



UTMACH

FACULTAD DE INGENIERÍA CIVIL

CARRERA DE TECNOLOGÍAS DE LA INFORMACIÓN

Diseño de un chatbot utilizando técnicas de procesamiento de lenguaje natural para soporte de servicios de información de una IES

**ROMERO TACURI JEANINA MABEL
INGENIERA EN TECNOLOGIAS DE LA INFORMACION**

**VELEZ ENCALADA TANIA PAMELA
INGENIERA EN TECNOLOGIAS DE LA INFORMACION**

**MACHALA
2023**



UTMACH

FACULTAD DE INGENIERÍA CIVIL

CARRERA DE TECNOLOGÍAS DE LA INFORMACIÓN

**Diseño de un chatbot utilizando técnicas de procesamiento de
lenguaje natural para soporte de servicios de información de una
IES**

**ROMERO TACURI JEANINA MABEL
INGENIERA EN TECNOLOGIAS DE LA INFORMACION**

**VELEZ ENCALADA TANIA PAMELA
INGENIERA EN TECNOLOGIAS DE LA INFORMACION**

**MACHALA
2023**



UTMACH

FACULTAD DE INGENIERÍA CIVIL

CARRERA DE TECNOLOGÍAS DE LA INFORMACIÓN

PROPUESTAS TECNOLÓGICAS

Diseño de un chatbot utilizando técnicas de procesamiento de lenguaje natural para soporte de servicios de información de una IES

**ROMERO TACURI JEANINA MABEL
INGENIERA EN TECNOLOGIAS DE LA INFORMACION**

**VELEZ ENCALADA TANIA PAMELA
INGENIERA EN TECNOLOGIAS DE LA INFORMACION**

RIVAS ASANZA WILMER BRAULIO

**MACHALA
2023**

tesisvelezromero

por Tania Velez

Fecha de entrega: 19-feb-2024 11:52a.m. (UTC-0500)

Identificador de la entrega: 2298858955

Nombre del archivo: Tesis_Velez_Romero.pdf (2.41M)

Total de palabras: 13924

Total de caracteres: 77033

tesisvelezromero

INFORME DE ORIGINALIDAD

5%

INDICE DE SIMILITUD

6%

FUENTES DE INTERNET

0%

PUBLICACIONES

1%

TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1

repositorio.utn.edu.ec

Fuente de Internet

1%

2

repositorio.upea.bo

Fuente de Internet

<1%

3

Submitted to Universidad Politécnica Estatal de Carchi

Trabajo del estudiante

<1%

4

www.educaplay.com

Fuente de Internet

<1%

5

1library.co

Fuente de Internet

<1%

6

www.dropbox.com

Fuente de Internet

<1%

7

documents.mx

Fuente de Internet

<1%

8

recerc.eu

Fuente de Internet

<1%

9

repositorio.utm.edu.ec

Fuente de Internet

CLÁUSULA DE CESIÓN DE DERECHO DE PUBLICACIÓN EN EL REPOSITORIO DIGITAL INSTITUCIONAL

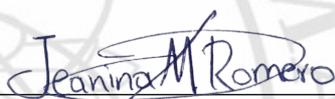
Las que suscriben, ROMERO TACURI JEANINA MABEL y VELEZ ENCALADA TANIA PAMELA, en calidad de autoras del siguiente trabajo escrito titulado Diseño de un chatbot utilizando técnicas de procesamiento de lenguaje natural para soporte de servicios de información de una IES, otorgan a la Universidad Técnica de Machala, de forma gratuita y no exclusiva, los derechos de reproducción, distribución y comunicación pública de la obra, que constituye un trabajo de autoría propia, sobre la cual tienen potestad para otorgar los derechos contenidos en esta licencia.

Las autoras declaran que el contenido que se publicará es de carácter académico y se enmarca en las disposiciones definidas por la Universidad Técnica de Machala.

Se autoriza a transformar la obra, únicamente cuando sea necesario, y a realizar las adaptaciones pertinentes para permitir su preservación, distribución y publicación en el Repositorio Digital Institucional de la Universidad Técnica de Machala.

Las autoras como garantes de la autoría de la obra y en relación a la misma, declaran que la universidad se encuentra libre de todo tipo de responsabilidad sobre el contenido de la obra y que asumen la responsabilidad frente a cualquier reclamo o demanda por parte de terceros de manera exclusiva.

Aceptando esta licencia, se cede a la Universidad Técnica de Machala el derecho exclusivo de archivar, reproducir, convertir, comunicar y/o distribuir la obra mundialmente en formato electrónico y digital a través de su Repositorio Digital Institucional, siempre y cuando no se lo haga para obtener beneficio económico.



ROMERO TACURI JEANINA MABEL

0705576650



VELEZ ENCALADA TANIA PAMELA

0706795671

UNIVERSITAS
MAGISTRORUM
ET SCHOLARIUM



UTMACH

FACULTAD DE INGENIERÍA CIVIL

CARRERA DE TECNOLOGÍAS DE LA INFORMACIÓN

Diseño de un chatbot utilizando técnicas de procesamiento de lenguaje natural para soporte de servicios de información de una IES.

ROMERO TACURI JEANINA MABEL.

INGENIERÍA EN TECNOLOGÍAS DE LA INFORMACIÓN

VELEZ ENCALADA TANIA PAMELA.

INGENIERÍA EN TECNOLOGÍAS DE LA INFORMACIÓN

MACHALA

2023



UTMACH

FACULTAD DE INGENIERÍA CIVIL

CARRERA DE TECNOLOGÍAS DE LA INFORMACIÓN

Diseño de un chatbot utilizando técnicas de procesamiento de lenguaje natural para soporte de servicios de información de una IES.

ROMERO TACURI JEANINA MABEL.

INGENIERÍA EN TECNOLOGÍAS DE LA INFORMACIÓN

VELEZ ENCALADA TANIA PAMELA.

INGENIERÍA EN TECNOLOGÍAS DE LA INFORMACIÓN

RIVAS ASANZA WILMER BRAULIO

MACHALA

2023

DEDICATORIA

El presente trabajo lo dedico a mis progenitores, Milton y Enma, quienes, con su invaluable paciencia y apoyo, me han ayudado a alcanzar este logro. Cada día, alentándome a seguir adelante, motivándome en cada paso y enfrentando con valentía cada situación que se presenta.

A mi hermano Klinton, quien siempre estuvo brindarme enseñanzas, además de ser mi inspiración que cualquier objetivo puede lograrse con dedicación, disciplina y, sobre todo, paciencia.

Por último, a ti abuelita que parecía que nunca me ibas hacer falta, quien me vio prácticamente crecer y me enseñó el verdadero significado de paciencia.

Romero Tacuri Jeanina Mabel

El presente trabajo se lo dedico con profundo amor y agradecimiento a mis queridos padres, que han sido mi pilar inquebrantable, me han brindado su apoyo incondicional y la valiosa lección de perseverancia. Siempre han sido mi fuente de motivación, impulsándome a alcanzar cada una de mis metas.

A mi hermana menor, por ser un apoyo constante y a toda mi familia por su aliento durante toda mi formación profesional.

Velez Encalada Tania Pamela

AGRADECIMIENTO

Agradezco a mis progenitores, Milton y Enma, quienes siempre me han apoyado tanto moral como económicamente. Asimismo, me han enseñado el verdadero significado del trabajo arduo y la independencia. No queda más que expresar mi sincero agradecimiento por todo lo que han hecho por mí.

Además, quiero expresar mi gratitud a mi amiga y compañera de tesis, Tania Velez, que gracias a su profesionalismo y dedicación a contribuido de manera constante a lo largo del desarrollo de este proyecto.

Y por último agradezco a los docentes que han formado parte de todo este proceso y nuestro tutor Ing. Wilmer Rivas Asanza, quien dedicó tiempo y su conocimiento para lograr cumplir con el objetivo.

Romero Tacuri Jeanina Mabel

Agradezco sinceramente a mis padres, Rosa Encalada y Bello Velez, por ser mi soporte durante mi formación profesional. Su apoyo emocional y económico ha sido primordial durante toda mi trayectoria académica. Gracias por estar siempre a mi lado, brindándome fuerza y recursos para alcanzar mis metas.

También, agradezco a mi tutor el Ing. Wilmer Rivas Asanza, por su orientación invaluable, contribuyendo significativamente al éxito de este trabajo.

Velez Encalada Tania Pamela

RESUMEN

En el panorama actual de la carrera de TI, la integración de tecnologías innovadoras se ha convertido en un pilar para mejorar la calidad de los servicios ofrecidos a los estudiantes y público en general, es así como nació la necesidad del desarrollo de un chatbot para TI en un perfil social de Facebook como prototipo para dar asistencia con información académica de la carrera sin restricción de horarios.

En cuanto a las tecnologías de desarrollo, existen diversos tipos de chatbots, desde aquellos basados en reglas hasta los impulsados por IA y aprendizaje profundo. Sin embargo, la tendencia predominante es el uso de tecnologías como ChatGPT, que se basan en modelos de lenguaje generativo y NLP para ofrecer respuestas naturales y contextuales, lo que los hace muy populares y novedosos.

La propuesta se centró en el desarrollo del prototipo de chatbot impulsado por el modelo de lenguaje largo GPT-3.5 Turbo, el propósito era optimizar la manera en que los usuarios interactúan con el chatbot y así obtener la respuesta eficaz agilizando las interacciones y ofreciendo respuestas precisas y la disponibilidad ininterrumpida.

La implementación de técnicas de procesamiento de lenguaje natural facilitó la contextualización de preguntas en el chatbot, siendo crucial para el apoyo en la gestión de información académica. Para fortalecer este enfoque, la realización de investigaciones sistemáticas se reveló como un paso esencial ya que contribuyeron al entendimiento de casos similares y orientaron la selección óptima de técnicas y herramientas en el desarrollo del chatbot. En el prototipo, se emplearon tecnologías como GPT-3.5, Pinecone y Langchain, creando un sistema especializado que ofrece respuestas precisas. La integración en Facebook fue crucial para pruebas exhaustivas y gestión de errores, mejorando la eficacia del algoritmo. Tras la evaluación, el chatbot alcanzó una sólida precisión del 94.5%, respaldando su eficacia en la plataforma.

PALABRAS CLAVES

chatbot, chatgpt, Inteligencia artificial, modelo de lenguaje largo, redes sociales.

SUMMARY

In the current IT career landscape, the integration of innovative technologies has become a pillar to improve the quality of services offered to students and the general public, this is how the need for the development of a chatbot for IT in a Facebook social profile was born as a prototype to provide assistance with academic information of the career without time restriction.

In terms of development technologies, there are several types of chatbots, from those based on rules to those driven by AI and deep learning. However, the predominant trend is the use of technologies such as ChatGPT, which are based on generative language models and NLP to provide natural and contextual responses, making them very popular and novel.

The proposal focused on the development of the chatbot prototype driven by the long language model GPT-3.5 Turbo, the purpose was to optimize the way users interact with the chatbot and thus obtain the effective response by streamlining interactions and providing accurate responses and uninterrupted availability.

The implementation of natural language processing techniques facilitated the contextualization of questions in the chatbot, being crucial for academic information management support. To strengthen this approach, conducting systematic research proved to be an essential step as it contributed to the understanding of similar cases and guided the optimal selection of techniques and tools in the development of the chatbot. In the prototype, technologies such as GPT-3.5, Pinecone and Langchain were employed, creating a specialized system that delivers accurate responses. Facebook integration was crucial for thorough testing and error management, improving the algorithm's efficiency. After evaluation, the chatbot achieved a solid 94.5% accuracy, supporting its effectiveness on the platform.

KEYWORDS

chatbot, chatgpt, artificial intelligence, long language model, social networks.

ÍNDICE DE CONTENIDO

DEDICATORIA.....	III
AGRADECIMIENTO.....	IV
RESUMEN.....	V
PALABRAS CLAVES	V
SUMMARY	VI
KEYWORDS	VI
ÍNDICE DE CONTENIDO.....	VII
GLOSARIO.....	XII
INTRODUCCIÓN	13
i. Declaración y formulación del Problema.....	13
ii. Objeto de estudio y Campo de acción.....	15
iii. Objetivos	16
iv. Hipótesis y variables o Preguntas de investigación.....	16
v. Justificación.....	18
vi. Organización del documento.....	18
CAPÍTULO I. MARCO TEÓRICO.....	19
1.1 Antecedentes de la Investigación.....	19
1.2 Antecedentes Teóricos.....	24
1.2.1 Inteligencia Artificial	26
1.2.2 Chatbots	28
1.2.3 Herramientas	32
1.2.4 Metodología de desarrollo	34
CAPÍTULO II. DESARROLLO DEL PROTOTIPO	39
2.1 Definición del prototipo.....	39
2.2 Metodología de desarrollo del prototipo.....	39
2.2.1 Enfoque, alcance y diseño de investigación	39
2.2.2 Unidades de análisis	40
2.2.3 Técnicas e instrumentos de recopilación de datos	40
2.2.4 Técnicas de procesamiento de datos para la obtención de resultados	41
2.2.5 Metodología o métodos específicos	41
2.2.6 Herramientas y/o Materiales	41
2.3 Desarrollo del prototipo.....	42
2.3.1 Metodología XP	42
2.4 Ejecución del prototipo.....	66
2.4.1 Ejecución Local	66
2.4.2 Ejecución en línea desde Facebook	68
2.4.3 Análisis de Resultados	69
CAPÍTULO III. EVALUACIÓN DEL PROTOTIPO	74

3.1	Plan de evaluación	74
3.1.1	Objetivo de la evaluación.....	74
3.1.2	Alcance	74
3.1.3	Técnica de Evaluación	75
3.1.4	Recursos	75
3.1.5	Cronograma.....	75
3.1.6	Matriz de confusión	76
3.2	Resultados de la evaluación.....	77
CONCLUSIONES		79
RECOMENDACIONES		80
BIBLIOGRAFÍA.....		81
ANEXOS.....		85

ÍNDICE DE FIGURAS

Figura 1. Árbol de problemas.....	14
Figura 2. Proceso de resultados de información.....	22
Figura 3. Estadística de países que destacan en publicaciones	22
Figura 4. Publicaciones acordes al área de estudio	23
Figura 5. Bibliométrico por palabras claves.....	23
Figura 6. Organigrama del marco teórico	25
Figura 7. Ejemplo de Arquitectura basada en microservicio.	30
Figura 8. Etapas del método Extreme Programming XP	35
Figura 9. Arquitectura del prototipo de chatbot	49
Figura 10. Primera etapa de arquitectura del chatbot.....	50
Figura 11. Segunda etapa de arquitectura del chatbot.....	50
Figura 12. Tercera etapa de arquitectura del chatbot	51
Figura 13. Cuarta etapa de arquitectura del chatbot.....	51
Figura 14. Propuesta de interfaz.....	52
Figura 15. Librerías empleadas	53
Figura 16. Integración en Visual Code.....	54
Figura 17. Base de datos Pinecone.....	54
Figura 18. Fragmentar Datos.....	55
Figura 19. Creando embeddings.....	55
Figura 20. Búsqueda de Índice.....	56
Figura 21. Ajuste de Precisión	56
Figura 22. Domino	57
Figura 23. Creación de entorno virtual.....	57
Figura 24. Ruta del entorno virtual	58
Figura 25. Subir archivos al directorio de Pythonanywhere	58
Figura 26. Recargando app.....	58
Figura 27. Crear nueva app en Facebook.....	59
Figura 28. Tipo de aplicación.....	59
Figura 29. Nombrando App	60
Figura 30. Selección de productos.	60
Figura 31. Suscripciones Webhook.....	60
Figura 32. Callback URL y Token	61
Figura 33. Conexión con el chatbot	61
Figura 34. Fragmentar	62
Figura 35. Prueba index pinecone	63
Figura 36. Consultas desde Flask.....	64
Figura 37. Prueba del chatbot integrado.....	65
Figura 38. Actualizar data	66
Figura 39. Activar entorno	66
Figura 40. Injecting para cargar la data.....	67
Figura 41. Ejecución del modelo.....	67
Figura 42. Ejecución y prueba local.....	67
Figura 43. Login de ingreso al administrador	67
Figura 44. Página de Prototipo.....	68
Figura 45. Interacción de prueba.....	68

Figura 46. Asignaturas por semestre	69
Figura 47. Cuantas y cuales materias veo por semestre	70
Figura 48. Proceso de ingreso a la carrea	70
Figura 49. Proceso de justificación de faltas.	71
Figura 50. Proceso de pago de matriculas	71
Figura 51. Coordinador de la carrera.....	72
Figura 52. Planta docente	72
Figura 53. Ubicación y a que Facultad pertenece la carrea.....	73
Figura 54. Anexo 2.....	87
Figura 55. Anexo 3.....	87
Figura 56. Anexo 4.....	88

ÍNDICE DE TABLAS

Tabla 1. Definición de variables y dimensiones.....	17
Tabla 2. Preguntas de investigación.....	19
Tabla 3. Criterios de inclusión y exclusión.....	21
Tabla 4. Comparación con modelos de código cerrado.....	31
Tabla 5. Comparativa LLAMA2 vs GPT-3.5.....	32
Tabla 6. Requisitos Funcionales.....	37
Tabla 7. Requisitos No Funcionales.....	38
Tabla 8. Población.....	40
Tabla 9 Técnica de recolección de datos.....	40
Tabla 10 Herramientas y Materiales.....	41
Tabla 11. Roles y Responsables.....	42
Tabla 12. Historia de Usuario N°1 - Estructura de datos a utilizar.....	42
Tabla 13. Fragmentar de Datos.....	43
Tabla 14. Integración al IDE.....	43
Tabla 15. Creación y carga de vectores.....	43
Tabla 16. Host y servidor web.....	44
Tabla 17. Generando Consultas.....	44
Tabla 18. Ajuste algoritmo de búsqueda.....	44
Tabla 19. Depuración de Data.....	45
Tabla 20. Integración a la red social.....	45
Tabla 21. Prueba del chatbot en la red social integrado.....	45
Tabla 22. Desarrollo de modulo para actualizar la data.....	46
Tabla 23. Aplicación de encuesta.....	46
Tabla 24. Plan de publicaciones.....	47
Tabla 25. Plan de duración de Iteraciones.....	48
Tabla 26. Prueba 1.....	62
Tabla 27. Prueba 2.....	62
Tabla 28. Prueba 3.....	63
Tabla 29. Prueba 4.....	63
Tabla 30. Prueba 5.....	64
Tabla 31. Prueba 6.....	64
Tabla 32. Prueba 7.....	65
Tabla 33. Prueba 8.....	65
Tabla 34. Cronograma de evaluación.....	75
Tabla 355. Tabla de predicción.....	76
Tabla 36. Predicción Malla Curricular.....	77
Tabla 37. Predicción estructura organizativa.....	78
Tabla 38. Predicción preguntas frecuentes.....	78
Tabla 39. Resultado Final.....	78
Tabla 40. Anexo 1.....	85
Tabla 41. Anexo 5.....	92
Tabla 41. Anexo 5.....	96
Tabla 41. Anexo 5.....	99
Tabla 41. Anexo 5.....	102

GLOSARIO

A

API (Application Programming Interface): Interfaz de programación de aplicaciones es una interfaz que permite a dos aplicaciones interactuar, definiendo los parámetros y reglas para el intercambio de datos.

I

Inteligencia Artificial (IA): Es la capacidad de las máquinas para aprender y actuar de forma similar a los humanos.

M

Machine Learning ML: Es una rama de la IA que permite a las máquinas aprender y mejorar sin ser explícitamente programadas.

N

Natural Language Processin (NPL): El Procesamiento de lenguaje natural es una rama de la IA que se ocupa de la lograr interactuar la computadora con un lenguaje natural, incluyendo el análisis, la comprensión y la generación de texto.

INTRODUCCIÓN

En la carrera de TI de la UTMACH, se propuso y desarrolló un innovador prototipo de chatbot basado en procesamiento de lenguaje natural con el propósito de optimizar el acceso a información relevante para los estudiantes. Este proyecto surgió con la meta primordial de agilizar la interacción entre los estudiantes, reduciendo significativamente los tiempos de espera y proporcionando asistencia de manera ininterrumpida

La implementación de la metodología XP (Programación Extrema) y la integración de la API de ChatGPT permitieron alcanzar una interacción fluida y eficiente en el chatbot. Además, se llevó a cabo una iniciativa cuasi-experimental que se focalizó específicamente en la población estudiantil de la carrera. La evaluación de este proyecto se ejecutó mediante la aplicación de la matriz de confusión, permitiendo así determinar la precisión de las respuestas proporcionadas por el chatbot y evaluando su idoneidad. Este enfoque sistemático no solo buscó medir la efectividad del sistema, sino también identificar áreas de mejora para perfeccionar la información recopilada además de ajustar el algoritmo de búsqueda y sus parámetros correspondientes.

i. **Declaración y formulación del Problema**

En la actualidad los chatbots como medio de comunicación con los usuarios y clientes a evolucionado, debido a la facilidad de uso y la disponibilidad; además, de la capacidad para procesar múltiples consultas al mismo tiempo; sin embargo, no todas las organizaciones han incorporado estas nuevas herramientas tecnológicas que ayudan a gestionar el servicio al cliente, ahorrando tiempo de espera y proporcionado información clara y precisa además de reducir la necesidad de mantener personal dedicado a esta función. A continuación, se analizan algunos casos de chatbots.

Según [1], describe que existen chatbots que se han implementado en varias plataformas, pero una de las opciones más populares es Facebook, ya que posee una gran base de usuarios activos diarios. En cuanto a la evaluación, se realizaron pruebas para valorar la funcionalidad del chatbot, arrojando resultados prometedores ya que el chatbot tuvo un alto porcentaje de éxito para las palabras clave predefinidas y pudo responder a la mayoría de los comentarios del público.

En el trabajo [2], Se menciona que, en el entorno educativo, la aparición de tecnologías innovadoras y la expansión de Internet han generado nuevas ideas que potencian la efectividad del sistema de enseñanza al ofrecer a los alumnos acceso inmediato a información sin requerir

desplazamientos, lo que ahorra tiempo y recursos a un costo reducido. En este trabajo se utilizó tecnología cognitiva tipo chatbot con el objetivo de mejorar la experiencia y tiempo de atención en el campo del proceso académico de los estudiantes del Politécnico Colombiano Jaime Isaza Cadavid.

En [3], [4], se mencionó que la implementación de un chatbot va a lograr la obtención de respuestas, reduciendo el tiempo de espera; es así que se puede lograr una aceptación mayoritaria debido a su efectividad y sencillez de su funcionalidad.

Es así que los chatbot no solo se los está utilizando en los ámbitos comerciales, sino que ahora pasamos a los sitios web de instituciones educativas ayudado a la gestión de información. Las técnicas de lenguaje natural han mejorado, siendo factible en los asistentes virtuales [5], [6] sin embargo, la carrera de Tecnologías de la Información, no cuentan con un servicio de chatbots que ayude a gestionar la atención a los estudiantes y aspirantes en línea, ayudando a despejar sus dudas; reduciendo el tiempo de espera hacia dichas preguntas e incluso evitando que los usuarios interesados deban movilizarse hacia las instalaciones. En la figura 1 se detalla las causas, el problema y los efectos.

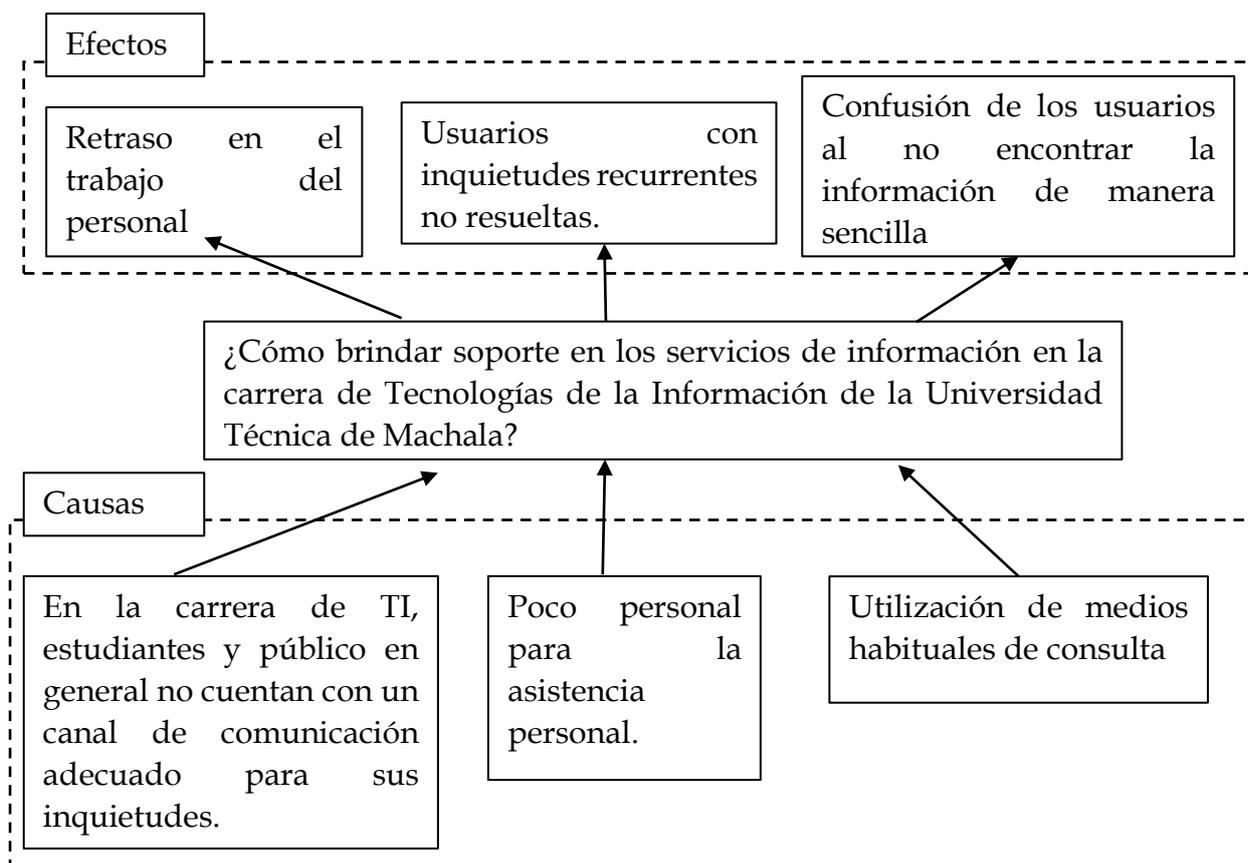


Figura 1. Árbol de problemas

Formulación del problema

- Problema principal
 - ¿Cómo brindar soporte en los servicios de información en la carrera de Tecnologías de la Información de la Universidad Técnica de Machala?
- Problemas específicos:
 - ¿Cuáles son los estudios y trabajos previos relacionados con el diseño y desarrollo de chatbots basados en procesamiento de lenguaje natural para brindar soporte en servicios de información en Instituciones de Educación Superior (IES)?
 - ¿Cuáles son las principales necesidades y requerimientos de los estudiantes de Tecnología de la Información (TI) de la Universidad Técnica de Machala en cuanto a la asesoría y orientación?
 - ¿Cómo se puede desarrollar un prototipo de chatbot que brinde asesoría y orientación a los estudiantes de la carrera de TI, utilizando técnicas de procesamiento de lenguaje natural?
 - ¿Cómo se puede integrar de manera efectiva el prototipo de chatbot en la página de Facebook para la carrera de TI?
 - ¿Cuáles son las métricas e indicadores adecuados para evaluar la precisión del chatbot en la asesoría y orientación de los estudiantes de TI en la página de Facebook de la carrera?

ii. Objeto de estudio y Campo de acción

Objeto de estudio

- Respuestas automatizadas de inquietudes de estudiantes y público sobre información de la carrera de Tecnología de la Información de la Universidad Técnica de Machala.

Campo de acción

- Chatbot con procesamiento del lenguaje natural (NLP)
- Aprendizaje Automático (Machine Learning)
- Modelo de Lenguaje Largo (MLL)

iii. Objetivos

Objetivo General

- Desarrollar un chatbot utilizando técnicas de procesamiento del lenguaje natural para soporte en los servicios de información requeridos por los estudiantes de la carrera de Tecnología de la Información de la Universidad Técnica de Machala y público en general.

