

**Escuela Superior Politécnica del Litoral**

**Facultad de Ingeniería en Electricidad y Computación**

Desarrollo de un asistente virtual para mejorar el acceso a la información  
académica y administrativa de la FIEC - ESPOL.

TECH-328

**Proyecto Integrador**

Previo la obtención del Título de:

**Nombre de la titulación**

**Ingeniero en Ciencias de la Computación**

Presentado por:

Johan Jair Gilces Reyes

Eduardo Damian Castillo Suarez

Guayaquil - Ecuador

Año: 2023 - 2024

## Dedicatoria

---

A mi querida madre, Roxana, y a mi hermana, Daniela, cuya motivación y amor incondicional han sido la luz guiadora en cada paso de este camino.

A Gabriela, Jean, Joshua y Daniela, gracias por ser el apoyo inquebrantable y mi colchón emocional.

**Johan Jair Gilces Reyes**

Dedico este proyecto a mis padres, hermano y a mi novia, cuyo apoyo inquebrantable, experiencia y motivación han sido fundamentales para mi desarrollo profesional. Sin ellos, este proyecto y mi formación como profesional no habrían sido posibles.

**Eduardo Damian Castillo Suarez**

## Agradecimientos

---

Mi más sincero agradecimiento a todos los que han contribuido a este proyecto.

A mi tutor Allan Avendaño, cuya dirección y consejos sabios han sido fundamentales en cada fase de este proyecto. No solo me ha guiado académicamente, sino que también ha enriquecido mi crecimiento personal y profesional. También, mi gratitud a mis profesores Eduardo Cruz y Carlos Mera, les agradezco por su inspiradora enseñanza y su incansable apoyo.

Un agradecimiento especial a los clubes Argumentum y Taws, por ser fundamentales en mi desarrollo.

Finalmente, a mi grupo de la carrera, Free Computagod, gracias por ser excelentes compañeros y amigos, compartiendo desafíos y éxitos juntos.

**Johan Jair Gilces Reyes**

## Agradecimientos

---

Expreso mi más sincero agradecimiento a mis padres por su inquebrantable apoyo en cada paso que me ha llevado hasta este momento, así como por su comprensión y apoyo incondicional en todo momento.

Agradezco a mi novia, Ana Muñoz, por ser un pilar de apoyo constante a lo largo de mi etapa como estudiante. También deseo agradecer a todos los profesores y compañeros que me han enseñado y guiado para convertirme en un profesional capacitado.

**Eduardo Damian Castillo Suarez**

## Declaración Expresa

---

Nosotros, Johan Jair Gilces Reyes y Eduardo Damian Castillo Suarez acordamos y reconocemos que:

La titularidad de los derechos patrimoniales de autor (derechos de autor) del proyecto de graduación corresponderá al autor o autores, sin perjuicio de lo cual la ESPOL recibe en este acto una licencia gratuita de plazo indefinido para el uso no comercial y comercial de la obra con facultad de sublicenciar, incluyendo la autorización para su divulgación, así como para la creación y uso de obras derivadas. En el caso de usos comerciales se respetará el porcentaje de participación en beneficios que corresponda a favor del autor o autores. La titularidad total y exclusiva sobre los derechos patrimoniales de patente de invención, modelo de utilidad, diseño industrial, secreto industrial, software o información no divulgada que corresponda o pueda corresponder respecto de cualquier investigación, desarrollo tecnológico o invención realizada por mí/nosotros durante el desarrollo del proyecto de graduación, pertenecerán de forma total, exclusiva e indivisible a la ESPOL, sin perjuicio del porcentaje que me/nos corresponda de los beneficios económicos que la ESPOL reciba por la explotación de mi/nuestra innovación, de ser el caso.

En los casos donde la Oficina de Transferencia de Resultados de Investigación (OTRI) de la ESPOL comunique los autores que existe una innovación potencialmente patentable sobre los resultados del proyecto de graduación, no se realizará publicación o divulgación alguna, sin la autorización expresa y previa de la ESPOL.

Guayaquil, 1 de enero del 2024.



Firmado electrónicamente por:  
**JOHAN JAIR  
GILCES REYES**

Johan Gilces

Eduardo Castillo

## **Evaluadores**

---

**Ph.D. Lucia Marisol Villacres Falconi**

Profesor de la Materia

---

**Msc. Allan Avendaño, MSIA**

Tutor de proyecto

## Resumen

La Facultad de Ingeniería en Electricidad y Computación (FIEC) de la ESPOL aborda el desafío de mejorar la distribución y accesibilidad de la información en su plataforma digital. Se implementó un asistente virtual basado en modelos de lenguaje de gran escala para simplificar la comunicación y el acceso a la información, con el objetivo de mejorar la experiencia del usuario. Se realizaron pruebas con diversos Large Language Models (LLM), utilizando técnicas de rendimiento de modelos grandes. Se llevó a cabo un proceso de scraping en la página web de FIEC para recopilar la información relevante. Los prototipos se evaluaron utilizando LangChain y LlamaIndex para el aprendizaje del modelo acerca de FIEC. Se probaron los modelos GPT, Llama2 y Bloom, siendo GPT y Llama2 los más efectivos. Se desarrollaron dos chatbots con estos modelos, según las necesidades y preferencias. Se realizaron encuestas que revelaron una experiencia satisfactoria para los usuarios, con un 54% muy satisfechos, un 40% más rápido en la obtención de información en comparación con la página web de FIEC, y una precisión del 96.4% en las respuestas.

**Palabras Clave:** Grande modelos de Lenguaje, Llama 2, GPT, LangChain, Retrieval Augmented Generation

## **Abstract**

The School of Electrical and Computer Engineering (FIEC) at ESPOL faces a challenge in distributing and accessing information on its digital platform. Despite data availability, the site's structure is intricate due to multiple pages and sections, hindering smooth navigation. The objective is to simplify communication and information access at ESPOL, providing users with a friendlier and more productive experience. An AI-based virtual assistant was implemented using large-scale language models, utilizing techniques for model performance enhancement. Web scraping was conducted on FIEC's webpage to gather pertinent information. Various Large Language Models (LLM) were tested to find the most suitable for the problem. Prototypes were evaluated using LangChain and LlamaIndex for model learning about FIEC. GPT, Llama2, and Bloom models were tested, with GPT and Llama2 proving most effective. Thus, two chatbots were developed with these models, as per chatbot needs or preferences. Surveys verified the chatbot models, yielding a highly satisfactory user experience with 54% very satisfied, 40% faster information retrieval compared to the FIEC webpage, and 96.4% accuracy in responses.

**Keywords:** Large Language Models, Llama 2, GPT, LangChain , Retrieval Augmented Generation



## Índice General

|  |     |
|--|-----|
| Resumen.....   | I   |
| Abstract.....  | II  |
| Abreviaturas.....  | VI  |
| Índice de Ilustraciones.....                                 | VII |
| Índice de tablas.....  | VII |
| Capítulo 1.....  | 1   |
| 1. Introducción.....   | 2   |
| 1.1 Descripción del Problema.....                            | 2   |
| 1.2 Justificación del Problema.....                          | 3   |
| 1.3 Objetivos.....   | 6   |
| 1.3.1 Objetivo general.....                                  | 6   |
| 1.3.2 Objetivos específicos.....                             | 6   |
| 1.4 Marco teórico.....                                       | 6   |
| 1.4.1 Problemas con la información en las universidades..... | 6   |
| 1.4.2 Otras soluciones.....                                  | 7   |
| 1.4.3 Ciencias de Datos.....                                 | 8   |
| Capítulo 2.....  | 11  |
| 2. Metodología.....  | 12  |
| 2.1 Investigación preliminar.....                            | 12  |
| 2.1.1 Expectativas del cliente.....                          | 12  |
| 2.1.2 Perspectiva de los usuarios.....                       | 13  |
| 2.1.3 Evaluación de soluciones tecnológicas.....             | 13  |

|       |   |    |
|-------|---|----|
| 2.2   | Diseño Tecnológico.....                           | 14 |
| 2.2.1 | Selección del modelo LLM.....                     | 14 |
| 2.2.2 | Selección de frameworks .....                     | 15 |
| 2.2.3 | Hiperparámetros .....                             | 15 |
| 2.3   | Obtención y validación de la información .....    | 16 |
| 2.4   | Pruebas de modelos de lenguaje y frameworks ..... | 17 |
| 2.5   | Implementación de modelos.....                    | 18 |
| 2.5.1 | Caso de Uso.....                                  | 18 |
| 2.5.2 | Diagrama de eventos .....                         | 19 |
| 2.5.3 | Arquitectura de RAG .....                         | 20 |
| 2.5.4 | Arquitectura del prototipo .....                  | 21 |
| 2.5.5 | Prototipo .....                                   | 22 |
| 2.6   | Pruebas de usuario.....                           | 24 |
| 2.6.1 | Métricas de evaluación.....                       | 24 |
| 2.6.2 | Plan de desarrollo .....                          | 24 |
|       | Capítulo 3.....                                   | 26 |
| 3.    | Análisis y resultados.....                        | 27 |
| 3.1   | Experimentos desarrollados .....                  | 27 |
| 3.2   | Entorno de Desarrollo .....                       | 27 |
| 3.3   | Parámetros del RAG.....                           | 27 |
| 3.4   | Resultados LangChain.....                         | 27 |
| 3.5   | Resultados Llama Index .....                      | 31 |
| 3.6   | Comparación Arquitecturas.....                    | 32 |

|       |                                     |    |
|-------|-------------------------------------|----|
| 3.7   | Pruebas de usuario.....             | 33 |
| 3.7.1 | Tiempos de respuesta .....          | 34 |
| 3.7.2 | Exactitud de las respuestas .....   | 35 |
| 3.7.3 | Satisfacción de las respuestas..... | 36 |
|       | Capítulo 4.....                     | 38 |
| 4.    | Conclusiones y recomendaciones..... | 39 |
| 4.1   | Conclusiones .....                  | 39 |
| 4.2   | Recomendaciones.....                | 40 |
| 5.    | Referencias.....                    | 42 |
|       | ANEXOS .....                        | 45 |
|       | Anexo A .....                       | 46 |
|       | Anexo B.....                        | 50 |
|       | Anexo C.....                        | 51 |

## **Abreviaturas**

|       |  |
|-------|--|
| ESPOL | Escuela Superior Politécnica del Litoral             |
| ODS   | Objetivos de Desarrollo Sostenible                   |
| FIEC  | Facultad de Ingeniería en Electricidad y Computación |
| LLM   | Large Language Model                                 |
| RAG   | Retrieval Augmented Generation                       |

## Índice de Ilustraciones

|   |    |
|---|----|
| Figura 1 Pasos de Metodología.....              | 12 |
| Figura 2 Diagrama Casos de uso.....             | 19 |
| Figura 3 Diagrama de Eventos.....               | 20 |
| Figura 4 Arquitectura de RAG.....               | 21 |
| Figura 5 Arquitectura del Prototipo .....       | 22 |
| Figura 6 Interfaz gráfica del prototipo 1 ..... | 22 |
| Figura 7 Interfaz gráfica del prototipo 2 ..... | 23 |
| Figura 8 Resultados LangChain 1 .....           | 29 |
| Figura 9 Resultados LangChain 2.....            | 30 |
| Figura 10 Resultados Llama Index 1 .....        | 31 |
| Figura 11 Resultados Llama Index 2 .....        | 32 |
| Figura 12 Pruebas de usuario .....              | 34 |
| Figura 13 Resultados de Encuestas 1 .....       | 35 |
| Figura 14 Resultados de Encuestas 2.....        | 36 |

## Índice de tablas

|   |    |
|---|----|
| Tabla 1 Métricas de evaluación .....            | 24 |
| Tabla 2 Preguntas de evaluación del modelo..... | 29 |

# Capítulo 1

## **1. Introducción**

### ***1.1 Descripción del Problema***

La comunicación efectiva es crucial en la era contemporánea, siendo necesario transmitir información de manera precisa, directa y clara. Aunque Internet ha facilitado significativamente este proceso, la comunicación a través de páginas web se ha vuelto menos efectiva comparada con otros medios. [1] Formatos como el video y el audio pueden ser más comprensibles que el texto, aunque pueden requerir más tiempo para transmitir grandes volúmenes de información.

