



Evaluación de la calidad de un chatbot educativo

*Joyce Huacón Morales
Carolina Pacheco Pozo
Bertha Naranjo Sánchez*

Introducción

Los chatbots son programas de software diseñados con un propósito de interacción dinámica con diferentes usuarios, lo que puede denominarse como una conversación con fines informativos y de resolución de problemas. Se pueden utilizar para complementar diferentes servicios como atención al cliente, educación, viajes y salud. Se ha vuelto muy común su uso al ser una herramienta tecnológica que se comunica con los usuarios de manera inteligente (Adamopoulou y Moussiades, 2020).

Para las empresas el uso de un chatbot puede ser considerado como una estrategia. Internet está tan saturado en estos días que puede ser difícil destacarse entre la multitud. En ese sentido, las empresas buscan canales digitales para llamar la atención y conectarse con los usuarios. Los consumidores aceptan este hecho y están más dispuestos a interactuar con las empresas a través de un chatbot (Presti *et al.*, 2021).

En cambio, de la forma tradicional, el contar con un equipo de servicio al cliente tiene más contras que pros. Algunos clientes valoran más el hecho de interactuar con una persona real antes que con un software. Sin embargo, muchas más personas valoran el obtener respuestas inmediatas,

claras, a toda hora, los 365 días del año. Volviendo con la forma tradicional, en muchos casos las empresas ven afectados sus propios ingresos ya que contratar un equipo de servicio al cliente es más costoso, a la vez que por el número de solicitudes recibidas o cualquier otro motivo, el personal puede tardar en contestar las solicitudes y eso puede afectar la imagen que los clientes tienen de la empresa (Følstad y Taylor, 2021). En ocasiones esto se debe a que los empleados abandonan la empresa o simplemente no están en el horario laboral. Por lo tanto, esas cuestiones no se resolverán hasta el comienzo del siguiente día hábil. Por otro lado, las interacciones entre clientes y el personal de servicio al cliente, pueden no desarrollarse de la mejor forma, si el personal no es tolerante a ciertas actitudes de los clientes.

Los chatbots se pueden usar en diferentes tipos de servicios como atención al cliente, educación, consulta en línea, viajes, salud, etc. Se han vuelto muy populares, convirtiéndose en herramientas inteligentes para comunicarse con los usuarios de una forma espontánea, donde el usuario muchas veces ni reconoce que quien le responde los mensajes es un software.

En un mundo actual donde el internet está tan saturado por el volumen de data que es información que se transmite y retransmite, puede ser difícil destacar entre la multitud. Es por eso que las empresas están recurriendo a los canales digitales para llamar la atención y conectarse con los usuarios. Los consumidores han aceptado este hecho y están dispuestos a conectarse con las empresas a través de estas nuevas tecnologías.

Pero no solo las empresas están recurriendo a este tipo de tecnologías, en lo que concierne al rubro educativo, el cual es el de interés para este estudio. Muchas instituciones educativas de diferentes niveles están empezando a usar chatbots educacionales para brindar apoyo y recursos educativos a estudiantes, profesores y otros usuarios interesados en aprender o adquirir conocimientos en una variedad de temas (Lopez y Qamber, 2022).

Dado que un chatbot es esencialmente un software que utiliza inteligencia artificial para aprender a medida que va interactuando con los humanos, es necesaria su evaluación para conocer aspectos relacionados a la efectividad del aprendizaje, es decir si cumple con su propósito educativo, para identificar además, posibles problemas relacionados a la experiencia de usuario, para saber si es adaptable a los diferentes estilos de aprendizaje y niveles de conocimiento de los usuarios, con el fin de obtener una retroalimentación e identificar áreas de mejora continua.

Para la evaluación del chatbot se utilizó los criterios SQUARE contenidos en la norma ISO 25000 en la cual se establece el estándar para las evaluaciones de la efectividad de productos de software como los chatbots. Esta evaluación se realizó en base con las características, métricas e indicadores establecidos en la norma y su propósito es permitir adquirir información que ayude a mejorar los chatbots educativos.

El objetivo general de este estudio fue evaluar un chatbot educativo mediante un modelo de calidad para determinar si cumple con las normas ISO 25000. Los objetivos específicos fueron: Identificar las características, métricas e indicadores para la evaluación del chatbot educativo, verificar el nivel de cumplimiento de la norma ISO 25000, y elaborar un informe de la evaluación del chatbot educativo con recomendaciones.

Revisión de literatura

Chatbot

Un chatbot es un programa de inteligencia artificial diseñado para interactuar y brindar información y apoyo a los estudiantes (Auqui, 2021). Estos chatbots están diseñados para responder preguntas frecuentes, brindar orientación académica, ofrecer recursos de aprendizaje y proporcionar asistencia en general a los estudiantes (Vega, 2021).

Los chatbots para estudiantes pueden utilizarse en entornos educativos como universidades, escuelas, instituciones de educación en línea y plataformas de aprendizaje. Pueden ser integrados en sitios web, aplicaciones móviles y plataformas de mensajería como WhatsApp, Facebook Messenger y Slack (Pisco, 2021).

Los chatbots educativos pueden ofrecer una variedad de funciones y servicios. Algunas características comunes incluyen:

Respuestas a preguntas frecuentes: al procesar datos en milisegundos, pueden generar información que son respuestas a usuarios de manera muy rápida, acerca de sus horarios, tasas de matrícula, requisitos del programa, fechas importantes, servicios para estudiantes y más.