Objetivos específicos

- Indagar trabajos relacionados respecto a chatbots y NPL, tecnologías y herramientas para el desarrollo del chatbot.
- Recabar las principales necesidades y requerimientos de la carrera de TI respecto a servicios de información.
- Programar el prototipo de chatbot para soporte en los servicios de información de estudiantes de la carrera de TI y público en general.
- Integrar el prototipo de chatbot a una página de prueba de Facebook.
- Evaluar el chatbot en la página de prueba de Facebook.

iv. Hipótesis y variables o Preguntas de investigación

Hipótesis principal

- La implementación de un chatbot basado en técnicas de procesamiento del lenguaje natural (NLP) para la carrera de Tecnologías de la Información generará respuestas con un 80% de aciertos al aprovechar la capacidad de comprensión contextual y análisis semántico de las consultas.

Variables y dimensionamiento (o categorización)

Tabla 1. Definición de variables y dimensiones

Variable	Categoría	Indicadores	Técnica
Variable independiente: Desarrollo de chatbot basado en MLL para la carrera de TI.	Procesamiento de lenguaje natural	Aprendizaje automático	Programación y adaptación de algoritmos de Modelo de Lenguaje Largo.
		Diseño de la base de datos	Selección de datos vectorial (Pinecone)
	Sistema de chatbot	Desarrollo del prototipo	Metodología XP.
		Programación del Chatbot	Lenguaje de programación Python.
		Estructuración de Data Informativa.	Recopilación de información de portales de la institución.
Variable dependiente: Eficiencia en el acceso de los usuarios a los servicios de información relevante de la carrera de TI.	Precisión de Respuesta.	Contextualización de respuesta	Matriz de confusión
	Pruebas del chatbot.	Rendimiento del chatbot.	Análisis de tiempo de respuesta.
		Facilidad de uso.	

v. **Justificación**

La implementación de nuevas tecnologías ayuda a la gestión de diversas tareas de la organización, una de estas herramientas son los denominados chatbots, los cuales se han implementado para proporcionar respuestas rápidas y efectivas con un público selecto con el objetivo de que el usuario logre mantener más interacción con la organización.

El estudio realizado se enfocó en la integración de un chatbot en un perfil social de Facebook, específicamente en uno de prueba. El objetivo era que este actuara como un asistente para la carrera de Tecnologías de la Información, con el fin de mejorar la eficiencia en la atención al usuario y reducir los tiempos de espera en las solicitudes, ofreciendo asistencia sin restricciones de horario. La metodología de programación extrema respaldó el desarrollo del proyecto, junto con la adopción de técnicas de procesamiento de lenguaje natural mediante la API de ChatGPT para retroalimentar la información, lo que facilitó mantener una interacción fluida y clara.

En conclusión, la creación e implementación del chatbot se planteó con la finalidad de mejorar el tiempo de espera además de la disponibilidad hacia los usuarios, brindando respuestas adecuadas a las preguntas planteadas en el chatbot.

vi. **Organización del documento**

El presente trabajo de titulación se estructura por capítulos los mismo que se detalla ac continuación

Capítulo I: La presente sección se enfoca por entender el contexto de la propuesta donde investigamos los antecedentes mediante la fundamentación bibliográfica.

Capítulo II: Se procede al desarrollo del chatbot, detallando las metodologías de desarrollo.

Capítulo III: Se procede a la evaluación del prototipo del chatbot, para la respectiva incorporación en el perfil social de Facebook de prueba.

CAPÍTULO I. MARCO TEÓRICO

1.1 Antecedentes de la Investigación

En esta sección se aplicó la Metodología de Revisión Sistemática de la Literatura (SRL: Systematic Review of the Literature).

a) Preguntas de investigación

Para el desarrollo de esta sección se detalla en la tabla 2, donde se formula las preguntas de investigación junto a la descripción.

Tabla 2. Preguntas de investigación

Pregunta de investigación	Descripción y motivación
RQ1. ¿Cuáles son los enfoques más utilizados para el diseño e implementación de chatbots basados en el procesamiento de lenguaje natural (PLN)?	Esta pregunta busca identificar los diferentes enfoques y metodologías empleadas en la creación de chatbots que hacen uso de técnicas de procesamiento de lenguaje natural, comprender los enfoques más utilizados ayudará a identificar las tendencias actuales en la implementación de estos sistemas.
RQ2. ¿Cuáles son las funcionalidades específicas que puede ejecutar el chatbot?	El objetivo de esta pregunta es identificar y analizar detalladamente las actividades y tareas particulares que el chatbot debe llevar a cabo.
RQ3. ¿Cuál es el funcionamiento del procesamiento de lenguaje natural en el chatbot?	Esta pregunta tiene como objetivo analizar en detalle el funcionamiento del procesamiento de lenguaje natural implementado en el chatbot
RQ4. ¿Cómo han evolucionado las tecnologías de PLN aplicadas a chatbots en el contexto?	Esta pregunta pretende entender cómo han avanzado las técnicas y herramientas de procesamiento de lenguaje natural durante los últimos cinco años, lo cual puede ofrecer información relevante para diseñar un chatbot actualizado y eficiente.

Tabla 2. Preguntas de investigación

Pregunta de investigación	Descripción y motivación
<p>RQ5. ¿Cuáles son las principales plataformas y frameworks de desarrollo utilizados para la implementación de chatbots basados en PLN?</p>	<p>Esta pregunta tiene como objetivo identificar las plataformas y frameworks más populares y ampliamente utilizados para desarrollar chatbots con capacidades de procesamiento de lenguaje natural.</p>
<p>RQ6. ¿Cuáles son las mejores prácticas para la integración de un chatbot basado en PLN con otros sistemas y plataformas utilizados en una IES?</p>	<p>El objetivo de esta pregunta es identificar las prácticas más efectivas para lograr una integración adecuada del chatbot con los sistemas y plataformas existentes en una institución de educación superior. Una integración exitosa es clave para asegurar un funcionamiento armonioso del chatbot y una experiencia de usuario coherente.</p>
<p>RQ7. ¿Qué métodos de evaluación y métricas se han utilizado para medir el rendimiento y la efectividad de los chatbots basados en PLN en el contexto educativo?</p>	<p>Esta pregunta tiene como objetivo identificar las diferentes métricas y métodos utilizados para evaluar la calidad y el rendimiento de los chatbots en el ámbito educativo. Conocer estas evaluaciones ayudará a determinar qué aspectos son fundamentales para medir el éxito del chatbot y qué áreas se pueden mejorar.</p>

b) Palabras claves y Cadena(s) de búsqueda

Para la búsqueda bibliográfica se establecieron términos principales de investigación.

Cadena de búsqueda en español:

- "CHATBOTS" AND ("Procesamiento de lenguaje natural (NLP)" OR "Aprendizaje automático (ML)" OR "LARGE LANGUAGE MODE") AND "FACEBOOK"

Cadena de búsqueda en inglés

- "CHATBOTS" AND ("NLP" OR "ML" OR "LLM") AND "FACEBOOK"

c) Criterios de inclusión y exclusión

Para el desarrollo investigativo se toma en cuenta fuentes como revistas y artículos científicos los mismos que estén enfocados en la temática de chatbots e inteligencia artificial, desarrollo e implementación, procesamiento de lenguaje natural.

El rango del tiempo de la información debe ser de los últimos cinco años (2019-2023), a nivel de artículos científicos como de fuentes primarias a continuación, en la tabla 3 se detalla los criterios de inclusión y exclusión.

Tabla 3. Criterios de inclusión y exclusión

Criterios de inclusión
<ol style="list-style-type: none">1. Estudios publicados desde el año 2019 en adelante2. Estudios que corresponden a machine learning3. Estudios que abordan procesamiento de lenguaje natural (LNP)4. Estudios que corresponden a chatbots asistentes virtuales.5. Estudios que abordan contexto de la inteligencia artificial6. Artículos en inglés y español
Criterios de exclusión
<ol style="list-style-type: none">1. Artículos cortos (< 3 páginas)2. Estudios duplicados3. Estudios que no se relacionan con las preguntas de investigación.4. Trabajos redundantes

d) Proceso y resultados de la búsqueda

Como se muestra en la **figura 2**, el proceso de búsqueda bibliográfica; se realizó mediante bases de datos conocidas, en las cuales se clasificaron los artículos por los 5 últimos al año de publicación, lo cual debe corresponder al rango del 2019 al 2023, así como otros factores como palabras clave y resumen del mismo.

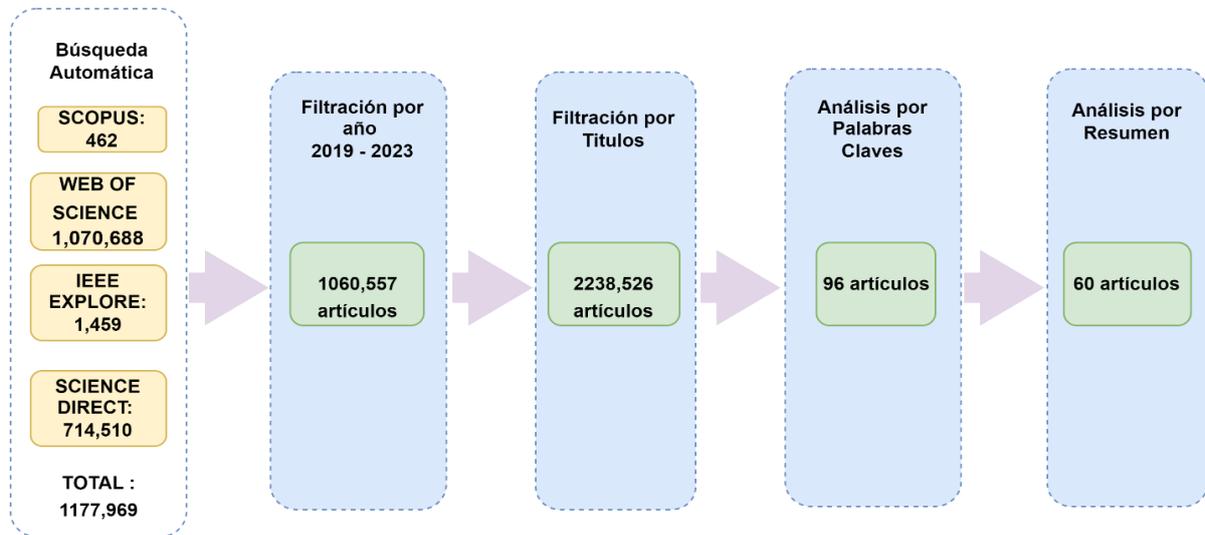


Figura 2. Proceso de resultados de información

De acuerdo con el análisis estadístico de Scopus, **figura 3** sobre la temática de chatbot con lenguaje de procesamiento natural. Se observa con el filtro de los últimos 5 años que los países que destacan son India, Estados Unidos, China y Marruecos.

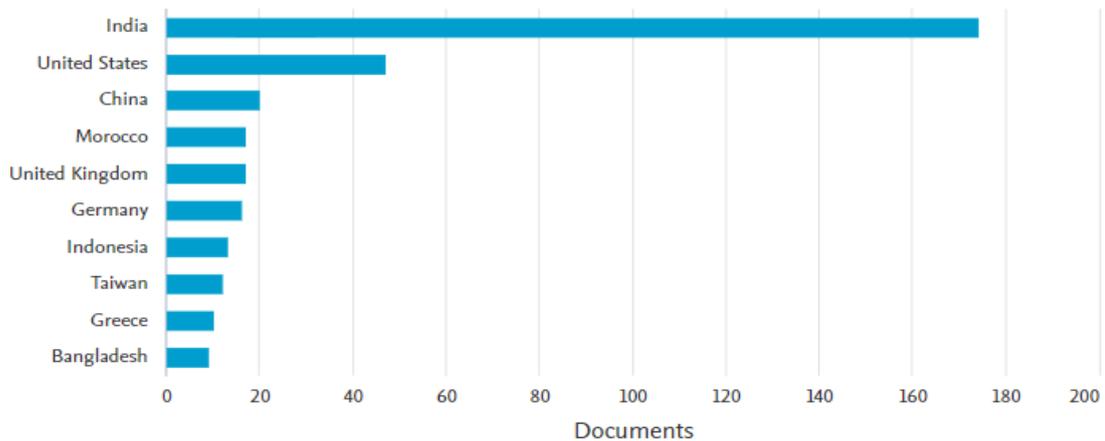


Figura 3. Estadística de países que destacan en publicaciones

Fuente: Análisis tomado de Scopus [7]

Además, los documentos pueden analizarse en función del área de enfoque, donde sobresale las ciencias de la computación con 371 documentos de los últimos 5 años e ingeniería con 195, los mismo que se detalla en la figura 4 tomada de Scopus.

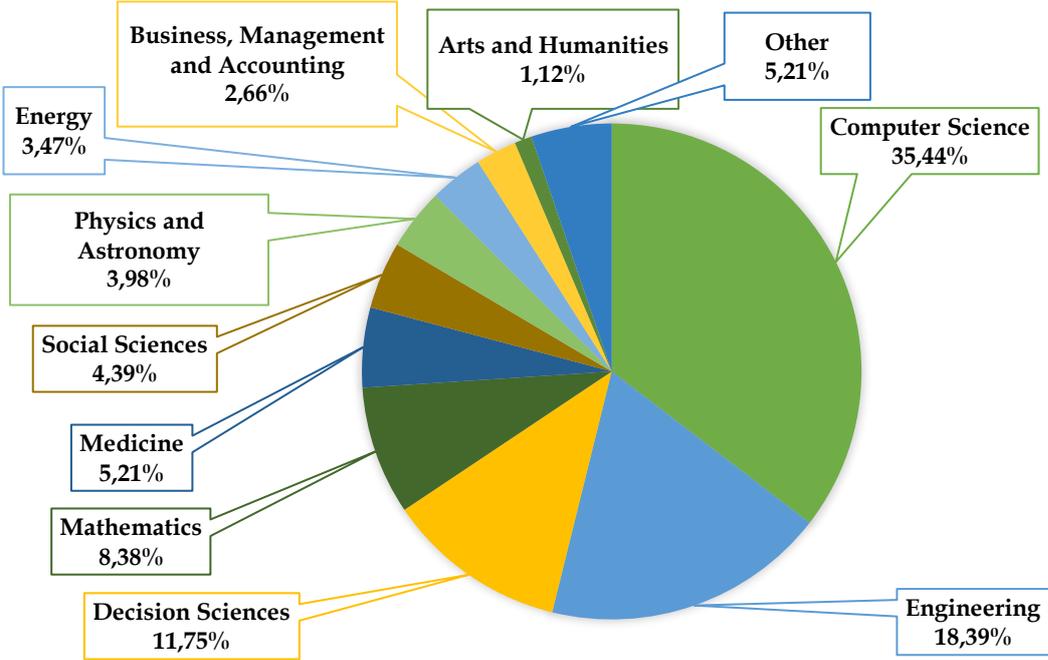


Figura 4. Publicaciones acordes al área de estudio
Fuente: Análisis tomado de Scopus [7]

Con el fin de mejorar la comprensión de la búsqueda, se llevó a cabo un análisis biométrico mediante la aplicación Vosviewer, con el fin de identificar las palabras claves que se mencionan en la figura 5, tales como chatbots, nlp, e-learning y machine-learning

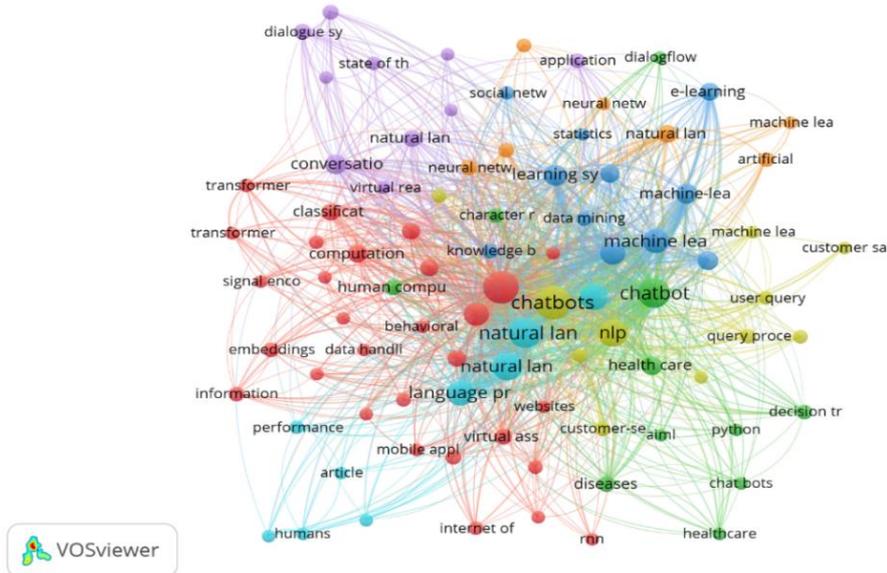


Figura 5. Bibliométrico por palabras claves

1.2 Antecedentes Teóricos

La Inteligencia Artificial (IA) ha evolucionado desde sus inicios, marcados por Alan Turing en 1950, precursor de la informática moderna. En 1971 surgió el Chatterbot, seguido en 1995 por el primer chatbot online, Alice, y en 2001, la revolución de los chatbots con SmarterChild [8], [9].

En la actualidad, los chatbots de texto han experimentado una significativa modernización con la introducción de tecnologías avanzadas como son los Modelos de Lenguaje Largo (LLM), siendo un ejemplo destacado el sistema ChatGPT desarrollado por OpenAI. Hasta el año 2023, Google ha lanzado Bard, una alternativa a ChatGPT [10].

En el ámbito del desarrollo, se ha optado por una metodóloga ágil debido a su estabilidad logrando abarcar a todas las áreas de dominio y en este caso se optó por eXtreme Programming (XP) debido a que demuestra ser altamente adaptable al proceso de construcción de chatbots. Su estructura de diseño y desarrollo se revela particularmente beneficiosa, permitiendo una personalización eficiente. Cabe destacar que esta metodología ha evolucionado con el tiempo, incorporando cambios que facilitan la adaptación precisa a los requisitos específicos de cada proyecto [11], [12].

En términos de implementación, se ha determinado que las empresas han debido adaptarse a los cambios en la Inteligencia Artificial e integrar chatbots en sus sitios web. Esto se debe a que la interacción con el usuario se intensifica y resulta más factible proporcionar respuestas inmediatas en comparación con la espera asociada a un asistente tradicional. Por lo tanto, en el desarrollo de un chatbot se tienen en cuenta características cruciales como la eficiencia, funcionalidad, efectividad, así como la naturaleza humanizada de la interacción. Es fundamental destacar que la interfaz debe ser intuitiva para el usuario [13], [14].

Para mantener una interacción casi humana, se implementan técnicas de procesamiento de Lenguaje Natural, las cuales surgieron con la IA. De esta manera, se definen arquitecturas de desarrollo, siendo la más general aquella que se centra en tres aspectos: el NLP, la secuencia del diálogo y la respuesta obtenida [14].

La creatividad al proporcionar respuestas ha mejorado significativamente, dando como resultado que el aprendizaje automático ha superado la respuesta humana. Es así que la base de conocimientos debe ser adecuada para lograr el éxito en la comprensión del mensaje y proporcionar respuestas acordes a lo solicitado. La experiencia del usuario se convierte en un elemento clave al momento de evaluar, ya que se obtiene el criterio del usuario final [15]. En la

figura 6, se detalla la organización de los antecedentes teóricos investigados mediante la investigación sistemática.

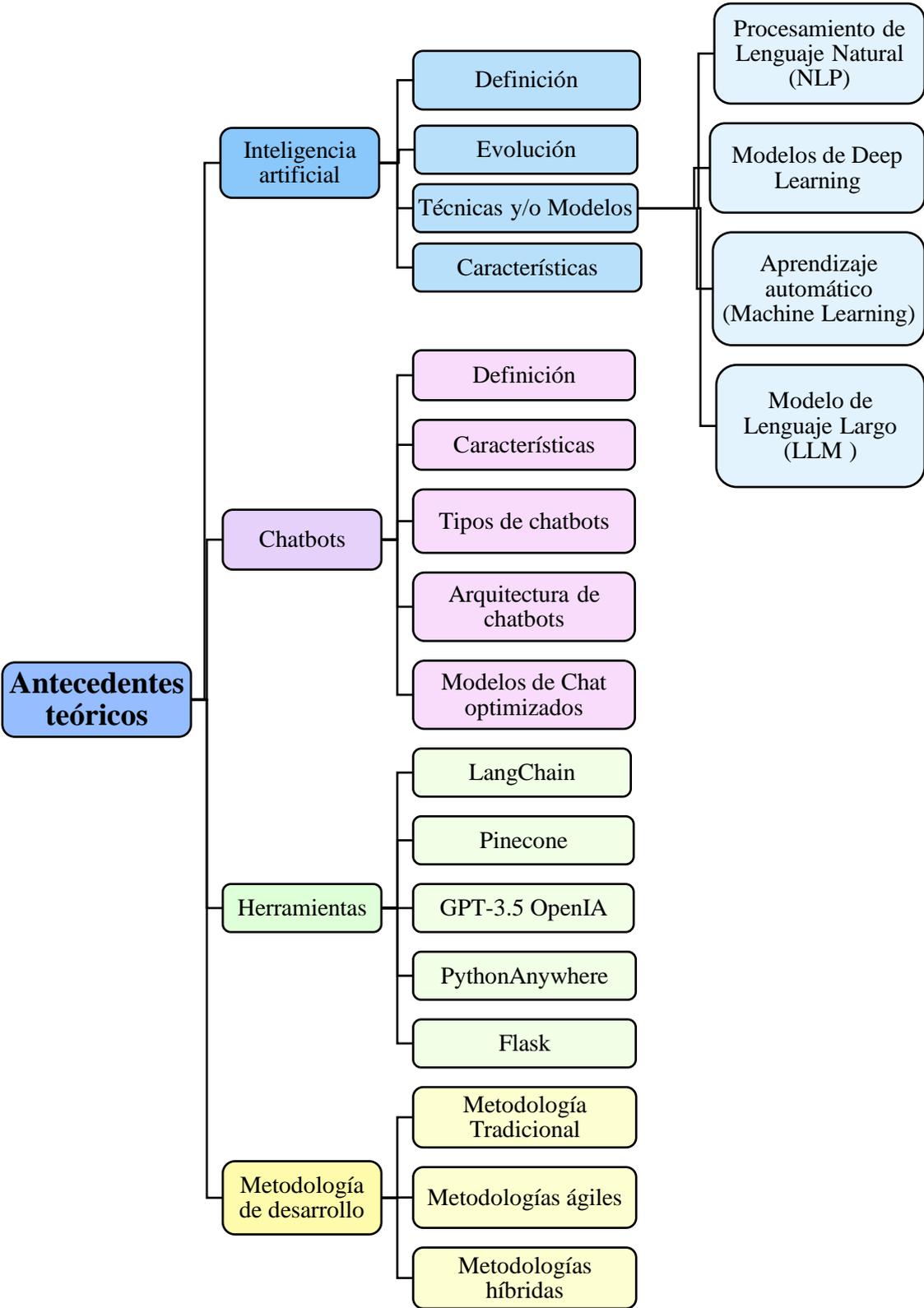


Figura 6. Organigrama del marco teórico

1.2.1 Inteligencia Artificial

1.2.1.1 Definición

La IA se define como sistemas informáticos capaces de realizar tareas que imitan la inteligencia humana [16], estos sistemas informáticos tienen la capacidad para aprender a partir de datos de entrada [17].

1.2.1.2 Evolución

La inteligencia artificial y sus sistemas aparecieron por primera vez en la década de los 1950 [18], oficialmente en 1956 por Marvin Minsky y John McCarthy, se acuñó la palabra Inteligencia Artificial. En 1950 Alan Turing, publicó su artículo “Computing Machinery and Intelligence” donde se detallaba el proceso de construcción de máquinas con capacidad intelectual, centrándose especialmente en la evaluación de su inteligencia. A día de hoy, esta evaluación de Turing sigue siendo un estándar utilizado para determinar la inteligencia de un sistema artificial. Si durante una interacción un ser humano no logra diferenciar entre otro ser humano y una máquina, se concluye que la máquina exhibe inteligencia. [9]. Desde que el concepto de aprendizaje profundo se hizo realidad a mediados de la década de 1960, la IA comenzó a convertirse en algo más que una serie de pruebas contra las que competir la humanidad [19]. El famoso programa informático ELIZA, desarrollado en el período entre 1964 y 1966 en el Instituto Tecnológico de Massachusetts (MIT) por Joseph Weizenbaum, este software representó una herramienta avanzada para el procesamiento de lenguaje natural. Tenía la capacidad de emular diálogos con personas, marcando uno de los primeros intentos de programa en pasar la prueba de Turing. [9].

1.2.1.3 Técnicas

Procesamiento de Lenguaje Natural (NLP)

El procesamiento del Lenguaje Natural (NLP) es un ámbito interdisciplinario que se enfoca en la informática, la inteligencia artificial y la lingüística que explora cómo los ordenadores pueden utilizar el lenguaje natural, inteligencia artificial (IA) y la lingüística, que estudia el uso de los ordenadores para comprender y manipular textos o discursos en lenguaje natural [20].

El procesamiento del lenguaje natural (NLP) ha ganado recientemente mucha atención por representar y analizar computacionalmente el lenguaje humano. Ha extendido sus aplicaciones en varios campos, como la traducción automática, la detección de spam de correo electrónico, la extracción de información, el resumen, la medicina y la respuesta a preguntas, etc [21].

Modelos de Deep Learning

El aprendizaje profundo, como un tipo de técnicas de aprendizaje supervisado, ha adquirido una atención creciente y se está desarrollando mucho en los últimos años [22], se podría definir como un conjunto de capas de procesamiento de datos, esta función está basada en algoritmos que permite que un computador aprenda de estos datos para posteriormente realizar una tarea determinada.

Aprendizaje automático

En lo que respecta al aprendizaje automático supervisado, es decir, redes neuronales artificiales, máquinas de vectores de soporte, clasificadores bayesianos ingenuos, algoritmos de regresión, etc., esos “entornos sociales” pueden entenderse, entre otros, como diferentes estímulos de entrenamiento que configuran el comportamiento de una máquina [23].

Modelo de lenguaje Largo (LLM)

Denominado como Modelo de lenguaje Largo, se basa en los grandes datos para su entrenamiento, siendo el modelo más avanzado del procesamiento de lenguaje natural que a su vez apareció con las nuevas tecnologías como son GTP, Bard y otros chatbots actuales, siendo estas herramientas de apoyo para la investigación y soporte debido a su gran tamaño de información.[24]

Actualmente, estos modelos se los implementa en el PLN, para lo cual se incluyen los chatbots, traductores de idiomas y más debido a su facilidad al momento de comprender el texto y proporcionar respuestas precisas.

1.2.1.4 Características

Las principales áreas de inteligencia artificial son los sistemas profesionales, el procesamiento del lenguaje natural, la comprensión del habla, la robótica y los sistemas sensoriales, la visión del sistema informático y el reconocimiento de escenas, la instrucción asistida por el sistema informático inteligente y la computación neuronal [25]. Dentro de la IA, los términos aprendizaje automático y minería de datos también son de uso mucho más común hoy en día, esto refleja los avances tecnológicos. El término de aprendizaje automático habitualmente se consideraba una expresión muy técnica, pero gracias a la difusión en noticias como la del sistema AlphaGo, la del sistema Watson de IBM, ha contribuido mucho a renovar su imagen [18].

1.2.2 Chatbots

1.2.2.1 Definición

Los chatbots son programas automatizados que se utilizan para comunicarse con humanos a través de intercambio de mensajes de texto o chat [26], en el trabajo [27], se conceptualizan los chatbots como asesores automatizados que pueden facilitar la toma de decisiones, ya que según [28], un chatbot puede emplear la lengua del ser humano para lograr una comunicación eficiente con el mismo.

En la revisión literaria encontramos otra definición como lo menciona el trabajo [29], donde se define un chatbot como una aplicación de software diseñado para mantener interacciones naturales y conversacionales con usuarios, abordando temas específicos o dominios particulares, utilizando tanto texto como voz.

1.2.2.2 Características

Según se menciona en [30], [31], los chatbots poseen las siguientes características.

