	37.86%
Medios de pago	3.47	26.35%	50.68%

Notas. Elaboración propia

A partir de los resultados obtenidos, se ajustó el entrenamiento del modelo mediante aprendizaje supervisado, incorporando nuevas configuraciones para mejorar la precisión en temas específicos. Este enfoque, basado en un corpus etiquetado de preguntas frecuentes, permitió una mejora continua guiada por retroalimentación estructurada.

3.4. Optimización y mantenimiento (spint 5 y 6)

Durante esta fase, se priorizó el entrenamiento y la integración de nuevos requerimientos, lo que elevó las intenciones de 30 a 114 en noviembre de 2024. La eficiencia también mejoró, alcanzando un 91,72% de respuestas satisfactorias, lo que refleja su creciente capacidad para adaptarse a las necesidades de los usuarios.

El análisis de la retroalimentación de los usuarios evidenció una disminución en el feedback negativo respecto al periodo anterior del 18.01%, pero manteniendo la baja tasa de implicación. Un dato que podría indicar que los usuarios consideran que el chatbot cumple con su propósito de manera satisfactoria y, por ende, no sienten la necesidad de comentar o por el contrario podría ser señal de insatisfacción o desinterés, que no se manifiesta de manera explícita.

Para resolver esta ambigüedad y avanzar en la mejora del sistema, en diciembre de 2024 se socializó una encuesta de satisfacción (tabla 1) dirigida a los 9,710 estudiantes activos durante el año académico 2024. Esta encuesta tuvo como objetivo recopilar opiniones y percepciones sobre el uso de la herramienta. De este total, 172 estudiantes respondieron, proporcionando información valiosa que combinada con los indicadores de la herramienta sirve para evaluar el desempeño del chatbot y orientar futuros ajustes en su diseño y funcionalidad.

Se optó por un análisis de componentes principales (ACP) que sirve para identificar los factores clave que influyen en la percepción de los usuarios, proporcionando información útil para optimizar el diseño y funcionalidad del chatbot Sofía. En este análisis se empleó la herramienta Google Colab que permite ejecutar código Python de manera interactiva sin necesidad de instalar software localmente, empleando librerías especializadas como pandas (para la

manipulación y limpieza de datos), numpy y scikit-learn (para realizar los cálculos numéricos y análisis PCA) y matplotlib (para la visualización de resultados).

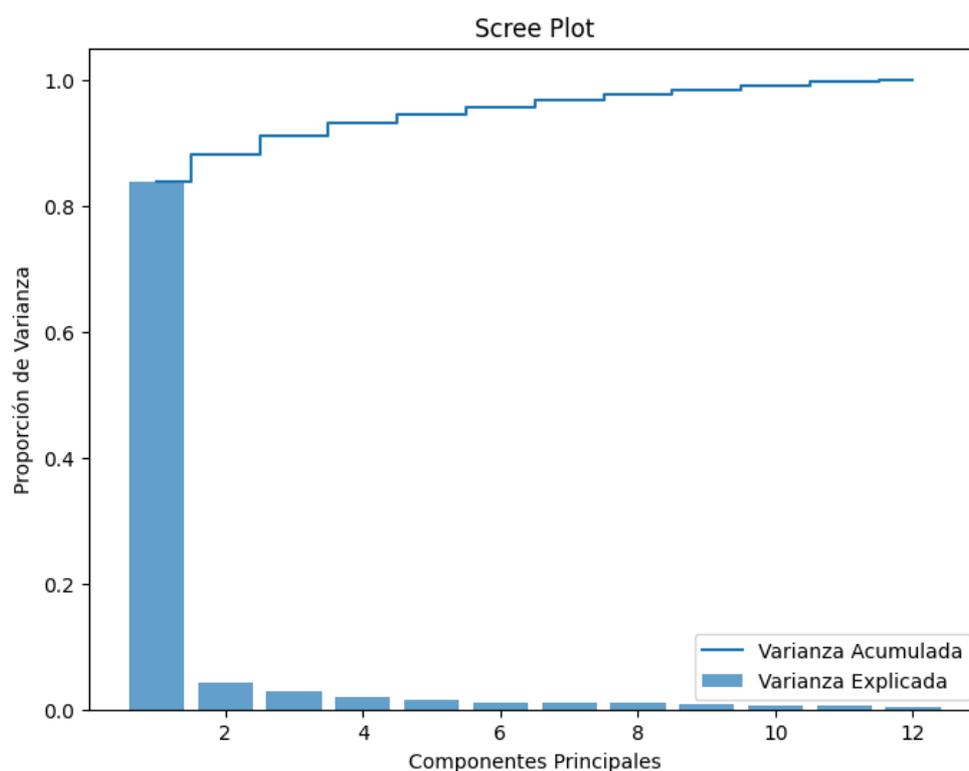
Los resultados del ACP, permitieron identificar los principales factores que influyen en la satisfacción de los estudiantes con la herramienta agrupándolas en las siguientes dimensiones:

