Intelligent Enterprise AI Assistant

Submitted in partial fulfillment of the requirements of the degree

BACHELOR OF ENGINEERING IN **COMPUTER ENGINEERING**

By

VANSH PILLAI 122A1117

VIGNESH PILLAI 122A1119

VEDANT BAWKAR 223A1128

YASH SHINDE 223A1135

Name of the Mentor

Prof. Ujwala Ravale

Department of Computer Engineering

SIES GRADUATE SCHOOL OF TECHNOLOGY

NERUL, NAVI MUMBAI – 400706

ACADEMIC YEAR

2024 – 2025

CERTIFICATE

This is to certify that the Mini Project entitled **"Intelligent Enterprise AI Assistant"** is a bonafide work of **VANSH PILLAI (122A1117), VIGNESH PILLAI (122A1119), VEDANT BAWKAR (223A1128), YASH SHINDE (223A1135)** submitted to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of **"Bachelor of Engineering"** in **"Computer Engineering".**

Prof. Ujwala Ravale

Mentor

Dr. Aparna Bannore

Head of Department

Dr. K. Lakshmisudha

Principal

Mini Project Approval

This Mini Project entitled "**Intelligent Enterprise AI Assistant"** by **VANSH PILLAI (122A1117), VIGNESH PILLAI (122A1119), VEDANT BAWKAR (223A1128), YASH SHINDE (223A1135)** is approved for the degree of **Bachelor of Engineering** in **Computer Engineering.**

Examiners

1……

(Internal Examiner Name & Sign)

2……

(External Examiner name & Sign)

Date:

Place:

Contents

ABSTRACT

Organizations require optimizing information and query resolutions in this fast-changing technological landscape. The Intelligent Enterprise AI Assistant fulfills those needs by automating the mundane work of large public sector organizations through an advanced chatbot system with latest AI and NLP capabilities. This assistant takes care of routine queries that include IT support requests, freeing up human resources and providing accurate answers to help streamline operations. It designs context-aware responses to enhance the user experience via a friendly and efficient interface personalized according to the specific requirements of employees. Its focus on integration with AI, document processing, and security translates into its ability to help organizations deal with rising workload and service demands, especially within HR and IT departments. The report reviews its impact on organizational performance with an account of architecture and features, along with its implementation strategy. Findings reveal that the assistant considerably improves efficiency in operation and fosters a more responsive agile culture. The project, therefore, opens up the transformational potential of AI to handle and support employees through information management in public sector organizations under changing conditions.

ACKNOWLEDGEMENT

We have taken efforts in this project. However, it would not have been possible without the kind support and help of many individuals and organizations. We would like to extend my sincere thanks to all of them.

We are highly indebted to our Guide Prof. Ujwala Ravale, HOD Dr. Aparna Bannore ,project coordinator Prof. Sunil Punjabi and Principal Dr. K Lakshmi Sudha for their guidance and constant supervision as well as for providing necessary information regarding the project & also for their support in completing the project.

We would like to express our gratitude to all for their kind cooperation and encouragement which helped us in completion of this project.
We would like to express our special gratitude and thanks to faculties for giving us such attention and time.

Thanks, and appreciation to all colleagues in developing the project and people who have willingly helped us out with their abilities.

List of Figures

List of Tables

List of Abbreviations

Sr. No.	Abbreviation	Expansion/Explanation
1.	AI	Artificial Intelligence
2.	2FA	Two-Factor Authentication
3.	NLP	Natural Language Processing
4.	LLM	Large Language Model
5.	OTP	One-Time Password
6.	FAQ	Frequently Asked Questions
7.	JWT	JSON Web Token
8.	Redis	Remote Dictionary Server (In-memory data structure store)
9.	FAISS	Facebook AI Similarity Search
10.	RAG	Retrieval-Augmented Generation
11.	UI	User Interface HTML: HyperText Markup Language
12.	CSS	Cascading Style Sheets
13.	Bootstrap	A front-end framework for developing responsive websites
14.	Scikit-learn	A machine learning library for Python
15.	LangChain	A framework for developing applications with LLMs
17.	SMTP	Simple Mail Transfer Protocol