- Modelo de autoaprendizaje: Poseen la capacidad de crear un modelo de autoaprendizaje a través de programas informáticos y cálculos matemáticos. Esto les permite mejorar su rendimiento y eficiencia a medida que interactúan más con los usuarios y reciben comentarios.
- Respuestas en tiempo real: Brindan respuestas y otra información relevante lo más cercana posible a las preguntas de los usuarios en tiempo real. Esta capacidad de respuesta instantánea es valiosa para ofrecer una experiencia rápida y satisfactoria.
- Integración en plataformas: Los chatbots han sido integrados en diferentes plataformas, incluyendo redes sociales permitiéndoles interactuar con los usuarios de manera más conveniente y accesible.
- Uso en organizaciones e instituciones educativas: Los chatbots han experimentado un aumento en los procesos de organizaciones e instituciones educativas ya que poseen la capacidad para brindar asistencia y respuestas eficientes a preguntas frecuentes.

1.2.2.3 Tipos de chatbots

De acuerdo con [32], la categorización de los chatbots se fundamenta en las siguientes variables: ámbito de conocimiento, propósito, función prestada, grado de automatización, estilo de comunicación, plataforma de creación y manejo de respuestas.

Desde la posición de [33], propone extender la siguiente clasificación de chatbots:

- **Dominio del conocimiento:** Estos aspectos consideran la información a la que un chatbot puede acceder, pudiendo ser de conocimiento amplio en el caso de tópicos generales, o de conocimiento limitado en el caso de enfoques especializados en un ámbito particular, lo que se denomina dominio cerrado.
- **Objetivo del chatbot:** Según su propósito central, los chatbots pueden ser clasificados en categorías informativas, orientadas al diálogo y centradas en tareas específicas. Aquellos de índole informativa suministran información al usuario, como ejemplifican los chatbots de preguntas frecuentes. En contraste, los chatbots enfocados en el diálogo se esfuerzan por establecer comunicación natural con el usuario y responder de manera precisa a sus inputs. Por último, los chatbots centrados en tareas específicas ejecutan una labor particular, como la reserva de un servicio.
- **Servicio provisto:** Se expone la relación del chatbot con el usuario desde una perspectiva emocional, la cual está determinada por la tarea que realiza el chatbot. En este sentido, se identifican distintos tipos de relaciones: interpersonales, que buscan obtener y brindar información al usuario sin retener sus datos; intrapersonales, capaces de desarrollar una conexión cercana con el usuario al aprender sus preferencias, a menudo presentes en aplicaciones de mensajería; por último, se encuentran los interagentes, siempre activos y necesitados de comunicación con otros asistentes inteligentes, siendo ejemplos notables Alexa y Cortana.
- **Nivel de automatización:** Se trata del grado de participación humana en la interacción del chatbot durante una conversación. Aquí se presentan dos enfoques: el semiautomatizado, en el que el humano interviene en ciertos momentos de la conversación, y el segundo es el completamente automatizado, donde no se requiere intervención humana.
- **Forma de comunicación:** Dentro de esta categoría de chatbots, se hace referencia a la forma en que el usuario introduce información para interactuar con el sistema. Esta clasificación abarca la comunicación a través de texto o mediante el uso de la voz.

- Plataforma de desarrollo: Esta característica posibilita la categorización de los chatbots en función de los permisos concedidos por la plataforma en la que han sido creados. Estas clasificaciones pueden ser de acceso abierto, tal como ocurre con RASA, o restringido, que hace referencia a plataformas comerciales desarrolladas por grandes empresas como Google o IBM, donde el código fuente no se encuentra disponible.

1.2.2.4 Arquitectura de chatbots

La arquitectura de un chatbot puede variar según su complejidad y funcionalidades requeridas dependiendo también del tipo de chatbot y su objetivo.

El trabajo [34], se menciona una arquitectura basada en microservicios, la cual está destinada a proporcionar nuevos servicios y funcionalidades para las necesidades cambiantes de los usuarios, este tipo de arquitectura emplea la modularidad lo que permite que se pueda construir un nuevo microservicio de forma independiente y sin necesidad de modificar el chatbot desarrollado. La figura 7 ilustra dicha arquitectura en detalle

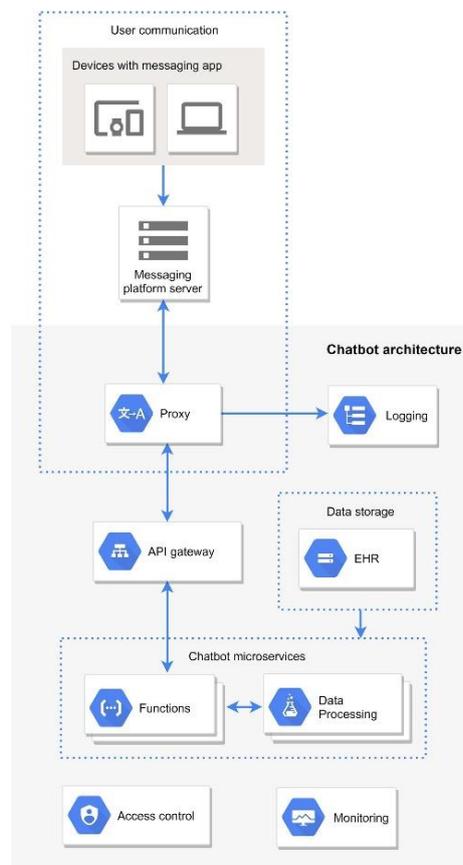


Figura 7. Ejemplo de Arquitectura basada en microservicio.

Fuente: Figura tomada de [34]

1.2.2.5 Modelos de Chat Optimizados

El procesamiento de información optimizado en los chatbots se basa en el Modelo de Lenguaje Largo, también conocido como LLM. Su amplia base de conocimientos abarca una variedad de campos, lo que le permite proporcionar respuestas creativas y versátiles [35].

Entre los chatbots más reconocidos se encuentran ChatGPT 3.5 y Google Bard, ambos ofreciendo opciones gratuitas, así como el modelo ChatGPT-4, disponible como versión de pago. Sin embargo, la pregunta sobre cuál es más óptimo depende del enfoque que se le dé. Un estudio oftalmológico reveló que la versión GPT-4 superó significativamente a ChatGPT-3.5 y Google Bard, obteniendo un 70% de calificaciones positivas, en comparación con el 40% y 45%, respectivamente [36].

En cuanto a otra opción gratuita, tenemos a LLAMA-2, propiedad de Meta, pero su integración requiere más herramientas para personalizar el entrenamiento y tiempo para alcanzar un rendimiento comparable con GPT3.5. Además, se necesita acceso a GPU para facilitar el proceso de entrenamiento [37].

Explorando el modelo de código abierto de Meta, LLAMA2 de 70B, y comparándolo con otros modelos cerrados desde una perspectiva académica, se observa que su arquitectura está cerca de GPT-3.5 en términos de comprensión masiva del lenguaje multitarea y razonamiento. A pesar de ello, persiste una brecha de rendimiento significativa entre LLAMA2 70B y GPT-4, así como PaLM-2-L [38], [39]. Tal y como mencionan en [35], los resultados de esta comparación se detallan en la Tabla 4, presentada a continuación.

Tabla 4. Comparación con modelos de código cerrado

Fuente: Tabla Obtenida de [38]

Benchmark (shots)	GPT-3.5	GPT-4	PaLM	PaLM-2-L	Llama 2
MMLU (5-shot)	70.0	86.4	69.3	73.3	68.9
TriviaQA (1-shot)	-	-	81.4	86.1	85.0
Natural Questions (1-shot)	-	-	29.3	37.5	33.0
HumanEval (0-shot)	48.1	67.0	26.2	-	29.9

De acuerdo a [40], [41], ChatGPT muestra un rendimiento superior a Llama 2-Chat 70B en lo que respecta a la asistencia lingüística. Por otro lado, Llama 2-Chat 70B supera a ChatGPT en preguntas relacionadas con hechos concretos es así como Llama 2 lo consideran un potencial software gracias a la capacidad de generar texto y la coherencia con la que ofrece las respuestas.

Además, cabe resaltar que ambos modelos dan respuestas correctas, pero la evaluación de sus resultados también se fundamenta en cómo se redactó la respuesta. A continuación, se presenta la tabla 5, comparativa entre LLAMA2 y GPT-3.5 Turbo.

Tabla 5. Comparativa LLAMA2 vs GPT-3.5

	LLAMA2	GPT 3.5 turbo
Modelo	Código abierto.	Código cerrado.
Arquitectura	LLM de modalidad única entrada de texto.	LLM de modalidad única entrada de texto.
Parámetros	<ul style="list-style-type: none"> • 7 mil millones de parámetros • 13 mil millones de parámetros • 70 mil millones de parámetros 	175 mil millones de parámetros
Compresión Masiva del Lenguaje Multitarea puntuación	68.9	70.0
Modalidad	Texto	Texto
Longitud máxima del contexto	4,096 tokens	4,096 tokens

1.2.3 Herramientas

1.2.3.1 LangChain

LangChain es un marco de desarrollo, o también denominado Framework, que facilita la creación de aplicaciones que interactúan con usuarios mediante modelos de lenguaje avanzados. Su propósito es impulsar el desarrollo eficiente de aplicaciones habilitadas para procesar el lenguaje natural [37]. Para lo cual provee componentes que viene a ser abstracciones modulares y cadenas personalizables definidas para cada asunto.

1.2.3.2 Pinecone

Pinecone es una base de datos vectorial que trabaja en la nube la misma que es implementada para la integración del modelo de lenguaje grande LLM de OpenAI. Es así que Pinecone es implementada

para la búsqueda semántica de incrustaciones vectorial obteniendo resultados más precisos es así que al integrarla los LLM de OpenAI se está combinando aprendizaje profundo para las respectivas incrustaciones [42].

1.2.3.3 GPT-3.5 OpenIA

Conocida como una herramienta para el procesamiento de información basado en aprendizaje automático y profundo, nos ayuda a entablar una conversación y generar textos como instructivos o traducciones acordes a lo que se le solicite al funcionar como un asistente inteligente [43].

También interpretada como la interfaz de programación o api, permite la comunicación de ChatGPT, este modelo permite comprender y generar lenguaje natural, además de no ser determinista es decir acorde al ajuste de la temperatura nos llega a variar la salida con pequeñas variantes [44].

1.2.3.4 Python

Python se distingue como un lenguaje de programación poderoso y de fácil asimilación. Se caracteriza por su eficiente manejo de estructuras de datos de alto nivel, cuenta con una sintaxis dinámica e interpretativa ideal para secuencias de comandos y desarrollo ágil de aplicativos con gran variedad de plataformas y dominios [45].

1.2.3.5 PythonAnywhere

Se trata de una plataforma que nos permite la ejecución de aplicaciones en línea, gracias al hosting que nos proporciona es así como este servicio web facilita la ejecución de la aplicación desde cualquier dispositivo, permitiendo el trabajo en conjunto con el desarrollo [46].

1.2.3.6 Flask

Flask también conocido como un microframework está basado en Python, lo cual es utilizado para el desarrollo de aplicaciones web, se lo caracteriza por ser ligero además de las pruebas unitarias, soporte de extensiones y la compatibilidad con WSGI (Web Server Gateway Interface) siendo el responsable de mantener la compatibilidad con el servidor web [47], [48].

1.2.4 Metodología de desarrollo

1.2.4.1 Metodología tradicional

La metodología tradicional nos ayuda al desarrollo del proyecto donde se seguirán etapas de análisis y diseño para garantizar resultados de calidad.

Por lo general se lo implementa en proyectos de grandes dimensiones donde no se producirá cambios debido a que se integra de procesos rígidos que demanda tiempos extensos de producción que llevará poca iteración con el cliente.

Entre las conocidas podemos nombrar la metodología cascada es un modelo lineal donde se realiza un seguimiento desde inicio a fin en la que se incluyen las etapas de análisis, diseño, construcción, pruebas, implementación y mantenimiento [49].

1.2.4.2 Metodologías ágiles

En la metodología ágil se destaca la flexibilidad para el manejo de proyectos, por la división de subprocesos en la que mantenemos constante comunicación para determinar cambios y poder adaptarlos, así que podemos comparar como un bucle donde se debe pasar por iteraciones para el análisis y retroalimentación de procesos. [49], [50].

Ayuda a reducir el riesgo de no culminar proyectos [47], porque se basa en entregables; divide el sistema en módulos o subsistemas para que el cliente pueda usar o probar estos módulos sin tener el sistema completo.

1.2.4.3 Metodología híbrida

Las metodologías híbridas se caracterizan por la integración de las dos metodologías, ágiles y tradicionales, lo que puede aplicarse total o parcialmente [48], por lo que el empleo de las híbridas se las determina como buena alternativa para lograr el objetivo del proyecto integrando procesos acordes a la evolución considerando el alcance, complejidad y restricciones [45].

Metodología XP para el desarrollo del chatbot

La metodología XP, la conocemos por ser de las más disciplinadas debido a su sistema de revisión donde realizamos pruebas frecuentes manteniendo la participación de los clientes con el objetivo de descubrir problemas del sistema [51].

En cuanto a las partes interesadas mencionamos a las siguientes [52], [53]:

- Clientes: encargados de definir los objetivos del sistema, se puede tratar de gerentes, directores financieros o marketing, etc.
- Usuarios: puede ser el público general o el personal administrativo, etc.
- Desarrolladores: aquí hablamos de los programadores, o medios de comunicación que tengan interés.
- Testers: encargado de las pruebas del sistema.
- Tracker: responsable de los puntos de control acorde a la planificación para mantener un control del sistema.
- Coach: encargado de mantener una orientación con los equipos de trabajo siendo su principal objetivo ser la guía del proyecto.
- Manager: es el coordinador del sistema además de impartir las ideas para el desarrollo.

Fases de la Metodología XP

Para cumplir los procesos se siguen una serie de etapas que inicia la planificación, que se revisará cada 2 semanas, luego el diseño donde se utiliza tarjetas CRC (Clase -Responsabilidad-colaboración), continuamos con la codificación, para luego pasar a las pruebas y al final con el lanzamiento cumpliendo las solicitudes del cliente y planteando un software útil incorporado a la versión principal [51]. En la figura 8 se presentan las fases de la metodología previamente mencionada.

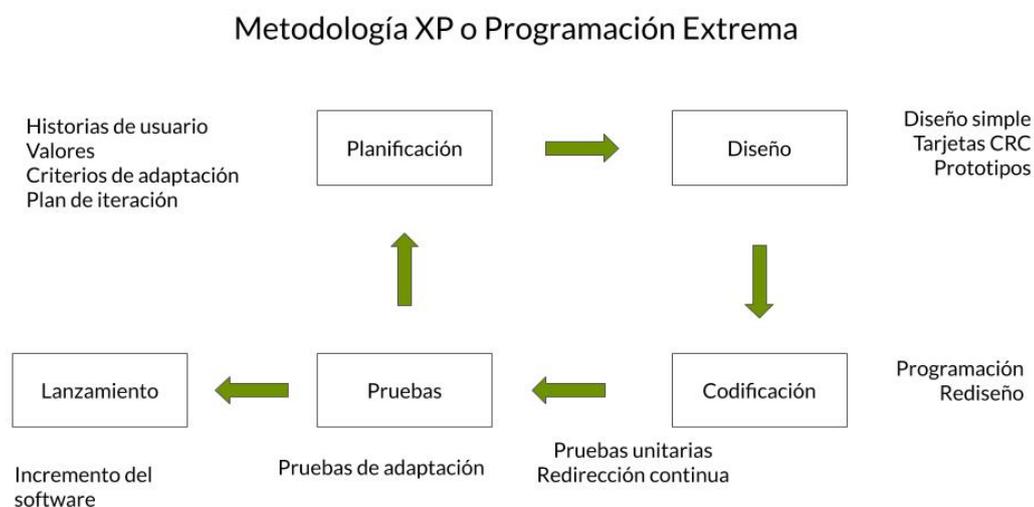


Figura 8. Etapas del método Extreme Programming XP

Fuente: [53].

1.2.4.4 Antecedentes Contextuales

Se desarrolla un chatbot para un perfil social de prueba que simule ser de la carrera de Tecnologías de la Información (TI) de la Universidad Técnica de Machala, situada en Machala, para que ayude a resolver inquietudes que se puede originar de parte de usuarios externos como de los mismos estudiantes de la carrera; el cual se fundamentará en técnicas de procesamiento de lenguaje natural.

En relación con la información, se estiman información como pasos para matrícula, justificación de inasistencia u otros, así como el proceso de admisión a la carrera, los horarios de atención presencial y las materias que se imparten, considerando que la carrera de TI tiene como visión la formación de profesionales para desarrollar soluciones tecnológicas.

1.2.4.5 Ámbito de aplicación

El chatbot personalizado desarrollado en este proyecto se enmarca en la carrera de Ingeniería en Tecnologías de la Información de la UTMACH. Este chatbot tiene como objetivo mejorar la experiencia informativa de los estudiantes al proporcionar respuestas pertinentes y en tiempo real sobre diversas áreas de interés. Su alcance abarca desde ofrecer detalles sobre programas académicos y demás información referente a la carrera, el chatbot actuará como una fuente de información para mantener a los estudiantes informados.

Esta solución también busca abordar el desafío de los servicios de información fragmentados y limitados que enfrenta actualmente la universidad. Mediante la integración del chatbot en la página de prueba de Facebook, se establecerá un punto centralizado para acceder a información confiable y actualizada. Impulsando la mejora constante. En última instancia, este chatbot personalizado no solo empodera a los estudiantes con información crucial para su progreso académico, sino que también descarga la carga administrativa al ofrecer respuestas ágiles a preguntas comunes, liberando recursos humanos para trabajos más complejos

1.2.4.6 Establecimiento de requerimientos

Requisitos Funcionales

Para las consultas planteadas, se han establecido los siguientes requisitos funcionales, considerándolos como información de uso cotidiano que podría llevar a cabo un usuario. A continuación, en la Tabla 6, se proporcionan detalles específicos al respecto.

Tabla 6. Requisitos Funcionales

#	Requerimientos	Descripción
RF1	Consultar asignaturas por semestre	Enlista las asignaturas ofrecidas para cada semestre específico.
RF2	Consulta el proceso para ingreso a la carrera	Detalla el proceso de inscripción que debe cumplir todo postulante.
RF3	Consulta el proceso de justificación de faltas.	Detalla el procedimientos y requisitos para justificar ausencias a clases
RF4	Consultas sobre el proceso de pago de matriculas	Ofrece detalles sobre cómo realizar el pago de matrículas, métodos de pago aceptados y # de cuenta.
RF5	Consulta del actual coordinador de la carrera	Proporciona el nombre del coordinador actual de la carrera.
RF6	Consulta la planta docente	Lista de profesores que forman parte de la planta docente titular de la carrera
RF7	Consulta la ubicación y a que Facultad pertenece la carrea.	Indica la ubicación física de la carrera dentro de la institución educativa y especifica a qué Facultad pertenece

Requisitos no funcionales

Para la creación del chatbot, se inicia el proceso definiendo los requisitos no funcionales que se encuentran detallados en la Tabla 7. Estos requisitos serán tenidos en cuenta a lo largo del desarrollo para garantizar la implementación efectiva del chatbot.

Tabla 7. Requisitos No Funcionales

#	Requerimientos	Descripción
RNF1	Base de datos vectorial.	Almacenamiento de los datos para la alimentación del chatbot.
RNF2	Personalización de código de entrenamiento en Python.	Codificación mediante la implementación del API ChatGPT.
RNF3	Integración IDE de desarrollo.	Integrar la codificación de bloques a un IDE de desarrollo.
RNF4	Configuración hacia un servidor web.	Disponibilidad de la conexión del servidor.
RNF5	Integración a la red social	Integrar la codificación a Facebook Messenger.
RNF6	Desarrollo de la interfaz para un módulo administrador.	Diseño del Interfaz grafico del módulo administrador que permitirá actualizar la base de datos.
RNF7	Desarrollo del interfaz login del módulo para actualizar información.	Diseño del Interfaz grafico del login para la sección administrador que permitirá actualizar la base de datos.
RNF8	Codificación del módulo administrador y login.	Realizar el algoritmo para el login de acceso y para la sección de actualizar la base de datos.
RNF9	Disponibilidad del chatbot.	Posibilitar que el chatbot se mantenga disponible para el usuario, sin restricción de horario.
RNF10	Bot por texto.	Entrada y salida de los datos por texto, no personalizado para voz.

A continuación, se detalla características que no cumplirá el chatbot.

- El chatbot no contara con la incorporación de audio Bot.
- La entrada y salida de información solo será mediante mensaje textual.

CAPÍTULO II. DESARROLLO DEL PROTOTIPO

2.1 Definición del prototipo

El prototipo de chatbot personalizado para la Carrera de Ingeniería en Tecnologías de la Información es una aplicación interactiva basada en la API de ChatGPT. Esta elección se basa en una investigación comparativa entre diversas APIs de aprendizaje automático, destacando la versión 3.5 Turbo de ChatGPT por su popularidad y extenso entrenamiento, la cual nos permite personalizar y adaptar para satisfacer las necesidades específicas de los estudiantes de la carrera. Aprovechando la tecnología avanzada de procesamiento de lenguaje natural, este chatbot tiene como objetivo ofrecer respuestas contextualmente relevantes y en tiempo real a las preguntas y consultas de los estudiantes, optimizando la accesibilidad y la calidad de la información académica.

Además, el prototipo del chatbot se integra de manera fluida en la página de prueba, para realizar la evaluación con la matriz de confusión. La integración en Facebook permite una mayor accesibilidad y conveniencia.

2.2 Metodología de desarrollo del prototipo

Como metodología de desarrollo se implementa XP, debido a su acoplamiento al tiempo de desarrollo y las fases de desarrollo del proyecto.

2.2.1 Enfoque, alcance y diseño de investigación

El presente trabajo está basado en un enfoque cuantitativo, donde tenemos como característica el diseño y desarrollo de un chatbot utilizando técnicas de procesamiento de lenguaje natural para brindar soporte en los servicios de información de la carrera de Ingeniería en Tecnología de la Información de la Universidad Técnica de Machala mejorando la accesibilidad de los usuarios a la información con respuestas precisas y rápidas.

En cuanto al alcance lo determinamos como descriptivo porque el objetivo de los chatbots es responder preguntas, para luego evolucionar a explicativo porque proporcionará respuestas más contextualizadas gracias a las técnicas de procesamiento de lenguaje natural.

Finalmente, el diseño de la investigación se establece como cuasi-experimental debido a que la herramienta está destinada a un grupo específico de usuarios que son los estudiantes de la carrera.

2.2.2 Unidades de análisis

Población

Para este trabajo se seleccionó como población a los estudiantes de Tecnologías de la Información de la UTMACH tabla 8, se consideran los estudiantes desde primer hasta décimo semestre, que se encuentran legalmente matriculados.

Tabla 8. Población

Estudiantes de Tecnologías de la Información de la UTMACH	400
--	-----

Muestra

Para obtener la muestra se empleó la herramienta web SurveyMonkey [52], la cual calculó, a partir de una población de 400 estudiantes, un tamaño de muestra de 197 con un margen de error del 5% y un nivel de confianza del 95%. Esta herramienta emplea la fórmula para la muestra en una población finita.

$$n = \frac{N * p * (1 - p)}{N * p * (1 - p) + Z^2 * \frac{E^2}{4}}$$

donde:

- n es el tamaño de la muestra,
- N es el tamaño de la población,
- p es la proporción estimada del atributo en la población (usaremos 0.5 para obtener el tamaño máximo de muestra).
- E es el margen de error (0.05 para un 5%),
- Z es el valor crítico z para el nivel de confianza del 95%.

2.2.3 Técnicas e instrumentos de recopilación de datos

A continuación, se detallan las técnicas e instrumentos de recolección de datos utilizados en el estudio, tal como se presenta en la Tabla 9.

Tabla 9 Técnica de recolección de datos.

Técnica	Instrumento
Herramienta Matriz de confusión.	Cuestionario de preguntas
Análisis de documentaciones o datos o procesos	Guía de análisis de documentos o de dato Lista de verificación.

2.2.4 Técnicas de procesamiento de datos para la obtención de resultados

En relación con al enfoque, alcance y modelo de investigación, se empleará el modelo de lenguaje largo que forma parte de las técnicas de procesamiento de lenguaje natural (NPL) para comprender las solicitudes que pueden plantearse y, de esta manera, entrenar el chatbot, además de utilizar estadísticas básicas para procesar encuestas de usuarios

2.2.5 Metodología o métodos específicos

Se ha optado por aplicar el método Extreme Programming XP para desarrollar el chatbot, dado su enfoque en la colaboración, las pruebas continuas y la entrega de valor incremental, lo cual es adecuado para un chatbot ágil e interactivo.

2.2.6 Herramientas y/o Materiales

En la siguiente tabla 10 se establecen las herramientas que serán empleadas para el desarrollo del proyecto:

Tabla 10 Herramientas y Materiales

Herramientas de software	<ul style="list-style-type: none">• Visual studio code• Git• GitHub• Facebook Graph API• Api ChatGpt• PythonAnywhere• Flask• Pinecone• Langchain• Jupyter
Herramientas de hardware	<ul style="list-style-type: none">• Laptop• Pc de escritorio
Lenguajes de programación	<ul style="list-style-type: none">• Python

2.3 Desarrollo del prototipo

El desarrollo del prototipo se basa en la metodología XP, el mismo que se alinea al desarrollo del mismo para la segmentación de información y entrenamiento del chatbot.

2.3.1 Metodología XP

2.3.1.1 FASE I: Planificación

Definición de roles

El desarrollo del prototipo se estableció acorde a diferentes roles, detallados a continuación en la tabla 11.

Tabla 11. Roles y Responsables

Rol	Responsable
Desarrollador	Romero Jeanina, Velez Tania
Analista	Romero Jeanina, Velez Tania
Tester	Romero Jeanina, Velez Tania
Cliente	Usuarios de TI de la UTMACH

Historias de Usuario

Se detalla los requerimientos no funcionales, es así como se especifica las historias de usuario siendo así la guía para el correcto desarrollo del chatbot.

Tabla 12. Historia de Usuario N°1 - Estructura de datos a utilizar

Historia de Usuario	
Número: 1	
Nombre de Historia: Estructuración de datos a utilizar.	Usuario: Desarrollador y Analista
Prioridad en Negocio: Alta	Iteración Asignada: 1
Programador Responsable: Velez Tania, Romero Jeanina	
Descripción: Desarrollar estructura la información a utilizar para el chatbot, tomando la información disponible en las páginas de la UTMACH.	
Observación: Ninguna	

Tabla 13. Fragmentar de Datos

Historia de Usuario	
Número: 2	
Nombre de Historia: Fragmentar datos	Usuario: Desarrollador y Analista
Prioridad en Negocio: Alta	Iteración Asignada: 2
Programador Responsable: Velez Tania, Romero Jeanina	
Descripción: Fragmentar la información del documento estructura determinado la dimensión del texto.	
Observación: Ninguna	

Tabla 14. Integración al IDE

Historia de Usuario	
Número: 3	
Nombre de Historia: Integración al IDE de desarrollo	Usuario/os: Desarrollador y Analista
Prioridad en Negocio: Alta	Iteración Asignada: 3
Programador Responsable: Velez Tania, Romero Jeanina	
Descripción: Luego de obtener pruebas de desarrollo por bloque integrar el código a un IDE de desarrollo para la estructuración de las capas de desarrollo del chatbot.	
Observación: Ninguna	

Tabla 15. Creación y carga de vectores

Historia de Usuario	
Número: 4	
Nombre de Historia: Creación y carga de vectores.	Usuario: Desarrollador y Analista
Prioridad en Negocio: Alta	Iteración Asignada: 4
Programador Responsable: Velez Tania, Romero Jeanina	
Descripción: Se realiza la creación del índice en Pinecone para la correspondiente creación y carga de los vectores	
Observación:	

Tabla 16. Host y servidor web

Historia de Usuario	
Número: 5	
Nombre de Historia: Host y servidor web	Usuario/os: Desarrollador y Analista
Prioridad en Negocio: Alta	Iteración Asignada: 6
Programador Responsable: Velez Tania, Romero Jeanina	
Descripción: Integrar el chatbot local a un servidor compatible con Python	
Observación: Tomar en cuenta que al ejecutar no tener errores sobre la admisión a producción.	