En grandes organizaciones o instituciones académicas como la ESPOL, surge una problemática central respecto a la gestión de la información. Hay un volumen significativo de información crucial que debe estar accesible al público, y es de interés primordial que dicha información pueda ser accedida y comprendida de manera eficaz. Sin embargo, la tarea se complica ante la sobrecarga de información presente en diferentes aspectos, que van desde el contenido académico, direcciones, hasta protocolos o formatos de documentos específicos [2]. En un contexto de transición digital donde la información en mayor medida se encuentra disponible de manera digital en lugar de impreso, los sitios webs donde existe sobrecarga de información se refleja en “demasiado por hacer” [3] lo que desincentiva en gran medida a seguir explorando la información que necesito y buscar otras soluciones.

En grandes organizaciones o instituciones académicas como la ESPOL, surge una problemática central respecto a la gestión de la información. Hay un volumen significativo de información crucial que debe estar accesible al público, y es de interés primordial que dicha información pueda ser accedida y comprendida de manera eficaz. Sin embargo, la tarea se complica ante la sobrecarga de información presente en diferentes aspectos, que van desde el contenido académico, direcciones, hasta protocolos

o formatos de documentos específicos [2]. En un contexto de transición digital donde la información en mayor medida se encuentra disponible de manera digital en lugar de impreso, los sitios webs donde existe sobrecarga de información se refleja en “demasiado por hacer” [3] lo que desincentiva en gran medida a seguir explorando la información que necesito y buscar otras soluciones.

La Facultad de Ingeniería en Electricidad y Computación (FIEC) de la ESPOL enfrenta un desafío en la distribución y acceso a la información en su plataforma digital. A pesar de la disponibilidad de datos, la estructura del sitio resulta intrincada debido a las múltiples páginas y secciones por lo que obstaculiza una navegación fluida, en entrevistas con estudiantes de esta facultad se reveló que encuentran la página muy formal y difícil de interactuar. Esta situación no solo afecta a la comunidad politécnica, sino que es reflejo de una problemática más amplia que podría identificarse en sitios web similares de instituciones educativas y otras organizaciones. La consecuencia directa resulta en que los usuarios optan por asistencia personalizada para obtener información, lo cual incurre en pérdida de tiempo y en eficiencia.

Para abordar estos desafíos y mejorar la experiencia de los usuarios, proponemos desarrollar un asistente virtual en forma de chatbot. Este chatbot estará diseñado para responder preguntas y proporcionar información de manera rápida y eficiente, ayudando a los usuarios a encontrar lo que necesitan sin complicaciones. Nuestro objetivo es simplificar la comunicación y el acceso a la información en la ESPOL, brindando a los usuarios una experiencia más amigable y productiva.

## **1.2 Justificación del Problema**

La estructura de navegación en plataformas digitales desempeña un papel crucial al determinar cómo los usuarios acceden a las unidades de información, lo cual, a su vez, reduce el esfuerzo necesario para navegar desde el punto de entrada hasta una unidad



específica. Una estructura de navegación eficaz brinda indicaciones claras sobre la ubicación de los usuarios y las opciones disponibles para avanzar [4]. Esto es especialmente relevante en el ámbito educativo, donde una navegación eficaz y fluida es imperativa para garantizar un acceso sencillo a información significativa.

La estructura de navegación en plataformas digitales desempeña un papel crucial al determinar cómo los usuarios acceden a las unidades de información, lo cual, a su vez, reduce el esfuerzo necesario para navegar desde el punto de entrada hasta una unidad específica. Una estructura de navegación eficaz brinda indicaciones claras sobre la ubicación de los usuarios y las opciones disponibles para avanzar [4]. Esto es especialmente relevante en el ámbito educativo, donde una navegación eficaz y fluida es imperativa para garantizar un acceso sencillo a información significativa.

La Facultad de Ingeniería en Electricidad y Computación de la ESPOL, al igual que muchas otras instituciones educativas, tiene la responsabilidad de proporcionar información accesible y comprensible a su comunidad. Sin embargo, una estructura de sitio web percibida como compleja o desorganizada puede ser un obstáculo significativo para los usuarios, dando lugar a posibles frustraciones y pérdida de tiempo. La accesibilidad fácil es esencial, ya que los aprendices en línea deben poder acceder a la información con facilidad, sin necesidad de dedicar mucho tiempo a dominar los controles de navegación.

Las plataformas digitales en instituciones educativas se han diseñado para permitir a las personas acceder de manera rápida a información relevante como horarios, oficinas, y tarifas. Los avances en las tecnologías digitales han permitido a las escuelas y a las instituciones educativas comprender el valor de los datos exhaustivos sobre el rendimiento de los estudiantes y los instructores a medida que amplían el uso de aulas

virtuales, plataformas de e-learning y exámenes en línea. Sin embargo, la tecnología educativa no está exenta de dificultades, especialmente en su implementación y uso [5].

Las plataformas digitales en instituciones educativas se han diseñado para permitir a las personas acceder de manera rápida a información relevante como horarios, oficinas, y tarifas. Los avances en las tecnologías digitales han permitido a las escuelas y a las instituciones educativas comprender el valor de los datos exhaustivos sobre el rendimiento de los estudiantes y los instructores a medida que amplían el uso de aulas virtuales, plataformas de e-learning y exámenes en línea. Sin embargo, la tecnología educativa no está exenta de dificultades, especialmente en su implementación y uso [5].

En un entorno educativo en línea efectivo, una planificación de navegación clara y consistente es vital para crear un ambiente de aprendizaje manejable y acogedor. La navegación no solo sirve como un marco y guía sobre cómo se presentará la información a los aprendices, sino que también dirige a los aprendices a través del proceso de aprendizaje, permitiéndoles conocer su progreso.

En definitiva, mejorar la estructura de navegación en la plataforma digital de la FIEC contribuiría a una mejor experiencia de usuario y alinearía a la ESPOL con las prácticas recomendadas en la educación digital, optimizando el acceso y la comprensión de la información crucial. La comunicación precisa y eficaz de la información es esencial para mejorar el rendimiento académico, y la falta de precisión en la información puede resultar en tiempos de consulta largos y una mala gestión de los sistemas de información en la educación [6].

## **1.3 Objetivos**

### ***1.3.1 Objetivo general***

Implementar un asistente virtual basado en modelos de lenguaje de gran escala utilizando técnicas de rendimiento de modelos grandes, para la mejora en la obtención de información del sitio web de la FIEC.

### ***1.3.2 Objetivos específicos***

- Diseñar y desarrollar una plataforma web para la obtención de información precisa sobre la FIEC mediante el asistente virtual
- Diseñar una arquitectura basados en un modelo LLM para crear un asistente virtual que sea capaz de responder las dudas de manera coherente sobre FIEC
- Validar la experiencia del usuario con el asistente virtual mediante pruebas y retroalimentación, asegurando que la interacción con el chatbot sea satisfactoria y responda eficazmente a las consultas relacionadas con la FIEC.

## **1.4 Marco teórico**

### ***1.4.1 Problemas con la información en las universidades***

En la era digital, las instituciones educativas enfrentan desafíos significativos en términos de accesibilidad y distribución de información. Uno de los problemas predominantes es la accesibilidad web en las páginas universitarias. Se ha destacado que la falta de texto alternativo para imágenes es una cuestión de accesibilidad común en los sitios web, incluidos los de educación superior [7]. Según el estudio por Macakoğlu, que evaluó los sitios web de más de 2,000 colegios y universidades utilizando las Pautas de Accesibilidad para el Contenido Web (WCAG) 2.1, se encontró que el 96% de los sitios web auditados no cumplían con los estándares WCAG 2.1 Nivel AA, lo que indica barreras potenciales para la accesibilidad [8].

## **1.4.2 Otras soluciones**

### **1.4.2.1 Atención por personal administrativo.**

Las universidades tradicionalmente han utilizado secretarías para brindar información a los estudiantes y personal. Estas profesionales desempeñan un papel crucial en la organización de reuniones, programación de citas y gestión de información. Sin embargo, la automatización de oficinas y la utilización de tecnologías han influido en la efectividad y eficiencia de las secretarías, proporcionando herramientas para una gestión más eficaz de los recursos y la información. A pesar de la ayuda de la tecnología, la interacción humana directa sigue siendo esencial para resolver dudas complejas o proporcionar información específica.

### **1.4.2.2 Chatbots no cognitivos.**

Los chatbots no inteligentes, también conocidos como chatbots tradicionales o basados en reglas, son programas diseñados para interactuar con los usuarios siguiendo una lógica predefinida y estructurada, operan basados en flujos de conversación predefinidos y una lógica condicional "si-entonces" [9]. Estos chatbots tradicionales son capaces de responder preguntas específicas y realizar tareas definidas proporcionando respuestas preprogramadas basadas en las entradas del usuario. Sin embargo, su capacidad para entender e interpretar las entradas del usuario es limitada y no pueden manejar variaciones en la formulación de las preguntas o entrar en conversaciones más complejas y dinámicas [10].

Los chatbots no inteligentes son más simples y están diseñados con una programación básica que les permite ejecutar ciertas tareas de manera eficaz, aunque su interacción es bastante estática y limitada en comparación con los chatbots inteligentes.

### **1.4.2.3 Asistentes virtuales inteligentes.**

En la actualidad chatbots inteligentes como CHAT-GPT<sup>1</sup>, han demostrado ser herramientas muy buenas, según el Journal Technology in Society, los chatbots a la hora de comunicar información que sea difícil de acceder o entender, de manera concisa para el usuario, evitando pérdidas de tiempo o inconvenientes. Además de haber sido analizadas que su aceptación por el público no solo se debe por lo novedoso de los chatbots, sino también por sus eficiencia y experiencia con el usuario [11]. Estos chatbots proporcionan una comunicación concisa para el usuario, evitando pérdidas de tiempo o inconvenientes. Su aceptación por parte del público no solo se atribuye a su novedad, sino también a su eficiencia y a la experiencia positiva que proporcionan al usuario. A diferencia de los chatbots tradicionales, que son efectivos para escenarios orientados a tareas con reglas predefinidas, ChatGPT sobresale en conversaciones más dinámicas y abiertas, permitiendo una interacción más natural y enriquecedora [12].