- Satisfacción General (Q1, Q2, Q3)
- Ahorro de Tiempo y Utilidad (Q4, Q5, Q6):
- Claridad y Precisión de las Respuestas (Q7, Q8, Q9)
- Usabilidad Técnica (Q10, Q11, Q12)

El análisis de la varianza indicó que los primeros dos componentes principales explican aproximadamente 80% de la varianza acumulada (figura 4) destacando que la mayoría de la información relevante sobre la satisfacción de los usuarios puede ser resumida en cuanto a la experiencia general y claridad de las respuestas y la utilidad y eficiencia del Chatbot.

Figura 4

Varianza explicada



Por otro lado, el PCA permitió identificar que la satisfacción general (PC2) estuvo principalmente influenciada por la percepción global de la experiencia (Q1, carga factorial = 0.870). La usabilidad técnica (PC3) destacó variables como la facilidad de uso (Q10, carga factorial = 0.446) y la rapidez de respuesta (Q12, carga factorial = 0.372). Asimismo, la claridad y relevancia de las respuestas (PC1) fueron representadas por preguntas relacionadas con la precisión y opciones suficientes del chatbot (Q8, Q9). Finalmente, la disposición a recomendar el chatbot (PC4) estuvo dominada por la variable Q2 (carga factorial = 0.808)

Tabla 3

Cargas factoriales por Componente Principales

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12
Q1	0,238	0,870	0,203	-	0,035	-	0,080	0,003	-	0,078	0,039	0,043
				0,349		0,035			0,069			
Q2	0,277	0,245	-	0,809	0,021	-	0,076	0,234	-	0,214	0,042	0,001
			0,206			0,135			0,201			
Q3	0,292	-	-	-	0,509	0,354	-	0,264	0,181	0,315	-	-
		0,062	0,281	0,178			0,459				0,004	0,063
Q4	0,294	0,131	-	0,112	0,066	0,008	0,133	-	0,480	-	-	-
			0,373					0,404		0,387	0,423	0,027
Q5	0,297	-	-	-	0,217	-	0,263	-	-	-	0,703	0,234
		0,148	0,294	0,166		0,232		0,078	0,042	0,222		
Q6	0,294	-	-	-	-	-	-	0,016	-	0,053	-	-
		0,124	0,283	0,290	0,247	0,527	0,224		0,349		0,214	0,417
Q7	0,296	-	-	-	-	0,279	0,487	-	-	0,581	-	0,201
		0,214	0,074	0,158	0,228			0,222	0,111		0,173	
Q8	0,300	-	0,064	0,022	-	0,550	0,102	0,197	-	-	0,240	-
		0,021			0,369				0,036	0,328		0,501
Q9	0,303	-	0,044	-	-	0,115	-	0,105	-	-	-	0,672
		0,053		0,003	0,363		0,424		0,169	0,272	0,120	
Q10	0,286	-	0,446	0,072	0,537	0,060	0,098	-	-	-	-	-
		0,182						0,184	0,470	0,248	0,244	0,063
Q11	0,291	-	0,428	0,182	-	-	-	-	0,308	0,258	0,305	-
		0,076			0,136	0,142	0,370	0,496				0,132
Q12	0,291	-	0,372	-	0,000	-	0,259	0,572	0,458	-	-	0,045
		0,190		0,069		0,325				0,005	0,153	

Los resultados obtenidos indican que, aunque la claridad y precisión de las respuestas (Q7, Q8), están bien evaluadas, aún existe margen de mejora para que sean percibidas como más claras y relevantes. La carga factorial del ítem Q9, relacionado con la continuidad de consultas, sugiere que las opciones disponibles para seguir interactuando con el chatbot no se perciben como suficientemente sólidas, lo que representa un área a mejorar. Esto evidencia que, aunque la herramienta es útil, podría ofrecer más rutas para profundizar en las respuestas, especialmente en consultas complejas.