CHAPTER 1

INTRODUCTION

1.1 Introduction

The deep need to handle information efficiently and provide query resolution quickly has existed in today's fast-paced organizational setting. Organizations are increasingly pressured to streamline their operations and optimize resource use. Thus, it is a great necessity to offer solutions that increase productivity and effectiveness. On this critical demand, solutions such as AI-driven chatbots have emerged to depict the future strategy in automating routine tasks and getting quick access to information. [1] [2] [5]

These intelligent systems will allow an organization to let human resources engage themselves in value-driven roles such as thinking, creativity, and human-to-human interaction, instead of wasting time answering the same and repetitive questions.

Intelligent Enterprise AI Assistant is one such entity that displays its sophistication since it forms the most advanced system that exploits cutting-edge Artificial Intelligence (AI) technologies to streamline and automate wide-ranging tasks. Other tasks include regular inquiries related to human resource queries, IT support, and other operational challenges. The interaction with the assistant is very precise due to using high-end natural language processing (NLP) and a highly context-aware response mechanism that may well meet individuals' needs and ensures user-friendly interactions, thereby improving the experience of employees who try to get things done. [3] [6]

The Intelligent Enterprise AI Assistant comes remarkably well suited to the unique needs of large public sector organizations, especially where high volume and complexity prevent old-fashioned support structures from being applied. The AI Assistant has features that integrate AI, document processing, and robust security measures to handle the increasingly high workload burdened on the various HR and IT departments. This innovative solution boosts the efficiency of organizations and enables them to maintain a high standard of service while adapting to the ever-evolving landscape of public sector operations through the automation of routine queries and immediate support. Through its ability to empower organizations to utilize their human resources more strategically, it ensures enhanced organizational performance and improved service delivery. [7] [4]

1.2 Motivation

In large organizations, several employees and teams of administrators are swamped with high workloads because routine, low-complexity tasks, including answering employee queries, have to be carried out. The Intelligent Enterprise AI Assistant is so created as to automate these routine tasks thereby freeing up valuable human resources for more critical tasks. Moreover, the increased demand for immediate and accurate access to organizational information has been so rapid that it pressures organizations into adopting technologies reducing manual dependency and hastening response times.

Since this AI is enabled with real-time access to data and document processing features, productivity is enhanced, and resource usage optimized while allowing the easy exchange of information between multiple departments. This leads to a work environment where tasks requiring critical thinking have a high premium because most routine tasks are well taken care of by the AI.

1.3 Problem statement & Objectives

As organizations expand, they often face significant challenges in managing employee inquiries, IT-related issues, and various routine operational tasks. The complexity of these challenges can lead to prolonged response times, elevated operational costs, and diminished employee satisfaction. In this context, the Intelligent Enterprise AI Assistant is proposed as a solution to address these critical pain points. This AI-driven system aims to enhance operational efficiency by automating the management of queries and providing secure, accurate, and real-time responses utilizing advanced artificial intelligence (AI) and natural language processing (NLP) technologies.

Main objectives of the system include:

- AI Chatbot Development: To create a robust and scalable chatbot capable of handling a high volume of queries efficiently.
- Efficiency Enhancement: To minimize the need for human intervention in query management, thereby accelerating response times and increasing overall operational efficiency.

- Real-Time Document Processing: To implement features that enable the summarization and analysis of documents in real time, facilitating quicker access to information.
- Security Implementation: To establish a Two-Factor Authentication (2FA) system that safeguards sensitive employee data and protects the organization's resources.

By focusing on these objectives, the Intelligent Enterprise AI Assistant is positioned to significantly improve the management of organizational queries, streamline operational processes, and enhance employee satisfaction.