### **1.4.3 Ciencias de Datos**

#### **1.4.3.1 Large Language Models.**

Los Modelos de Lenguaje de Gran Tamaño, conocidos como Large Language Models (LLMs), son una clase de algoritmos de inteligencia artificial diseñados para comprender y generar texto de forma autónoma. Estos modelos son entrenados con vastos conjuntos de datos lingüísticos, lo que les permite aprender patrones complejos en el lenguaje natural. Un ejemplo destacado de un LLM es GPT-3, desarrollado por OpenAI, que cuenta con 175 mil millones de parámetros, otorgándole una capacidad notable para comprender y generar texto en una variedad amplia de estilos y contextos.

---

<sup>1</sup><https://chat.openai.com>

Los LLMs han demostrado un rendimiento sobresaliente en una gama de tareas de procesamiento de lenguaje natural (NLP), convirtiéndose en una herramienta altamente buscada en la comunidad científica [13].

Los Modelos de Lenguaje de Gran Tamaño, conocidos como Large Language Models (LLMs), son una clase de algoritmos de inteligencia artificial diseñados para comprender y generar texto de forma autónoma. Estos modelos son entrenados con vastos conjuntos de datos lingüísticos, lo que les permite aprender patrones complejos en el lenguaje natural. Un ejemplo destacado de un LLM es GPT-3, desarrollado por OpenAI, que cuenta con 175 mil millones de parámetros, otorgándole una capacidad notable para comprender y generar texto en una variedad amplia de estilos y contextos. Los LLMs han demostrado un rendimiento sobresaliente en una gama de tareas de procesamiento de lenguaje natural (NLP), convirtiéndose en una herramienta altamente buscada en la comunidad científica [13].

Los modelos de lenguaje de gran tamaño han exhibido una capacidad notable para abordar eficazmente tareas de procesamiento de lenguaje natural, acercándose al rendimiento del estado del arte. Estas potentes herramientas se han demostrado eficaces en diversas aplicaciones, que van desde la generación de texto hasta la traducción automática y más, como se detalla en su preimpresión [14].

Los modelos de lenguaje de gran tamaño han exhibido una capacidad notable para abordar eficazmente tareas de procesamiento de lenguaje natural, acercándose al rendimiento del estado del arte. Estas potentes herramientas se han demostrado eficaces en diversas aplicaciones, que van desde la generación de texto hasta la traducción automática y más, como se detalla en su preimpresión [14].

### **1.4.3.2 Generación Aumentada por Recuperación (RAG)**

La Generación Aumentada por Recuperación (RAG, por sus siglas en inglés) es una técnica de procesamiento de lenguaje natural que combina modelos de recuperación y generación para mejorar la precisión y relevancia de las respuestas generadas, utilizando fuentes externas de información o conocimiento. RAG mejora las capacidades de un LLM como ChatGPT, añadiendo un sistema de recuperación de información que proporciona los datos necesarios para formular una respuesta. Este enfoque híbrido integra modelos de recuperación y modelos generativos para producir texto que es contextualmente preciso y rico en información. RAG ha sido fundamental en tareas intensivas de conocimiento en NLP, proporcionando un marco híbrido que integra modelos de recuperación y generativos para producir texto que no solo es contextualmente preciso sino también rico en información [15].

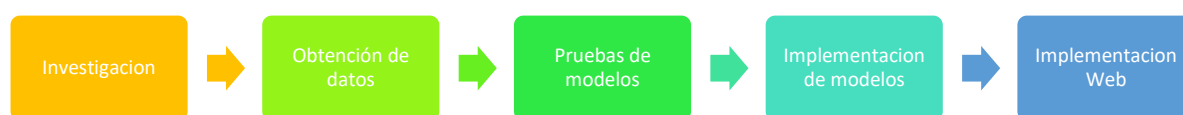
## **Capítulo 2**



Este capítulo detalla la metodología empleada para el desarrollo del chatbot, que facilitará el acceso a la información a través de la página web de la FIEC. Esta metodología comprende cinco etapas secuenciales, donde cada etapa subsiguiente se basa en los resultados y conclusiones de la anterior.

## 2. METODOLOGÍA

La metodología empleada consta en 5 etapas secuenciales, es decir las etapas posteriores dependen de las etapas pasadas, por ejemplo, la etapa Pruebas de Modelos depende de la etapa de Investigación y Scraping.



**Figura 1** Pasos de Metodología

### 2.1 Investigación preliminar

#### 2.1.1 *Expectativas del cliente*

En la fase inicial de este proyecto, se sostuvieron reuniones claves con el cliente, para comprender sus expectativas y necesidades. A través de estas discusiones, se identificó una demanda principal: la creación de una plataforma que mejore significativamente la interacción del usuario, específicamente orientada a resolver dudas y consultas relacionadas con la facultad. Esta necesidad subraya la importancia de un sistema interactivo y eficiente que facilite el acceso a información relevante para estudiantes y personal.

Una consideración importante en este proceso fue la familiaridad del cliente con soluciones basadas en machine learning, esta familiaridad orientó nuestra exploración hacia tecnologías avanzadas en este campo tomando en cuenta la limitación de recursos, tanto en términos de tiempo como económicos, por parte de la ESPOL.

### **2.1.2 *Perspectiva de los usuarios***

Con el objetivo de comprender profundamente las necesidades y preocupaciones de los usuarios finales, se llevaron a cabo entrevistas con estudiantes de la FIEC. Esta investigación reveló un desconocimiento notable sobre la existencia y el contenido de la página web de la facultad. Muchos estudiantes no estaban informados sobre datos fundamentales disponibles en el sitio, como la ubicación de las aulas de los profesores o la identidad del decano. Además, los que conocían la página enfrentaban dificultades en su navegación, lo que indicaba problemas en la accesibilidad y usabilidad del sitio web actual.

Ante estos hallazgos, los estudiantes propusieron la implementación de un sistema que se distinga por su facilidad de uso y la claridad en sus respuestas. Recomendaron incluir un conjunto de preguntas predeterminadas que faciliten el acceso a información relevante.

### **2.1.3 *Evaluación de soluciones tecnológicas***

Para la creación de un chatbot que ofrezca información contextualizada sobre la FIEC, se exploraron varias alternativas:

- Uso de modelos de lenguaje preentrenados.
- Creación de un Large Language Model (LLM).
- Plataformas para el desarrollo de chatbots no cognitivos.

- Aplicación de técnicas específicas de machine learning, como Retrieval-Augmented Generation (RAG), diseñadas para optimizar los modelos de generación de lenguaje.

Debido a las limitaciones de tiempo y recursos, reentrenar un modelo LLM o desarrollar uno desde cero resulta complicado. Los chatbots no cognitivos, aunque ofrecen respuestas estáticas y limitadas, presentaban una solución viable. Sin embargo, para maximizar la efectividad y relevancia de las respuestas, se optó por seleccionar un modelo LLM y utilizar RAG. Esta técnica permite al chatbot responder preguntas de manera precisa, extrayendo información directamente del sitio web de la FIEC. La elección de RAG se fundamenta en su capacidad para enriquecer las respuestas generadas por el LLM con información contextual relevante, lo que resulta esencial dada la naturaleza específica y dinámica de las consultas relacionadas con la FIEC.

## **2.2 Diseño Tecnológico**

Esta sección se enfoca en la selección del lenguaje de gran escala (LLM) , el framework de desarrollo y la configuración de hiperparámetros, todos ellos fundamentales para la implementación del sistema de chatbot con Retrieval-Augmented Generation (RAG).

### **2.2.1 Selección del modelo LLM**

Para la selección del modelo LLM de nuestro chatbot, se consideraron varias opciones teniendo en cuenta las limitaciones de recursos. Entre los modelos de código abierto y ejecutados localmente, destacó Llama2, desarrollado por Meta, por su disponibilidad gratuita y eficiencia en el uso de recursos. Además, se exploraron modelos Open Source adicionales, incluyendo Open Orca y Bloom, conocidos por su robustez y accesibilidad.

Paralelamente, se evaluaron alternativas de pago, como GPT-3.5 de OpenAI, que ofrecen capacidades avanzadas en el manejo del lenguaje. Sin embargo, la elección final de Llama2 se basó en su equilibrio entre rendimiento y accesibilidad, siendo una opción efectiva y sin costo para el proyecto.

### ***2.2.2 Selección de frameworks***

Para la implementación de la técnica Retrieval-Augmented Generation (RAG) en los modelos seleccionados, se hizo necesaria la elección de frameworks adecuados que pudieran integrarse de manera efectiva con los modelos de lenguaje de gran escala. Después de una revisión detallada de las opciones disponibles, se seleccionaron Langchain y Llamaindex como los frameworks más apropiados para nuestro proyecto.

Langchain y Llamaindex son frameworks especializados, diseñados para trabajar con modelos de lenguaje y facilitar la aplicación de técnicas avanzadas de machine learning como RAG. Estos frameworks fueron elegidos por su compatibilidad con los modelos LLM, su capacidad para integrar procesos de recuperación y generación de información, y su flexibilidad para adaptarse a los requisitos específicos de nuestro chatbot.

### ***2.2.3 Hiperparámetros***

La personalización de hiperparámetros juega un rol crucial en la optimización del rendimiento de RAG dentro de nuestro chatbot. Estos ajustes permiten afinar el modelo para adaptarlo a las necesidades específicas y las limitaciones de recursos del proyecto. Entre los hiperparámetros más significativos que se pueden ajustar se encuentran:

- **Tamaños de Chunk:** Este hiperparámetro determina el tamaño de los segmentos de texto que el modelo utiliza durante el proceso de recuperación y generación. Ajustar el tamaño del chunk puede impactar directamente en la precisión y la relevancia de la información recuperada, así como en la fluidez y coherencia de las respuestas generadas.
- **Umbral de Similitud:** En el proceso de recuperación de información, se puede ajustar el umbral de similitud que determina qué tan cercanos deben ser los documentos recuperados con respecto a la consulta del usuario. Esto afecta la relevancia del contenido recuperado.
- **Número de Documentos Recuperados:** Este hiperparámetro define la cantidad de documentos que el modelo consulta para generar una respuesta. Un número mayor puede proporcionar más contexto, pero también aumentar el tiempo de procesamiento.

### **2.3 Obtención y validación de la información**

En esta sección, se describe el proceso de recolección y preparación de datos del sitio web de la FIEC, crucial para alimentar y evaluar los modelos de LLM. Primero, se realizó un proceso de web scraping para extraer información relevante de la página web de la FIEC. Esta extracción se centró en datos directamente accesibles, facilitando así su integración en los modelos de LLM para pruebas variadas. Para llevar a cabo esta tarea, se emplearon herramientas como Selenium y Jupyter Notebook.