Tabla 17. Generando Consultas

Historia de Usuario	
Número: 6	
Nombre de Historia: Generando Consultas	Usuario: Desarrollador y Analista
Prioridad en Negocio: Alta	Iteración Asignada: 5
Programador Responsable: Velez Tania, Romero Jeanina	
Descripción: Se realiza la codificación para poder realizar las consultas además de mantener memoria en el chat para no perder el hilo de la conversación.	
Observación: Ninguna	

Tabla 18. Ajuste algoritmo de búsqueda

Historia de Usuario	
Número: 7	
Nombre de Historia: Ajuste algoritmo de búsqueda	Usuario/os: Desarrollador y Analista
Prioridad en Negocio: Alta	Iteración Asignada: 7
Programador Responsable: Velez Tania, Romero Jeanina	
Descripción: Ajustar la temperatura para optimizar el tiempo y la interpretación de las respuestas.	
Observación: Ninguna	

Tabla 19. Depuración de Data

Historia de Usuario	
Número: 7	
Nombre de Historia: Optimización de Data	Usuario/os: Desarrollador y Analista
Prioridad en Negocio: Alta	Iteración Asignada: 8
Programador Responsable: Velez Tania, Romero Jeanina	
Descripción: Mejorar la información evitando redundancia de palabras.	
Observación: Actualizar la información sobre el actual coordinador	

Tabla 20. Integración a la red social

Historia de Usuario	
Número: 8	
Nombre de Historia: Integración a la red social	Usuario: Desarrollador y Analista
Prioridad en Negocio: Alta	Iteración Asignada: 9
Programador Responsable: Velez Tania, Romero Jeanina	
Descripción: Una vez realizada la estructuración por capas se procede a integrar a la red social establecida.	
Observación: Ninguna	

Tabla 21. Prueba del chatbot en la red social integrado.

Historia de Usuario	
Número: 9	
Nombre de Historia: Prueba del chatbot integrado	Usuario: Desarrollador y Analista
Prioridad en Negocio: Alta	Iteración Asignada:10
Programador Responsable: Velez Tania, Romero Jeanina	
Descripción: Realizar las respectivas consultas al chatbot integrado en el perfil social y determinar que este correctamente funcionando.	
Observación: Ninguna	

Tabla 22. Desarrollo de modulo para actualizar la data

Historia de Usuario	
Número: 10	
Nombre de Historia: Modulo para actualizar la data	Usuario: Desarrollador y Analista
Prioridad en Negocio: Alta	Iteración Asignada: 11
Programador Responsable: Velez Tania, Romero Jeanina	
Descripción: Desarrollar un módulo administrador que permita actualizar la data.	
Observación: Ninguna	

Tabla 23. Aplicación de encuesta

Historia de Usuario	
Número: 11	
Nombre de Historia: Aplicación de matriz de confusión	Usuario: Desarrollador y Analista
Prioridad en Negocio: Alta	Iteración Asignada: 12
Programador Responsable: Velez Tania, Romero Jeanina	
Descripción: Aplicación de matriz de confusión acorde a los criterios que se establezcan.	
Observación: Ninguna	

Análisis del sistema

Plan de publicación. El desarrollo del prototipo se realiza junto con las historias de usuario, lo que permite definir la duración y las tareas específicas para cada etapa. Los detalles de planificación se presentan en la Tabla 24, asegurando una coordinación efectiva entre las funcionalidades requeridas y el tiempo asignado para su implementación.

Tabla 24. Plan de publicaciones

N°	Historia de usuario	Tarea	Prioridad	Iteración
1	Estructuración de datos a utilizar.	Selección de información	Alta	1
2	Fragmentar datos	Desarrollo de función para establecer la dimensión adecuada del texto en el chatbot.	Alta	2
3	Integración al IDE	Integración de la programación de bloques a una de capas.	Alta	3
4	Creación y carga de vectores.	Desarrollo de función para vectorizar la información para cargar a Pinecone acorde al índice establecido.	Alta	4
5	Host y servidor web	Integrar el chatbot local a un servidor compatible con Python	Alta	5
6	Generando Consultas	Pruebas de funcionamiento desde el servidor	Alta	
7	Ajuste del algoritmo de búsqueda	Ajustar variables como temperatura del api.	Media	7
8	Optimización de la data	Depurar información para actualizar la base de datos	Alta	8
9	Integración a la red social	Integración a la red social establecida.	Alta	9
10	Prueba del chatbot integrado	Prueba del chatbot integrado.	Alta	10
11	Módulo para actualizar la data	Desarrollo de un módulo administrador que permita la actualización de la data.	Alta	11
12	Aplicación de matriz de confusión	Evaluar la precisión del chatbot al realizar consultas	Alta	12

Plan de duración de iteraciones. - Las iteraciones se adaptan a cada proceso y se documentan en la Tabla 25, alineadas con las historias de usuario. Esto asegura una planificación detallada y eficiente para cada etapa del proyecto.

Tabla 25. Plan de duración de Iteraciones

Iteración	Orden de las historias de usuario	Duración de la iteración
1	Selección de la información que proporcionara el chatbot personalizado.	15 - días
2	Segmentar la longitud de los textos que proporcionara el chatbot.	2 - días
3	Desarrollo de la codificación de bloques a una por capas en un IDE	4 - días
4	Desarrollo para vectorizar los datos y cargarlos a sistema de Pinecone y realizar las consultas.	3 - días
5	Integración al host y servidor web compatible con Python.	20 - días
6	Pruebas de ejecución a las consultas estimadas.	7 -días
7	Ajustar variables como temperatura del api para optimizar la búsqueda.	7 -días
8	Mejorar la información que servirá como data para personalizar el chatbot.	7 - días
9	Integración a la red social establecida.	15 - días
10	Pruebas de funcionamiento del prototipo ya integrado en el chat del Facebook.	30 - días
11	Implementación del módulo administrador para la gestión de actualizar la información del chatbot.	15 - días
12	Implementación de la matriz de con función como técnica de evaluación	7 - días

2.3.1.2 FASE II: Diseño

Diseño arquitectónico

El diseño del chatbot estará dividido en tres capas, a continuación, la descripción de cada una:

Capa de datos: En esta capa codificamos los métodos que cargaran el documento a la plataforma que los fragmentara y los convertirá en vectores los cuales serán enviados al modelo de lenguaje que emplea el chatbot

Capa de lógica de negocios: Aquí relacionamos a través de api keys el modelo de lenguaje y la base de datos vectorial donde se encuentran nuestros datos, para luego definir los métodos que capturarán las peticiones las cuales obtendrán su respectiva respuesta.

Capa de presentación: Esta es la capa final, donde el usuario ya tendrá acceso al chatbot mediante una interfaz en la cual podrá realizar consultas en lenguaje natural, recibiendo las respuestas en el mismo lenguaje.

Diseño arquitectónico del chatbot

El diseño arquitectónico del chatbot se basa en un gran modelo de lenguaje largo llamado gpt-3.5-turbo el cual será implementado mediante un apikey, utilizaremos la librería Langchain la cual tiene herramientas que nos permite comunicarnos con este modelo de lenguaje, también se hará uso de la base de datos vectorial Pinecone la cual permite cargar un documento con extensión pdf y lo que hará es fragmentar la información en vectores de los cuales el modelo de lenguaje obtendrá la información correspondiente. En la Figura 9 Se visualiza el funcionamiento del chatbot.

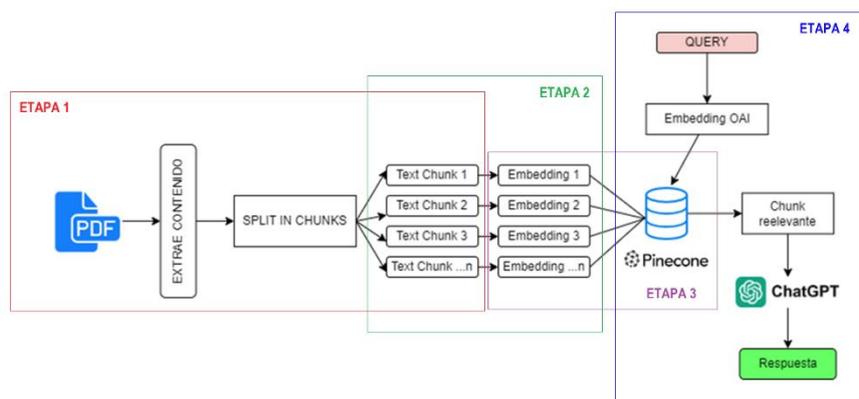


Figura 9. Arquitectura del prototipo de chatbot

Etapas de la arquitectura del chatbot

Primera Etapa: Carga de documento y fragmentación

Esta primera etapa implica la segmentación de la información contenida en el archivo PDF para su posterior procesamiento en el chatbot.

Como se explica en la figura 10, esta segmentación se realiza mediante una función de la librería de Langchain, RecursiveCharacterTextSplitter. Se indica el archivo donde está la información y se establece el parámetro chunk_size, que determina la cantidad de caracteres que contendrá cada fragmento generado durante el proceso de división.

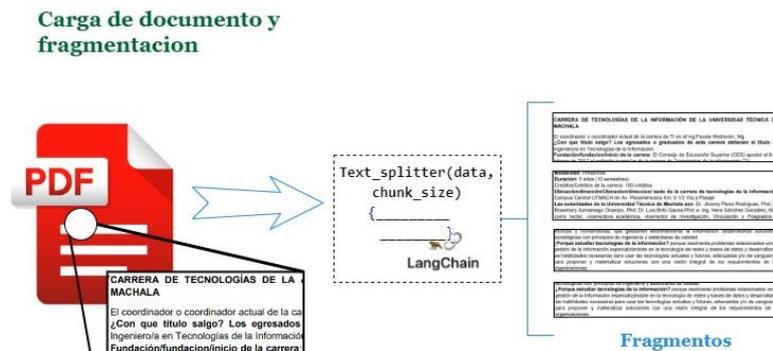


Figura 10. Primera etapa de arquitectura del chatbot

Segunda Etapa: Creación de vectores

En esta etapa del proceso figura 11, se procede a la creación de vectores que permitirán al chatbot comprender y procesar la información de manera más efectiva. Para lograrlo, se emplea la función OpenAIEmbeddings, la cual convierte cada uno de los fragmentos previamente generados en una representación binaria creando así los embeddings (vectores).

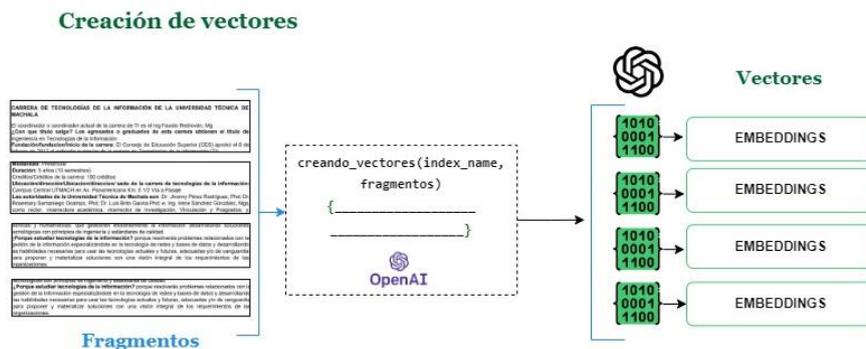


Figura 11. Segunda etapa de arquitectura del chatbot

Tercera Etapa: Subir vectores a DB

En esta fase figura 12, los vectores previamente generados se cargan en una base de datos de tipo vectorial. Para esta tarea, se emplea una Api Key para conectarse a la base de datos Pinecone. Este paso asegura que la información se almacene de manera eficiente en la base de datos, facilitando su posterior acceso y manipulación dentro del sistema del chatbot.

Subir Vectores a DB

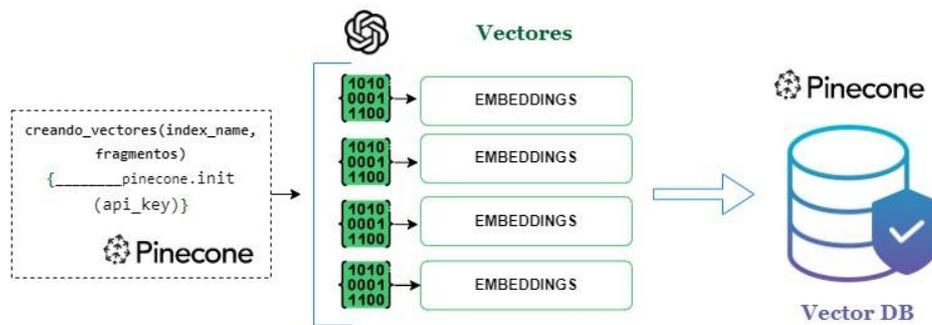


Figura 12. Tercera etapa de arquitectura del chatbot

Cuarta Etapa: Consulta y Respuesta

Esta es la fase final del proceso figura 13. Cuando se realiza una consulta al chatbot, la pregunta ingresa y se convierte primero en un vector binario, luego este se introduce en la DB, donde se compara con la información almacenada. A partir de esta comparación, se identifica el vector más relevante que responde a la pregunta. Este vector se somete al proceso de recuperación mediante la función RetrievalQA, donde el LLM interviene para proporcionar una respuesta en contexto con la pregunta planteada.

Consulta y Respuesta

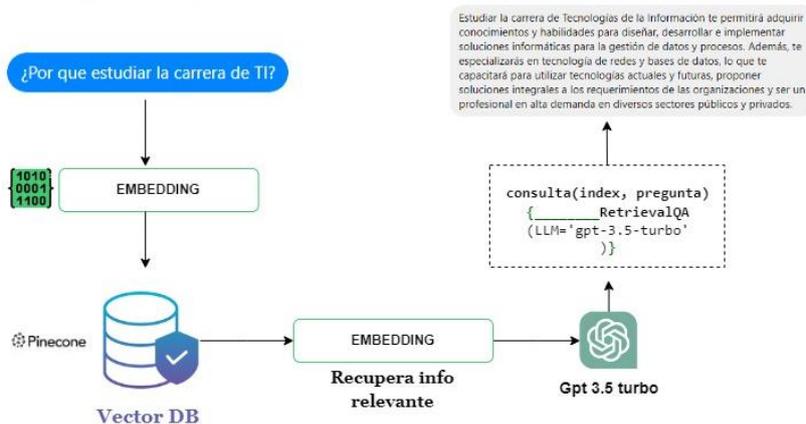


Figura 13. Cuarta etapa de arquitectura del chatbot

Diseño de interfaz

El diseño de la interfaz del chatbot se creó utilizando la herramienta en línea Canva y busca emular la apariencia de chat de un perfil de Facebook relacionado con la carrera de Tecnologías de la Información (TI). Aunque Canva no permite la funcionalidad de simulación de perfiles, este diseño representa cómo podría verse visualmente el chat en el perfil en la red social. La figura 14 muestra este diseño, facilitando la conceptualización y planificación de la interfaz del chatbot para un perfil de carrera en redes sociales, lo que resulta útil para la presentación visual de la propuesta de diseño de usuario.



Figura 14. Propuesta de interfaz

2.3.1.3 FASE III: Codificación

La codificación del chatbot se realizó en el entorno de desarrollo integrado (IDE) Visual Studio Code. Se estableció un entorno virtual para gestionar las dependencias esenciales. Esta práctica garantiza un ambiente controlado y eficiente para la programación y desarrollo del chatbot.

Lenguaje de programación

Python fue seleccionado como el lenguaje de programación central para este proyecto de chatbot debido a su versatilidad, facilidad de aprendizaje y abundante conjunto de bibliotecas y recursos disponibles. Python facilita un desarrollo ágil y eficiente, fundamental para implementar la compleja lógica conversacional del chatbot y su integración sin problemas con diversas tecnologías y sistemas, lo que asegura un proyecto eficaz y adaptable.

Base de datos

Se empleó la base de datos vectorial Pinecone, permite el almacenamiento de datos en forma de vectores, un formato esencial para la comunicación efectiva con modelos de lenguaje. Esto facilita la transmisión de información y optimiza el procesamiento en el contexto de la tarea en cuestión.

Librerías

Se empleó Langchain, una biblioteca de Python, que ofrece diversas capacidades para crear y concatenar prompts de manera eficiente. Esto facilita la transición entre proveedores, estandariza las interacciones con modelos de lenguaje de gran escala LLMS y brinda agentes para tareas como el razonamiento simbólico.

Desarrollo del Sistema

Para integrar el chatbot, se consideraron las bibliotecas necesarias para garantizar un funcionamiento correcto, las cuales se aprecia en figura siguiente.

```
· # from injecting_data import vectores
· import os
· from dotenv import load_dotenv, find_dotenv
· import pinecone
· from langchain.vectorstores import Pinecone
· from langchain.embeddings.openai import OpenAIEmbeddings
· from flask import Flask, render_template, request
```

Figura 15. Librerías empleadas

Entre las bibliotecas visualizadas, podemos apreciar Pinecone, que posibilita la conexión a la base de datos, y principalmente OpenAI, la cual sirve como interfaz de comunicación con GPT-3.5 para obtener la interacción del chatbot.

El chatbot se integró en un entorno de desarrollo de Visual Studio Code, utilizando la API de GPT-3.5 en su versión de texto. Esta elección se consideró apropiada para el tipo de chatbot que se estaba desarrollando. Además, se incorporó una base de datos vectorial que contribuye a personalizar el chatbot, aprovechando la información recopilada, figura 16 integración realizada.

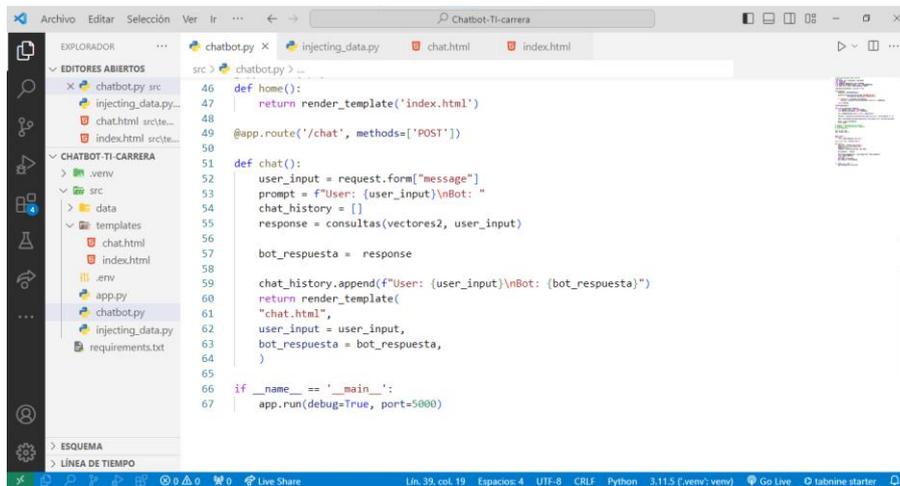


Figura 16. Integración en Visual Code

La base de datos Pinecone opera generando vectores a través de la fragmentación de la información, y luego indexa estos vectores en la base de datos. Esta metodología implica dividir la información en fragmentos, cada uno representado por un vector numérico. Este enfoque de representación vectorial facilita la eficiente organización y recuperación de datos, ya que los vectores resultantes permiten comparaciones para identificar similitudes y relaciones en la base de datos figura 17 implementación de la indexación en Pinecone.

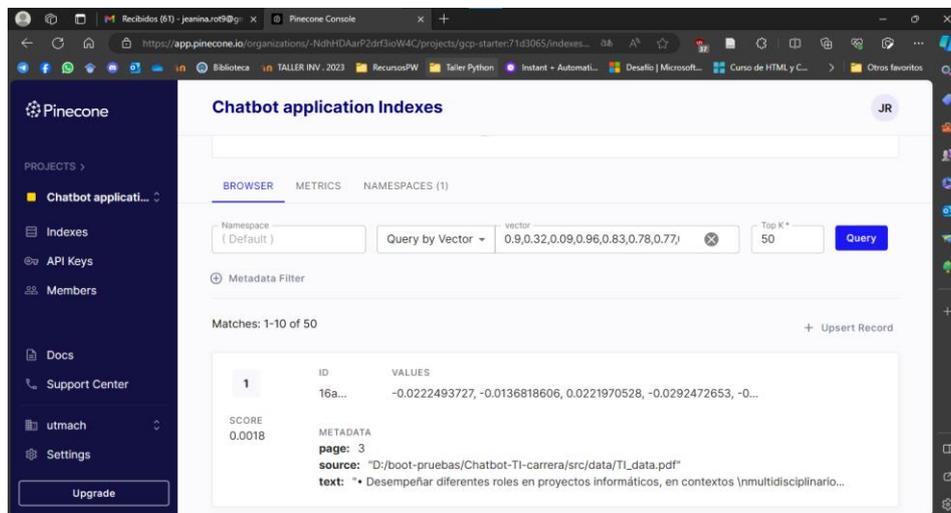


Figura 17. Base de datos Pinecone

Codificación capa de datos

En la capa de datos figura 18, se ha desarrollado un método que carga un archivo y lo divide en fragmentos, lo cual es crucial para agilizar las consultas de respuestas. Con un overlap del 100 y un chunk size que representa el tamaño del texto, este proceso implica la apertura del archivo, la lectura de su contenido y la fragmentación en secciones más pequeñas. Este enfoque proporciona una funcionalidad esencial para la manipulación eficiente de datos almacenados en archivos, asegurando tiempos de respuesta optimizados durante las consultas.

```
def fragmentar(data, chunk_size=800):
    from langchain.text_splitter import RecursiveCharacterTextSplitter
    text_splitter = RecursiveCharacterTextSplitter(chunk_size=chunk_size, chunk_overlap=100)
    fragmentos = text_splitter.split_documents(data)
    return fragmentos
```

Figura 18. Fragmentar Datos

Luego se define como se crearán los vectores a partir de los fragmentos de información, estos vectores se subirán a la base de datos vectorial Pinecone visualizado en la figura 19.

```
def creando_vectores(index_name):
    embeddings = OpenAIEmbeddings()

    pinecone.init(api_key=os.environ.get('PINECONE_API_KEY'),
                 environment=os.environ.get('PINECONE_ENV'))

    if index_name in pinecone.list_indexes():
        print(f'El índice {index_name} ya existe. Cargando los embeddings ... ', end='')
        vectores = Pinecone.from_existing_index(index_name, embeddings)
        print('Ok')
    else:
        print(f'Creando el índice {index_name} y los embeddings ...', end='')
        pinecone.create_index(index_name, dimension=1536, metric='cosine')
        vectores = Pinecone.from_documents(fragmentos, embeddings, index_name=index_name)
        print('Ok')

    return vectores
```

Figura 19. Creando embeddings

Codificación capa de lógica de negocio

En la capa de negocios, se utilizan los agentes proporcionados por Langchain para establecer la conexión con el modelo de lenguaje GPT-3.5 Turbo. Este modelo se comunica con la base de datos vectorial para obtener los vectores correspondientes. Para llevar a cabo este proceso, se realiza una búsqueda del índice utilizando la API key. El método realiza la búsqueda del índice con el nombre propuesto y lo guarda. El método de

consulta principal utiliza el índice obtenido y la pregunta. A continuación, la figura 20 presenta la lógica utilizada para la comparación de índices.

```
# import os
from dotenv import load_dotenv, find_dotenv
import pinecone
from langchain.vectorstores import Pinecone
from langchain.embeddings.openai import OpenAIEmbeddings
from pinecone.core.client.configuration import Configuration as OpenApiConfiguration

load_dotenv(find_dotenv(), override = True)

def getIndex():
    embeddings = OpenAIEmbeddings()

    openapi_config = OpenApiConfiguration.get_default_copy()
    openapi_config.proxy = "http://proxy.server:3128"
    pinecone.init(api_key='0391eb4d-108f-43f6-8925-254600e43140',
                 environment='gcp-starter',
                 openapi_config=openapi_config)

    if 'chatbot-ti1' in pinecone.list_indexes():
        vectores2 = Pinecone.from_existing_index('chatbot-ti1', embeddings)

    return vectores2

vectores2=getIndex()
```

Figura 20. Búsqueda de Índice

En cuanto a la precisión, se enfoca en la temperatura para obtener respuestas más precisas. A mayor temperatura, se busca aumentar la diversidad en las respuestas, mientras que a menor temperatura se persigue una respuesta más precisa y específica. Este enfoque permite ajustar la precisión de las respuestas según las necesidades del sistema. A continuación, la Figura 21 presenta la lógica utilizada para ajustar la precisión de las respuestas.

```
vectores2=getIndex()

def consultas(vectores2, pregunta):
    from langchain.chains import RetrievalQA
    from langchain.chat_models import ChatOpenAI

    llm = ChatOpenAI(model='gpt-3.5-turbo', temperature=0.3)

    retriever = vectores2.as_retriever(search_type='similarity', search_kwargs={'k': 3})

    chain = RetrievalQA.from_chain_type(llm=llm, chain_type="stuff", retriever=retriever)

    answer = chain.run(pregunta)
    return answer
```

Figura 21. Ajuste de Precisión

Configuración del Servidor PythonAnywhere

En la plataforma pythonanywhere escogemos el framework que con él se va a trabajar en este caso el escogido es flask, luego se procede hacer las siguientes configuraciones:

Establecemos el dominio tal y como se aprecia en la figura 22.

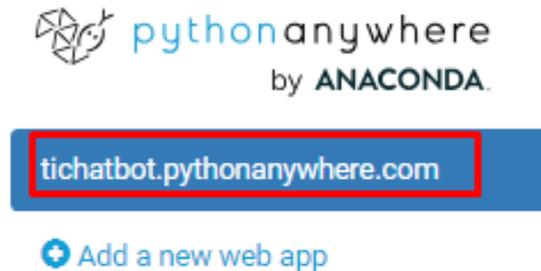


Figura 22. Dominio

Se configura un entorno virtual mediante el comando **workon** con el nombre del entorno, en este caso, **.venv**, para así instalar las dependencias necesarias y asegurar el correcto funcionamiento del proyecto, como se ilustra en la figura 23.

```
20:26 ~ $ workon .venv
(.venv) 20:26 ~ $ pip list
Package          Version
-----
aiohttp          3.9.0
aiosignal        1.3.1
annotated-types  0.6.0
anyio            3.7.1
async-timeout    4.0.3
attrs            23.1.0
blinker          1.7.0
certifi          2023.11.17
charset-normalizer 3.3.2
click            8.1.7
dataclasses-json 0.6.2
distro           1.8.0
dnspython        2.4.2
exceptiongroup   1.2.0
Flask            3.0.0
frozenlist       1.4.0
greenlet         3.0.1
h11              0.14.0
httpcore         1.0.2
httpx            0.25.1
idna             3.4
itsdangerous     2.1.2
Jinja2           3.1.2
jsonpatch        1.33
jsonpointer      2.4
langchain        0.0.340
langsmith        0.0.66
loguru           0.7.2
MarkupSafe       2.1.3
marshmallow      3.20.1
multidict        6.0.4
mypy-extensions 1.0.0
numpy            1.26.2
```

Figura 23. Creación de entorno virtual

La ruta del entorno virtual se configura como se muestra en la figura 24.

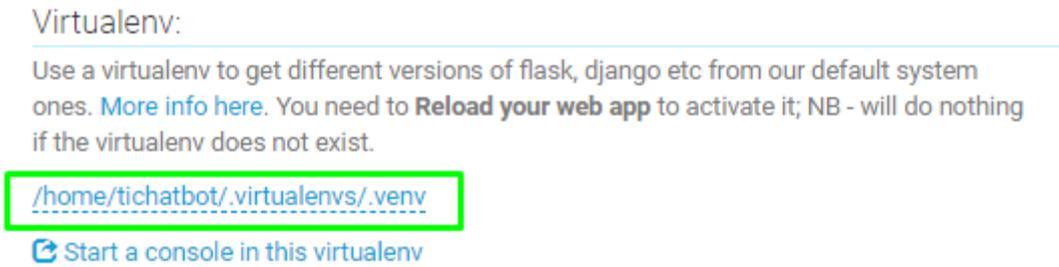


Figura 24. Ruta del entorno virtual

Todos los archivos **.py** y las plantillas HTML se cargan en el directorio desde el cual la aplicación los ejecutará, tal como se muestra en la figura 25.