Recordatorios y notificaciones: los chatbots pueden enviar recordatorios y notificaciones a los estudiantes sobre fechas límite de tareas, fechas de exámenes, eventos y otras actividades académicas

Solicitud de servicios estudiantiles: los chatbots pueden ser útiles para los estudiantes en la atención de solicitudes, entre otros servicios estudiantiles, como asesoramiento académico, tutorías, cambio de horarios y servicios de apoyo.

Seguimiento del progreso del estudiante: los chatbots pueden ayudar a los estudiantes a realizar un seguimiento de su progreso académico, como notas, calificaciones y requisitos de graduación (Mrsic *et al.*, 2020).

Los chatbots contribuyen en el mejoramiento de la comunicación y la interacción entre los estudiantes y las instituciones educativas. Pueden proporcionar respuestas rápidas y precisas a las consultas de los estudiantes, de forma que son una herramienta útil que no solo sirve para autoeducarse sino para mejorar su socialización hasta cierto punto (Vega, 2021).

Tipos de chatbots

Chatbots basados en reglas

Los chatbots de este tipo siguen un patrón predefinido durante la conversación. El chatbot presenta al usuario una lista de estas opciones y ajusta el flujo de la conversación de tal manera que el usuario no pueda interactuar libremente con este tipo de chatbots debido a las opciones que presenta. El usuario debe elegir una para obtener la respuesta (Saldarini *et al.*, 2022).

Chatbots conversacionales

Los chatbots son bots que permiten a los usuarios publicar libremente preguntas y respuestas relacionadas con conversaciones en un chat. Este tipo de chatbots utilizan técnicas como el procesamiento del lenguaje natural para comprender el contexto de las conversaciones y el aprendizaje automático para comprender y aprender el comportamiento, las emociones y las intenciones del usuario para hacer las conversaciones más complejas y realizar funciones que las hagan naturales (Lucana y Roldan, 2023).

Chatbots contextuales

También son bots cognitivos, lo que significa que su inteligencia les permite comprender cuáles son las intenciones del usuario y responder de acuerdo con el contexto solicitado. Pueden ser fluidos y capaces de expresarse de manera similar a una persona, lo que facilita una comunicación

cómoda para el usuario. No dependen de comandos que limitan y cierran la discusión, sino que permiten una interacción más abierta y natural (Lucana y Roldan, 2023).

Chatbot en la educación

Los chatbots en la educación son herramientas que interactúan con los estudiantes y docentes para brindarles información y apoyo en su proceso educativo de aprendizaje/enseñanza. Estas herramientas automatizadas pueden compartir contenido educativo, responder preguntas frecuentes, ofrecer asistencia en tiempo real y comunicarse de forma rápida y eficiente.

Los chatbots en educación ofrecen beneficios como la disponibilidad las 24 horas del día, los siete días de la semana, la capacidad de proporcionar respuestas rápidas y precisas, y la posibilidad de ofrecer una experiencia personalizada para cada estudiante. También pueden ayudar a las instituciones educativas a automatizar tareas administrativas y liberar tiempo para los educadores. Algunas funciones comunes que los chatbots educativos pueden desempeñar incluyen dar respuestas a preguntas frecuentes.

Los chatbots pueden responder de manera instantánea preguntas comunes sobre información académica, horarios, requisitos del programa, entre otros. Dar soporte técnico. Los chatbots pueden ayudar a los estudiantes con problemas técnicos relacionados con plataformas y herramientas digitales utilizadas en el proceso educativo. Brindar orientación académica. Los chatbots pueden proporcionar orientación sobre planes de estudio, opciones de cursos, requisitos de graduación y oportunidades de aprendizaje adicional.

El uso de chatbots agiliza la organización de estudiantes, docentes y administración y optimiza la gestión del tiempo en las instituciones educativas. Así, se puede invertir esfuerzo en otros procesos y se puede mejorar la calidad de la enseñanza (Saldarini *et al.*, 2022).

El fácil acceso al contenido para los estudiantes definitivamente aumenta el compromiso con la institución y el plan de estudios. Además de la comunicación entre estudiantes y profesores, los chatbots deben notificar a los estudiantes sobre tareas, fechas límite y eventos importantes (García *et al.*, 2023).

ISO 25000

ISO (Organización Internacional para la Estandarización) ha desarrollado muchos estándares enfocados en las TIC (tecnologías de la información

y la comunicación), pero ISO 25000, también llamado SQuaRE (requisitos de calidad y evaluación de sistemas y software), es la única familia de estándares que se centran en la calidad del software (Morochó y Gómez, 2019).

La familia de normas ISO 25000 está destinada a empresas de software y se utiliza para evaluar la calidad de los productos de software. En este estándar se especifican criterios de evaluación, así como sus correspondientes métricas que en conjunto forman un modelo de calidad cuyo fin es unificar las diferentes definiciones de calidad, tanto las del cliente como la de los atributos desarrollados del producto (Naranjo, 2015). El estándar contiene métricas de calidad, tanto internas como externas. El propósito de crear un modelo basado en SQuaRE es identificar correctamente las características que se deben considerar en el desarrollo de software, la posibilidad de obtener una ventaja competitiva, los requisitos del software son la base de las medidas de calidad, planificación, desarrollo de objetivos, coordinación, aprendizaje y adaptación (Dávila y Angeleri, 2022).

La norma ISO 25000 consta de las siguientes secciones:

Figura 1. Secciones de la ISO 25000



Nota. Tomado de ISO 25000 (ISO/IEC, 2019).