Una vez recopilados, los datos pasaron por un proceso de curación y formateo. Esta etapa fue esencial para limpiar y estructurar los datos, eliminando información irrelevante o incorrecta, y preparándolos para su uso eficiente. El resultado fue un conjunto de datos estructurado, listo para entrenar y probar los modelos. Estos datos

estructurados se almacenaron en formatos adecuados para su posterior uso con diversas arquitecturas de RAG, como LangChain. Adicionalmente, se realizó una revisión exhaustiva de la información para asegurar su precisión y relevancia, evitando la inclusión de datos erróneos que pudieran afectar la calidad de las respuestas del chatbot.

## **2.4 Pruebas de modelos de lenguaje y frameworks**

La evaluación de los modelos de lenguaje y los frameworks seleccionados fue una etapa fundamental en el desarrollo de nuestro chatbot. Se diseñó un proceso de evaluación estructurado para garantizar la efectividad de las herramientas elegidas. Los modelos evaluados incluyeron GPT-3.5, Llama 2 , Open Orca y Bloom, y los frameworks considerados fueron LlamaIndex y LangChain.

Para realizar una evaluación rigurosa, se creó un conjunto de preguntas basadas en la información obtenida de la página de FIEC recopilada en documentos. Estas preguntas se clasificaron en diferentes categorías, como comprensión de contenido, preguntas de inferencia y preguntas para identificar respuestas falsas positivas. Cada pregunta fue diseñada para probar la capacidad de los modelos de comprender y responder de manera precisa y coherente.

Durante las pruebas, se registró la respuesta de cada modelo a las preguntas, así como el tiempo estimado de respuesta. Posteriormente, las respuestas fueron almacenadas en un archivo .csv y comparadas con las respuestas esperadas. Este proceso permitió generar un campo de "correctness" para cada respuesta, evaluando cuán cercana estuvo cada una de la respuesta ideal.

Los criterios de evaluación se centraron en dos aspectos fundamentales:

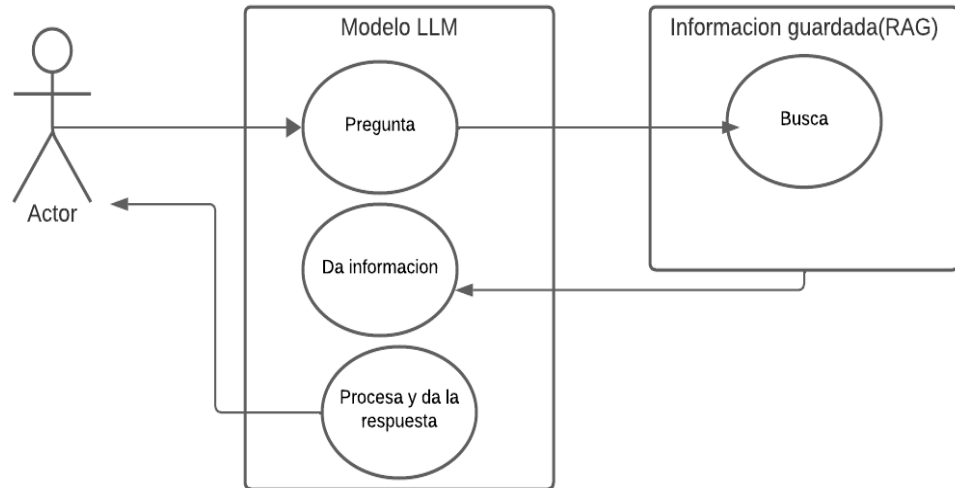
- Corrección de la Respuesta: Se evaluó si la respuesta era correcta en términos de contenido y relevancia con respecto a la pregunta planteada.
- Claridad y Comprensibilidad: Se analizó la claridad de la respuesta, asegurándose de que fuera entendible, libre de ambigüedades y sin errores gramaticales.

Esta metodología de evaluación proporcionó una base sólida para seleccionar la combinación más efectiva de modelo de lenguaje y framework, garantizando así que el chatbot proporcionara respuestas precisas y coherentes a las consultas de los usuarios relacionadas con la FIEC.

## **2.5 Implementación de modelos**

### **2.5.1 Caso de Uso**

Al ser un servicio que tiene una sola función principal solo se obtiene un solo caso de uso para el usuario o actor, que es comunicarse o preguntar al chatbot, lo cual se desarrolla como el diagrama a continuación.



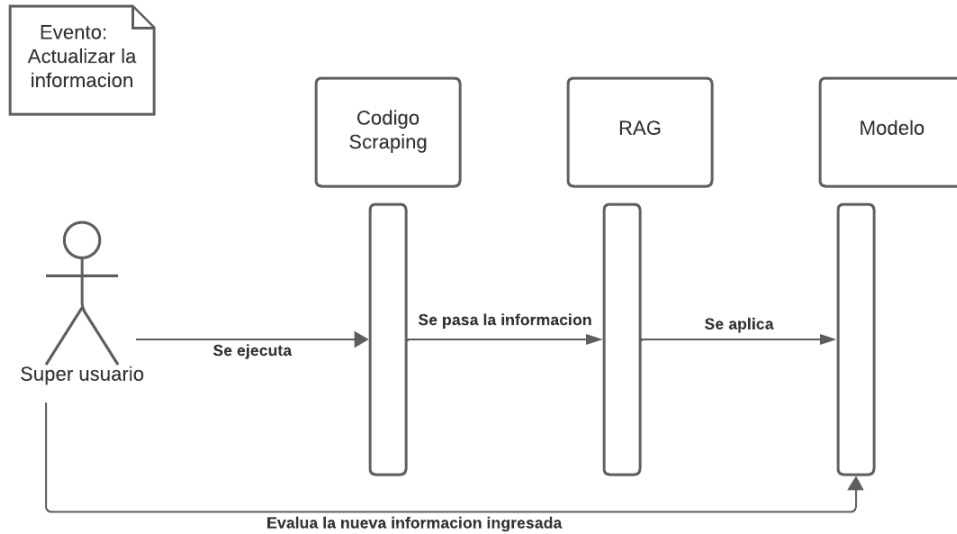
**Figura 2** Diagrama Casos de uso

Como se observa RAG solo se encarga de mantener la información de FIEC y facilitarla al modelo, mientras que el modelo se encarga de entender esta información y responder de la manera más adecuada según la pregunta hecha.

### 2.5.2 Diagrama de eventos

Como evento principal del super usuario es el actualizar la información del chatbot, es decir cuando la página web FIEC tenga nueva información esta necesita ser actualizada en los conocimientos del chatbot.





**Figura 3** Diagrama de Eventos

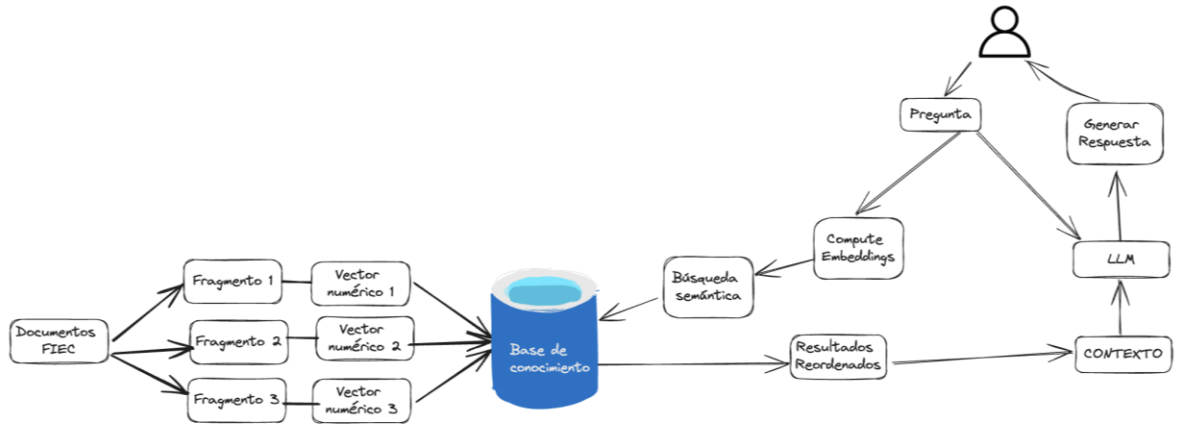
La forma que se maneja es simple dado que es modularizado el proyecto, se necesita ejecutar el código de Scraping, que generara los documentos con toda la nueva información de FIEC, de ahí solo se ejecutara el RAG y de ahí se activara le modelo.

En caso de haber una nueva sección de información solo se necesitará hacer modificaciones en el Código de Scraping, mas no en los demás módulos.

### 2.5.3 *Arquitectura de RAG*

La arquitectura Retrieval-Augmented Generation (RAG) se ilustra en el diagrama proporcionado, donde los documentos de la FIEC se segmentan en fragmentos y se convierten en vectores numéricos para su inclusión en una base de conocimiento. Al recibir una consulta, RAG calcula vectores similares para la pregunta, ejecuta una búsqueda semántica en la base para hallar coincidencias y reordena los resultados según su relevancia. Estos pasos preparan el terreno para que el modelo de lenguaje de gran escala (LLM) genere una respuesta informada,

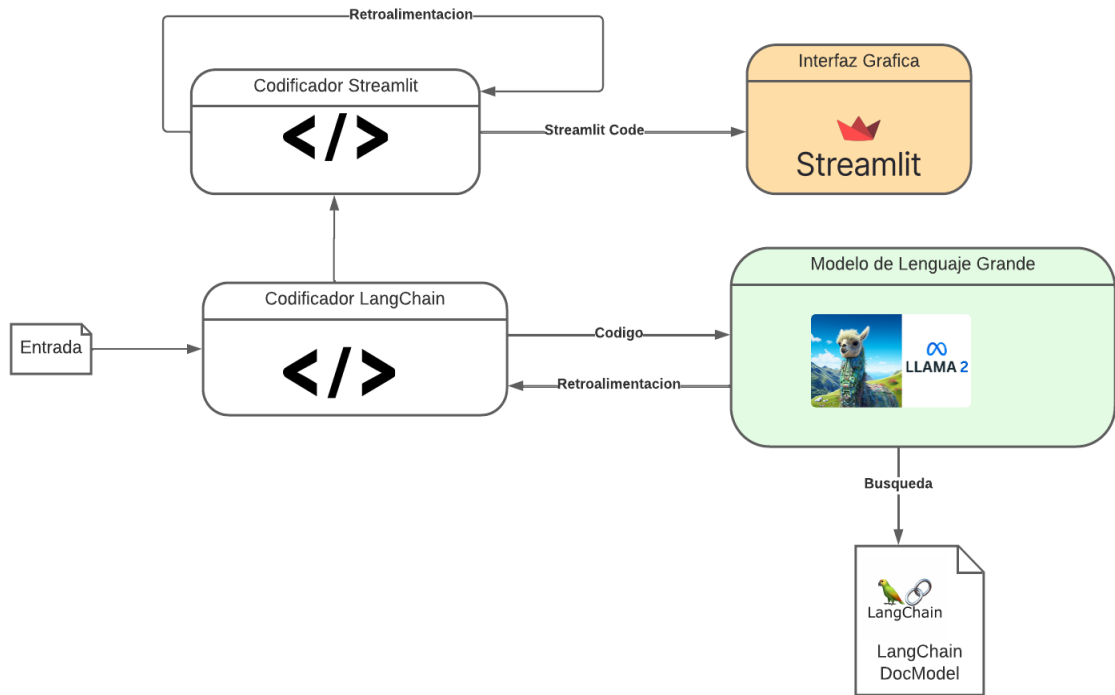
aprovechando el contexto y los datos recuperados, lo que culmina en un proceso integrado que mejora la calidad y precisión de las interacciones del chatbot.