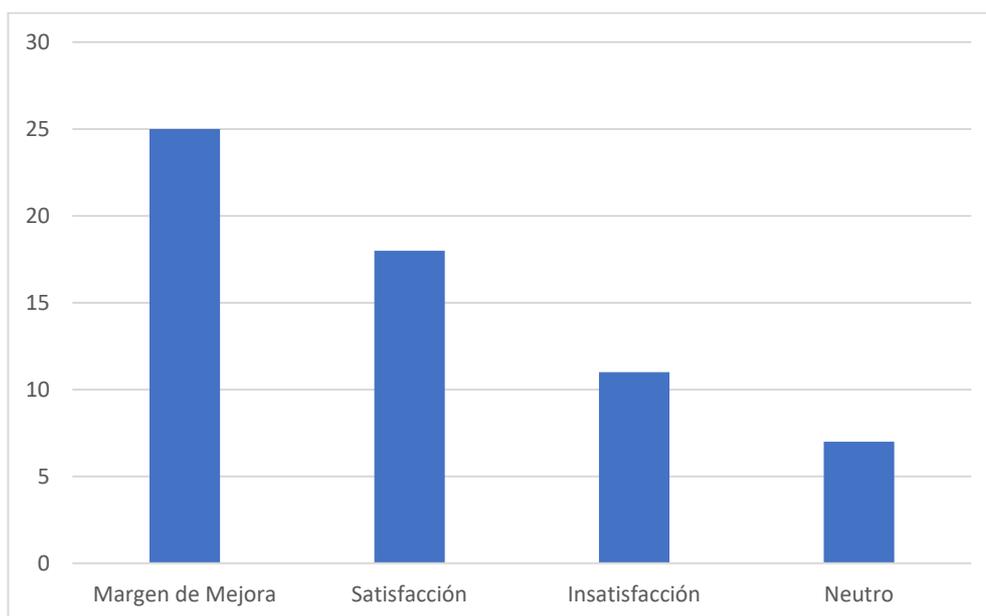
La carga factorial de Q10 y Q11 indican que la facilidad de uso es un aspecto destacado, pero sigue siendo clave monitorearlo para garantizar una experiencia fluida. En este sentido podría ser necesario ajustar la interfaz o el flujo conversacional para que sea aún más intuitivo, especialmente para nuevos usuarios o aquellos menos familiarizados con herramientas tecnológicas y reforzarse con un entrenamiento más profundo en PLN para gestionar errores comunes o regionalismos que los usuarios puedan incluir en sus consultas.

Estos datos se ven respaldados por el análisis de contenido realizado a partir de la pregunta abierta incluida en la encuesta. La información obtenida fue categorizada en función de temas

comunes identificados: Satisfacción, Margen de Mejora, Insatisfacción y Neutro. El conteo por categorías refleja la distribución de las respuestas de los estudiantes, permitiendo identificar los principales focos de atención en sus opiniones (Figura 5). Se destaca que persiste un alto nivel de peticiones que advierten áreas de mejora, especialmente en las consultas más complejas, donde la herramienta no logra resolver de manera eficiente las necesidades planteadas por los estudiantes.

Figura 5

Frecuencia de respuestas abiertas



Los resultados alcanzados revelan que el uso del aprendizaje supervisado no solo permitió una mejora progresiva de las respuestas del chatbot, sino que también ofreció un marco metodológico claro para su entrenamiento iterativo y validación en entornos reales. Este enfoque resalta cómo la aplicación de técnicas de inteligencia artificial puede ser eficaz cuando se combina con procesos pedagógicos estructurados y una comprensión profunda de las necesidades de los usuarios.

4. DISCUSIÓN Y CONCLUSIONES

El presente trabajo ha sido un ejercicio de aplicación que ha recopilado y analizado durante un año los datos derivados de la implementación y funcionamiento de un chatbot de atención estudiantil. Esta experiencia ha permitido evaluar tanto las ventajas como los retos asociados a su uso, proporcionando una visión integral sobre su impacto en la atención al estudiante y los procesos institucionales. Los hallazgos obtenidos no solo destacan el potencial de esta herramienta para optimizar la experiencia estudiantil, sino también las áreas de mejora necesarias para garantizar una integración más efectiva y sostenible en el entorno educativo coincidiendo con las conclusiones de Huang et al. (2021) o Zhang et al. (2023), entre otros.