1.4 Organization of the Report

This project report is structured into five key chapters. Chapter 1 introduces the project's focus on developing an AI-powered system to enhance organizational efficiency by automating routine queries, with objectives like scalability and security (2FA). Chapter 2 reviews existing AI chatbot systems, identifying limitations in scalability, document processing, and security. Chapter 3 explains the proposed system's architecture, detailing AI/NLP models, tools like ReactJS and Python, and evaluation methods. Chapter 4 presents experimental results, discussing system testing, performance, and key features like real-time responses and 2FA. Chapter 5 concludes by summarizing the project's success and future enhancements.

CHAPTER 2

LITERATURE SURVEY

2.1 Survey of Existing system

More recent systems and research works have focused on utilizing AI and NLP to enhance the working mechanisms of an organization. Much work has been done in developing chatbots to manage customer-related queries in the commercial field including customer care bots, FAQ bots, and virtual assistants. The systems are all quite effective for their intended purposes but usually do not satisfy the particular needs or requirements of organizations working in the public sector, especially with respect to HR and IT-related work activities.

Przegalinska et. al. [1] provides an extensive survey of chatbot usage in customer service, focusing on the integration of AI for enhancing customer experience. It discusses the cost efficiency and availability benefits of chatbots, alongside challenges such as user dissatisfaction and security concerns. However, it lacks detailed exploration of industry-specific applications and real-time user feedback, limiting its practical utility. See the table 2.1 for comparison of existing research paper.

Gao et. al. [2] reviews the application of deep learning in chatbots, showcasing how techniques like NLP and neural networks improve conversational accuracy and adaptability. However, it primarily remains theoretical, with minimal discussion on real-world deployments or the operational challenges businesses might face. Table 2.1 provides a detailed comparison of the current research studies.

Joulin et. al. [3] focuses on the design principles for conversational agents. His research highlights essential features for improving user interaction, such as dialogue flow and response relevance. Despite these insights, the study does not extensively cover user testing processes or validation methodologies to refine the design further. For a side-by-side comparison of the reviewed research papers, see Table 2.1.

Akkiraju et. al. [4] examines how reinforcement learning can create more responsive and

adaptive chatbots. While the use of RL shows promise in dynamic learning environments, the paper highlights issues in scalability and the intensive training data requirements, making large-scale deployment challenging. You can find the comparison of the existing research work in Table 2.1.

Tewari et. al. [5] presents a hybrid chatbot approach that blends rule-based and machine learning techniques. His study proposes a model that combines the strengths of both to enhance chatbot accuracy and versatility. However, it calls for more empirical validation and real-world case studies to substantiate the hybrid model's effectiveness. Refer to Table 2.1 for an analysis and comparison of the related research papers.

Table 2.1 Comparison of Existing Research Paper

Sr. No.	Authors/Year	Context	Quality Factors/Key Findings	Gaps/Issues
1	Przegalinska et al., 2019	A Survey on Chatbot Implementation in Customer Service	Comprehensive overview of chatbot applications in customer service, highlighting benefits and challenges.	Limited focus on specific industries and user feedback.
2	Gao et al., 2020	Deep Learning for Chatbots: A Review	Discusses advancements in deep learning techniques applied to chatbots, showing improved performance.	Lack of practical case studies and real-world applications.
3	Joulin et al., 2020	Chatbot Design for Conversational Agents	Explores design principles for effective chatbot interactions, emphasizing user experience.	Insufficient exploration of user testing methodologies.
4	Akkiraju et al., 2019	Building Conversational Agents with Reinforcement Learning	Examines the use of reinforcement learning for creating adaptive conversational agents.	Challenges in scalability and training data requirements.

5	Tewari et al., 2020	A Hybrid Approach to Chatbot Development	Presents a model combining rule-based and machine learning techniques for enhanced chatbot functionality.	Need for more empirical validation of the hybrid approach.

2.2 Limitations in existing system or research gap

- Efficiency in Simple Queries: Excel primarily in simple question answering, struggle with complex tasks.
- Real-Time Document Handling: Significant limitations in supporting real-time document management.
- Advanced Security Integration: Lack robust security features for handling sensitive information.
- Scalability Issues: Designed for a limited number of users (a few tens), resulting in slow performance and incorrect responses under heavier loads.