Los beneficios que tiene la ISO/IEC 25000

Los beneficios clave de ISO/IEC 25000 incluyen evitar ineficiencias en el proceso de evaluación de software, maximizar la rentabilidad y la calidad del producto, evaluar las características de calidad del producto software en un marco que permite obtener valores cuantitativos de este proceso con el fin de determinar el nivel de calidad del producto, objetivo del presente trabajo (ISO/IEC, 2019).

Instaurar el modelo de calidad

El modelo de calidad del software que se pretende crear se basa en el estándar internacional ISO/IEC 25000, el cual se ha aplicado en trabajos previos como en (Naranjo y González, 2018; Naranjo *et al.*, 2020) y modelos similares como en (Ríos, 2023; Tentra, 2022; Romilla *et al.*, 2023), satisfaciendo no solo los requisitos básicos de calidad del software, sino que además ayudan al evaluador a alcanzar sus objetivos. Estos niveles son el tipo de calidad, características, sub-características, métricas e indicadores.

Características y sub-características de la ISO 25000

Según el modelo de calidad de la ISO 25000, como se muestra en la figura 1 y para propósitos de este estudio se evaluará la característica de usabilidad con sus correspondientes subcaracterísticas que son: capacidad de aprendizaje, protección frente a errores y estética, capacidad de aprendizaje: evalúa si el sistema presentado cuenta con elementos que faciliten su uso, desde aprender a utilizarlo, si se cuenta con un manual o documento de referencia, hasta una breve descripción de sus funciones que permitan al usuario administrar de manera óptima la terapia. Capacidad de respuesta: evalúa el software para ver si tiene elementos que faciliten su uso y permitan controles simples y adecuados. Protección frente a errores: evalúa la capacidad del software para detectar la presencia de errores, informarlos y emitir advertencias, ya sea una entrada de datos incorrecta o una falla del sistema. Accesibilidad: evalúa la capacidad del software para adaptarse a los requisitos del usuario, espacio de color, tamaño de fuente, teclas de acceso rápido o acceso desde cualquier dispositivo (Debnath *et al.*, 2021).

Tabla 1. Métricas de la usabilidad según norma ISO 25000

| Subcaracterísticas | Métricas |
|--------------------------|------------------------------------------------|
| Operabilidad | Recuperabilidad de errores operacionales |
| | Tiempo de respuestas más eficientes |
| Capacidad de aprendizaje | Tolerancia a fallos del sistema |
| | Simplicidad para el usuario al usar el sistema |
| Estética | Complacencia al usuario con el sistema |
| | Diseño integral de interfaces al usuario |
| Accesibilidad | Consideración de acceso a los usuarios |
| | Facilidad de ingreso en el sistema |

| Subcaracterísticas | Métricas |
|----------------------------------------|--------------------------------------|
| Inteligibilidad | Evaluación y aprendizaje del sistema |
| | Facilidad de entendimiento |
| Protección Frente a Errores de Usuario | Capacidad funcional del sistema |
| | Copias de seguridad del sistema |

Nota. Elaborado por los autores a partir de la norma ISO 25000.

ISO 25010

En cuanto a la usabilidad del software, la norma ISO/IEC 25010 ofrece pautas y subcaracterísticas específicas que se pueden utilizar para evaluar la usabilidad de un producto de software (Lucas-Franco y Naranjo Sánchez, 2023). La norma ISO/IEC 25010 define la calidad del producto de software en términos de características de calidad, y para la usabilidad, se centra en las siguientes subcaracterísticas:

La adecuación, que evalúa si el software satisface las necesidades del usuario en términos de tareas específicas y objetivos. La adecuación se refiere a si el software es apropiado y útil para el usuario.

La exactitud, que mide la precisión con la que el software realiza sus funciones. Un software preciso es aquel que produce resultados correctos y confiables. La interoperabilidad evalúa la capacidad del software para interactuar y funcionar de manera efectiva con otros sistemas o componentes de software.

La facilidad de uso se refiere a la facilidad con la que los usuarios pueden aprender a utilizar el software y realizar tareas dentro de él. Incluye aspectos como la intuitividad de la interfaz de usuario y la capacidad de los usuarios para realizar tareas sin esfuerzo excesivo. La estética de la interfaz de usuario evalúa la apariencia visual y la presentación de la interfaz de usuario. Esto puede afectar la percepción del usuario influyendo sobre este para calificarlo como un software de calidad o no. La accesibilidad evalúa si el software es accesible para usuarios con discapacidades, como aquellos con problemas de visión o movilidad. Se centra en hacer que el software sea utilizable por una amplia gama de personas (ISO/IEC, 2019).

Evaluación de chatbots

Trabajos previos como el de Nevado (2019), han considerado las métricas eficacia, eficiencia, satisfacción y calidad de la ISO 25000 para su

evaluación de un chatbot, para ello desarrolló un estudio experimental en el que comparó la usabilidad alcanzada por el chatbot SOCIO y Creately, ambas, herramientas para la elaboración de diagramas de clase, sin embargo, SOCIO tenía la capacidad de interpretar el lenguaje natural. En el experimento, de 30 estudiantes que conforman un curso de ingeniería informática, la mitad usó primero la herramienta SOCIO y luego Creately, mientras que la otra mitad lo hizo a la inversa, primero Creately y luego SOCIO. Evaluó la usabilidad respecto a la eficiencia, satisfacción y eficacia. Sus resultados no mostraron diferencias significativas entre la usabilidad de una herramienta con otra, sin embargo, si se mide la eficiencia tomando en cuenta el tiempo necesario para obtener el propósito de creación de diagramas y también considerando los mensajes de discusión generados, resultando SOCIO más eficiente que su competidor, dicho de otra forma, el esfuerzo de creación de diagramas fue menor al que se utilizó en la otra herramienta.