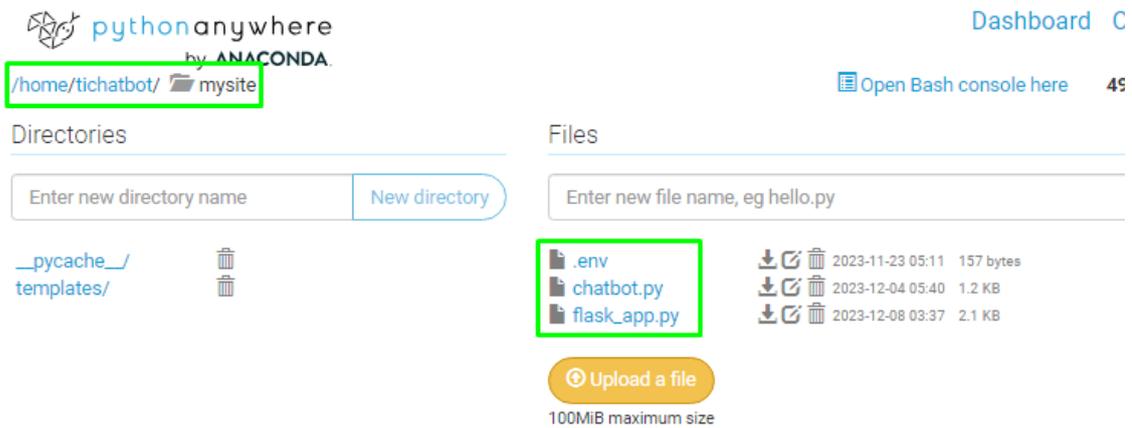


Figura 25. Subir archivos al directorio de Pythonanywhere

Finalmente, se recarga la aplicación para que los cambios surtan efecto, conforme se ilustra en la figura 26.

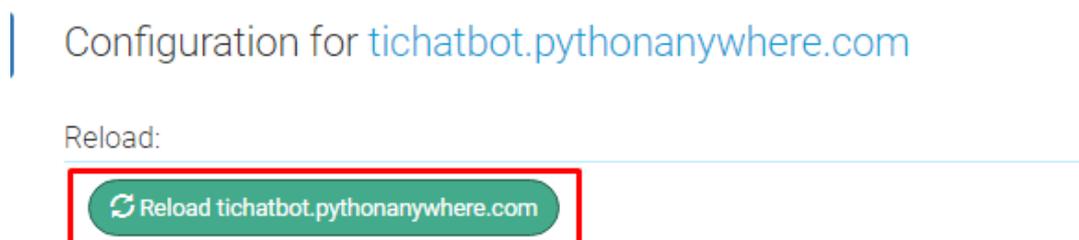


Figura 26. Recargando app

Configuración de Facebook Developer

En la plataforma, se procede a crear una nueva aplicación, siguiendo los pasos indicados en la figura 27.

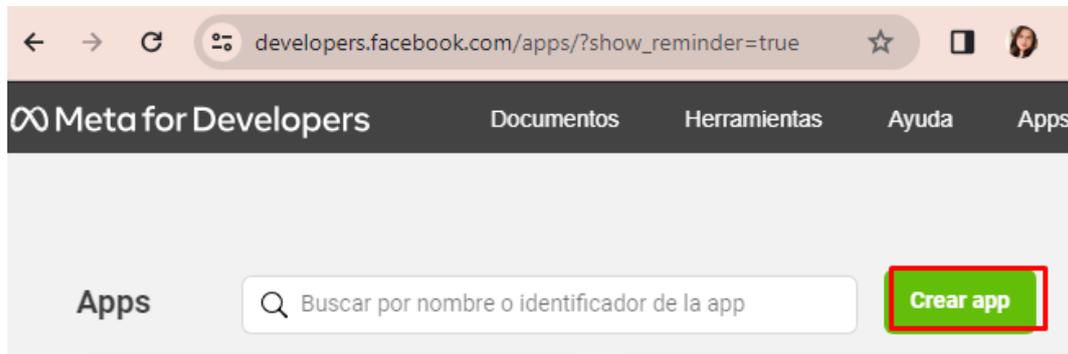


Figura 27. Crear nueva app en Facebook

Se selecciona el tipo de aplicación, en este caso, se elige "Negocios", siguiendo la indicación de la figura 28.

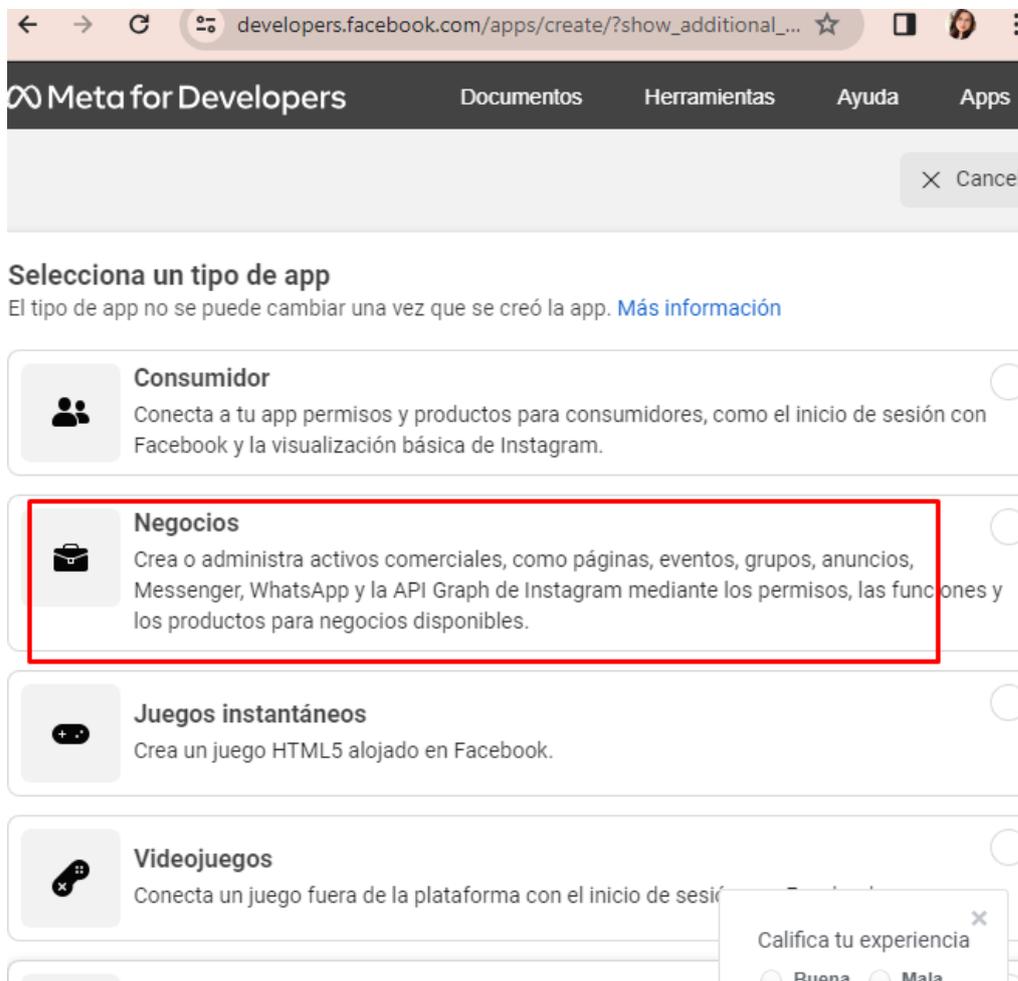


Figura 28. Tipo de aplicación

Se asigna un nombre a la aplicación y se finaliza el proceso de creación, siguiendo el procedimiento descrito en la figura 29.

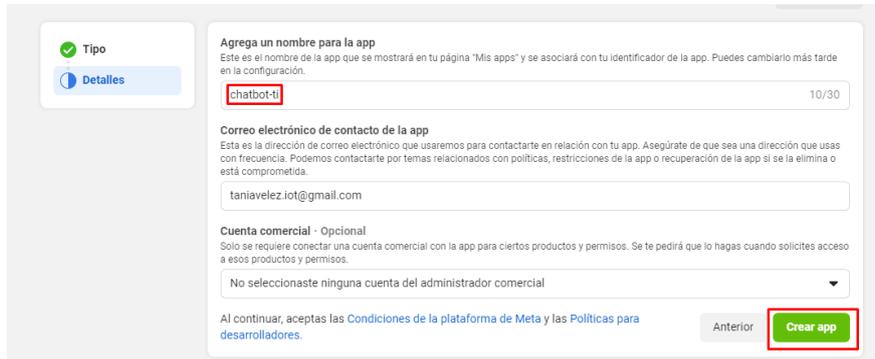


Figura 29. Nombrando App

Una vez creada la aplicación, es necesario seleccionar los productos de Facebook necesarios, en este caso, Messenger y Webhook, como se muestra en la figura 30.

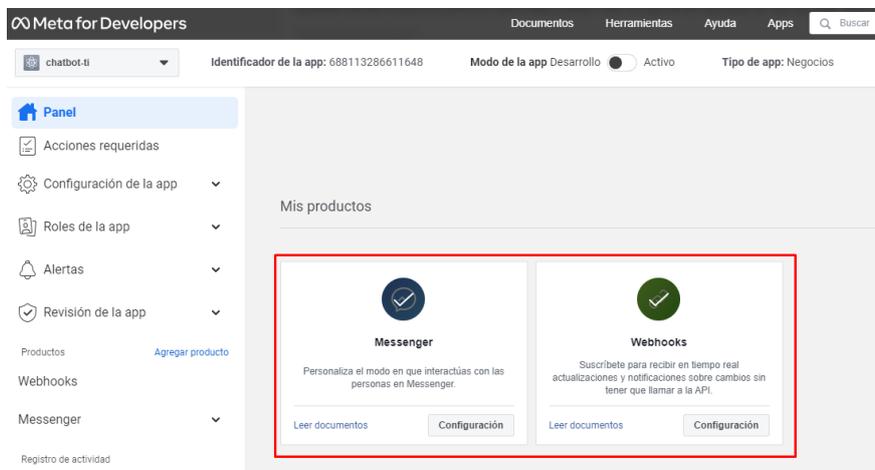


Figura 30. Selección de productos.

En la sección de Webhook, se configuran las suscripciones detallado en la figura 31.

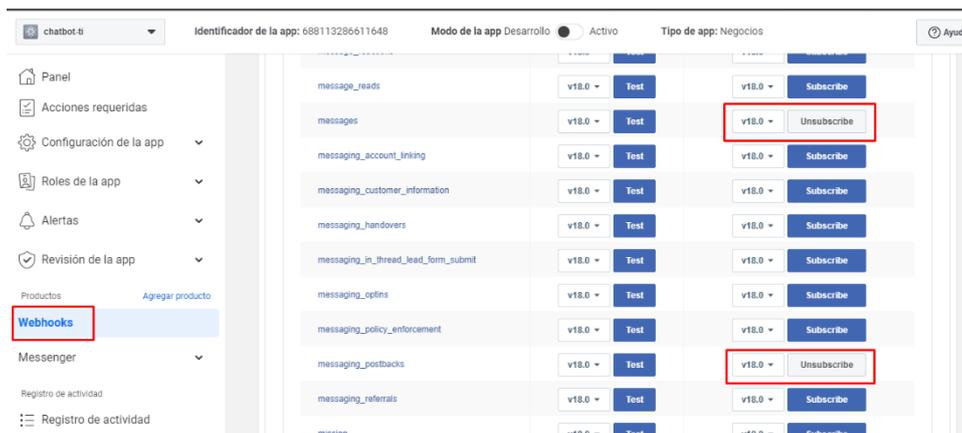


Figura 31. Suscripciones Webhook

En la configuración de Messenger, la primera acción consiste en establecer el API, donde se introduce la Callback URL y el Token de verificación, tal como se muestra en la figura 32.

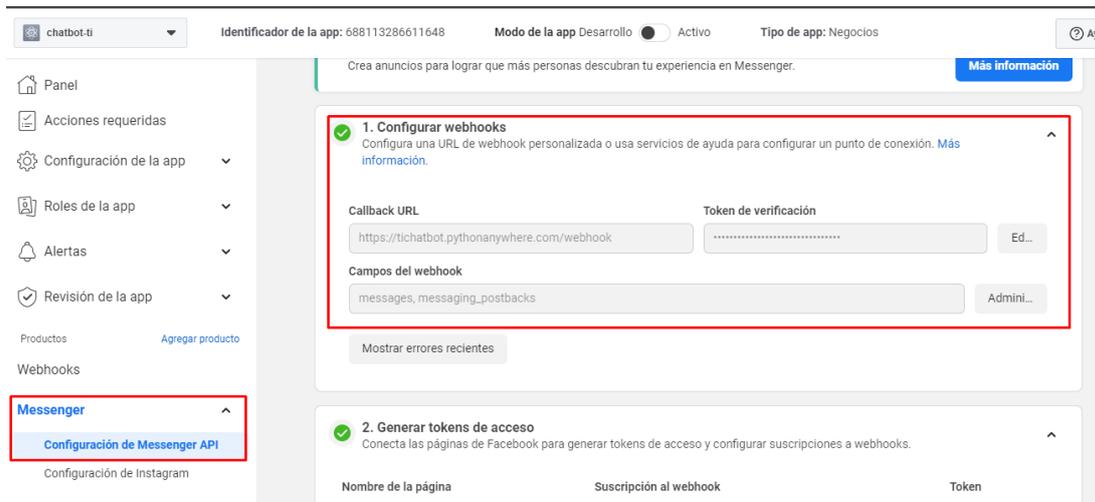


Figura 32. Callback URL y Token

Posteriormente, se procede a configurar la página de Facebook a la cual se conectará la aplicación, generando un token de seguridad para la conexión entre la aplicación de Facebook y nuestra aplicación web, como se detalla en la figura 33

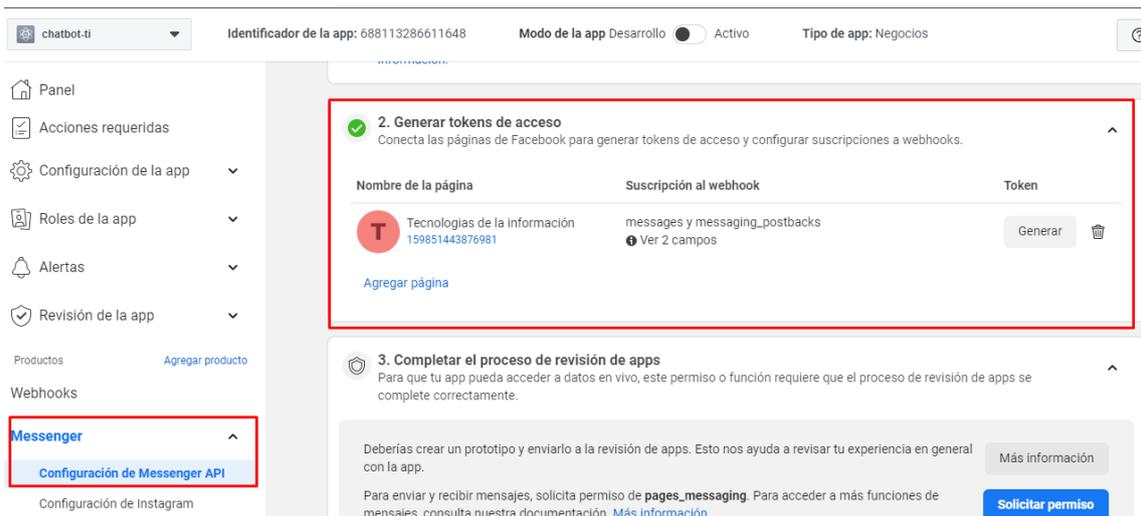


Figura 33. Conexión con el chatbot

2.3.1.4 FASE IV: Prueba

La presente etapa se basa en la realización de las pruebas pertinentes para determinar que el prototipo funcione adecuadamente, el desarrollo de esta etapa es primordial para verificar la funcional del chatbot y así lograr cumplir lo propuesto.

Tabla 26. Prueba 1

Prueba de Aceptación	
Número:1	Historia de usuario 1
Nombre de Historia: Estructuración de datos a utilizar.	
Descripción: Seleccionar la información a necesitar para el entrenamiento del chatbot.	
Condiciones de Ejecución: Establecer información destacada	
Interfaz: Ninguna	
Resultados esperados: Proporcionar información precisa.	
Evaluación de la prueba: Satisfactorio	

Tabla 27. Prueba 2

Prueba de Aceptación	
Número:1	Historia de usuario 2
Nombre de Historia: Fragmentar datos.	
Descripción: Se realizar los embeding de la información para la generación de la base vectorial tomando en cuenta las características de búsqueda de los datos	
Condiciones de Ejecución: Establecer la dimensión del embedding soportado	
Interfaz: Ninguna	
Resultados esperados: Generación de los embenddings en la plataforma de Pinecone	
Evaluación de la prueba: Satisfactorio	

A continuación, en la figura 34 se visualiza la lógica de la fragmentación de los datos recolectados

```
def fragmentar(data, chunk_size=800):
    from langchain.text_splitter import RecursiveCharacterTextSplitter
    text_splitter = RecursiveCharacterTextSplitter(chunk_size=chunk_size, chunk_overlap=100)
    fragmentos = text_splitter.split_documents(data)
    return fragmentos

fragmentos = fragmentar(contenido)
```

Figura 34. Fragmentar

Tabla 28. Prueba 3

Prueba de Aceptación	
Número:3	Historia de usuario 3
Nombre de Historia: Integración al IDE.	
Descripción: Se comprueba la integración a un IDE Visual Code utilizar Python como lenguaje.	
Condiciones de Ejecución: Ninguna	
Interfaz: Ninguna	
Resultados esperados: Algoritmo integrados acorde a la lógica establecida	
Evaluación de la prueba: Satisfactorio	

Tabla 29. Prueba 4

Prueba de Aceptación	
Número:1	Historia de usuario 4
Nombre de Historia: Creación y carga de vectores.	
Descripción: Se comprueba los vectores en la PINECONE	
Condiciones de Ejecución: Ninguna	
Interfaz: Ninguna	
Resultados esperados: Index establecido en Pinecone	
Evaluación de la prueba: Satisfactorio	

A continuación, se visualiza los detalles del index de la base de datos figura 35.

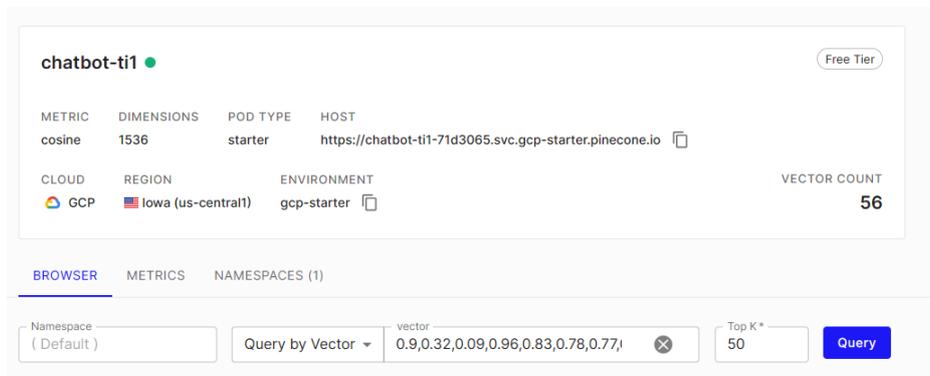


Figura 35. Prueba index pinecone

Tabla 30. Prueba 5

Prueba de Aceptación	
Número:5	Historia de usuario 5
Nombre de Historia: Host y servidor web	
Descripción: Se que el algoritmo este integrado a falsk y PythonAnywhere	
Condiciones de Ejecución: Ninguna	
Interfaz: Ninguna	
Resultados esperados: Integrar a Flask.	
Evaluación de la prueba: Satisfactorio	

Tabla 31. Prueba 6

Prueba de Aceptación	
Número:6	Historia de usuario 6
Nombre de Historia: Generando Consultas	
Descripción: Pruebas de funcionamiento desde el servidor o host	
Condiciones de Ejecución: Ninguna	
Interfaz: Ninguna	
Resultados esperados: Respuesta correcta y precisa	
Evaluación de la prueba: Satisfactorio	

A continuación, se visualiza una prueba de consulta del chatbot integrado en flask figura 36.

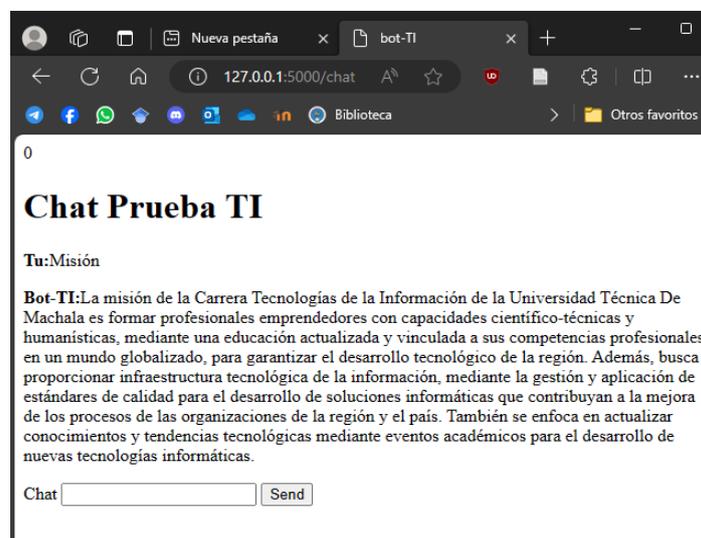


Figura 36. Consultas desde Flask

Tabla 32. Prueba 7

Prueba de Aceptación	
Número:7	Historia de usuario 9
Nombre de Historia: Prueba del chatbot integrado	
Descripción: Prueba del chatbot integrado	
Condiciones de Ejecución: Ninguna	
Interfaz: Ninguna	
Resultados esperados: Chatbot integrado en Facebook y respuestas correctas	
Evaluación de la prueba: Satisfactorio	

A continuación, en la figura 37, se visualiza el chatbot integrado en Facebook correctamente funcionando.

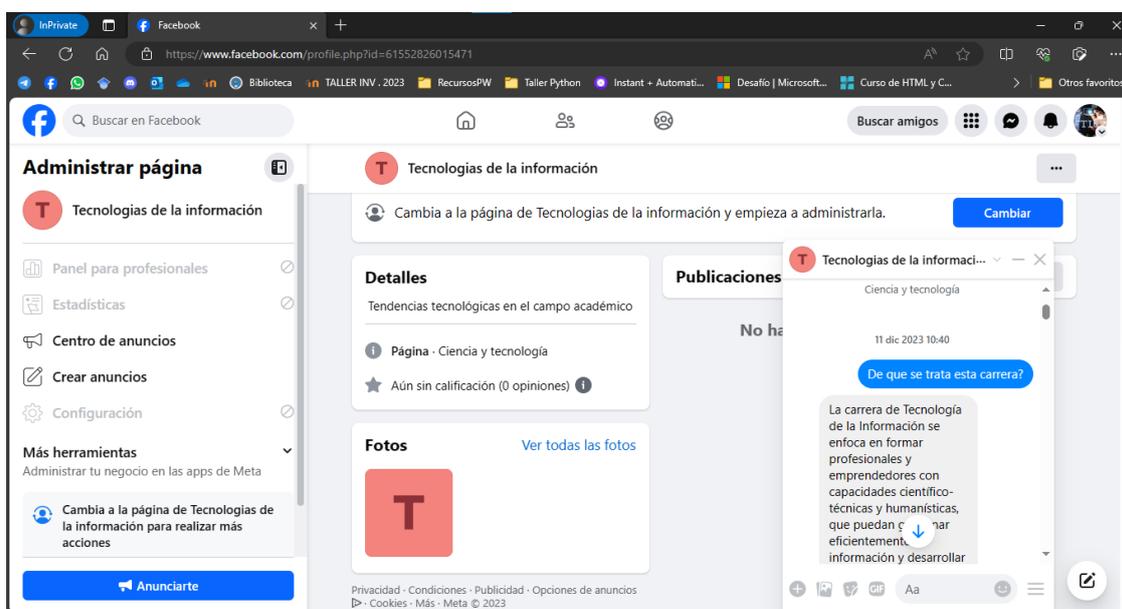


Figura 37. Prueba del chatbot integrado

Tabla 33. Prueba 8

Prueba de Aceptación	
Número:8	Historia de usuario 11
Nombre de Historia: Modulo Administrador	
Descripción: Prueba de actualizar data.	
Condiciones de Ejecución: Ninguna	
Interfaz: Ninguna	
Resultados esperados: Modulo administrador	
Evaluación de la prueba: Satisfactorio	

Módulo administrador diseñado para gestionar la actualización de información del chatbot. Este módulo se observa en la imagen 38.



Figura 38. Actualizar data

2.4 Ejecución del prototipo

2.4.1 Ejecución Local

Se realiza el prototipo del chatbot, seguido de rigurosas pruebas locales para lo cual se requiere de lo siguiente.

- Api GPT-3.5 turbo
- Pinecone
- Langchain
- Python
- Flask
- tiktoken

Una vez realizado la obtención y configuración de api key y configuración de respectivo entorno de desarrollo se lo activa con el siguiente parámetro `.venv/Scripts/activate` para lo cual se agregará un `(.venv)` de otro color como se aprecia en la siguiente imagen figura 39.

```
PROBLEMAS  SALIDA  CONSOLA DE DEPURACIÓN  TERMINAL  PUERTOS  COMENTARIOS

PS D:\boot-pruebas\Chatbot-TI-carrera>
● PS D:\boot-pruebas\Chatbot-TI-carrera> .\venv\Scripts\activate
○ (.venv) PS D:\boot-pruebas\Chatbot-TI-carrera> |
```

Figura 39. Activar entorno

Se lleva a cabo la inyección del archivo **injecting_data.py**. Este archivo realiza una verificación para determinar si la base de datos ya existe. En caso contrario, procede con la respectiva indexación, tomando en cuenta el proceso fragmentar la data que se subirá para esta comparativa se realiza mediante la búsqueda del índice, donde se determina si

ya existe uno con el nombre definido chatbot-ti1, si no se encuentra, carga los datos correspondientes para su posterior utilización figura 40.

```
PS D:\boot-pruebas\Chatbot-TI-carrera> & d:/boot-pruebas/Chatbot-TI-carrera/.venv/Scripts/Activate.ps1
(.venv) PS D:\boot-pruebas\Chatbot-TI-carrera> & d:/boot-pruebas/Chatbot-TI-carrera/.venv/Scripts/python.exe d:/boot-pruebas/Chatbot-TI-carrera/src/injecting_data.py
Cargando D:/boot-pruebas/Chatbot-TI-carrera/src/data/TI_data.pdf...
El índice chatbot-ti1 ya existe. Cargando los embeddings ... Ok
```

Figura 40. Injecting para cargar la data

Ahora se ejecuta el archivo chatbot.py el mismo que contiene la configuración respectiva de flask y del funcionamiento del chatbot figura 41.

```
(.venv) PS D:\boot-pruebas\Chatbot-TI-carrera> & d:/boot-pruebas/Chatbot-TI-carrera/.venv/Scripts/python.exe d:/boot-pruebas/Chatbot-TI-carrera/src/chatbot.py
* Serving Flask app 'chatbot'
* Debug mode: on
WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
* Restarting with stat
* Debugger is active!
* Debugger PIN: 928-082-512
127.0.0.1 - - [28/Dec/2023 01:14:57] "GET / HTTP/1.1" 200 -
```

Figura 41. Ejecución del modelo

Ejecución y prueba desde un entorno local donde el mismo que nos sirve para realizar las respectivas pruebas correspondientes para luego integrar a Facebook, figura 42.

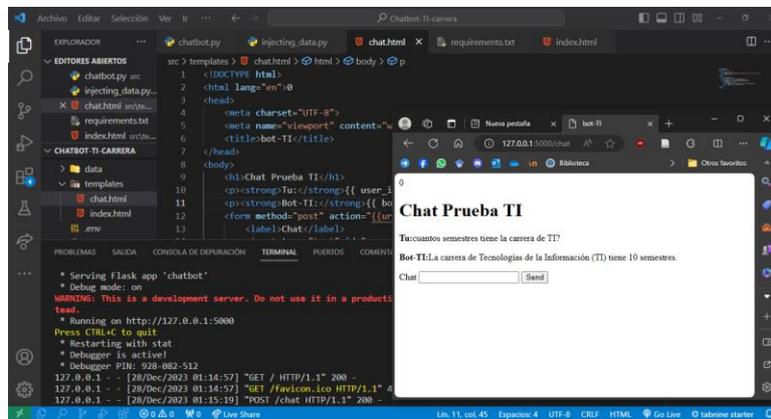


Figura 42. Ejecución y prueba local

Diseño de Ingreso al panel administrador que permite la actualización a la base de datos, figura 43.