Los datos obtenidos a lo largo de las fases de implementación permitieron identificar factores determinantes en la satisfacción estudiantil durante la interacción con la herramienta. Entre

ellos destacan la claridad de las respuestas y la fluidez conversacional, aspectos que favorecen una experiencia positiva y refuerzan su valor percibido como recurso institucional. Esta dimensión subjetiva —centrada en la percepción del usuario— constituye uno de los principales aportes del estudio, al permitir valorar el impacto del asistente virtual más allá de su rendimiento técnico. Considerar las valoraciones estudiantiles ha resultado esencial para la evaluación global del sistema, dado que, como señalan Isa et al. (2024), son los propios usuarios quienes legitiman su utilidad en función de las expectativas y necesidades que logre satisfacer.

Los resultados evidencian que la calidad de las respuestas ofrecidas por el asistente virtual depende no solo de la tecnología subyacente, sino también del mantenimiento continuo, que requiere supervisión constante, ajustes periódicos y una optimización sostenida de los flujos conversacionales (Chaves & Gerosa, 2020; Haqimi & Kusuma, 2023; Lin et al., 2023; Pawlik et al., 2022). Este proceso se desarrolló bajo un enfoque de aprendizaje supervisado, mediante el entrenamiento progresivo del sistema con ejemplos etiquetados y retroalimentación estructurada, lo cual favoreció su precisión y capacidad de adaptación ante consultas recurrentes (Aleedy et al., 2022; Balderas et al., 2023; Okonkwo & Ade-Ibijola, 2021).

El análisis de las interacciones evidenció que, aunque el sistema responde eficazmente a preguntas frecuentes y estructuradas, presenta limitaciones frente a lenguaje informal, ambigüedad o conversaciones complejas, como señalan Sudiatmika y Ariantini (2021). Para mitigarlas, se implementaron botones de redirección que ayudaron a acotar intenciones y mejorar la tasa de respuestas, aunque con posible impacto en la fluidez (Nguyen et al., 2021). A pesar de los avances, persisten dificultades ante consultas que requieren una comprensión contextual más profunda o el manejo de expresiones no convencionales.

Aunque este modelo ha demostrado ser eficaz para responder consultas estructuradas y frecuentes, en línea con Balderas et al. (2023), surge la necesidad de avanzar hacia modelos de lenguaje más sofisticados, como GPT, que posibiliten interacciones más fluidas, personalizadas y contextualmente relevantes. Para ello, las instituciones deben considerar no solo los aspectos técnicos, económicos y de privacidad, sino también su potencial pedagógico, explorando estrategias como el fine-tuning para adaptar estas herramientas a distintos entornos. Los hallazgos de este estudio evidencian no solo el impacto funcional y experiencial de los chatbots en la atención al estudiante, sino también su valor como base para aplicaciones más amplias, como el desarrollo de tutores virtuales inteligentes adaptados a diversos perfiles de aprendizaje, en línea con estudios que destacan su eficacia y personalización en contextos educativos (Belkacem & Hireche, 2024). Todo ello refuerza la importancia de adoptar enfoques de entrenamiento adecuados en escenarios que exigen precisión, control y mejora continua, impulsando un uso más estratégico y formativo de la IA en la educación superior.

5. FINANCIACIÓN

Este estudio no recibió financiación para su realización.

6. CONTRIBUCIÓN DE LOS AUTORES

Conceptualización, N.S-G.; curación de datos, N.S-G.; análisis formal, N.S-G., A.G.R.; investigación, N.S-G.; metodología, N.S-G.; administración del proyecto, N.S-G., A.G.R.; recursos, N.S-G.; software, N.S-G.; supervisión, A.G.R.; validación, N.S-G., A.G.R.; visualización, A.G.R.; redacción-preparación del borrador original, N.S-G., A.G.R.; redacción-revisión y edición, N.S-G., A.G.R.