Ren *et al.* (2022) realizaron un mapeo sistemático de 28 estudios primarios para identificar las preguntas de investigación, las características y las métricas utilizadas para evaluar la usabilidad de los chatbots en dichos estudios experimentales. Encontraron que la eficiencia, la satisfacción y la eficacia son características de usabilidad que se utilizan para determinar hasta qué punto los usuarios pueden aprender y utilizar los chatbots para alcanzar sus objetivos, y hasta qué punto están satisfechos durante la interacción. En general, los resultados experimentales revelaron que los chatbots tienen varias ventajas (por ejemplo, proporcionan una respuesta en tiempo real y mejoran la facilidad de uso) y algunas deficiencias (por ejemplo, el procesamiento del lenguaje natural, que se califica como el punto débil que más necesita mejorar).

Bacilio (2021) desarrolló un estudio con el objetivo de evaluar la usabilidad de un chatbot para el seguimiento de adherencia al tratamiento, y efectos adversos en un ensayo clínico en el Perú. Implementó el chatbot y finalizado su estudio, invitó a los estudiantes y al equipo de monitoreo, un total de 40 participantes, a evaluar el sistema con la “Escala de Usabilidad del Sistema”, al mismo tiempo que recolectó métricas del uso del chatbot. Se realizaron 3526 interacciones de las cuales 86,6 % fueron satisfactorias. El puntaje promedio de usabilidad fue de 78.25 (“Aceptable”), mientras, no encontró diferencias estadísticas significativas acerca de las variables sociodemográficas. El equipo de monitoreo calificó el sistema con un valor similar, con un 71,7 (“Aceptable”).

De acuerdo con lo investigado en la revisión de la literatura se puede concluir que luego de un análisis pormenorizado, que las métricas necesarias para evaluar el chatbot son las comunes a los trabajos previos, esto es: eficiencia, efectividad, satisfacción y calidad.

El objetivo de este estudio es evaluar un chatbot educativo mediante un modelo de calidad para determinar si cumple con las normas de calidad ISO 25000.

Metodología

Esta investigación trata de un estudio con diseño no experimental, ya que no se realizó manipulación de variables. De acuerdo con el nivel de conocimiento es un estudio descriptivo ya que primero se detalla el problema en torno a la necesidad de evaluación de los chatbots en el ámbito educativo. Además, se realizó un estudio documental en el que se revisaron artículos científicos para recolectar información referente al uso de chatbots en el ámbito educativo y su evaluación.

Se aplicó el método analítico sintético para conocer todo el contexto relacionado al uso de los chatbots en la educación y realizar un resumen adecuado. De acuerdo con el enfoque, es un estudio cuantitativo ya que los resultados de la aplicación del instrumento, son cuantificables y proporcionales a una muestra de usuarios.

Se diseñó un modelo en base con el estándar ISO 25000 con el fin de evaluar la calidad en uso de un chatbot educacional. El modelo establecido evalúa la usabilidad del chatbot de la asignatura Arquitectura de computadoras a partir de 9 métricas.

Además, se efectuó la evaluación de la percepción de los usuarios sobre el chatbot creado por estudiantes de la Universidad Politécnica Salesiana, para ello se diseñó un cuestionario de ocho preguntas de opción múltiple el cual fue aplicado sobre una muestra de 50 estudiantes de cuarto nivel de la carrera de Computación de la Universidad Politécnica Salesiana sede Guayaquil, quienes hicieron uso del chatbot para obtener información acerca de la materia de Arquitectura de computadoras.

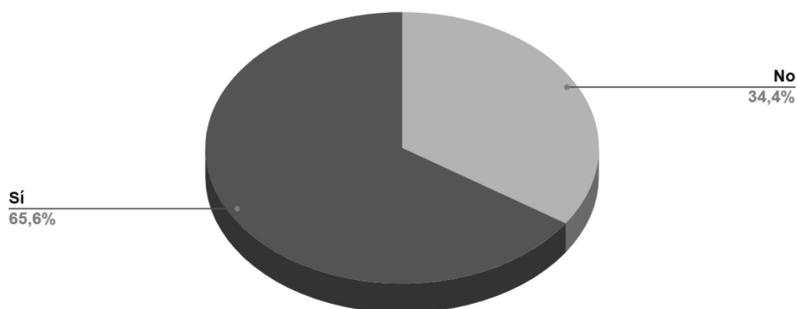
Análisis y resultados

De la aplicación de la encuesta a los estudiantes se obtuvo el siguiente análisis:

a. ¿Ha tenido alguna una experiencia previa con un chatbot?

El 65,6 % sí ha tenido experiencia con el uso de chatbots, mientras que el 34,4 % no, lo que indica que más del 50 % de los estudiantes encuestados han tenido experiencias usando chatbots, lo cual es favorable (ver figura 2).