**Figura 4** Arquitectura de RAG

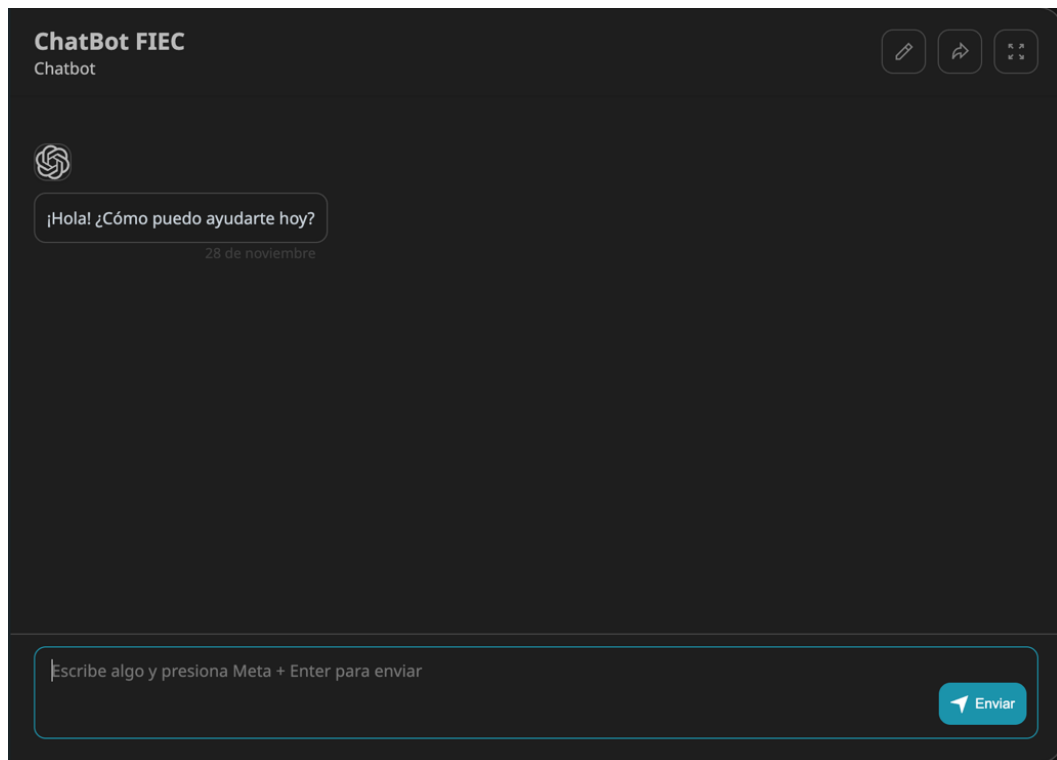
#### 2.5.4 *Arquitectura del prototipo*

La plataforma web está formada por un frontend y un backend que es administrado por la misma tecnología: Streamlit, Gradio.

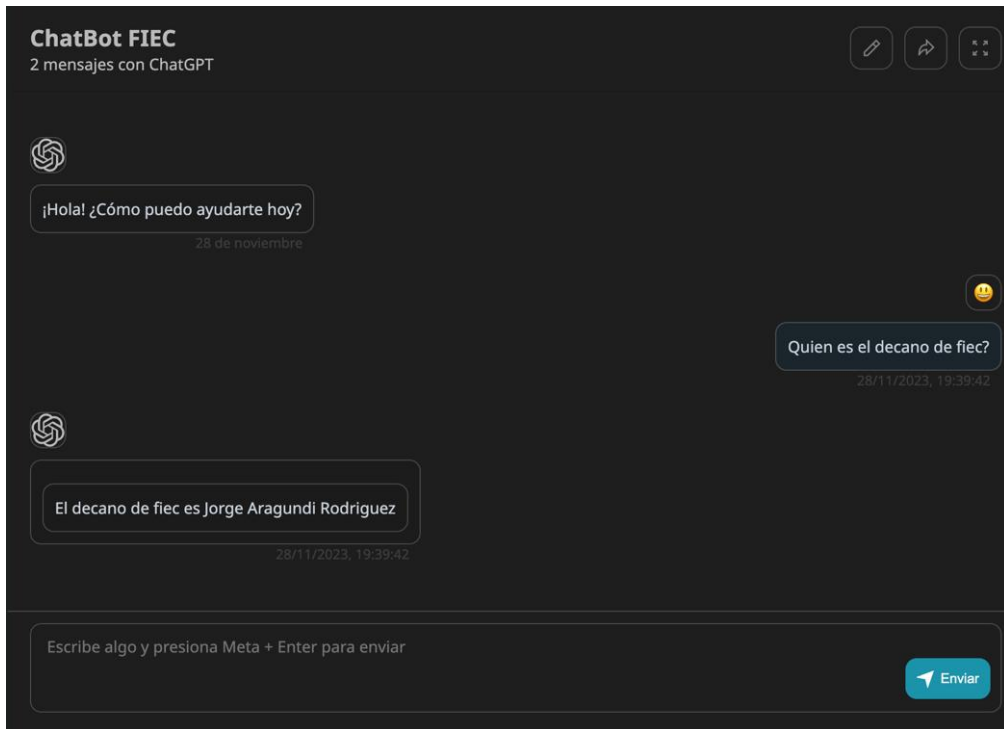


**Figura 5** Arquitectura del Prototipo

### 2.5.5 Prototipo



**Figura 6** Interfaz gráfica del prototipo 1



**Figura 7** Interfaz gráfica del prototipo 2

## 2.6 Pruebas de usuario

### 2.6.1 Métricas de evaluación

Con el objetivo de medir la efectividad del asistente virtual se realizaron pruebas de usuario donde se compara la pagina actual de FIEC y nuestro asistente virtual, para ellos planteamos 2 metas de nuestro producto.

| Meta del Producto  | Ideas de Evaluación  | Perfil del Usuario  | Cantidad de Usuarios | Duración de la Actividad |
|--|--|---------------------|----------------------|--------------------------|
| Ofrecer un sistema eficiente de consultas para los Estudiantes | Realizar consultas al chatbot para conocer sobre su facultad.      | Estudiantes de FIEC | 30                   | 30 minutos               |
|  | Medir el tiempo de respuesta a Consultas de Estudiantes            |                     |                      |                          |
|  | Brindar información precisa (correcta) sobre FIEC.                 |                     |                      |                          |
| Proveer un producto satisfactorio para los Estudiante.         | Conocer la satisfacción de los estudiantes al usar la herramienta. | Estudiantes de FIEC | 30                   | 10 minutos               |
|  | Evaluación de la facilidad de uso y accesibilidad                  |                     |                      |                          |

**Tabla 1** Métricas de evaluación

### 2.6.2 Plan de desarrollo

Para estas pruebas de usabilidad se diseñaron dos formularios con sus respectivos casos de uso, uno medir el tiempo de respuesta de el asistente virtual y otro para medir la satisfacción de los usuarios al interactuar con el chatbot. Las preguntas del formulario estaban enfocadas en la obtención de información de FIEC, la mitad de los participantes usaron el asistente virtual para obtener la respuesta y la otra mitad debió buscar dicha

información en la página web de la FIEC. Las plantillas de las preguntas utilizadas se encuentran en el Apéndice B.

A través del formulario de satisfacción que se implementó luego de realizar las pruebas correspondientes, se busca comparar a través de una escala likert el nivel de satisfacción de los usuarios.

## **Capítulo 3**

### **3. ANÁLISIS Y RESULTADOS**

En esta sección se vera reflejado lo resultados a las pruebas con los modelos mencionados en el capítulo anterior.

#### **3.1 Experimentos desarrollados**

Se desarrollo pruebas con dos arquitecturas LangChain y LlamaIndex, usando los modelos Llama2, Bloom, Open Orca y ChatGPT, se evaluó ambos en base a como respondieran preguntas predefinidas, donde ChatGPT representaba el modelo de pago a comparar y Bloom era la alternativa libre a Llama2.

#### **3.2 Entorno de Desarrollo**

Se desarrollo en Python usando las librerías correspondientes de LangChain y Llama Index, en dos entornos locales, una laptop con procesador 17 de 6th generación, tarjeta gráfica ENVIDIA RTX 940 y una MAC.

#### **3.3 Parámetros del RAG**

Dado que la información utilizada en estos experimentos fue corta no se tomaron todos lo parámetro posibles para estos, los más relevantes fueron:

- El tamaño de los Chunks fuera de 30
- El número de documentos recuperados sea 4

#### **3.4 Resultados LangChain**

Se creó una lista de preguntas para la evaluación del modelo, y se lo alimento con información básica de FIEC no toda la data disponible.