7. REFERENCIAS

- Aleedy, M., Atwell, E., & Meshoul, S. (2022). Using AI chatbots in education: Recent advances challenges and use case. *Algorithms for Intelligent Systems* (pp. 661-675). https://doi.org/10.1007/978-981-19-1653-3_50
- Ali, F., Zhang, Q., Tauni, M. Z., & Shahzad, K. (2023). Social chatbot: My friend in my distress. *International Journal of Human-Computer Interaction*, 1-11. <https://doi.org/10.1080/10447318.2022.2150745>
- Balderas, A., García-Mena, R. F., Huerta, M., Mora, N., & Doderó, J. M. (2023). Chatbot for communicating with university students in emergency situation. *Heliyon*, 9(9), e19517. <https://doi.org/10.1016/j.heliyon.2023.e19517>
- Belda-Medina, J., & Kokošková, V. (2023). Integrating chatbots in education: Insights from the chatbot-human interaction satisfaction model (CHISM). *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00432-3>
- Belkacem, A. N., & Hireche, A. (2024). Exploring student engagement and learning preferences: A comparative study between virtual- and robot-based tutors. *Measurement Sensors*, 101704. <https://doi.org/10.1016/j.measen.2024.101704>
- Caldarini, G., Jaf, S., & McGarry, K. (2022). A literature survey of recent advances in chatbots. *Information*, 13(1), 41. <https://doi.org/10.3390/info13010041>
- Chaves, A. P., & Gerosa, M. A. (2020). How should my chatbot interact? A survey on social characteristics in human–chatbot interaction design. *International Journal of Human-Computer Interaction*, 37(8), 729-758. <https://doi.org/10.1080/10447318.2020.1841438>
- Croes, E., & Antheunis, M. L. (2022). "I am in your computer while we talk to each other" a content analysis on the use of language-based strategies by humans and a social chatbot in initial human-chatbot interactions. *International Journal of Human-Computer Interaction*, 39(10), 2155-2173. <https://doi.org/10.1080/10447318.2022.2075574>

- García-Peñalvo, F. J. (2023). La percepción de la inteligencia artificial en contextos educativos tras el lanzamiento de ChatGPT: Disrupción o pánico. *Education in the Knowledge Society*, 24, e31279. <https://doi.org/10.14201/eks.31279>
- Haqimi, N. A., & Kusuma, R. K. (2023). Timeline reminder system bot and telegram assistant chatbot for a university student and lecturer. *Journal of Soft Computing Exploration*, 4(4), 186-194. <https://doi.org/10.52465/josce.v4i4.221>
- Huang, W., Hew, K. F., & Fryer, L. K. (2021). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237-257. <https://doi.org/10.1111/jcal.12610>
- Ifelebuegu, A., Kulume, P., & Cherukut, P. (2023). Chatbots and AI in education (AIEd) tools: The good, the bad, and the ugly. *Journal of Applied Learning and Teaching*, 6(2). <https://doi.org/10.37074/jalt.2023.6.2.29>
- Isa, N. A. N. M., Jawaddi, S. N. A., & Ismail, A. (2024). Experimental evaluation of machine learning models for goal-oriented customer service chatbot with pipeline architecture. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2409.18568>
- Keerthana, R. R., Fathima, G., & Florence, L. (2021, May). Evaluating the performance of various deep reinforcement learning algorithms for a conversational chatbot. In *2021 2nd International Conference for Emerging Technology (INCET)* (pp. 1-8). IEEE. <https://doi.org/10.1109/INCET51464.2021.9456321>
- Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00426-1>
- Latif, E., & Zhai, X. (2024). Fine-tuning ChatGPT for automatic scoring. *Computers and Education: Artificial Intelligence*, 6, 100210. <https://doi.org/10.1016/j.caeai.2024.100210>
- Lin, C., Huang, A. Y., & Yang, S. J. (2023). A review of AI-driven conversational chatbots implementation methodologies and challenges (1999–2022). *Sustainability*, 15(5), 4012. <https://doi.org/10.3390/su15054012>
- Medina, G. T., Lengua, M. A. C., & Medrano, H. V. (2024). Design of a chatbot in a mobile application for managing payments and controlling activities in a fast school organization. *Indonesian Journal of Electrical Engineering and Computer Science*, 35(2), 1271. <https://doi.org/10.11591/ijeecs.v35.i2.pp1271-1286>
- Mekić, E., Jovanović, M., Kuk, K., Prlinčević, B., & Savić, A. (2024). Enhancing educational efficiency: Generative AI chatbots and DevOps in education 4.0. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2406.15382>
- Mou, Y., & Xu, K. (2017). The media inequality: Comparing the initial human-human and human-AI social interactions. *Computers in Human Behavior*, 72, 432-440. <https://doi.org/10.1016/j.chb.2017.02.067>