2.3 Mini Project Contribution

Our Mini project team consists of four people. The 2FA Authentication module was created by Vedant Bawkar , Chatbot was created Yash Shinde and document processing RAG application was created by Vansh Pillai and Vignesh Pillai.

CHAPTER 3

PROPOSED SYSTEM

3.1 Overview of Proposed System

The proposed system, Intelligent Enterprise AI Assistant, will aim at maintaining the efficiency in the organization through auto-allocation of routine mundane tasks, streamlined workflow, and providing a secure scalable query handling platform. As the system will integrate the process of document processing and two-factor authentication, therefore can be applied to various departments with an organization. Hybrid AI uses the engine of both custom-trained models and fine-tuned LLMs to provide a wide variety of queries accurately.

This assistant, therefore, reflects the famous phrase from Alan Turing, "Machines take me by surprise with great frequency." The processing of documents and queries in real time in this system is indeed surprising since it could mechanize previous manual processes.

3.1 Architecture/ Framework

Architecture modular, scalable and flexible and, lastly, secure

1. Frontend: This is front-end development using ReactJS, allowing for intuitive user interface with text as well as voice inputs.
2. Backend: Powered by both Python and NodeJS. The backend has JWT, or JSON Web Tokens, for secure authentication.
3. AI Model: The AI system takes user inputs into its process to understand and generate accurate responses using gemini. For document processing, LangChain combines use with FAISS for vector similarity searches.
4. Security: 2FA ensures secure user access with login verification through OTPs sent via email.
5. Caching: Redis is for caching frequently asked queries and responses. It is because the response time will be reduced and, hence the system efficiency increases.

It has balanced real-time data access and has ensured a safe system to operate. Hence, there won't be any lagging of the chatbot even when there is a high workload. The architecture diagram for Intelligent Enterprise AI Assistant is as shown in Figure 3.1