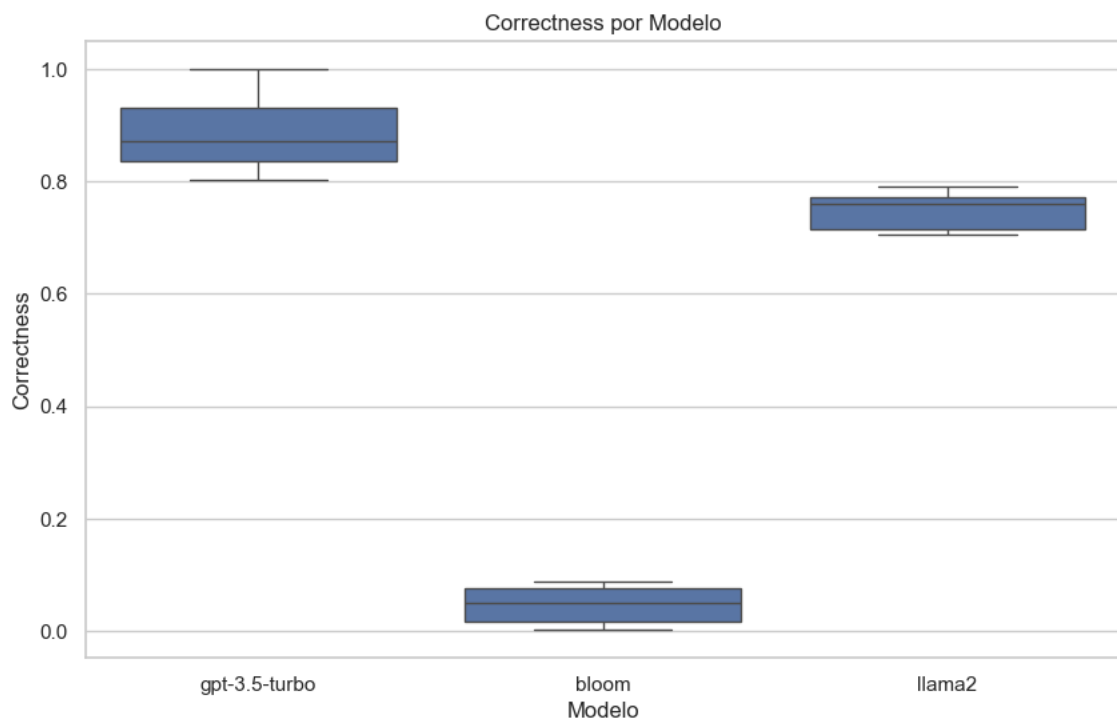


| Tipo de Pregunta                       | Pregunta   | Respuesta  |
|--|--|--|
| <b>Content Understanding Questions</b> | ¿Qué hay en el Edificio 11-A?                              | En el Edificio Principal de la FIEC, se encuentran las oficinas administrativas, Decanato, Subdecanato, Secretarías, la Sala de Sesiones, Auditorio, Sala de Consejo Unidad Académica, Sala de Eventos Múltiples, oficinas de profesores, aulas de clase   |
| <b>Content Understanding Questions</b> | ¿Qué laboratorios hay en el Edificio 11-B?                 | Laboratorio de Computación 2<br>Laboratorio de Computación 3<br>Laboratorio de Computación 4<br>Laboratorio de Computación 5<br>Laboratorio de Computación 6<br>Laboratorio de Computación 7<br>Laboratorio Móvil  |
| <b>Content Understanding Questions</b> | ¿En qué edificio se encuentra el grupo estudiantil KOKOA?  | Edificio 11 - D  |
| <b>Content Understanding Questions</b> | ¿Cuáles son las áreas de investigación de FIEC?            | Automation Systems<br>Control Systems<br>Data Science and Artificial Intelligence<br>Electric Energy Systems<br>Electric Power Systems<br>Electronics<br>HCC Human-Centered Computing<br>Networking and Distributed Systems<br>Optical Communications<br>Signal Processing<br>Smart Environments and Telematics Systems<br>Software Engineering<br>Systems Security<br>Wireless Communications |
| <b>Inference Questions</b>             | ¿CIDIS es un centro de investigación?                      | SI   |
| <b>Inference Questions</b>             | ¿En qué edificio se encuentran la mayoría de laboratorios? | Edificio 11 - C (Planta Alta)  |

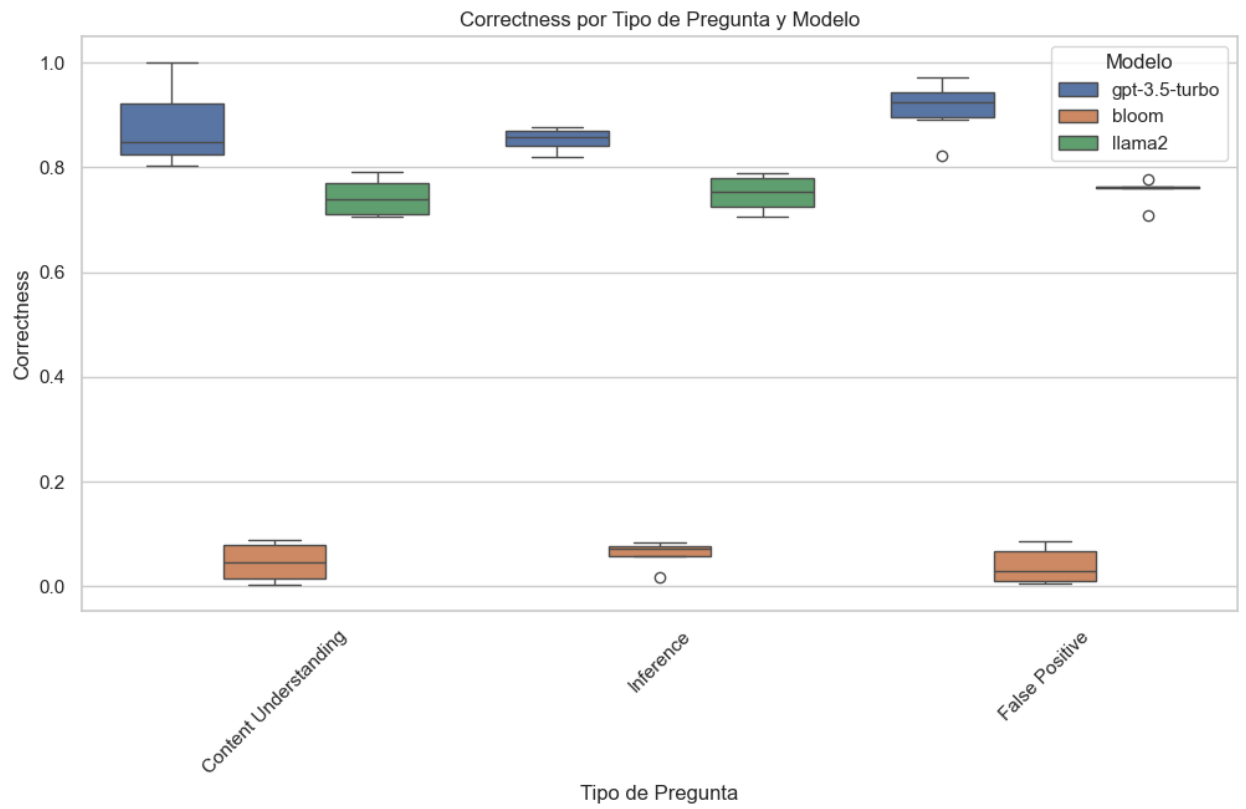
| Inference Questions      | ¿En qué edificio se encuentran menos laboratorios? | Edificio 11f |
|--------------------------|--|--------------|
| False Positive Questions | ¿Qué laboratorios hay en el Edificio 15?           | No.          |
| False Positive Questions | ¿CEPROEM es un centro de investigación?            | No.          |
| False Positive Questions | ¿El decanato está ubicado en el Edificio 15?       | No.          |

**Tabla 2.** Preguntas de evaluación del modelo

Con esta información se tabulo las respuestas y se obtuvo los siguientes resultados, donde se comparó cada respuesta con su respuesta esperada para obtener un nuevo campo de correctness para identificar que tan cercano estuvo de la respuesta.



**Figura 8.** Resultados LangChain 1



**Figura 9.** Resultados LangChain 2

### 3.5 Resultados Llama Index

Se utilizo el mismo procedimiento en Llama Index, donde se realizó las mismas preguntas con esta arquitectura.

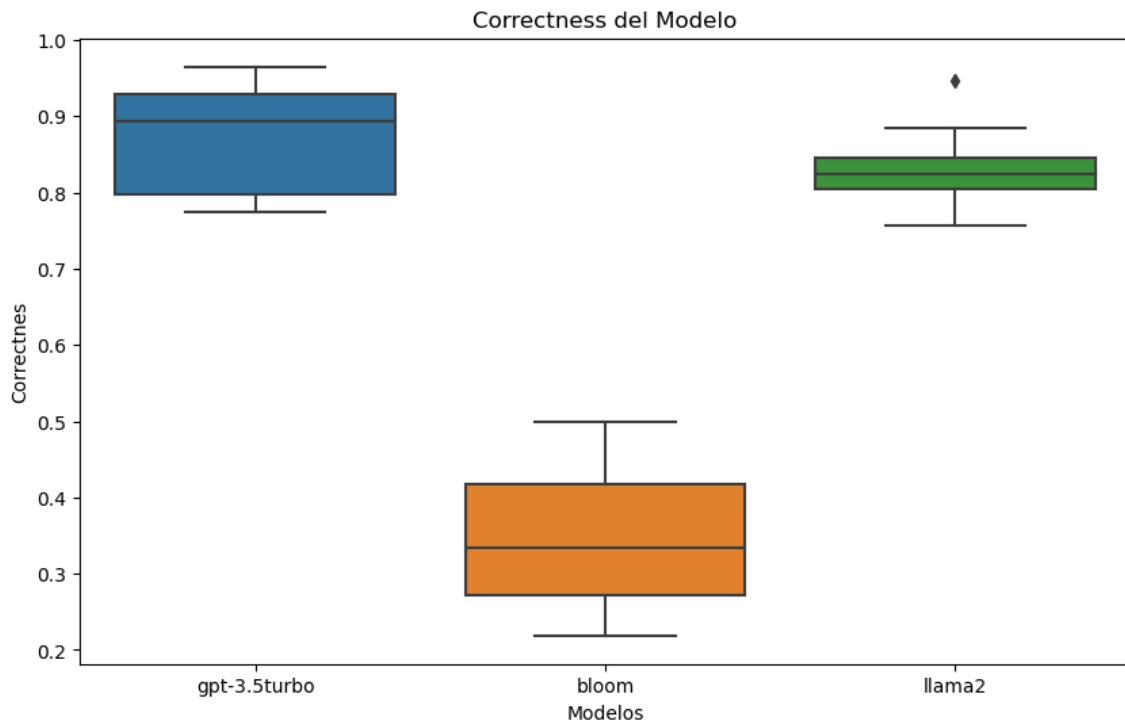
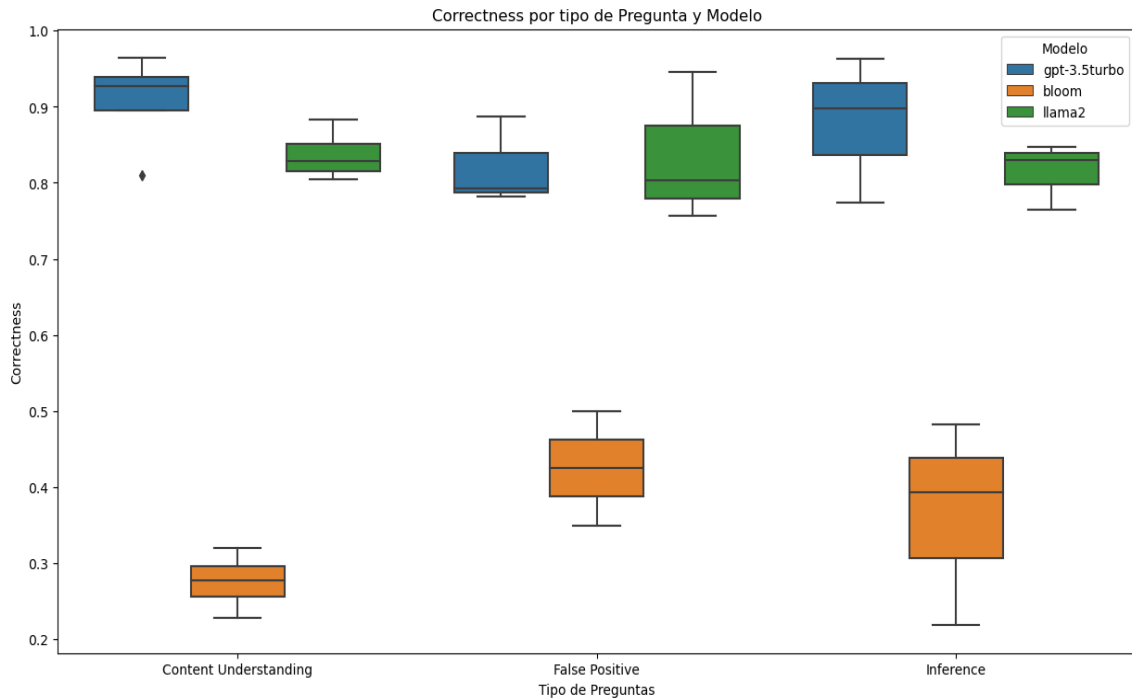


Figura 10. Resultados Llama Index 1



**Figura 11.** Resultados Llama Index 2

### 3.6 Comparación Arquitecturas

En los resultados obtenidos, se demostró que Bloom, como alternativa de código abierto, resultó impreciso en ambas arquitecturas. GPT, por otro lado, se destacó como superior en todos los tipos de preguntas, lo cual era esperado en una alternativa de pago. Sin embargo, Llama2 mostró resultados por encima del 80%, lo que sugiere que podría ser considerada como una alternativa viable a Chat GPT. Esto podría cambiar al utilizar versiones más robustas de Llama2. En conclusión, tanto Chat GPT como Llama2 son modelos competentes, considerando a Llama2 como una alternativa de código abierto.