- Nguyen, Q. N., Sidorova, A., & Torres, R. (2021). User interactions with chatbot interfaces vs. menu-based interfaces: An empirical study. *Computers in Human Behavior*, 128, 107093. <https://doi.org/10.1016/j.chb.2021.107093>
- Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, 2, 100033. <https://doi.org/10.1016/j.caeai.2021.100033>
- Pawlik, Ł., Płaza, M., Deniziak, S., & Boksa, E. (2022). A method for improving bot effectiveness by recognising implicit customer intent in contact centre conversations. *Speech Communication*, 143, 33-45. <https://doi.org/10.1016/j.specom.2022.07.003>
- Pratama, Y. A., & Kristiana, T. (2023). Design of cloud-based chatbot application at Pt. Traveloka Singapore using the agile method. *Jurnal Pilar Nusa Mandiri*, 19(1), 19-26. <https://doi.org/10.33480/pilar.v19i1.3055>
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI Blog*, 1(8), 9. https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf
- Rojas, E. M. (2020). Machine learning: Análisis de lenguajes de programación y herramientas para desarrollo. *Revista Ibérica de Sistemas e Tecnologías de Informação*, 28, 586–599. <https://www.proquest.com/docview/2388304894/1866F7F83224F28PQ/3?sourcetype=Scholarly%20Journals>
- Segovia-García, N., & Said-Hung, E. M. (2021). Factores de satisfacción de los alumnos en e-learning en Colombia. *Revista Mexicana de Investigación Educativa*, 26(89), 595-621. <https://ojs.rmie.mx/index.php/rmie/article/view/186>
- Segovia-García, N. (2023). Percepción y uso de los chatbots entre estudiantes de posgrado online: Un estudio exploratorio. *Revista de Investigación en Educación*, 21(3), 335-349. <https://doi.org/10.35869/reined.v21i3.4974>
- Segovia-García, N., & Segovia-García, L. (2024). Chatbot de IA para prevenir el abandono de la educación superior: Una revisión de la literatura. *Education in the Knowledge Society*, 25, e31416. <https://doi.org/10.14201/eks.31416>
- Sharma, V., Goyal, M., & Malik, D. (2017). An intelligent behaviour shown by chatbot system. *International Journal of New Technology and Research*, 3(4), 52-54. <https://media.neliti.com/media/publications/263312-an-intelligent-behaviour-shown-by-chatbo-7020467d.pdf>
- Skuridin, A., & Wynn, M. (2024). Chatbot design and implementation: Towards an operational model for chatbots. *Information*, 15(4), 226. <https://doi.org/10.3390/info15040226>
- Stokel-Walker, C. (2022). AI bot ChatGPT writes smart essays — should professors worry? *Nature*. <https://doi.org/10.1038/d41586-022-04397-7>

- Sudiatmika, I. P. G. A., & Ariantini, M. (2021). Digital marketing chatbot using API dialog flow case study ITB Stikom Bali, Jimbaran campus. <https://doi.org/10.4108/eai.19-12-2020.2309122>
- Tabone, W., & De Winter, J. (2023). Using ChatGPT for human–computer interaction research: A primer. *Royal Society Open Science*, 10(9). <https://doi.org/10.1098/rsos.231053>
- Torres Martínez, D. M., & Cruz Guerrero, S. A. (2020). ¿Qué tipos de agentes virtuales pueden usar las pequeñas empresas para mejorar su publicidad? *Revista Vínculos*, 17(2), 140-155. <https://doi.org/10.14483/2322939X.16874>
- Uc-Cetina, V., Navarro-Guerrero, N., Martín-González, A., Weber, C., & Wermter, S. (2022). Survey on reinforcement learning for language processing. *Artificial Intelligence Review*, 56(2), 1543-1575. <https://doi.org/10.1007/s10462-022-10205-5>
- Zhang, R., Zou, D., & Cheng, G. (2023). A review of chatbot-assisted learning: Pedagogical approaches, implementations, factors leading to effectiveness, theories, and future directions. *Interactive Learning Environments*, 32(8), 4529–4557. <https://doi.org/10.1080/10494820.2023.2202704>

Para citar este artículo:

Segovia García, N., Guzmán Rincón, A. (2025). Evaluación de un chatbot basado en aprendizaje supervisado: impacto en la satisfacción y propuestas de mejora. *Eduotec, Revista Electrónica de Tecnología Educativa*, (92), 234-252. <https://doi.org/10.21556/edutec.2025.92.3849>