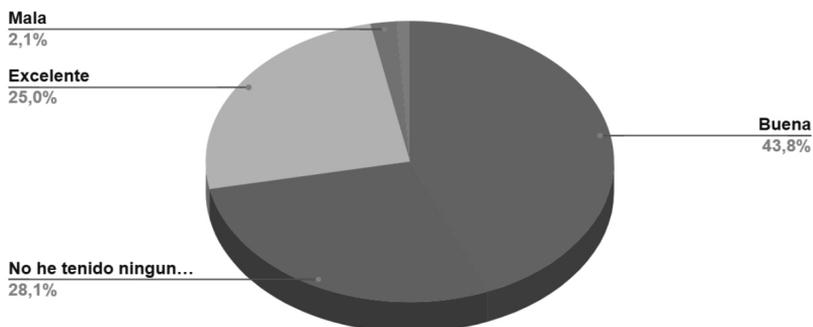
Figura 2. *Experiencia previa con chatbots*



b. ¿Está de acuerdo en implementar un chatbot para la educación en general?

Se obtuvo que el 83,4 % está de acuerdo con la implementación de un chatbot para mejorar la educación y solo el 16,7 % no está de acuerdo con esta implementación. Estos resultados indican que la mayoría de las estudiantes tienen una percepción positiva sobre los chatbots para el uso en la educación (ver figura 3).

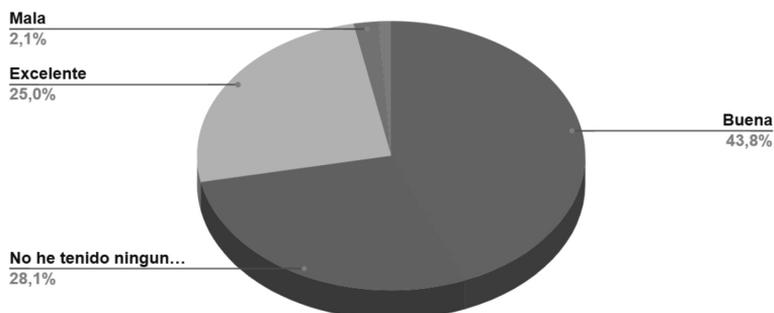
Figura 3. *Percepción sobre la implementación de un chatbot para la educación*



c. ¿Qué tan satisfecho se encuentra con la experiencia que ha tenido con el chatbot?

Al 2,1 % le pareció una mala experiencia, al 43,8 % le pareció una buena experiencia, al 25% le pareció excelente, mientras que el 28,1 % no ha tenido experiencia con un chatbot (ver figura 4).

Figura 4. Percepción sobre la experiencia con el chatbot

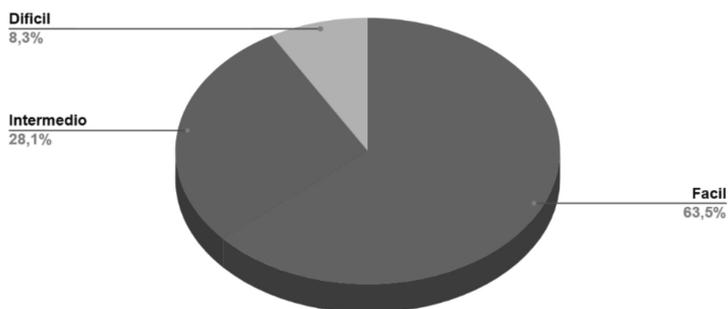


Estos resultados demuestran que más del 50 % de estudiantes han tenido experiencias positivas con el uso del chatbot.

d. ¿Con que nivel calificaría su interacción con el chatbot?

Al 8,3 % de los encuestados le pareció difícil interactuar con el chatbot, mientras que al 28,1 % le pareció una dificultad intermedia, y al 63,5 % le pareció fácil. Esto muestra que interactuar con el chatbot es fácil para los estudiantes (ver figura 5).

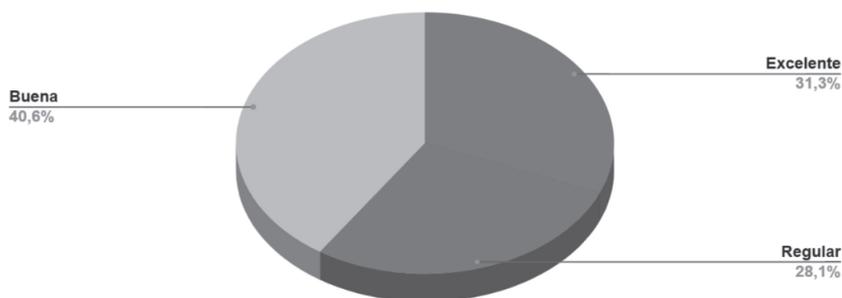
Figura 5. Percepción sobre la facilidad para interactuar con el chatbot



e. ¿Considera que el chatbot fue útil para resolver su problema o brindarle la información que necesitaba?

Según la Figura 6, al 31,3 % de los estudiantes encuestados le pareció que el chatbot fue excelente para resolver sus necesidades de información, al 40,6 % le pareció buena la forma en que el chatbot contribuyó en sus conocimientos, mientras que al 28,1 % le pareció regular (ver figura 6).

Figura 6. Percepción sobre la utilidad del chatbot para resolver un problema

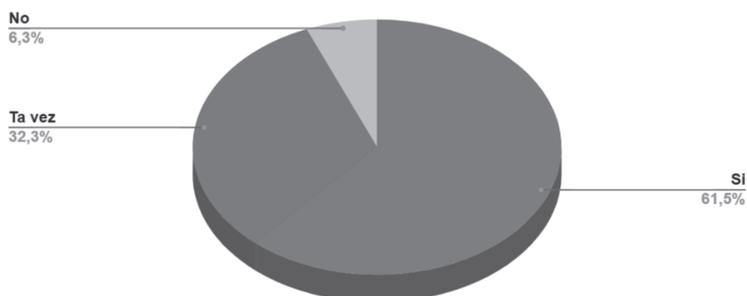


En general, la mayoría de estudiantes consideran que es muy útil usar un chatbot para resolver necesidades de conocimiento.

f. ¿Recomendaría el uso del chatbot a otras personas?