Figura 43. Login de ingreso al administrador

2.4.2 Ejecución en línea desde Facebook

Para llevar a cabo la ejecución desde Facebook, debemos dirigirnos a la página del prototipo y abrir el chat haciendo clic en el icono de la burbuja, tal y como se muestra en la figura 44.

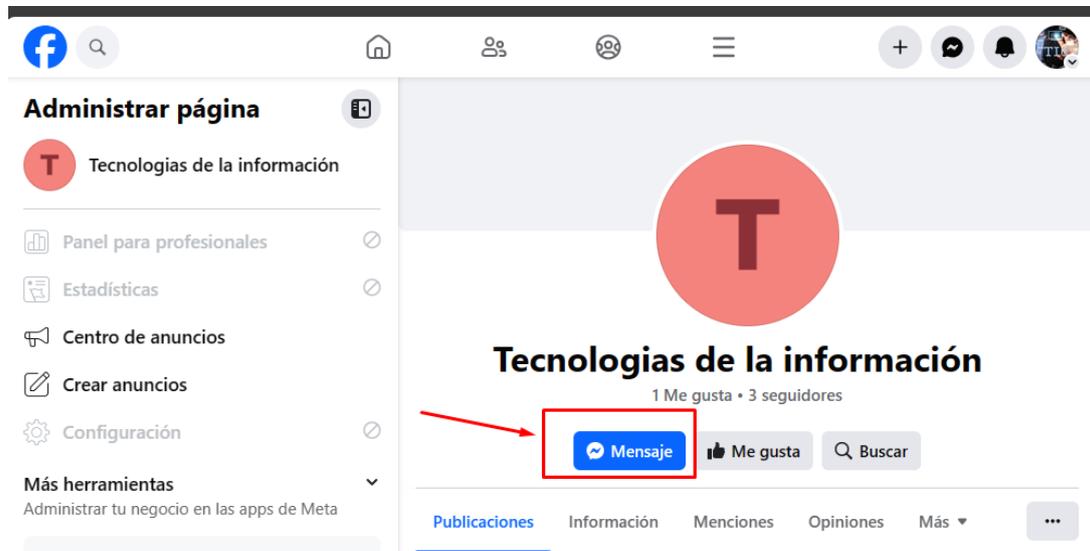


Figura 44. Página de Prototipo

Una vez ubicado en el chat, realizamos las pruebas respectivas recordando que solo se admite ingreso de texto tal como se aprecia en la figura 45.

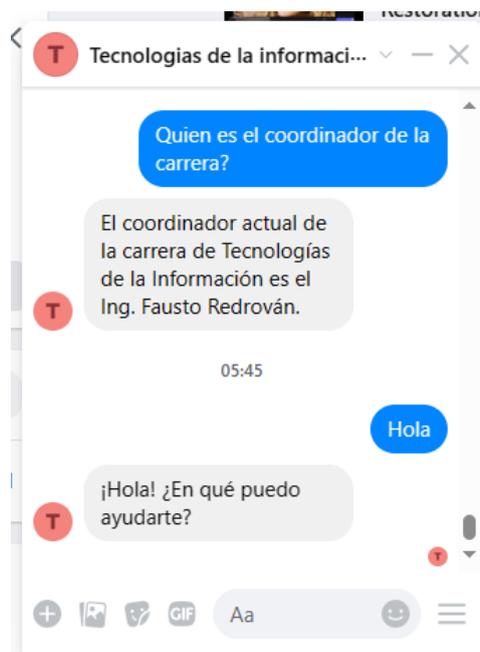


Figura 45. Interacción de prueba

Prueba de los requisitos funcionales, que consiste en poner a prueba las preguntas estimadas que se espera que el sistema pueda responder, tomando en cuenta que cada usuario no va a escribir por igual al momento de consultar y siempre va a variar.

2.4.3 Análisis de Resultados

Prueba de los requisitos funcionales, la cual implica evaluar la capacidad del sistema para responder a preguntas estimadas, considerando la variabilidad en la forma en que cada usuario formula sus consultas. Teniendo en cuenta que la redacción de las preguntas puede variar entre usuarios, se busca verificar la flexibilidad y robustez del sistema frente a distintos estilos de consulta.

Consulta de asignaturas por semestre, como podemos observar en la figura 46 nos ha proporcionado una información exacta del total de asignaturas, pero ahora esto puede cambiar si es que formulamos de forma distinta como en la figura 47.

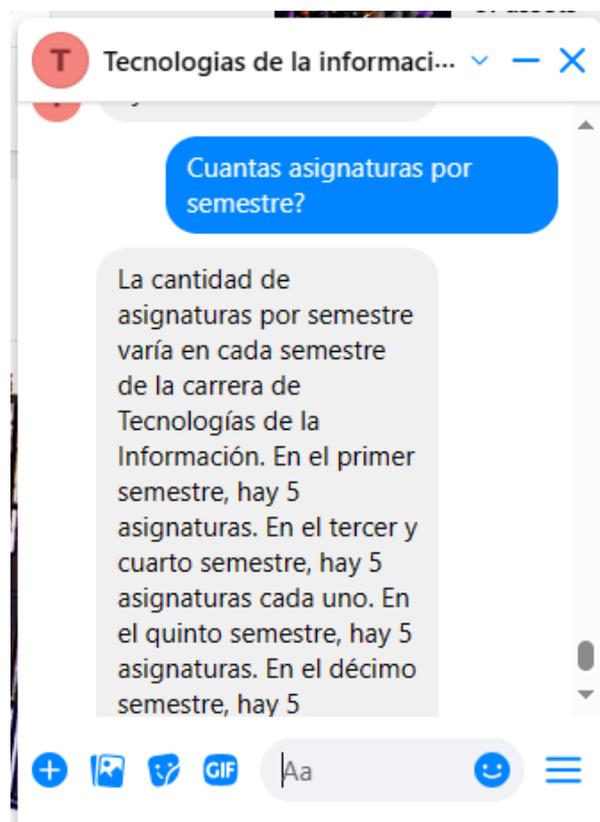


Figura 46. Asignaturas por semestre

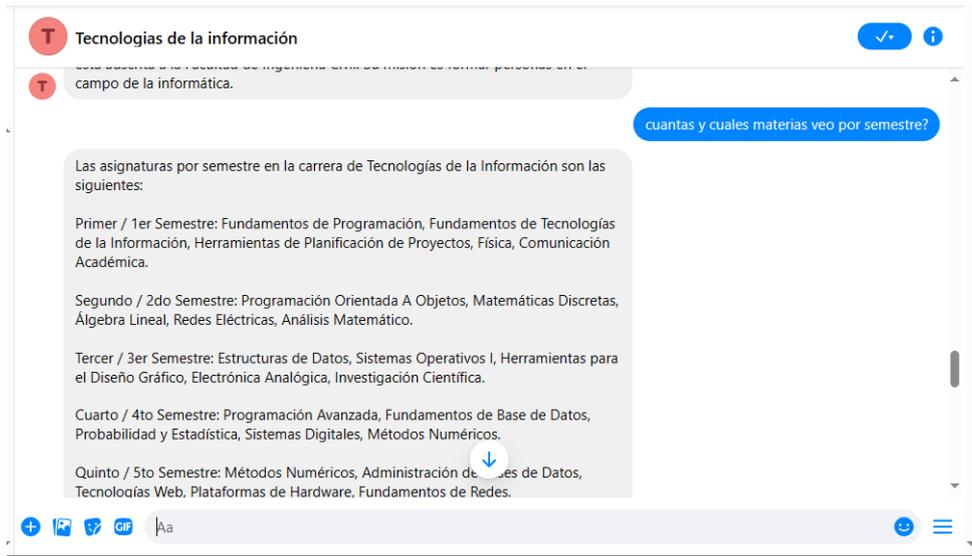


Figura 47. Cuantas y cuales materias veo por semestre

Prueba de consulta el proceso para ingreso a la carrera nos proporciona los pasos a seguir respuesta obtenida en la figura 48.

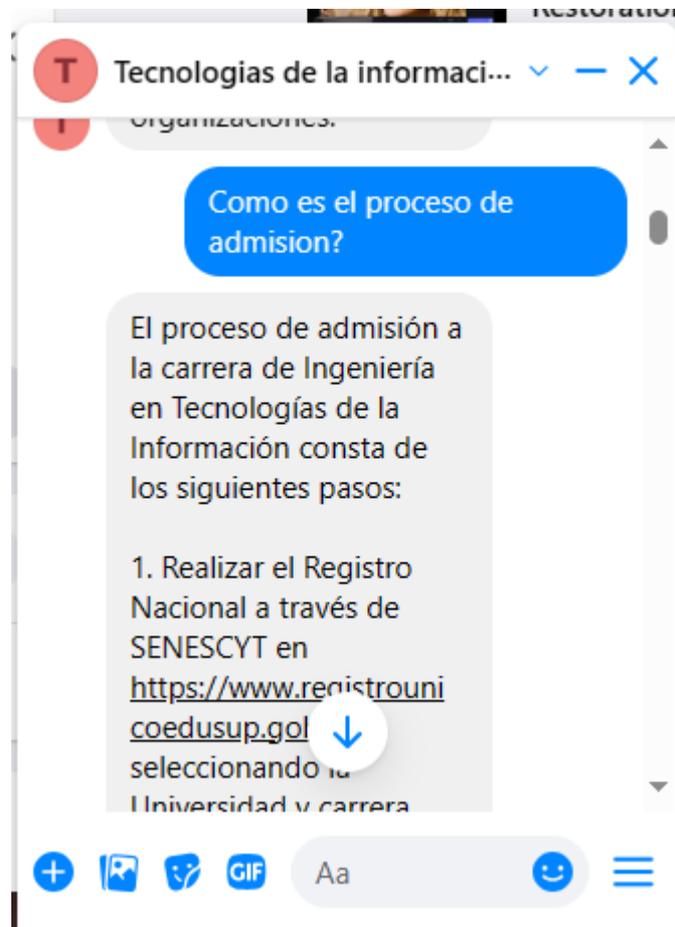


Figura 48. Proceso de ingreso a la carrea

La prueba consulta el proceso de justificación de faltas nos proporciona una guía de Los pasos a realizar tal y como se aprecia en la figura 49.

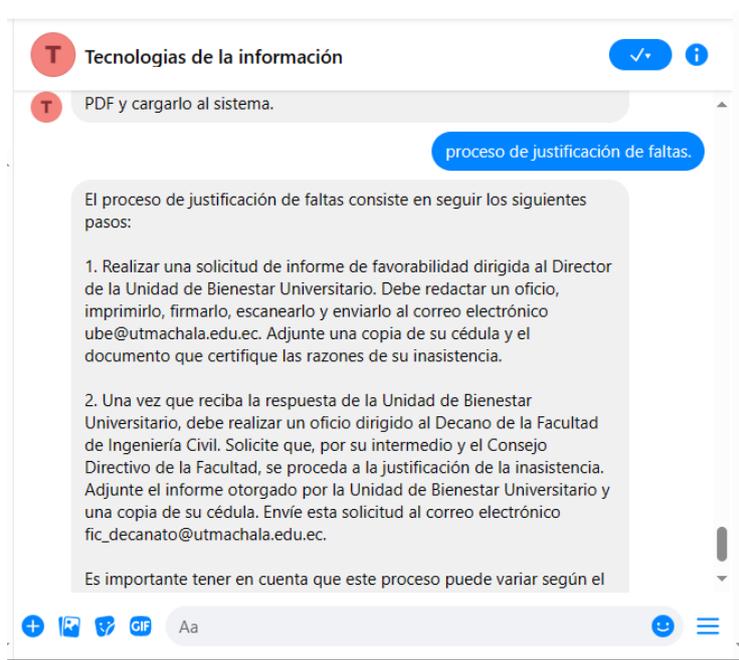


Figura 49. Proceso de justificación de faltas.

En la siguiente prueba se consulta sobre el proceso de pago de matrículas, a continuación, figure 50.

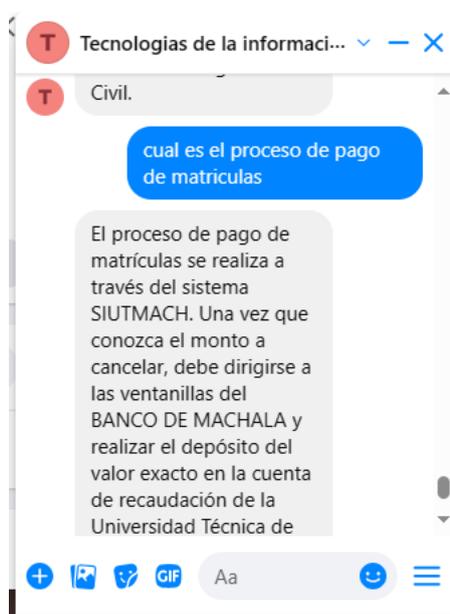


Figura 50. Proceso de pago de matrículas

Se visualiza en la figura 51, la consulta del actual coordinador de la carrera.

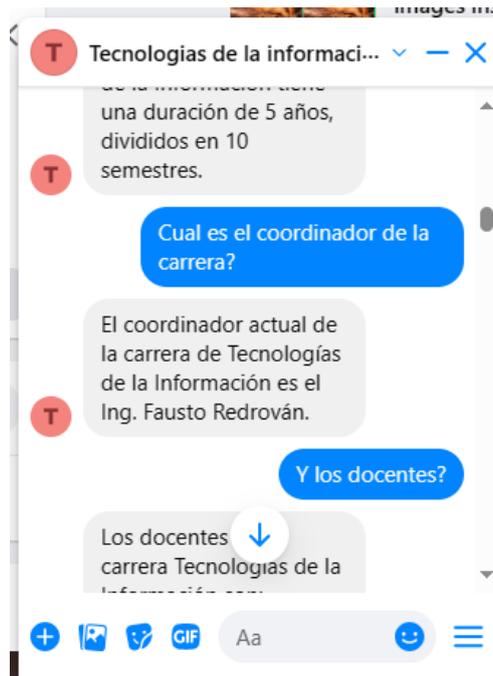


Figura 51. Coordinador de la carrera

Prueba Consulta la planta docente proporciona una lista de los docentes titulados de la carrera de tecnologías de la información Figura 52.

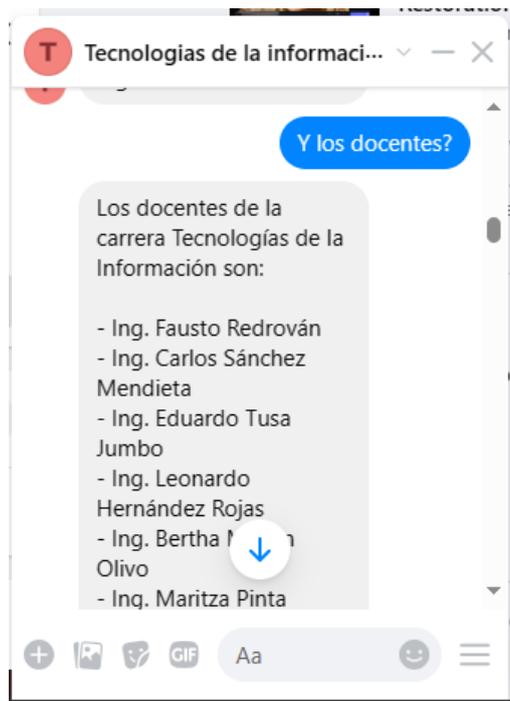


Figura 52. Planta docente

Prueba de obtener la ubicación y a que Facultad pertenece la carrera, vista proporcionada en la figura 53.

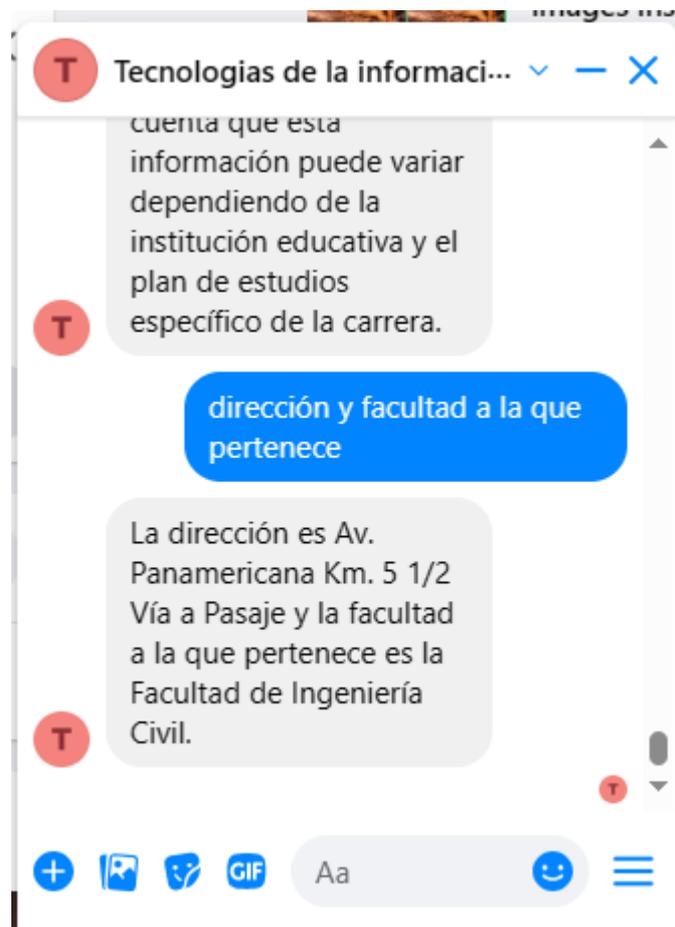


Figura 53. Ubicación y a que Facultad pertenece la carrera

CAPÍTULO III. EVALUACIÓN DEL PROTOTIPO

3.1 Plan de evaluación

El uso de chatbot hoy en día se han convertido en herramientas que ayudan a la reducción de tiempo con respecto a la atención al cliente, es así que el propósito del presente chatbot a evaluar es determinar el nivel de aceptación por parte de la comunidad y su eficiencia con respecto a las respuestas que proporcione debido a que se implementa técnicas de procesamiento de lenguaje natural (PLN), la misma que se encuentra en el modelo de Lenguaje Largo (MLL), y eso es gracias al api de Open IA con la versión 3.5 Turbo que nos ayuda con el lenguaje de texto y actuar como un asistente virtual.

Al emplear la matriz de confusión, se busca evaluar la eficacia del sistema, identificar posibles desafíos y perfeccionar su rendimiento. Este enfoque permite explorar las capacidades del chatbot en la comprensión contextual, anticipación de consultas y respuesta precisa, contribuyendo así al avance de las tecnologías de conversación automatizada.

3.1.1 Objetivo de la evaluación.

- Evaluar la eficacia y rendimiento del chatbot mediante el análisis de una matriz de confusión, con el propósito de validar la hipótesis de que el sistema de conversación automatizada es capaz de brindar respuestas precisas y coherentes en un amplio rango de consultas, identificando áreas de mejora y fortalezas en su desempeño.

3.1.2 Alcance

El alcance de la evaluación se centra en la efectividad y desempeño del chatbot implementado en el perfil de prueba en Facebook, con la finalidad de operar como un asistente virtual especializado en el ámbito académico. La evaluación abarcará la capacidad del chatbot para comprender y responder preguntas relacionadas con información académica, tales como matrícula, procesos de admisión, asignaturas por semestre.

Por lo tanto, se analizará la calidad de las respuestas generadas por el chatbot, considerando la coherencia, relevancia y precisión

3.1.3 Técnica de Evaluación

La técnica central de evaluación para el chatbot en Facebook implica la aplicación de la matriz de confusión. Este enfoque analiza la clasificación de respuestas en verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos. Se realiza una evaluación de la capacidad del chatbot para ofrecer respuestas académicas precisas.

3.1.4 Recursos

La evaluación del chatbot se realizará a través la matriz de confusión, por lo que se evalúa el rendimiento del chatbot en la plataforma de Facebook, además se utilizará un conjunto de datos de prueba representativo. Este conjunto incluirá diversas consultas académicas que permitirán medir la capacidad del chatbot para proporcionar respuestas precisas y coherentes.

3.1.5 Cronograma

Para la aplicación de evaluación se planteó un cronograma que se distribuye en cinco semanas donde iniciamos desde la planificación del plan donde se define la metodología o herramienta a implementas además de las preguntas a evaluar, la aplicación con la respectiva recopilación de datos y los resultados, lo cual se detalla en la tabla 34.

Tabla 34. Cronograma de evaluación

Semana	Actividad	Descripción
	Planificación del plan evaluación.	<ul style="list-style-type: none">Definir estructura para el plan de evaluación y metodología o instrumento.
Semana 9	Diseño del plan de evaluación.	<ul style="list-style-type: none">Revisión y ajuste del plan de evaluación.Diseño de las preguntas a evaluar.Matriz de confusión.

Tabla 34. Cronograma de evaluación

Semana	Actividad	Descripción
Semana 10 - 11	Aplicación de la matriz de confusión.	<ul style="list-style-type: none"> • Aplicar las pruebas necesarias para determinar los falsos positivos y negativos.
	Recopilación de Datos	<ul style="list-style-type: none"> • Tabulación de las respuestas obtenidas durante la etapa de prueba aplicando la matriz de confusión.
	Análisis de Datos	<ul style="list-style-type: none"> • Análisis de los datos recopilados.
Semana 12-13	Resultados de evaluación.	<ul style="list-style-type: none"> • Elaboración del informe integrando los datos obtenidos de la etapa de prueba.
	Revisión final	<ul style="list-style-type: none"> • Presentación de hallazgos, conclusiones.

3.1.6 Matriz de confusión

Para aplicar la matriz de confusión se establecen 4 criterios a evaluar para lo cual se plantea 20 preguntas por cada criterio, resultado un total de 80 preguntas a evaluar entre los criterios establecidos tenemos a los siguientes:

- Perfil de la Carrera
- Malla Curricular
- Estructura organizativa
- Preguntas Frecuentes

Es así que mediante dichos criterios se evalúa la precisión del chatbot y las predicciones se basan en falsos positivo, negativos y verdaderos positivos, negativos tal y como se representa en la tabla 35.

Tabla 355. Tabla de predicción.

		Predicción	
		Positivos	Negativos
Actual	Positivo	Verdadero Positivo	Falso Negativo
	Negativo	Falso Positivo	Verdadero Negativo

Formula a emplear:

$$\text{Precisión} = \frac{\text{Verdadero Positivo}}{\text{Verdadero Positivo} + \text{Falso Positivo}}$$

3.2 Resultados de la evaluación

Perfil de la Carrera

Tabla 36. Predicción Perfil de la Carrera

		Predicción	
		Positivos	Negativos
Actual	Positivo	14	1
	Negativo	1	4

Se lograron 14 aciertos en relación con la información del perfil de la carrera, lo que representa el 93.3% de precisión. El 7% restante corresponde a falsos positivos y negativos, que surgieron al no proporcionar una respuesta concreta o al dar respuestas incorrectas.

Malla Curricular

Tabla 36. Predicción Malla Curricular

		Predicción	
		Positivos	Negativos
Actual	Positivo	11	2
	Negativo	1	6

Se obtuvo 11 aciertos en relación con la información de la malla curricular lo cual abarca las materias de todos los semestres, obteniendo un 92% de precisión. El 8% restante corresponde a falsos positivos y negativos, que surgieron al no proporcionar una respuesta concreta o al dar respuestas incorrectas, lo cual va ha tiene gran dependencia de la forma en cómo se formula la consulta tomando en cuenta que se comete errores ortográficos.

Estructura organizativa

Tabla 37. Predicción estructura organizativa

		Predicción	
		Positivos	Negativos
Actual	Positivo	13	2
	Negativo	1	4

Del criterio estructura organizativa se obtuvo un total de 13 acierto representando el 93% de precisión, es así que la predicción se la ve con normalidad debido a que la contextualización de respuesta en gran parte depende de cómo cada usuario formule su pregunta.

Preguntas Frecuentes

Tabla 38. Predicción preguntas frecuentes

		Predicción	
		Positivos	Negativos
Actual	Positivo	12	4
	Negativo	0	4

Con respecto a las preguntas frecuentes como números de teléfono, información básica sobre estructura de justificativos o entre otras cosas se obtuvo un 100% de precisión teniendo 12 aciertos verdaderos positivos como 4 de verdaderos negativos logrando respuestas correctas.

Total de resultados

Tabla 39. Resultado Final

	Resultados Finales		
	Total Preguntas	Aciertos	Precisión
Perfil de la Carrera	20	14	93%
Malla Curricular	20	11	92%
Estructura organizativa	20	13	93%
Preguntas Frecuentes	20	12	100%

Al culminar las pruebas de evaluación mediante el análisis de la matriz de confusión y la consideración de la precisión en cada uno de los criterios evaluados, se logró obtener una precisión global del chatbot del 94.5%, al promediar los resultados obtenidos. Este porcentaje sólido indica un desempeño consistente y confiable en la comprensión y respuesta a las consultas, evidenciando la efectividad del prototipo.

CONCLUSIONES

- La implementación de técnicas de procesamiento de lenguaje natural ayuda a la contextualización de las preguntas que se formules siendo así una adecuada técnica para que el chatbot sirva de apoyo para la gestión de información académica.
- La realización de investigaciones sistemáticas aporta significativamente al entendimiento de casos similares y a la estructuración de la investigación. Además, facilita la selección de técnicas y herramientas que se ajusten de manera óptima al trabajo a realizar.
- En el desarrollo del prototipo de chatbot para soporte en los servicios de información de estudiantes de la carrera de TI, se emplearon tecnologías clave como GPT-3.5 para procesamiento de lenguaje natural, Pinecone como base de datos vectorial y la librería Langchain. Esta combinación ha resultado en un sistema especializado que proporciona respuestas precisas a las consultas de los estudiantes de la carrera de TI.
- La integración del chatbot en la página de Facebook es fundamental para llevar pruebas exhaustivas de su funcionamiento, permitiendo así la identificación y gestión de posibles errores de configuración. Además, facilita la optimización del tiempo de respuesta en el algoritmo de búsqueda para mejorar la eficacia del chatbot.
- Al finalizar la evaluación del chatbot en la página de prueba de Facebook, se observó un desempeño destacado. Analizando la matriz de confusión y considerando la precisión en los criterios evaluados, se logró alcanzar una sólida precisión global del 94.5%. Este resultado respalda la efectividad del prototipo al comprender y responder de manera consistente a las consultas. La evaluación satisfactoria confirma que el chatbot ha cumplido de manera confiable con el objetivo establecido de responder eficientemente en la plataforma de Facebook.

RECOMENDACIONES

- Continuar refinando las técnicas de procesamiento de lenguaje natural implementadas en el chatbot, con el objetivo de mejorar su capacidad de contextualización de preguntas en el ámbito académico.
- Mantener una investigación continua para estar al tanto de nuevas metodologías, herramientas y casos de estudio, asegurando así que el chatbot evolucione con las mejores prácticas y enfoques más actuales.
- Regularmente evaluar nuevas tecnologías y actualizaciones en procesamiento de lenguaje natural y bases de datos vectoriales para asegurarse de que el chatbot aproveche las innovaciones más recientes y mantenga su especialización.
- Establecer un protocolo de pruebas periódicas en la página de Facebook y recopilar retroalimentación de información para ajustar de manera proactiva la configuración y mejorar la eficiencia del chatbot
- Aplicar una encuesta de satisfacción para determinar la aceptación del chatbot por parte de la comunidad académica y público en general.