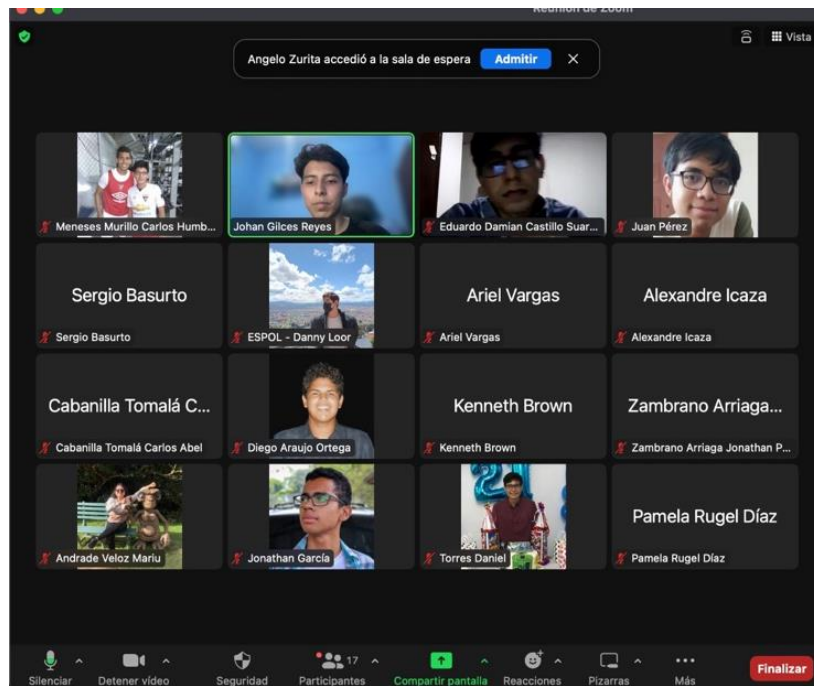
En cuanto a las arquitecturas, no se observó un cambio considerable de una a otra en GPT y Llama. En todo caso, se optó por considerar otras características de cada una. Tomamos en cuenta a Lang Chain como la mejor opción debido a su mayor tiempo en el mercado, la robustez de su comunidad y su facilidad para ser actualizada o implementar características adicionales en el futuro. En otras palabras, Lang Chain

proporciona mayor fiabilidad a largo plazo en términos de mantenimiento y actualización.

Dado el excelente rendimiento demostrado por GPT y Llama2, se optó por desarrollar dos chatbots empleando ambos modelos. La selección del modelo a utilizar se determinaría de acuerdo con los requisitos específicos del cliente. GPT se destacó por su precisión, aunque al tratarse de una API de pago, los costos de mantenimiento podrían ser significativos. Por otro lado, Llama2 ofrece una alternativa gratuita; sin embargo, la implementación del modelo más robusto requerirá recursos adicionales en el servidor.

### **3.7 Pruebas de usuario**

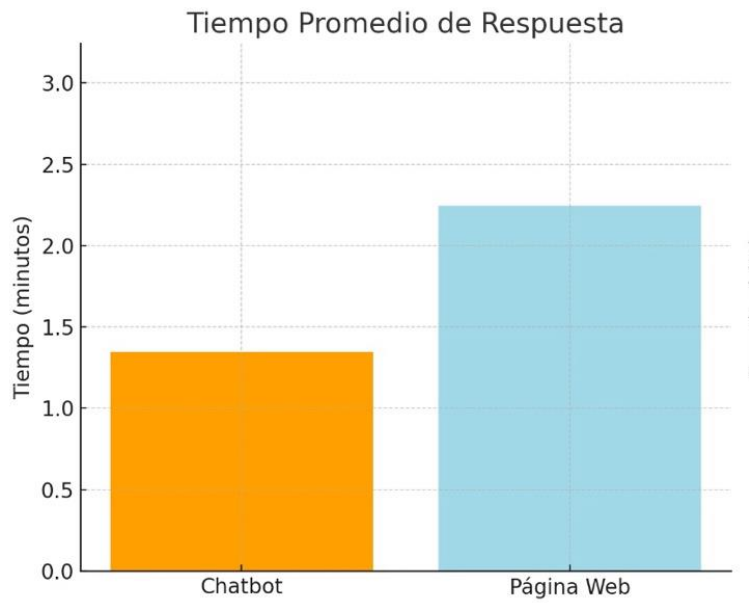
Se realizaron las pruebas de usuario a través de la plataforma Zoom, para luego dividir en grupos la mitad de los participantes. En la sala 1 se desarrollaron las pruebas con el asistente virtual. Se realizaron dos encuestas una evaluando 5 preguntas en las cuales se mide el tiempo y exactitud, y una encuesta acerca de la experiencia de usuario. Las encuestas se encuentran en el apéndice A.



**Figura 12.** Pruebas de usuario

### ***3.7.1 Tiempos de respuesta***

En la comparativa del tiempo de respuesta, el chatbot mostró una eficiencia superior. Se registró un tiempo promedio de aproximadamente 1.35 minutos para el chatbot, en contraste con los 2.24 minutos en promedio para la página web. Esta diferencia, ilustrada en el siguiente gráfico, resalta la rapidez del chatbot en proveer información requerida.

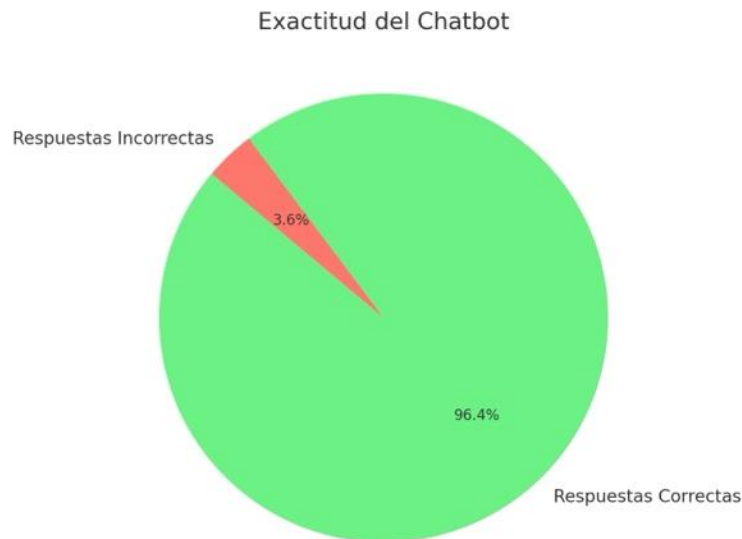


**Figura 13.** Resultados de Encuestas 1

### 3.7.2 *Exactitud de las respuestas*

Respecto a la exactitud, el chatbot también aventajó a la página web. El chatbot alcanzó una tasa de exactitud del 96.36%, mientras que la página web tuvo una tasa del 86.67%. Este hallazgo sugiere que el chatbot no solo es más rápido, sino también más preciso en la entrega de información correcta. El gráfico de pastel siguiente muestra la proporción de respuestas correctas e incorrectas proporcionadas por el chatbot, evidenciando su alta fiabilidad.

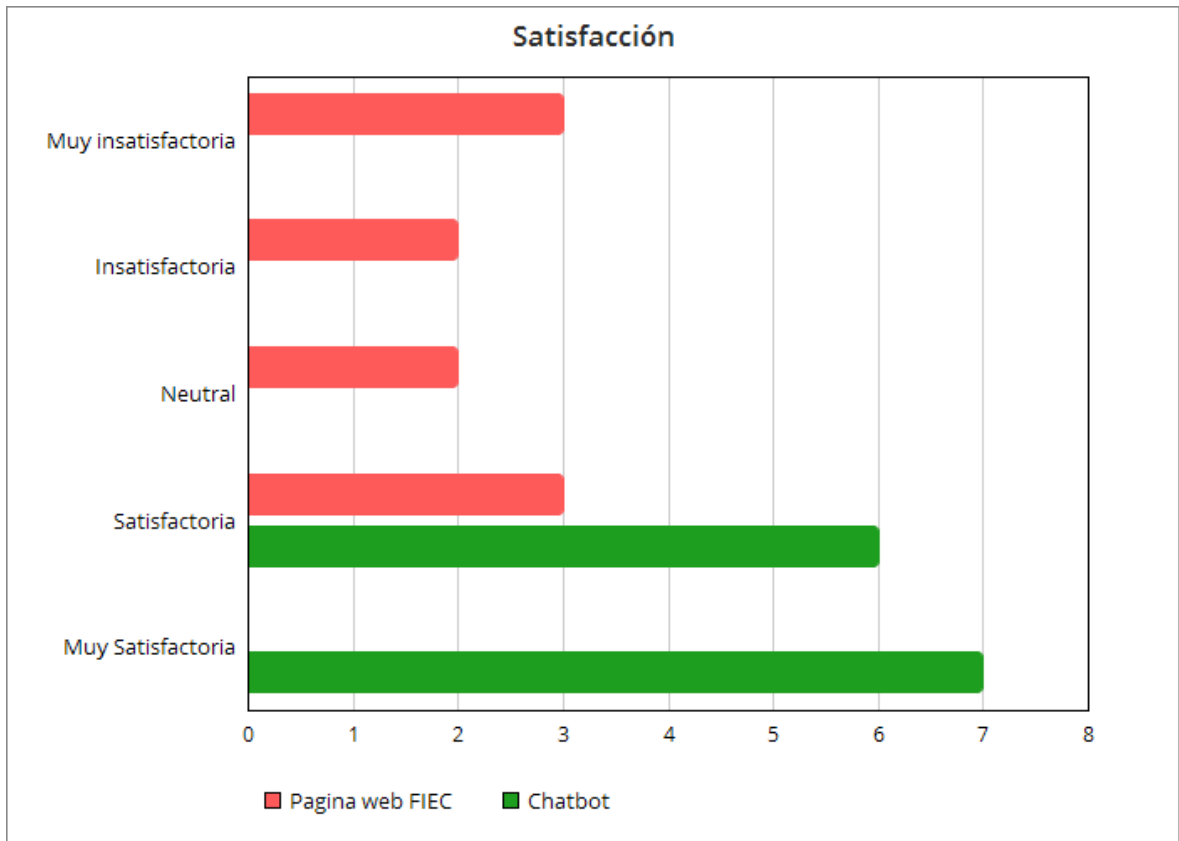




**Figura 14.** Resultados de Encuestas 2

### 3.7.3 *Satisfacción de las respuestas*

La encuesta reveló que los usuarios del chatbot reportan niveles significativamente más altos de satisfacción y perciben una mayor velocidad en encontrar información en comparación con los usuarios de la página web. El chatbot obtuvo una media de 4.54 en satisfacción, acercándose a "Muy Satisfactoria", y una media de 4.21 en velocidad, situándose entre "Rápido" y "Muy rápido". Por otro lado, la página web mostró una media de 3.38 en satisfacción y 3.00 en velocidad, lo que indica una experiencia más neutral o moderadamente satisfactoria. Además, se observó una fuerte correlación entre la satisfacción y la velocidad, subrayando la importancia de la eficiencia en la experiencia del usuario.