Al consultar si recomendarían el uso del chatbot a otras personas, el 61,5 % indicó que sí lo haría, el 32,3% tal vez lo haría, mientras que solo el 6,3 % no lo haría. Esto refleja lo positivo que ha sido el uso del chatbot para sus objetivos educativos, es así que desean recomendarlo a su círculo social (ver figura 7).

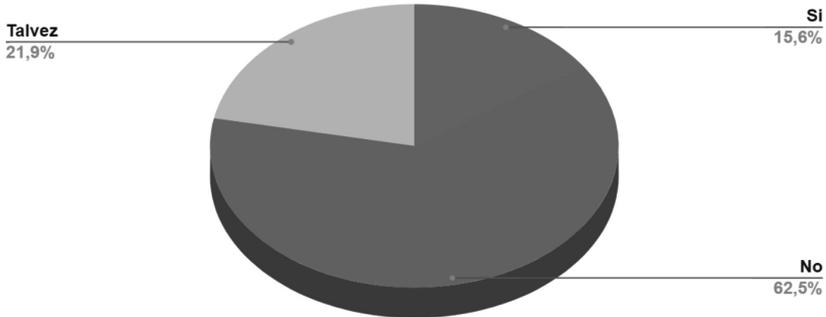
Figura 7. Percepción sobre recomendar el chatbot a otras personas



g. ¿Hubo algún problema al momento de usar el chatbot?

Referente a si han tenido algún problema al momento de usar el chatbot, el 62,5 % mencionó que no, mientras que el 21,9 % indicó que tal vez ha tenido inconvenientes, y el 15,6 % adujo sí haber tenido inconvenientes (ver figura 8).

Figura 8. Percepción sobre problemas al usar el chatbot

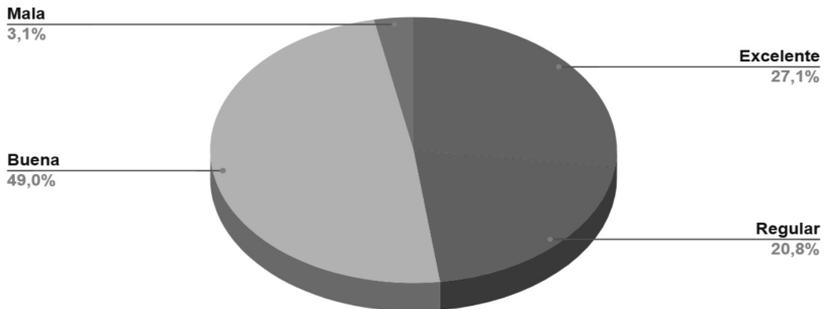


La mayoría de estudiantes no han tenido problemas cuando han usado el chatbot, lo cual es favorable pues indica que está cumpliendo la función para la que fue creado.

h. ¿Considera que el tiempo de respuesta del chatbot fue adecuado?

Se obtuvo que el 27,1 % de estudiantes encuestados considera que el tiempo de respuesta del chatbot fue excelente, el 49 % considera que fue bueno, el 20,8 % lo considera regular y solo para el 3,1 % fue malo (ver figura 9).

Figura 9. Percepción sobre el tiempo de respuesta del chatbot



Estos resultados muestran que la mayoría de estudiantes están de acuerdo con los tiempos de respuesta del chatbot.

Modelo de evaluación establecido

A continuación, se presenta el modelo de evaluación de la usabilidad establecido. El modelo fue elaborado a partir de lo que indica la norma ISO 25000. El modelo establecido evalúa la usabilidad del chatbot a través de 9 métricas agrupadas en tres distintas subcaracterísticas. La escala utilizada para la evaluación de la usabilidad se muestra en la tabla 2.

Tabla 2. Escala para evaluación de usabilidad

| Escala | Valor |
|----------|--------------------|
| 0-24.99 | Insatisfactorio |
| 25-49.99 | Poco satisfactorio |
| 50-74.99 | Satisfactorio |
| 75-100 | Muy satisfactorio |

La tabla 3 muestra la usabilidad como característica evaluada para determinar la calidad del chatbot, las subcaracterísticas de eficacia, eficiencia y satisfacción que componen a la usabilidad y cada una con su puntaje esperado.

Tabla 3. Puntaje por subcaracterísticas del chatbot

| Característica | Subcaracterísticas | Puntaje |
|----------------|--------------------|---------|
| Usabilidad | Eficacia | 40 |
| | Eficiencia | 30 |
| | Satisfacción | 30 |
| | TOTAL | 100 |

La tabla 4 muestra las métricas evaluadas con su correspondiente valor deseado en porcentaje. Se evaluaron en total 9 métricas.