BIBLIOGRAFÍA

- [1] R. M. Alkoudmani, G. S. Ooi, y M. L. Tan, «Implementing a Chatbot on Facebook to Reach and Collect Data from Thousands of Healthcare Providers: PharmindBot as a Case», *Journal of the American Pharmacists Association*, vol. 63, n.º 5, pp. 1634-1642, sep. 2023, doi: 10.1016/j.japh.2023.06.007.
- [2] R. Manjarres y M. Echeverri-Torres, «Asistente virtual académico utilizando tecnologías cognitivas de procesamiento de lenguaje natural», *Revista Politécnica*, vol. 16, pp. 85-95, may 2020, doi: 10.33571/rpolitec.v16n31a7.
- [3] M. D. Illescas-Manzano, N. Vicente López, N. Afonso González, y C. Cristofol Rodríguez, «Implementation of Chatbot in Online Commerce, and Open Innovation», *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 7, n.º 2, p. 125, jun. 2021, doi: 10.3390/joitmc7020125.
- [4] M. Adam, M. Wessel, y A. Benlian, «AI-based chatbots in customer service and their effects on user compliance», *Electron Markets*, vol. 31, n.º 2, pp. 427-445, jun. 2021, doi: 10.1007/s12525-020-00414-7.
- [5] E. Nhái, M. Lee, M. Luo, P. Chan, y T. Liang, «An intelligent knowledge-based chatbot for customer service-Web of Science Core Collection», vol. 50, nov. 2021, doi: 10.1016/j.elerap.2021.101098.
- [6] T. T. Nguyen, A. D. Le, H. T. Hoang, y T. Nguyen, «NEU-chatbot: Chatbot for admission of National Economics University», *Computers and Education: Artificial Intelligence*, vol. 2, p. 100036, ene. 2021, doi: 10.1016/j.caeai.2021.100036.
- [7] Elsevier B.V., «Scopus - Analyze search results», www-scopus-com.basesdedatos.utmachala.edu.ec. Accedido: 13 de agosto de 2023. [En línea]. Disponible en: <https://www-scopus-com.basesdedatos.utmachala.edu.ec/term/analyzer.uri?sort=plf-f&src=s&sid=1ce79f34d5960d3562a26f29282d1177&sot=a&sdt=a&cluster=scolang%2c%22English%22%2ct&sl=35&s=TITLE-ABS-KEY%28%22CHATBOTS%22+AND+%22NLP%22%29&origin=resultslist&count=10&analyzeResults=Analyze+results>
- [8] E. Adamopoulou y L. Moussiades, «Chatbots: History, technology, and applications», *Machine Learning with Applications*, vol. 2, p. 100006, dic. 2020, doi: 10.1016/j.mlwa.2020.100006.
- [9] M. Haenlein y A. Kaplan, «A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence», *California Management Review*, vol. 61, n.º 4, pp. 5-14, ago. 2019, doi: 10.1177/0008125619864925.
- [10] «Historia de ChatGPT: una cronología del ascenso meteórico de los chatbots generativos de IA | HatumSEO». Accedido: 27 de julio de 2023. [En línea]. Disponible en: <https://hatumseo.com/noticias/historia-de-chatgpt-una-cronologia-del-ascenso-meteorico-de-los-chatbots-generativos-de-ia/>
- [11] Khan, Sadia *et al.*, «The Essence of Software Engineering Framework-based Model for an Agile Software Development Method», *International Journal of Advanced Computer Science and Applications*, vol. 14, n.º 7, pp. 802-811, 2023, doi: 10.14569/IJACSA.2023.0140788.
- [12] S. Khan *et al.*, «Latest Transformations of XP Process Model: A Systematic Literature Review», *International Journal of Computer Science and Network Security*, vol. 21, n.º 6, pp. 143-150, jun. 2021, doi: 10.22937/IJCSNS.2021.21.6.19.
- [13] Sari, Hasrini y Adinda, Rachmasari, «Examining Customer Experience in Using a Chatbot-Web of Science Core Collection», *INTERNATIONAL JOURNAL OF*

- ASIAN BUSINESS AND INFORMATION MANAGEMENT*, vol. 14, n.º 1, pp. 1-16, sep. 2023, doi: 10.4018/IJABIM.322438.
- [14] Johari, Nurul Muizzah, Nohuddin, Puteri N. E., Baharin, Ahmad Hanif Ahmad, Yakob, Noor Azuddin, y Ebadi, Mohammad Javad, «Features requirement elicitation process for designing a chatbot application-Web of Science Core Collection», nov. 2022, doi: 10.1049/ntw2.12071.
- [15] G. A. Santos, G. G. de Andrade, G. R. S. Silva, F. C. M. Duarte, J. P. J. D. Costa, y R. T. de Sousa, «A Conversation-Driven Approach for Chatbot Management», *IEEE Access*, vol. 10, pp. 8474-8486, ene. 2022, doi: 10.1109/ACCESS.2022.3143323.
- [16] F. Rosa, D. Buccicardi, A. Romano, F. Borda, M. C. D'Auria, y A. Gastaldo, «Artificial intelligence and pelvic fracture diagnosis on X-rays: a preliminary study on performance, workflow integration and radiologists' feedback assessment in a spoke emergency hospital», *European Journal of Radiology Open*, vol. 11, dic. 2023, doi: 10.1016/j.ejro.2023.100504.
- [17] R. Gupta, D. Srivastava, M. Sahu, S. Tiwari, R. K. Ambasta, y P. Kumar, «Artificial intelligence to deep learning: machine intelligence approach for drug discovery», *Mol Divers*, vol. 25, n.º 3, pp. 1315-1360, ago. 2021, doi: 10.1007/s11030-021-10217-3.
- [18] Y. Duan, J. S. Edwards, y Y. K. Dwivedi, «Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda», *International Journal of Information Management*, vol. 48, pp. 63-71, oct. 2019, doi: 10.1016/j.ijinfomgt.2019.01.021.
- [19] Z. M. Najmi, «Artificial Intelligence: A Comprehensive and Systematic Literature Review of Applications and Comparative Technologies», *International Journal of Computer and Information Engineering*, vol. 17, n.º 7, pp. 402-420, jul. 2023.
- [20] Y. Li, M. A. Thomas, y D. Liu, «From semantics to pragmatics: where IS can lead in Natural Language Processing (NLP) research», *Eur. J. Inform. Syst.*, vol. 30, n.º 5, pp. 569-590, sep. 2021, doi: 10.1080/0960085X.2020.1816145.
- [21] D. Khurana, A. Koli, K. Khatter, y S. Singh, «Natural language processing: state of the art, current trends and challenges», *Multimed. Tools Appl.*, vol. 82, n.º 3, pp. 3713-3744, ene. 2023, doi: 10.1007/s11042-022-13428-4.
- [22] C. Chen, Y. Liu, M. Kumar, J. Qin, y Y. Ren, «Energy consumption modelling using deep learning embedded semi-supervised learning», *Computers & Industrial Engineering*, vol. 135, pp. 757-765, sep. 2019, doi: 10.1016/j.cie.2019.06.052.
- [23] T. Hagendorff, «Linking Human And Machine Behavior: A New Approach to Evaluate Training Data Quality for Beneficial Machine Learning», *Minds & Machines*, vol. 31, n.º 4, pp. 563-593, dic. 2021, doi: 10.1007/s11023-021-09573-8.
- [24] V. Sorin *et al.*, «Large language model (ChatGPT) as a support tool for breast tumor board», *npj Breast Cancer*, vol. 9, n.º 1, p. 44, may 2023, doi: 10.1038/s41523-023-00557-8.
- [25] B. Swathi, S. S. Babu, y M. Ayyavaraiah, «Artificial Intelligence: Characteristics, Subfields, Techniques and Future Predictions», *J. Mech. Contin. Math. Sci.*, vol. 14, n.º 6, pp. 127-135, dic. 2019, doi: 10.26782/jmcms.2019.12.00010.
- [26] M. Ashfaq, J. Yun, S. Yu, y S. M. C. Loureiro, «I, Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents», *Telematics and Informatics*, vol. 54, p. 101473, nov. 2020, doi: 10.1016/j.tele.2020.101473.

- [27] B. Sheehan, H. S. Jin, y U. Gottlieb, «Customer service chatbots: Anthropomorphism and adoption», *Journal of Business Research*, vol. 115, pp. 14-24, jul. 2020, doi: 10.1016/j.jbusres.2020.04.030.
- [28] M. Crespo Miguel y B. Domínguez Cabrera, «Perspectivas de las tecnologías de Chatbot y su aplicación a las entrevistas de evaluación del lenguaje», *Pragmalinguística*, n.º Monográfico 2, pp. 100-113, 2020, doi: 10.25267/Pragmalinguística.2020.iextra2.06.
- [29] P. Smutny y P. Schreiberova, «Chatbots for learning: A review of educational chatbots for the Facebook Messenger», *Comput. Educ.*, vol. 151, p. 103862, jul. 2020, doi: 10.1016/j.compedu.2020.103862.
- [30] S. Yu, Y. Chen, y H. Zaidi, «AVA: A Financial Service Chatbot Based on Deep Bidirectional Transformers», *Frontiers in Applied Mathematics and Statistics*, vol. 7, 2021, Accedido: 19 de julio de 2023. [En línea]. Disponible en: <https://www.frontiersin.org/articles/10.3389/fams.2021.604842>
- [31] S. Hwang y J. Kim, «Toward a Chatbot for Financial Sustainability», *Sustainability*, vol. 13, n.º 6, Art. n.º 6, ene. 2021, doi: 10.3390/su13063173.
- [32] E. Adamopoulou y L. Moussiades, «An Overview of Chatbot Technology», en *Artificial Intelligence Applications and Innovations*, I. Maglogiannis, L. Iliadis, y E. Pimenidis, Eds., en IFIP Advances in Information and Communication Technology. Cham: Springer International Publishing, 2020, pp. 373-383. doi: 10.1007/978-3-030-49186-4_31.
- [33] B. A. González-Beltrán, M. A. Vázquez-García, J. A. Reyes-Ortiz, y R. García-Ruiz, «Una revisión de chatbots en la salud», 2022.
- [34] S. Roca, J. Sancho, J. García, y Á. Alesanco, «Microservice chatbot architecture for chronic patient support», *Journal of Biomedical Informatics*, vol. 102, p. 103305, feb. 2020, doi: 10.1016/j.jbi.2019.103305.
- [35] Y. Kumar, P. Morreale, P. Sorial, J. Delgado, J. J. Li, y P. Martins, «A Testing Framework for AI Linguistic Systems (testFAILS)», en *2023 IEEE International Conference On Artificial Intelligence Testing (AITest)*, jul. 2023, pp. 51-54. doi: 10.1109/AITest58265.2023.00017.
- [36] Lim, Zhi Wei *et al.*, «Benchmarking large language models' performances for myopia care: a comparative analysis of ChatGPT-3.5, ChatGPT-4.0, and Google Bard-Web of Science Core Collection», *4*, vol. 95, n.º 104770, sep. 2023, doi: 10.1016/j.ebiom.2023.104770.
- [37] Y. Li, Z. Li, K. Zhang, R. Dan, S. Jiang, y Y. Zhang, «ChatDoctor: A Medical Chat Model Fine-Tuned on a Large Language Model Meta-AI (LLaMA) Using Medical Domain Knowledge», *Cureus*, vol. 15, n.º 6, p. e40895, doi: 10.7759/cureus.40895.
- [38] Hugo Touvron *et al.*, «Llama 2: Open Foundation and Fine-Tuned Chat Models», CONVERSATIONAL AI. [En línea]. Disponible en: https://scontent.fmch2-1.fna.fbcdn.net/v/t39.2365-6/10000000_662098952474184_2584067087619170692_n.pdf?_nc_cat=105&ccb=1-7&_nc_sid=3c67a6&_nc_ohc=fXYJ34-K-LEAX_GxkYq&_nc_ht=scontent.fmch2-1.fna&oh=00_AfAmuBuVTYCB9P4rBvQp4gH9vSBwL9rz8ogS6fq6AA9LCw&oe=6591FFBF
- [39] Diana Cheung, «Meta Llama 2 vs. OpenAI GPT-4: A Comparative Analysis of an Open-Source vs. Proprietary LLM». Accedido: 27 de diciembre de 2023. [En línea]. Disponible en: <https://www.codesmith.io/blog/meta-llama-2-vs-openai-gpt-4-a-comparative-analysis-of-an-open-source-vs-proprietary-llm>

- [40] «Llama 2: Open Foundation and Fine-Tuned Chat Models | Research - AI at Meta». Accedido: 27 de diciembre de 2023. [En línea]. Disponible en: <https://ai.meta.com/research/publications/llama-2-open-foundation-and-fine-tuned-chat-models/>
- [41] J. Insuasti, F. Roa, y C. M. Zapata-Jaramillo, «Computers' Interpretations of Knowledge Representation Using Pre-Conceptual Schemas: An Approach Based on the BERT and Llama 2-Chat Models», *BDCC*, vol. 7, n.º 4, p. 182, dic. 2023, doi: 10.3390/bdcc7040182.
- [42] Pinecone Systems, Inc., «Pinecone», Pinecone. Accedido: 25 de septiembre de 2023. [En línea]. Disponible en: <https://docs.pinecone.io/docs/openai>
- [43] W. H. Walters y E. I. Wilder, «Fabrication and errors in the bibliographic citations generated by ChatGPT», *Sci Rep*, vol. 13, n.º 1, p. 14045, sep. 2023, doi: 10.1038/s41598-023-41032-5.
- [44] OpenAI Platform, «OpenAI Introduction», OpenAI Platform. Accedido: 15 de noviembre de 2023. [En línea]. Disponible en: <https://platform.openai.com>
- [45] Python Software Foundation, «El tutorial de Python», Python documentation. Accedido: 25 de septiembre de 2023. [En línea]. Disponible en: <https://docs.python.org/3/tutorial/index.html>
- [46] PythonAnywhere LLP, «Sobre nosotros: PythonAnywhere». Accedido: 15 de noviembre de 2023. [En línea]. Disponible en: https://www.pythonanywhere.com/about/company_details/
- [47] M. Bonney *et al.*, «Development of a digital twin operational platform using Python Flask», *Cambridge University Press*, vol. 3, n.º e1, jun. 2022, doi: 10.1017/dce.2022.1.
- [48] B. Ahmed, «Flask vs Django: Elijamos Tu Próximo Framework Python», Kinsta®. Accedido: 15 de noviembre de 2023. [En línea]. Disponible en: <https://kinsta.com/es/blog/flask-vs-django/>
- [49] «Hacia la selección de una metodología adecuada de gestión de proyectos de TI - ProQuest», ProQuest. Accedido: 19 de julio de 2023. [En línea]. Disponible en: <https://www.proquest.com/openview/20e2980cb405f5cbbf96869d912d58d3/1?pq-origsite=gscholar&cbl=1006393>
- [50] P. Lean, «¿Qué es la metodología Agile y por qué está de moda?», Progressa Lean. Accedido: 19 de julio de 2023. [En línea]. Disponible en: <https://www.progressalean.com/metodologia-agile/>
- [51] R. Hightower, W. Onstine, y P. Visan, *Professional Java Tools for Extreme Programming: Ant, XDoclet, JUnit, Cactus, and Maven*. Wiley, 2004. Accedido: 22 de julio de 2023. [En línea]. Disponible en: <https://elibro.net/es/ereader/utmachala/180424>
- [52] M. Holcombe, *Running an Agile Software Development Project*. Wiley, 2008. Accedido: 23 de julio de 2023. [En línea]. Disponible en: <https://elibro.net/es/ereader/utmachala/178468?page=36>
- [53] Sinnaps, «Metodología XP o Programación Extrema: ¿Qué es y cómo aplicarla?», Sinnaps. Accedido: 23 de julio de 2023. [En línea]. Disponible en: <https://www.sinnaps.com/blog-gestion-proyectos/metodologia-xp>

ANEXOS

Anexo 1: Matriz de consistencia

Tabla 40. Anexo 1

Problema, objeto y campo	Objetivo	Marco Teórico	Hipótesis	Variables	Metodología
<p>Problema: ¿Cómo brindar soporte en los servicios de información en la carrera de Tecnologías de la Información de la Universidad Técnica de Machala?</p> <p>Problemas específicos (Opcional):</p> <ul style="list-style-type: none"> - ¿Cuáles son las principales necesidades y requerimientos de los estudiantes de Tecnología de la Información (TI) de la Universidad Técnica de Machala en cuanto a la asesoría y orientación? - ¿Cómo se puede desarrollar un prototipo de chatbot que brinde asesoría y orientación a los estudiantes de la carrera de TI, utilizando técnicas de procesamiento de lenguaje natural? - ¿Cuáles son las métricas e indicadores adecuados para evaluar la precisión del chatbot en la asesoría y orientación de los estudiantes de TI en la página de Facebook de la carrera? <p>Objeto de estudio:</p> <ul style="list-style-type: none"> - Respuestas automatizadas de inquietudes de 	<p>Objetivo General:</p> <ul style="list-style-type: none"> - Desarrollar un chatbot utilizando técnicas de procesamiento del lenguaje natural para soporte en los servicios de información requeridos por los estudiantes de la carrera de Tecnología de la Información de la Universidad Técnica de Machala y público en general. <p>Objetivos Específicos:</p> <ul style="list-style-type: none"> - Indagar trabajos relacionados respecto a chatbots y NPL, tecnologías y herramientas para el desarrollo del chatbot. - Recabar las principales necesidades y requerimientos de la carrera de TI respecto a servicios de información. - Programar el prototipo de chatbot para soporte en los servicios de información de estudiantes de la carrera de TI y público en general. - Integrar el prototipo de chatbot a la página de Facebook de la carrera de TI. - Evaluar el chatbot en la página de Facebook de la carrera de TI. 	<p>Antecedentes históricos a nivel internacional y nacional del objeto, campo: Historia y Evolución de los chatbot</p> <p>Fundamentos Teóricos de objeto, campo y variables:</p> <p>Inteligencia Artificial Definición Evolución Tecnologías y/o modelo Procesamiento de Lenguaje Natural (NLP) Modelos de Deep Learning Aprendizaje Automático (Machine Learning) Modelo de Lenguaje Largo (LLM)</p> <p>Chatbots Definición Características Tipos de chatbots Arquitectura de chatbots Modelos de chat optimizado</p> <p>Herramientas LangChain Pinecone GPT-3.5 OpenIA PythonAnywhere Flask</p> <p>Metodología de desarrollo Metodología tradicional Metodologías ágiles Metodologías híbridas</p>	<p>Hipótesis General:</p> <ul style="list-style-type: none"> - La implementación de un chatbot basado en técnicas de procesamiento del lenguaje natural (NLP) para la carrera de Tecnologías de la Información generará respuestas con un 80% de aciertos al aprovechar la capacidad de comprensión contextual y análisis semántico de las consultas <p>Hipótesis específicas o preguntas científicas (opcional):</p>	<p>Variable 1 / Independiente:</p> <ul style="list-style-type: none"> • Desarrollo de chatbot basado en MLL para la carrera de TI. <p>Dimensiones o categorías:</p> <ul style="list-style-type: none"> • Precisión de Respuesta. • Pruebas del chatbot <p>Variable 2/ Dependiente:</p> <ul style="list-style-type: none"> • Eficiencia en el acceso de los usuarios a los servicios de información relevante de la carrera de TI. <p>Dimensiones o categorías:</p> <ul style="list-style-type: none"> • Precisión de respuestas • Pruebas del chatbot. 	<p>Enfoque: Cuantitativo</p> <p>Alcance: Descriptivo</p> <p>Diseño: Cuasi - experimental</p> <p>Unidades de análisis: Población: 400 (Estudiantes de la carrera de TI de la universidad Técnica de Machala)</p> <p>Muestra: 197 usuarios</p> <p>Técnicas e instrumentos de recolección de datos: Matriz de confusión</p> <p>Análisis de documentos o datos o procesos.</p> <p>Técnicas de procesamiento de datos: Procesamiento de Lenguaje Natural</p>

<p>estudiantes y público sobre información de la carrera de Tecnología de la Información de la Universidad Técnica de Machala</p> <p>Campo de Acción:</p> <ul style="list-style-type: none"> - Chatbot con procesamiento del lenguaje natural (NLP), implementado en Facebook - Aprendizaje Automático (Machine Learning) - Modelo de Lenguaje Largo (MLL) 					
---	--	--	--	--	--

Anexo 2: Fragmentar datos

```
def fragmentar(data, chunk_size=800):
    from langchain.text_splitter import RecursiveCharacterTextSplitter
    text_splitter = RecursiveCharacterTextSplitter(chunk_size=chunk_size, chunk_overlap=100)
    fragmentos = text_splitter.split_documents(data)
    return fragmentos
```

Figura 54. Anexo 2

Anexo 3: Creando embeddings o vectores

```
def creando_vectores(index_name, fragmentos):

    borrar_indices(index_name)

    embeddings = OpenAIEmbeddings()

    pinecone.init(api_key=os.environ.get('PINECONE_API_KEY'),
                 environment=os.environ.get('PINECONE_ENV'))

    if index_name in pinecone.list_indexes():
        print(f'El índice {index_name} ya existe. Cargando los embeddings ... ', end='')
        vectores = Pinecone.from_existing_index(index_name, embeddings)
        print('Ok')
    else:
        print(f'Creando el índice {index_name} y los embeddings ...', end='')
        pinecone.create_index(index_name, dimension=1536, metric='cosine')
        vectores = Pinecone.from_documents(fragmentos, embeddings, index_name=index_name)
        print('Ok')

    return vectores
```

Figura 55. Anexo 3

Anexo 4: Consultas

```
def consultas(vectores2, pregunta):  
    from langchain.chains import RetrievalQA  
    from langchain.chat_models import ChatOpenAI  
  
    llm = ChatOpenAI(model='gpt-3.5-turbo', temperature=1)  
  
    retriever = vectores2.as_retriever(search_type='similarity', search_kwargs={'k': 3})  
  
    chain = RetrievalQA.from_chain_type(llm=llm, chain_type="stuff", retriever=retriever)  
  
    answer = chain.run(pregunta)  
    return answer
```

Figura 56. Anexo 4

Anexo 4: Preguntas de Evaluación

Evaluación del chatbot preguntas por cada criterio

Perfil de la Carrera

1. ¿Quién es el coordinador de la carrera de TI?
2. ¿Cuál es el título que se obtiene?
3. ¿Es cierto que la carrera tiene 151 créditos?
4. ¿Qué conocimientos y habilidades voy a adquirir en la carrera de TI?
5. ¿Al graduarme de esta carrera puedo trabajar como director de TI de alguna empresa?
6. ¿Es necesario realizar el servicio comunitario?
7. ¿Cuál es el objetivo de la carrera?
8. ¿Por qué se recomienda estudiar la carrera?
9. ¿Es necesario cumplir con todas las horas de prácticas laborales para graduarme?
10. Cumplir con todos los niveles de ingles no es necesario
11. ¿Cuáles son las horas de la jornada?
12. ¿Dónde se encuentra ubicada?
13. ¿Cuál es la misión de la carrera?
14. ¿Cuál es la visión de la carrera?
15. ¿Al estudiar esta carrera podre trabajar en organizaciones públicas y privadas?
16. ¿La carrera de ti no pertenece a la facultad de ciencia sociales?
17. La fecha de fundación de la carrera no es el 7 de febrero del 2017
18. Es correcto que la modalidad de la carrera de ti es presencial
19. La carrera de TI no dura 4 años
20. ¿Es correcto que la carrera perteneció a la escuela de informática?

Malla Curricular

1. ¿Cuántas asignaturas tiene el primer semestre?
2. La malla curricular consta de 9 semestre
3. ¿En qué semestre se cursa la asignatura Tecnologías Web?
4. ¿Cuántas materias hay por semestre?
5. ¿En noveno semestre se ve Titulación I?
6. ¿El número de materias que se ven por semestre varía?
7. ¿Cuántas horas de prácticas preprofesionales se realizan en el séptimo semestre?
8. ¿Cuántos semestres son en total?

9. ¿Como está conformada la malla curricular?
10. ¿Cuáles son las asignaturas de 10mo semestre?
11. ¿Cuántos semestres son en la carrera de TI?
12. ¿Administración de Base de datos no se ve en octavo semestre?
13. ¿Sistemas Operativos se cursa en segundo semestre?
14. ¿Nombra las asignaturas de 9no?
15. ¿En qué semestre se cursa Titulación II?
16. ¿Sexto semestre está conformado por 6 materias?
17. ¿Nombra las asignaturas de 7mo ?
18. ¿En qué semestre veo Métodos Numéricos?
19. ¿En el séptimo semestre se aborda la materia de Programación Móvil?
20. ¿La asignatura de Probabilidad y Estadística se debe cursar en el tercer semestre?

Estructura organizativa

1. ¿Quién es el actual coordinador de la carrera de Tecnologías de la Información?
2. ¿Cuál es el cargo del Ing. Fausto Redrován?
3. ¿Quiénes son docentes de la carrera de Tecnologías de la Información con el título de Doctor?
4. ¿Quién es el decano de la Facultad de Ingeniería Civil?
5. ¿Quién es la vicerrectora administrativa de la Universidad Técnica de Machala?
6. El Ing. Freddy Leonardo Espinoza Urgilés es el decano de la Facultad de Ingeniería Civil
7. ¿La Ing. Irene Sánchez González es la rectora de la Universidad?
8. ¿Quién es el jefe de la UMMOG en la Facultad de Ingeniería Civil?
9. ¿Quién es el vicerrector de Investigación, Vinculación y Posgrados?
10. ¿El Ing. Wilmer Rivas es docente de la carrera de Ingeniería Civil?
11. ¿Quién rector de la UTMACH?
12. No existe subdecano en la facultad
13. ¿El Ing. Eduardo Tusa Jumbo es docente de la carrera de TI?
14. ¿La Dra. Rosemary Samaniego Ocampo es la vicerrectora administrativa?
15. ¿La Ing Maritza Pinta forma parte de la planta docente de la carrera de TI?
16. ¿Cuántos docentes tiene la carrera?
17. ¿Listar los docentes que tiene la carrera?
18. El Ing. Joofre Honores es coordinador de la carrera

19. El Decano de la Facultad de Ingeniería civil es el Lic. César Peñaherrera
20. ¿Cuántos PhD hay en la lista de maestros de la carrera de TI?

Preguntas Frecuentes

21. ¿Cuál es la extensión de secretaria?
22. ¿Cuál es correo de la carrera?
23. ¿Cómo hacer un cambio de paralelo?
24. ¿Cuál es el teléfono de la facultad?
25. ¿Cómo justifico las faltas?
26. gratuita?
27. ¿El proceso de admisión no requiere de una prueba?
28. ¿La carrera de TI dispone de un correo?
29. ¿Cuántos paralelos tiene la carrera?
30. ¿Para el proceso de matrícula solo necesito mi título de bachiller?
31. ¿Cuál es el número de teléfono de secretaria?
32. ¿Cuál es la extensión de coordinación?
33. ¿Cuál es el número de teléfono del coordinador?
34. ¿La extensión de la UMMOG es 432?
35. ¿El proceso de justificar insistencia varía?
36. ¿Cómo se solicita el retiro voluntario?
37. ¿Qué es SIGEDA?
38. ¿El pago de matrícula es a través de transferencias electrónicas o tarjetas?
39. ¿Cuál es la red oficial de la facultad?
40. ¿La gratuidad se aplica al estudiar más carreras?