**Figura 15.** Resultados de Encuestas 3

## **Capítulo 4**

## 4. CONCLUSIONES Y RECOMENDACIONES

En el desarrollo del proyecto, se emplearon diversos modelos, a saber: Llama2, Chat GPT y Bloom. La elección de estos modelos se basó en una exhaustiva investigación de los más utilizados y adecuados para nuestro contexto. Se utilizó Python en conjunto con la biblioteca Selenium para la recopilación de información, llevando a cabo el scraping en la página web de FIEC, donde se recopiló toda la información necesaria para alimentar a los modelos.

Para el desarrollo de los prototipos, se consideraron dos posibles arquitecturas para utilizar RAG: Lang Chain y Llama Index. Ambas versiones del prototipo fueron implementadas en Python, y los modelos fueron obtenidos de la plataforma Hugging Face, con la excepción de Chat GPT, donde se utilizó la API de OpenAI. En el caso de Llama2, se emplearon versiones del modelo con menos información para facilitar el desarrollo.

Se diseñó un prototipo web final que utiliza Lang Chain y GPT para evaluar a estudiantes en relación con su experiencia con el chatbot versus la página web de FIEC, siendo esta última la alternativa actual al chatbot.

### 4.1 Conclusiones

- Se diseñó una arquitectura basada en Lang Chain, considerada la mejor opción debido a su extensa comunidad y mantenimiento. Se identificaron los modelos de mejor rendimiento, siendo Chat GPT y Llama2 seleccionados por asegurar una tasa de corrección superior al 80%. La información necesaria se obtuvo mediante el scraping de la página web de FIEC.
- Se desarrolló una página web utilizando Streamlit como interfaz visual para el público. Esta interfaz conecta el modelo, y la misma plataforma se empleó para desplegar el servicio en línea para el prototipo final.

- Se llevaron a cabo encuestas con el prototipo final para validar la experiencia de los usuarios con el chatbot y comparar sus ventajas con respecto a la página web de FIEC. Los resultados demostraron no solo ser una alternativa más rápida en un 40%, sino también una plataforma con una experiencia de usuario superior.

## 4.2 Recomendaciones

Se sugiere llevar a cabo una investigación exhaustiva inicial sobre los modelos y arquitecturas que mejor se adapten al problema a resolver. Esta práctica reduce significativamente las posibles alternativas a probar, generando ahorros tanto en tiempo como en costos de desarrollo.

Además, se recomienda tener en consideración la longevidad de las tecnologías utilizadas. Este fue uno de los motivos para la elección de Llama2 y Lang Chain, ya que representan tecnologías contemporáneas con un historial sólido de mantenimiento y actualizaciones a lo largo del tiempo, además de provenir de compañías líderes en el sector.

Un aspecto crítico en el desarrollo de proyectos que involucren Large Language Model (LLM) es la disponibilidad de recursos. Debido a estas limitaciones, no fue posible trabajar con modelos más robustos. Esto también es crucial al desplegar el proyecto en la web, dado que puede presentar complicaciones significativas en las etapas finales del desarrollo.

Las aplicaciones potenciales en el ámbito comercial del proyecto son diversas. Estas van desde la inclusión de otras facultades de la ESPOL hasta la integración de información adicional, como estatutos para profesores, reglas, entre otros. Además, el proyecto podría ser empleado por empresas que manejen una gran cantidad de información para sus usuarios o personal, habiéndose demostrado

como una opción más amigable con el usuario y una forma efectiva de centralizar la información de cualquier entidad.

## 5. REFERENCIAS

- [1] U. o. S. A. G. Z. R. W. U. o. G. G. Geissler, «Web home page complexity and communication effectiveness,» *J. Assoc*, vol. 2, n° 1, p. 1–48, 2001.
- [2] D. B. a. L. Robinson, «Information Overload: An Overview,» *Oxford Encyclopedia of Political Decision Making*, 2020.
- [3] D. B. a. L. Robinson, «The dark side of information: Overload, anxiety and other paradoxes and pathologies,» *Journal of Information Science*, vol. 2, n° 35, pp. 180-191, 2009.
- [4] B. V. Ridge, «La importancia de encontrar el sitio web más útil en la era digital,» *MEDIUM Multimedia Agencia de Marketing Digital*, 20 9 2023. [En línea]. Available: <https://www.mediummultimedia.com/web/cual-es-el-sitio-web-mas-util/>. [Último acceso: 24 10 2023].
- [5] M. J. M. A. Q. a. R. S. A. Haleem, «Understanding the role of digital technologies in education: A review,» *Sustainable Operations and Computers*, vol. 3, n° 3, p. 275–285, 2022.
- [6] K. Ratheeswari, «Information Communication Technology in Education,» *Journal of Applied and Advanced Research*, vol. 3, n° s1, p. 45, 2018.
- [7] [www.boia.org](http://www.boia.org), «Common accessibility issues for higher education websites,» [En línea]. Available: <https://www.boia.org/blog/common-accessibility-issues-for-higher-education-websites>. [Último acceso: 22 10 2023].

- [8] S. P. a. İ. T. M. S. Macakoğlu, «Accessibility, usability, and security evaluation of universities' prospective student web pages: a comparative study of Europe, North America, and Oceania,» *Universal Access in the Information Society*, vol. 22, n° 2, p. 671–683, 2023.
- [9] N. Sharma, «AI Chatbots vs,» Medium, 06 Jul 2023. [En línea]. Available: <https://medium.com/@snavi2506/hatboyai-chatbots-vs-1af2ff0b498d>. [Último acceso: 22 10 2023].
- [10] «Conversational AI vs. Chatbots: What's the difference?,» Capacity, 10 2 2023. [En línea]. Available: <https://capacity.com/learn/ai-chatbots/conversational-ai-vs-chatbot/>. [Último acceso: 24 10 2023].
- [11] X. Ma y Y. Huo, «Are users willing to embrace ChatGPT? Exploring the factors on the acceptance of chatbots from the perspective of AIDUA framework,» *Technol. Soc*, vol. 75, n° 102362, p. 102362, 2023.
- [12] N. Kumar, «Chatbot vs ChatGPT Differences & Similarities,» [En línea]. Available: <https://sparkbyexamples.com/machine-learning/chatbot-vs-chatgpt/#:~:text=Similarities%20between%20ChatGPT%20and%20ChatBot>. [Último acceso: 22 10 2023].
- [13] L. L. Z. M. S. L. H. Y. y. L. H. L. Fan, «A bibliometric review of large language models research from 2017 to 2023,» 2023.
- [14] B. M. N. R. M. S. J. K. P. D. e. a. T. B. Brown, «Language Models are Few-Shot Learners,» 2020.



- [15] P. L. e. al., «Retrieval-augmented generation for knowledge-intensive NLP tasks,» 2023. <sup>2</sup>

## ANEXOS

## Anexo A

# Encuesta de Proyecto Integrador

Esta evaluación sirve para conocer la manera de obtener información sobre FIEC-ESPOL.

\* Obligatoria

\* Este formulario registrará su nombre, escriba su nombre.

## Información del estudiante

1. ¿Cuál es su matrícula? \*

El valor debe ser un número.

2. ¿Cuál es su carrera? \*

- Ingeniería en Electricidad
- Ingeniería en Electrónica y Automatización
- Ingeniería en Telecomunicaciones
- Ingeniería en Telemática
- Ingeniería en Computación

3. Herramienta de información

- Pagina web FIEC
- CHATBOT

¿Quién es el Decano actual de la FIEC?

4. ¿Cuántos minutos te tomó obtener la respuesta? \*

|   |   |   |   |   |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|

5. ¿Obtuviste la respuesta correcta?  
Dr. Jorge Aragundi Rodríguez. \*

- Si
- No

¿Qué profesores son miembros del Consejo de la FIEC?

6. ¿Cuántos minutos te tomó obtener la respuesta? \*

|   |   |   |   |   |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|

7. ¿Obtuviste la respuesta correcta?  
Otto Ismael Alvarado Moreno, Mgtr.  
Vanessa Ines Cedeño Mieles, Ph.D.  
Javier Alejandro Tibau Benitez, Ph.D.

\*

Sí

No

¿Cuál es la oficina de Carmen Vaca?

8. ¿Cuántos minutos te tomó obtener la respuesta? \*

|   |   |   |   |   |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|

9. ¿Obtuviste la respuesta correcta?  
Es la oficina 11A-027 \*

Sí

No

¿Quién es Vanessa Cedeño Mieles?

10. ¿Cuántos minutos te tomó obtener la respuesta? \*

|   |   |   |   |   |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|

11. ¿Obtuviste la respuesta correcta?  
Subdecana \*

Si

No

¿Quién es el coordinador de telemática?

12. ¿Cuántos minutos te tomó obtener la respuesta? \*

|   |   |   |   |   |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|

13. ¿Obtuviste la respuesta correcta?  
José E. Córdova García, Ph.D. \*

Si

No

## Anexo B

### Encuesta de Satisfacción

Esta evaluación es para evaluar la experiencia con ambas plataformas

\* Obligatoria

\* Este formulario registrará su nombre, escriba su nombre.

#### Información del estudiante

##### 1. Herramienta de información

- Pagina web FIEC
- CHATBOT

##### 2. ¿Cómo evaluaría su experiencia al buscar información? \*

|              | Muy insatisfactoria   | Insatisfactoria       | Neutral               | Satisfactoria         | Muy Satisfactoria     |
|--------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Satisfacción | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

##### 3. ¿Cómo fue el tiempo que le tomo encontrar la información? \*

|           | Muy rápido            | Rápido                | Neutral               | Lento                 | Muy Lento             |
|-----------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Velocidad | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

##### 4. Comenta acerca de tu experiencia

## Anexo C

### Capturas del funcionamiento del sistema