Tabla 4. Valor deseado por métricas

| Sub-características | Nombre de las métricas | Valor deseado |
|---------------------|--------------------------------|---------------|
| Eficacia | Completitud de la tarea | 70 % |
| | Número de errores | 15 % |
| | Precisión | 15 % |
| Eficiencia | Tiempo para completar la tarea | 70 % |

| Sub-características | Nombre de las métricas | Valor deseado |
|---------------------|-------------------------------|---------------|
| Satisfacción | Experiencia de uso en general | 70 % |
| | Facilidad de uso | 15 % |
| | Comodidad | 15 % |

La tabla 5 presenta los resultados de la evaluación de las métricas con su puntaje esperado y obtenido. Se observa que los puntajes más altos se obtuvieron en las métricas de Facilidad de uso 4/4.5, Comodidad 4/4.5, Número de errores, 5/6, Precisión 5/6 y Calidad de respuesta 3.5/4.5. Mientras que los puntajes más bajos se obtuvieron en las métricas de Completitud de la tarea 24/28 y Experiencia de uso en general 19/21.

De acuerdo con las subcaracterísticas, los totales obtenidos fueron los siguientes: eficacia alcanzó un puntaje total de 34/40, eficiencia alcanzó el 25/30 y satisfacción logró el 27/30, lo que refleja que la subcaracterística de satisfacción alcanzó el más alto puntaje y la de eficacia el más bajo. De forma global, la usabilidad del chatbot alcanzó el puntaje total de 86/100 lo que representa una evaluación satisfactoria del uso del chatbot y de su calidad.

Los resultados muestran que para los estudiantes fue muy fácil y cómodo hacer uso del chatbot. Además, sus respuestas eran acordes a lo que los estudiantes preguntaban. Mientras tanto, el chatbot debe mejorar principalmente en el cumplimiento de la tarea, es decir en cumplir con su propósito que es de educar, a la vez debe mejorar el tiempo de respuesta para completar la tarea, para que su uso sea más atractivo para los usuarios.

Tabla 5. Puntaje obtenido y esperado de la evaluación de métricas

| Subcaracterísticas | Métricas | Puntaje esperado | Puntaje obtenido |
|--------------------|--------------------------------|------------------|------------------|
| Eficacia | Completitud de la tarea | 28 | 24 |
| | Número de errores | 6 | 5 |
| | Precisión | 6 | 5 |
| Eficiencia | Tiempo para completar la tarea | 21 | 18.5 |
| | Esfuerzo de comunicación | 4.5 | 3 |
| | Calidad de respuesta | 4.5 | 3.5 |
| Satisfacción | Experiencia de uso en general | 21 | 19 |
| | Facilidad de uso | 4.5 | 4 |
| | Comodidad | 4.5 | 4 |
| Total | | 100 | 86 |

Conclusiones

La evaluación de la calidad de un chatbot a través de la norma ISO 25000 es esencial en el panorama actual de la tecnología y la interacción con usuarios. La norma ISO 25000, que se basa en estándares internacionales reconocidos, proporciona un marco sólido y estructurado para evaluar varios aspectos de la calidad de un software, incluyendo la usabilidad.

En este estudio se hizo una evaluación de la usabilidad de un chatbot educativo desarrollado por estudiantes de la Universidad Politécnica Salesiana sede Guayaquil, a través de 9 métricas repartidas en tres subcaracterísticas de la usabilidad: eficacia, eficiencia y satisfacción del usuario.

Los resultados de las encuestas aplicadas a los estudiantes indican que el 65,5 % de estudiantes sí había tenido experiencias previas con un chatbot, el 83,4 % está de acuerdo con la implementación de un chatbot para mejorar la educación, al 43,8 % le pareció una buena experiencia, al 63,5 % le pareció fácil su uso, al 40,6 % le pareció buena la forma en que el chatbot contribuyó en sus conocimientos, el 61,5 % recomendaría el chatbot a otras personas, el 62,5 % no tuvo inconvenientes con el chatbot, el 49 % considera que el tiempo de respuesta del chatbot fue bueno, lo que permite obtener una buena percepción de los usuarios respecto al chatbot.

El puntaje total de 86/100 representa una evaluación muy satisfactoria del uso del chatbot en relación con su calidad acorde al estándar ISO 25000, arrojando puntajes bastante cercanos al total esperado en cada una de sus subcaracterísticas por lo que se recomienda su uso en un entorno educativo para el aprendizaje tanto de estudiantes como de docentes.

De acuerdo con los resultados, los criterios en los que se debe mejorar la usabilidad del software son: en el cumplimiento de la tarea, es decir en cumplir con su propósito que es de educar, a la vez debe mejorar el tiempo de respuesta.

Se recomienda que las instituciones educativas y los docentes consideren la evaluación de la usabilidad de los chatbots, basados en el estándar ISO 25000, para garantizar así la calidad del producto y su utilidad en el proceso educativo.

Agradecimiento

Al grupo TICAD por el apoyo en la realización de este artículo dentro del proyecto interdisciplinario MEMOTECH del grupo de innovación educativa GIE IDI.