Anexo 5: Desarrollo de la Evaluación

Tabla 41. Anexo 5

Perfil de la Carrera	Respuesta Chatbot	Respuesta Correcta	
1. ¿Quién es el coordinador de la carrera de TI?	El coordinador actual de la carrera de Tecnologías de la Información es el Ing. Fausto Redrován, Mg.	Ing. Fausto Redrován, Mg	VP
2. ¿Cuál es el título que se obtiene?	El título que se obtiene al graduarse de la carrera de Tecnologías de la Información es el de Ingeniero/a en Tecnologías de la Información.	Ingeniero/a en Tecnologías de la Información.	VP
3. ¿Es cierto que la carrera tiene 151 créditos?	No, según la información proporcionada, la carrera de Tecnologías de la Información tiene 150 créditos.	La carrera de Tecnologías de la Información tiene 150 créditos	VN
4. ¿Qué conocimientos y habilidades voy a adquirir en la carrera de TI?	Diseño, desarrollo e implementación de soluciones informáticas para la gestión de datos y procesos. Uso de tecnología de redes y bases de datos. Habilidades para utilizar tecnologías actuales y futuras de manera adecuada y/o vanguardista. Solución de problemas relacionados con la gestión de la información. Coordinación de procesos de planificación, adquisición, implementación, integración y operación de hardware, dispositivos de comunicación y software. Conocimientos en programación, bases de datos, redes de datos, matemáticas, sistemas	Diseño, desarrollo e implementación de soluciones informáticas para la gestión de datos y procesos. Uso de tecnología de redes y bases de datos. Habilidades para utilizar tecnologías actuales y futuras de manera adecuada y/o vanguardista. Solución de problemas relacionados con la gestión de la información. Coordinación de procesos de planificación, adquisición, implementación, integración y operación de hardware, dispositivos de comunicación y software. Conocimientos en programación, bases de datos, redes de datos, matemáticas, sistemas	VP

	web, interacción hombre-máquina, electricidad y electrónica, infraestructura de tecnologías, seguridad y aseguramiento de la información	web, interacción hombre-máquina, electricidad y electrónica, infraestructura de tecnologías, seguridad y aseguramiento de la información	
5. ¿Al graduarme de esta carrera puedo trabajar como director de TI de alguna empresa?	Sí, al graduarte de la carrera de Tecnologías de la Información podrías tener la posibilidad de trabajar como director de TI de alguna empresa.	SI	VP
6. ¿Es necesario realizar el servicio comunitario?	ND	Si es necesario cumplir con un total de 96 horas de Servicio comunitario	FN
7. ¿Cuál es el objetivo de la carrera?	El objetivo de la carrera de Tecnologías de la Información es formar profesionales capacitados en el campo de la tecnología de la información, con competencias para gestionar y optimizar los procesos organizacionales utilizando principios de ingeniería y estándares de calidad.	Formar profesionales y emprendedores con capacidades científico-técnicas y humanísticas, que gestionen eficientemente la información desarrollando soluciones tecnológicas con principios de ingeniería y estándares de calidad	VP
8. ¿Por qué se recomienda estudiar la carrera?	Se recomienda estudiar la carrera de Tecnologías de la Información porque te permitirá resolver problemas relacionados con la gestión de la información, especializándote en tecnología de redes y bases de datos. Además, adquirirás las habilidades necesarias para utilizar tecnologías actuales y futuras, lo que te permitirá proponer y materializar soluciones integrales a los requerimientos de las organizaciones. Esta carrera te brindará conocimientos y	porque resolverás problemas relacionados con la gestión de la información especializándote en la tecnología de redes y bases de datos y desarrollando las habilidades necesarias para usar las tecnologías actuales y futuras, adecuadas y/o de vanguardia para proponer y materializar soluciones con una visión integral de los requerimientos de las organizaciones	VP

	habilidades para diseñar, desarrollar e implementar soluciones informáticas para la gestión de datos y procesos.		
9. ¿Es necesario cumplir con todas las horas de practicas laborales para graduarme?	Sí, es necesario cumplir con todas las horas de prácticas laborales para graduarse.	Si	VP
10. Cumplir con todos los niveles de ingles no es necesario	No es necesario cumplir con todos los niveles de inglés para aprobar la carrera de Tecnologías de la Información.	Si es necesario cumplir con los niveles de ingles	FP
11. ¿Cuáles son las horas de la jornada?	El horario de clases es desde las 7:30 am hasta las 13:30.	7:30 am hasta las 13:30	VP
12. ¿Dónde se encuentra ubicada?	La carrera de Tecnologías de la Información se encuentra ubicada en el Campus Central de la Universidad Técnica de Machala, en la Avenida Panamericana Km. 5 1/2 Vía a Pasaje	Avenida Panamericana Km. 5 1/2 Vía a Pasaje	VP
13. ¿Cuál es la misión de la carrera?	La misión de la carrera de Tecnologías de la Información es formar profesionales emprendedores con capacidades científico-técnicas y humanísticas, que gestionen eficientemente la información, desarrollando soluciones tecnológicas con principios de ingeniería y estándares de calidad, comprometidos con el desarrollo socioeconómico de la población en su área de influencia.	Formar profesionales emprendedores con capacidades científico-técnicas y humanísticas, que gestionen eficientemente la información, desarrollando soluciones tecnológicas con principios de ingeniería y estándares de calidad, comprometidos con el desarrollo socioeconómico de la población en su área de influencia	VP
14. ¿Cuál es la visión de la carrera?	La visión de la carrera de Tecnologías de la Información de la Universidad Técnica De Machala es ser líder	Ser líder regional y nacional en la formación de profesionales de Tecnologías de la Información, y reconocida	VP

	regional y nacional en la formación de profesionales de Tecnologías de la Información.	por su excelencia académica e investigación	
15. ¿Al estudiar esta carrera podre trabajar en organizaciones públicas y privadas?	Sí, al estudiar la carrera de Tecnologías de la Información podrás trabajar tanto en organizaciones públicas como privadas. La formación en esta carrera te capacitará para gestionar la infraestructura tecnológica y los sistemas de información, habilidades que son necesarias en todos los sectores y áreas de interés público. Tanto las organizaciones públicas como las privadas, sin importar su tamaño o su objetivo	Si	VP
16. ¿La carrera de ti no pertenece a la facultad de ciencia sociales?	No, la carrera de Tecnologías de la Información no pertenece a la Facultad de Ciencias Sociales. Según la información proporcionada, la carrera de TI pertenece a la Facultad de Ingeniería Civil.	No, la carrera de Tecnologías de la Información pertenece a la Facultad de Ingeniería Civil	VN
17. La fecha de fundacion de la carrera no es el 7 de febrero del 2017	La fecha de fundación de la carrera de Tecnologías de la Información de la Universidad Técnica de Machala es el 8 de febrero de 2017.	No, fa fecha de fundación de la carrera de Tecnologías de la Información es el 8 de febrero de 2017	VN
18. Es correcto que la modalidad de la carrera de ti es presencial	Sí, es correcto. La modalidad de la carrera de Tecnologías de la Información es presencial.	Si	VP

19. La carrera de TI no dura 4 años	No, La duración de la carrera de Tecnologías de la Información en la Universidad Técnica de Machala es de 5 años (10 semestres).	No, Dura 5 años	VN
20. Es correcto que la carrera perteneció a la escuela de informática?	Sí, es correcto. La carrera de Tecnologías de la Información perteneció a la Escuela de Informática en la Universidad Técnica de Machala	SI	VP

Tabla 42. Anexo 5

Malla curricular	Respuesta Chatbot	Respuesta Correcta	
¿Cuántas asignaturas tiene el primer semestre?	El primer semestre tiene 5 asignaturas.	5 asignaturas	VP
La malla curricular consta de 9 semestre	No, la malla curricular consta de 10 semestres.	La malla curricular consta de 10 semestres	VN
¿En qué semestre se cursa la asignatura Tecnologías Web?	La asignatura "Tecnologías Web" se cursa en el quinto semestre.	5to semestre	VP
¿Cuántas materias hay por semestre?	En cada semestre, hay un total de 5 materias.	5 materias por semestre	VP
¿En noveno semestre se ve Titulación I?	Sí, en el noveno semestre se ve la materia de Seminario de Titulación I.	Si se ve	VP
¿El número de materias que se ven por semestre varía?	Sí, el número de materias que se ven por semestre varía. En el primer semestre se ven 5 materias, en el segundo semestre se ven 5 materias, en el tercer semestre se ven 5 materias, en el cuarto semestre se ven 5	Son cinco materias por semestre	FP

	materias, en el quinto semestre se ven 5 materias, en el sexto semestre se ven 5 materias, en el séptimo semestre se ven 5 materias, en el octavo semestre se ven 5 materias, y en el noveno semestre se ven 3 materias.		
¿Cuántas horas de prácticas preprofesionales se realizan en el séptimo semestre?	En el séptimo semestre se realizan 96 horas de prácticas preprofesionales.	96 en el septimo	VP
¿Cuántos semestres son en total?	En total, hay nueve semestres en la carrera.	Son 10 semestres	FN
¿Como está conformada la malla curricular?	La malla curricular de la carrera de Tecnologías de la Información (TI) está conformada por asignaturas o materias desde el primer hasta el décimo semestre de la carrera. Puedes revisar el contenido detallado de la malla curricular accediendo al siguiente enlace: https://app.utmachala.edu.ec/siutmach/public/administrativo/creardisenomalla/fac/62/esc/41/car/551/dis/460/red/2	La malla curricular de la carrera de Tecnologías de la Información (TI) está conformada por asignaturas o materias desde el primer hasta el décimo semestre	VP
¿Cuáles son las asignaturas de 10mo semestre?	Las asignaturas del décimo semestre son: Seminario de Titulación II, Gestión de TI II, Inteligencia Artificial, Normativas en TI y Auditoría de TI.	Las asignaturas del décimo semestre son: Seminario de Titulación II, Gestión de TI II, Inteligencia Artificial, Normativas en TI y Auditoría de TI.	VP
¿Cuántos semestres son en la carrera de TI?	La carrera de Tecnologías de la Información tiene una duración de 10 semestres.	Tiene 10 semestre	VP

¿Administración de Base de datos no se ve en octavo semestre?	No, la materia de Administración de Bases de Datos se encuentra en el quinto semestre, no en el octavo semestre.	No, Administración de Bases de datos se ve en quinto semestre.	VN
¿Sistemas Operativos se cursa en segundo semestre?	No, Sistemas Operativos se cursa en el tercer semestre.	Se cursa en tercero, septimo y octavo semestre	VN
¿Nombra las asignaturas de 9no?	Seminario de Titulación I, Sistemas de Información Gerencial.	Seminario de Titulación I, Sistemas de Información Gerencial, Gestión de Ti I, Gestión de Calidad en Ti, y Digital Business.	FN
¿En qué semestre se cursa Titulación II?	La materia de Titulación II se cursa en el décimo semestre.	Décimo semestre	VP
¿Sexto semestre está conformado por 6 materias?	No, el sexto semestre está conformado por 5 materias.	No	VN
¿Nombra las asignaturas de 7mo ?	Las asignaturas de 7mo semestre son: Gestión de Proyectos de Software, Seguridad de Sistemas de Bases de Datos, Inteligencia de Negocios, Sistemas Operativos II, Tecnologías de Conmutación II.	Gestión de Proyectos de Software, Seguridad de Sistemas de Bases de Datos, Inteligencia de Negocios, Sistemas Operativos II, Tecnologías de Conmutación II.	VP
¿En que semestre veo Métodos Numéricos?	Métodos Numéricos se ve en el quinto semestre	Quinto semestre	VP
¿En el séptimo semestre se aborda la materia de Programación Móvil?	No, en el séptimo semestre no se aborda la materia de Programación Móvil.	No	VN
¿La asignatura de Probabilidad y Estadística se debe cursar en el tercer semestre?	No, la asignatura de Probabilidad y Estadística se debe cursar en el cuarto semestre.	Cuarto Semestre	VN

Tabla 43. Anexo 5

Estructura organizativa	Respuesta Chatbot	Respuesta Correcta	
1. ¿Quién es el actual coordinador de la carrera de Tecnologías de la Información?	El actual coordinador de la carrera de Tecnologías de la Información es el Ing. Fausto Redrován, Mg.	Ing. Fausto Redrován, Mg.	VP
2. ¿Cuál es el cargo del Ing. Fausto Redrován?	El Ing. Fausto Redrován es el coordinador de la carrera de Tecnologías de la Información.	Es el coordinador de la carrera de Tecnologías de la Información.	VP
3. ¿Quiénes son docentes de la carrera de Tecnologías de la Información con el título de Doctor?	Los docentes de la carrera de Tecnologías de la Información con el título de Doctor son: - Ing. Eduardo Tusa Jumbo, Phd. - Ing. Dixys Leonardo Hernández Rojas, Phd. - Ing. Bertha Mazón Olivo, Phd. - Ing. Maritza Pinta, Phd. - Ing. Wilmer Rivas Asanza, Phd. Mg.	Ing. Eduardo Tusa Jumbo, Phd. Ing. Dixys Leonardo Hernández Rojas, Phd. Ing. Bertha Mazón Olivo, Phd. Ing. Maritza Pinta, Phd. Ing. Wilmer Rivas Asanza, Phd. Mg.	VP
4. ¿Quién es el decano de la Facultad de Ingeniería Civil?	El decano de la Facultad de Ingeniería Civil es el Arq. Luisana Campusano Vera, Mg. Sc.	Arq. Luisana Campusano Vera, Mg. Sc.	VP
5. ¿Quién es la vicerrectora administrativa de la Universidad Técnica de Machala?	La vicerrectora administrativa de la Universidad Técnica de Machala es la Ing. Irene Sánchez González, Mgs	Ing. Irene Sánchez González, Mgs	VP
6. El Ing. Freddy Leonardo Espinoza Urgilés es el decano de la Facultad de Ingeniería Civil	No sé quién es el decano de la Facultad de Ingeniería Civil.	El decano de la Facultad de Ingeniería Civil es el Arq. Luisana Campusano Vera, Mg. Sc.	FN
7. ¿La Ing. Irene Sánchez González es la rectora de la Universidad?	No, el rector de la Universidad Técnica de Machala es el Dr. Jhonny Pérez Rodríguez. La Ing. Irene Sánchez González es la vicerrectora administrativa.	No	VP

8. ¿Quién es el jefe de la UMMOG en la Facultad de Ingeniería Civil?	El jefe de la UMMOG en la Facultad de Ingeniería Civil es el Lic. César Peñaherrera.	Lic. César Peñaherrera	VP
9. ¿Quién es el vicerrector de Investigación, Vinculación y Posgrados?	El vicerrector de Investigación, Vinculación y Posgrados es el Dr. Luis Brito Gaona, Phd.	Dr. Luis Brito Gaona, Phd.	VP
10. ¿El Ing. Wilmer Rivas es docente de la carrera de Ingeniería Civil?	No, el Ing. Wilmer Rivas no es docente de la carrera de Ingeniería Civil.	No	VN
11. ¿Quién rector de la UTMACH?	El rector de la UTMACH es el Dr. Jhonny Pérez Rodríguez, Phd.	Dr. Jhonny Pérez Rodríguez, Phd.	VP
12. No existe subdecano en la facultad	No se menciona específicamente si existe un subdecano en la facultad de Ingeniería Civil. No puedo confirmar si hay un subdecano o no basándome en la información proporcionada.	Si, si existe es el Ing. Freddy Leonardo Espinoza Urgilés	FP
13. ¿El Ing. Eduardo Tusa Jumbo es docente de la carrera de TI?	Sí, el Ing. Eduardo Tusa Jumbo es docente de la carrera de Tecnologías de la Información.	Si	VP
14. ¿La Dra. Rosemary Samaniego Ocampo es la vicerrectora administrativa?	No, la Dra. Rosemary Samaniego Ocampo es la vicerrectora académica, no la vicerrectora administrativa.	No	VN
15. ¿La Ing Maritza Pinta forma parte de la planta docente de la carrera de TI?	Sí, la Ing. Maritza Pinta forma parte de la planta docente de la carrera de Tecnologías de la Información.	Si	VP
16. ¿Cuántos docentes tiene la carrera?	La carrera de Tecnologías de la Información cuenta con un total de 13 docentes.	14 docentes	FN
17. ¿Listar los docentes que tiene la carrera?	Los docentes de la carrera Tecnologías de la Información son:	Son: - Ing. Fausto Redrován, Mg.	VP

	<ul style="list-style-type: none"> - Ing. Fausto Redrován, Mg. - Ing. Eduardo Tusa Jumbo, Phd. - Ing. Dixys Leonardo Hernández Rojas, Phd. - Ing. Bertha Mazón Olivo, Phd. - Ing. Maritza Pinta, Phd. - Ing. Mariuxi Zea Ordóñez, Mg. - Ing. Luis Loján Cueva, Mgs. - Ing. Wilmer Rivas Asanza, Phd. Mg - Ing. Rodrigo Morocho Román, Mgs. - Ing. Milton Valarezo Pardo, Mg. Sc. - Ing. Nancy Loja Mora, Mg. Sc. - Ing. Joffre Cartuche Calva, Mgs. - Ing. Joofre Honores Tapia, Mgs. - Ing. Johnny Novillo Vicuña, Mgs. 	<ul style="list-style-type: none"> - Ing. Eduardo Tusa Jumbo, Phd. - Ing. Dixys Leonardo Hernández Rojas, Phd. - Ing. Bertha Mazón Olivo, Phd. - Ing. Maritza Pinta, Phd. - Ing. Mariuxi Zea Ordóñez, Mg. - Ing. Luis Loján Cueva, Mgs. - Ing. Wilmer Rivas Asanza, Phd. Mg - Ing. Rodrigo Morocho Román, Mgs. - Ing. Milton Valarezo Pardo, Mg. Sc. - Ing. Nancy Loja Mora, Mg. Sc. - Ing. Joffre Cartuche Calva, Mgs. - Ing. Joofre Honores Tapia, Mgs. - Ing. Johnny Novillo Vicuña, Mgs. 	
18. El Ing. Joofre Honores es coordinador de la carrera	No, el Ing. Fausto Redrován es el coordinador actual de la carrera de Tecnologías de la Información. No se menciona que el Ing. Joofre Honores sea el coordinador.	No	VN
19. El Decano de la Facultad de Ingeniería civil es el Lic. César Peñaherrera	No, el Decano de la Facultad de Ingeniería Civil es la Arq. Luisana Campusano Vera, Mg. Sc.	No	VN
20. ¿Cuántos PhD hay en la lista de maestros de la carrera de TI?	En la lista de maestros de la carrera de Tecnologías de la Información, hay un total de 5 profesores con el título de PhD.	5 Profesores	VP

Tabla 44. Anexo 5

Preguntas Frecuentes	Respuesta Chatbot	Respuesta Correcta	
¿Cuál es la extensión de secretaria?	La extensión de la Secretaría de la Facultad de Ingeniería Civil es 314.	Es 314	VP
¿Cuál es correo de la carrera?	El correo de la carrera de Tecnologías de la Información es fic_ing.tecnologias@utmachala.edu.ec.	fic_ing.tecnologias@utmachala.edu.ec.	VP
¿Cómo hacer un cambio de paralelo?	Para hacer un cambio de paralelo, debes solicitarlo mediante un oficio dirigido al Coordinador de la carrera.	Solicitar el cambio de paralelo mediante oficio dirigido al Coordinador.	VP
¿Cuál es el teléfono de la facultad?	El teléfono de la facultad es 2981860 – 2983895 – 2983710.	2981860, 2983895 y 2983710.	VP
¿Como justifico las faltas?	<p>Si te encuentras dentro de las 48 horas desde tu última falta, puedes justificar directamente con el docente enviando el respectivo justificativo. De lo contrario, debes seguir el siguiente proceso:</p> <p>1. Realiza una solicitud de INFORME DE FAVORABILIDAD dirigida al Lic. Jorge Villacis Salcedo, DIRECTOR DE LA UNIDAD DE BIENESTAR UNIVERSITARIO. Para ello, debes hacer un oficio, imprimirlo, firmarlo, escanearlo y enviarlo al correo electrónico ube@utmachala.edu.ec adjuntando una copia de tu cédula y el respectivo documento que certifique las razones de tu</p>	<p>Si se encuentra dentro de las 48 horas puede justificar directamente con el docente enviado el respectivo justificativo caso contrario realizar el siguiente proceso</p> <p>Basado en el Art. 176 del Régimen Académico de la UTMACH.</p> <p>Realizar una solicitud de INFORME DE FAVORABILIDAD dirigida al Lic. Jorge Villacis Salcedo DIRECTOR DE LA UNIDAD DE BIENESTAR UNIVERSITARIO. Para ello debe hacer un oficio, imprimirlo, firmarlo, escanearlo y enviarlo al correo electrónico ube@utmachala.edu.ec adjuntando una copia de su cédula y el respectivo documento que certifique</p>	VP

	<p>inasistencia.</p> <p>2. Una vez que recibas la respuesta de la Unidad de Bienestar Universitario, debes realizar un oficio dirigido al Sr. Decano de la Facultad de Ingeniería Civil, solicitando que por su digno intermedio y el Consejo Directivo de la Facultad se proceda a la justificación de la inasistencia. Adjunta el informe otorgado por la Unidad de Bienestar Universitario y una copia de tu cédula. Esta solicitud debe enviarla al correo electrónico fic_decanato@utmachala.edu.ec.</p>	<p>las razones de su inasistencia.</p> <p>Una vez que reciban la respuesta de la Unidad de Bienestar Universitario debe realizar un oficio dirigido al Sr Decano; Ing. Nombres y Apellidos DECANO DE LA FACULTAD DE INGENIERÍA CIVIL solicitando que por su digno intermedio y el Consejo Directivo de la Facultad se proceda a la justificación de la inasistencia. Adjuntar el informe otorgado por la Unidad de Bienestar Universitario y copia de cédula. Esta solicitud debe enviarla al correo electrónico fic_decanato@utmachala.edu.ec</p>	
gratuita?	<p>La carrera de Tecnologías de la Información (TI) es gratuita para estudiantes regulares que estén cursando carreras de nivel técnico y tecnológico, así como sus equivalentes de grado a nivel universitario. Esta gratuidad abarca las primeras matrículas y la escolaridad en todas las asignaturas, cursos o equivalentes de carrera, desde la nivelación hasta la obtención del título. Sin embargo, si ya posees un título universitario de otra carrera, cursar la carrera de TI requerirá el pago correspondiente de todas las asignaturas, cursos o equivalentes de carrera, desde la nivelación hasta</p>	<p>es gratuita para estudiantes regulares que estén cursando carreras de nivel técnico y tecnológico, así como sus equivalentes de grado a nivel universitario. Esta gratuidad abarca las primeras matrículas y la escolaridad en todas las asignaturas, cursos o equivalentes de carrera, desde la nivelación hasta la obtención del título. Cabe destacar que esta oferta se aplica a una sola carrera por estudiante y únicamente una vez para aquellos que cambien de carreras. En caso de que ya poseas o tengas un título universitario de otra carrera, cursar esta carrera requerirá el pago</p>	VP

	la obtención del título. Además, existen casos especiales en los que se pueden aplicar pagos de matrícula, como perder la gratuidad de una asignatura o no cumplir con el 60% de créditos en el nivel respectivo de matrícula.	correspondiente de la todas las asignaturas, cursos o equivalentes de carrera, desde la nivelación hasta la obtención del título	
¿El proceso de admisión no requiere de una prueba?	El proceso de admisión a la carrera de Ingeniería en Tecnologías de la Información puede incluir una prueba de admisión de capacidades y competencias. Sin embargo, también existe la posibilidad de pasar un curso de admisión para ingresar a la carrera. Por lo tanto, es posible que se requiera una prueba, pero también puede haber otras opciones de ingreso.	Si requiere	FN
¿La carrera de Ti dispone de un correo?	No se menciona en el texto si la carrera de Tecnologías de la Información dispone de un correo específico.	fic_ing.tecnologias@utmachala.edu.ec.	FN
cuantos paralelos tiene la carrera?	Actualmente, la carrera tiene 15 paralelos. En el primer año, hay 3 paralelos. Desde el segundo hasta el cuarto año, hay paralelos identificados como A y B. A partir del quinto año hasta el décimo año, hay un único paralelo designado como A.	Actualmente tenemos 15 paralelos. En el primer año contamos con 3 paralelos. Para los grados de segundo a cuarto, tenemos paralelos identificados como A y B. Desde quinto hasta décimo año, mantenemos un único paralelo designado como A.	VP
Para el proceso de matrícula solo necesito mi título de bachiller?	No, además del título de bachiller, también se deben subir otros documentos como la copia de la cédula, certificado de votación, una copia notariada y refrendada del título de	Copia de cédula, certificado de votación, copia notariada y refrendada del título de bachiller o del acta de grado dichos documentos deben ser en formato pdf y una foto a color tamaño	VN

	bachiller o del acta de grado, una foto a color tamaño carnet con fondo blanco, y en caso de aplicar, pasaporte vigente con visa de estudiante extranjero, carnet de discapacidad, Resolución del Consejo Directivo de aprobación del Informe de Reconocimiento u Homologación de estudios. Todos estos documentos deben estar en formato PDF.	carnet con fondo blanco en formato jpg. En caso de que aplique: Pasaporte vigente con visa de estudiante extranjero, carnet de discapacidad, Resolución del Consejo directivo de aprobación del Informe de Reconocimiento u Homologación de estudios	
¿Cuál es el número de teléfono de secretaria?	El número de teléfono de la Secretaria de decano/decanato de la Facultad de Ingeniería Civil es 2981860 – 2983895 – 2983710 extensión 314.	El número de teléfono de la Secretaria es 2981860 – 2983895 – 2983710 extensión 314.	VP
¿Cuál es la extensión de coordinación?	No se menciona la extensión de coordinación en el texto proporcionado.	Extensión 403.	FN
¿Cuál es el número de teléfono del coordinador?	El número de teléfono del coordinador de la carrera de Tecnologías de la información es 2981860 – 2983895 – 2983710 extensión 403.	2981860 – 2983895 – 2983710 extensión 403.	VP
La extensión de la UMMOG es 432?	No, la extensión de la UMMOG es 430, no 432.	la extensión de la UMMOG es 430	VN
¿El proceso de justificar insistencia varia?	No, el proceso para justificar la inasistencia a clases no varía según la información proporcionada. Se debe seguir el proceso descrito en el Art. 176 del Régimen Académico de la UTMACH, que incluye enviar una solicitud de informe de favorabilidad al Director de la Unidad	Si se encuentra dentro de las 48 horas puede justificar directamente con el docente enviado el respectivo justificativo caso contrario realizar el siguiente proceso Basado en el Art. 176 del Régimen Académico de la UTMACH. Realizar una solicitud de	FN

	de Bienestar Universitario.	<p>INFORME DE FAVORABILIDAD dirigida al Lic. Jorge Villacis Salcedo DIRECTOR DE LA UNIDAD DE BIENESTAR UNIVERSITARIO. Para ello debe hacer un oficio, imprimirlo, firmarlo, escanearlo y enviarlo al correo electrónico ube@utmachala.edu.ec adjuntando una copia de su cédula y el respectivo documento que certifique las razones de su inasistencia.</p> <p>Una vez que reciban la respuesta de la Unidad de Bienestar Universitario debe realizar un oficio dirigido al Sr Decano; Ing. Nombres y Apellidos DECANO DE LA FACULTAD DE INGENIERÍA CIVIL solicitando que por su digno intermedio y el Consejo Directivo de la Facultad se proceda a la justificación de la inasistencia. Adjuntar el informe otorgado por la Unidad de Bienestar Universitario y copia de cédula. Esta solicitud debe enviarla al correo electrónico fic_decanato@utmachala.edu.ec</p>	
¿Cómo se solicita el retiro voluntario?	Para solicitar el retiro voluntario de la carrera, se debe realizar mediante un oficio dirigido al Decano de la Facultad.	Solicitar el retiro voluntario se lo realiza mediante un oficio dirigido al Decano de la Facultad.	VP

<p>¿Qué es SIGEDA?</p>	<p>SIGEDA es el Sistema de Gestión de Documentos y Archivos de la Universidad Técnica de Machala (UTMACH). Es una unidad administrativa que se encarga de administrar los documentos físicos y electrónicos de la universidad. Permite el ingreso de trámites en formato digital y agiliza los procesos de seguimiento.</p>	<p>es el Sistema de Gestión de Documentos y Archivos de la Universidad Técnica de Machala (UTMACH). Es una unidad administrativa que administra los documentos físicos y electrónicos de la universidad. Permite el ingreso de trámites en formato digital y agiliza los procesos de seguimiento</p>	<p>VP</p>
<p>¿El pago de matrícula es a través de transferencias electrónicas o tarjetas?</p>	<p>No es posible realizar el pago de matrícula a través de transferencias electrónicas o tarjetas de crédito/débito. Debe realizar el depósito del valor exacto en la cuenta sugerida por la Universidad en las ventanillas del BANCO DE MACHALA.</p>	<p>No es posible recibir pagos de transferencias electrónicas o tarjetas de crédito/débito.</p>	<p>VN</p>
<p>¿Cual es la red oficial de la facultad?</p>	<p>La red oficial de la Facultad de Ingeniería Civil es https://www.facebook.com/uaic.utmach.</p>	<p>https://www.facebook.com/uaic.utmach.</p>	<p>VP</p>
<p>la gratuidad se aplica al estudiar más carreras?</p>	<p>No, la gratuidad solo se aplica a una sola carrera por estudiante. Si decides cambiar de carrera, deberás pagar por todas las asignaturas, cursos o equivalentes de carrera, desde la nivelación hasta la obtención del título.</p>	<p>Se aplica a una sola carrera por estudiante</p>	<p>VN</p>