Referencias bibliográficas

- Adamopoulou, E. y Moussiades, L. (2020). An overview of chatbot technology. In IFIP international conference on artificial intelligence applications and innovations. *IFIP International Conference on Artificial Intelligence Applications and Innovations*, 373-383. https://doi.org/10.1007/978-3-030-49186-4_31
- Auqui, J. (2021). Chatbot del proceso de aprendizaje universitario: una revisión sistemática. *Revista de Investigación Científica y Tecnológica Alpha Centauri*, 2(2), 29-43. <https://bit.ly/4fODEWX>
- Bacilio, A. (2021). *Evaluación del uso de un chatbot para el seguimiento en un ensayo clínico de profilaxis frente al Covid-19 en personal de salud*. Universidad peruana Cayetano Heredia.
- Dávila, A. y Angeleri, P. (2022). Software Product Evaluation-MyFEPS: A past and future reflective analysis. *11th International Conference On Software Process Improvement (CIMPS)*, 37-45. <https://doi.org/10.1109/CIMPS57786.2022.10035711>
- Debnath, N., Peralta, M., Salgado, C., Baigorria, L., Riesco, D., Montejano, G. y Mazzi, M. (2021). Digital transformation: A quality model based on ISO 25010 and user experience. *The 33rd International Conference on Computer Applications in Industry and Engineering*, 75, 11-21.
- Følstad, A. y Taylor, C. (2021). Investigating the user experience of customer service chatbot interaction: a framework for qualitative analysis of chatbot dialogues. *Quality and User Experience*, 6(1), 6. <https://doi.org/10.1007/s41233-021-00046-5>
- Lucas Franco, G. S. y Naranjo Sánchez, B. A. (2023). Análisis del Sistema Sipadi en el aprendizaje de niños con discapacidad intelectual. En *Estudiantes universitarios por la inclusión* (pp. 37-51). Ediciones Abya-Yala.
- García, H., Hernández, L. y Gámez, R. (2023). Perspectivas del uso de chatbots en la educación superior: caso de estudio de la Universidad Tecnológica de Gutiérrez. *Revista Multidisciplinaria*, 1, 32(37).
- ISO/IEC. (2019). *ISO/IEC 25000*. <https://bit.ly/3MaK6jU>
- Lopez, T. y Qamber, M. (2022). *The benefits and drawbacks of implementing chatbots in higher education: A case study for international students at Jönköping University*. Jönköping International Business School. <https://bit.ly/3MaKb7cv>
- Lucana, Y. y Roldan, W. (2023). Chatbot basado en inteligencia artificial para la educación escolar. *Horizontes Revista de Investigación en Ciencias de la Educación*, 7(29), 1580-1592. <https://doi.org/10.33996/revistahorizontes.v7i29.614>

- Mrsic, L., Mesic, T. y Balkovic, M. (2020). Cognitive Services Applied as Student Support Service Chatbot for Educational Institution. *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2019*, 1(1), 417-424. https://doi.org/10.1007/978-981-15-1286-5_35
- Morocho, E. y Gómez, E. (2019). Application of ISO 25000 for Comparative Analysis of Cross-Platform Mobile Application Development Tools. *International Conference on Information Systems and Computer Science (INCISCOS)*, 319-325. <https://doi.org/10.1109/INCISCOS49368.2019.00057>
- Naranjo, A. y González, J. (2018). Comparative study of the ebrai and quick braille tools for people with visual disabilities. *2018 International Conference on Information Systems and Computer Science (INCISCOS)*, 284-293. <https://doi.org/10.1109/INCISCOS.2018.00048>
- Naranjo, B., Tinoco, M. y Vega, D. (2020). Análisis de la usabilidad del sistema web de terapias cognitivas sanamentics. *Revista Redipe*, 9(5), 175-87. <https://doi.org/10.36260/rbr.v9i5.986>
- Naranjo Sánchez, B. A. (2015). Calidad del software educativo: Metodología de Evaluación de software educativo para determinar el que cumple con las especificaciones basadas en estándares internacionales.
- Nevado, A. (2019). *Evaluación empírica de la usabilidad de un chatbot*. Universidad Autónoma de Madrid.
- Pisco, E. (2021). *Desarrollo de un chatbot como soporte al proceso de aprendizaje de la programación en la UPS*. Universidad Politécnica Salesiana. Obtenido de <https://bit.ly/3SSMgsb>
- Presti, L., Maggiore, G. y Marino, V. (2021). The role of the chatbot on customer purchase intention: towards digital relational sales. *Italian Journal of Marketing*, 3, 165-188. <https://doi.org/10.1007/s43039-021-00029-6>
- Ren, R., Zapata, M., Castro, J., Dieste, O. y Acuña, S. (2022). Experimentation for chatbot usability evaluation: a secondary study. *IEEE Access*, 10, 12430-12463. <https://doi.org/10.1109/ACCESS.2022.3145323>
- Ríos, G. (2023). *Aplicación ISO 25000 para el proceso de desarrollo de software en el área de TI en una financiera*. Universidad César Vallejo. <https://bit.ly/46NVKL0>
- Romilla, P., Matuguinas, J., Santiago, A., Cortez, D., Sison, A. y Mercado, M. (2023). Children in-conflict chatbot system using natural language processing technique. *World Journal of Advanced Research and Reviews*, 18(3), 425-429. <https://doi.org/10.30574/wjarr.2023.18.3.1107>
- Saldarini, J., Carrizo, C., Sánchez, A., Salgado, C. y Peralta, M. (2022). Software público: Una propuesta para la mejora de la calidad a través de ISO/IEC 25000. *Sociedad Argentina de Informática e Investigación Operativa*, 8(12), 46-46. <https://bit.ly/3M6ICHc>

- Tentra, A. (2022). ISO/IEC 25000 and AI Product Quality Measurement Perspectives. *Proceedings of 4th International Workshop on Experience with SQuaRE series and its Future Direction*.
- Vega, S. (2021). *Desarrollo de un modelo de calidad externa de software para mejorar la eficiencia y portabilidad del producto final en aplicaciones móviles en la Empresa ADS - Software en el 2020*. Universidad de las Fuerzas Armadas.