3.2 Algorithm and Process Design

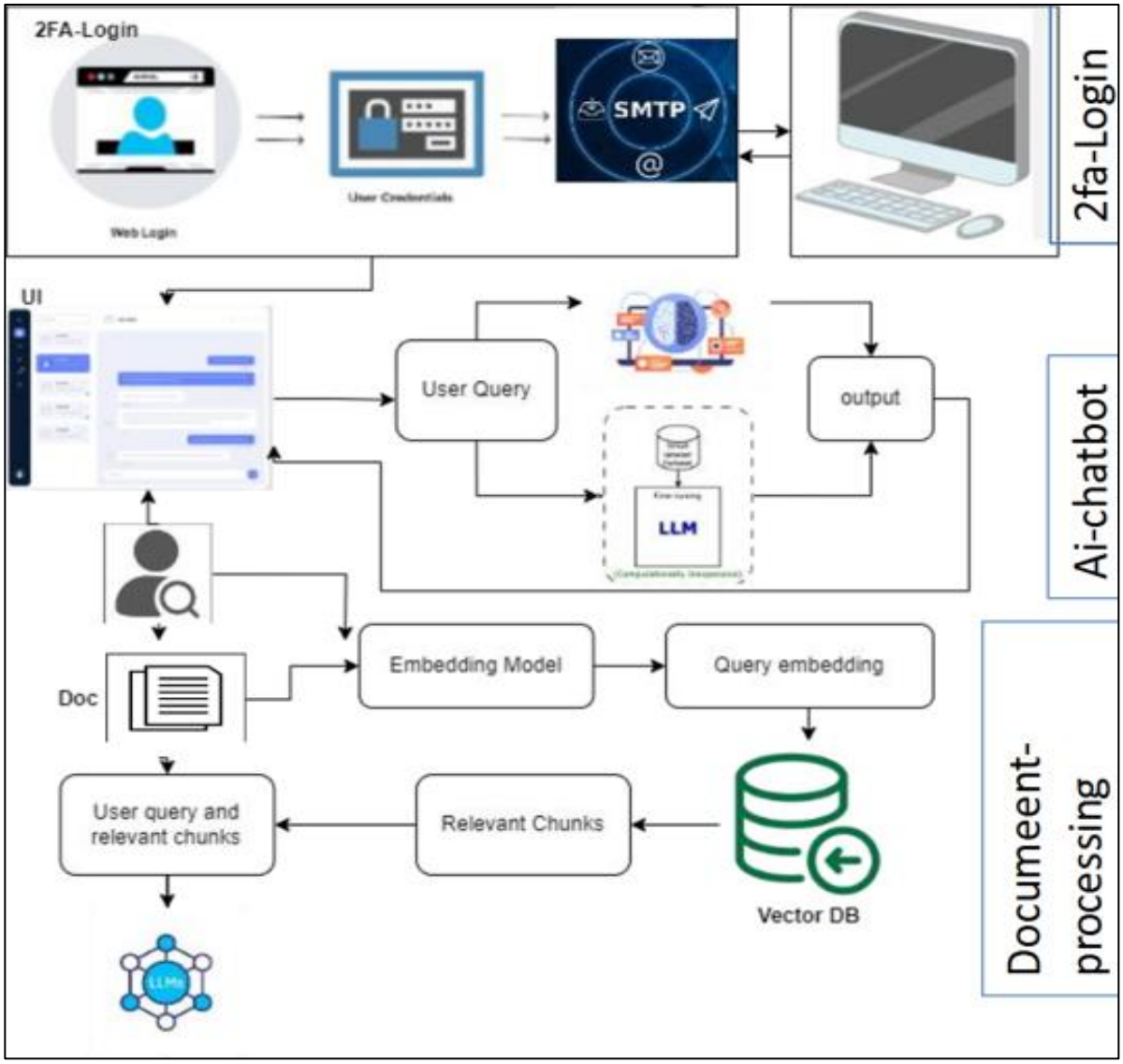

Fig 3.1 Block diagram of proposed system

It is a hybrid system in which user queries are processed as such. It follows:

1. Query Classification: Once the system receives the query, it first identifies as whether simple or complex and uses a custom-built model based on machine learning.
2. Simple Query Handling: In case of FAQs and regular questions, it uses in-memory Redis caching to easily cope up with the response.

3. Handling Complex Query: The system utilizes a fine-tuned large language model (LLM) for contextual accuracy on complex queries.
4. Document Processing: Users can upload documents that are processed in real-time for summarization and insights through a combination of LangChain and FAISS.
5. Two-Factor Authentication: The two-factor authentication lends much verification to the users' log-ins through OTP, which prevents unauthorized attempts at the sign-in procedure.

Hybrid design has proven to be efficient as this system has both fast response caching together with deep AI processing when complex tasks require more execution. It is great teamwork as Henry Ford said, "If everyone is moving forward together, then success takes care of itself." Therefore, given the structure of the collaborative system, each module works in coordination, making the AI assistant successful.

3.3 Details of Hardware & Software

Software Specifications:

- Operating System: Windows 10 or Linux-based systems
- Back-End: Python, NodeJS
- Frontend: ReactJS, Streamlit, HTML, CSS, JS

Tech Stack:

- ReactJS
  ReactJS is a popular JavaScript library used for building user interfaces, particularly single-page applications. It allows developers to create large web applications that can update and render efficiently in response to changing data. React's component-based architecture makes it ideal for creating a responsive and modular frontend for the Intelligent Enterprise AI Assistant.
- Streamlit
  Streamlit is an open-source app framework specifically for building and sharing data applications. It allows developers to rapidly create and deploy applications for machine learning and data science purposes with minimal effort. In this project, Streamlit is used to create interactive user interfaces that facilitate document uploads and chat functionality for the AI assistant.
- Python
  Python is the primary programming language used for the backend processing. It handles the AI models, document processing, and interaction between the system's various components. Python's extensive libraries, including LangChain for document handling and PyPDF2 for PDF parsing, make it ideal for this project.
- NodeJS
  NodeJS is used for building scalable server-side applications. It handles

asynchronous input/output operations in the system's backend and facilitates interaction between the user interface, the AI engine, and the database.

- Redis
  Redis is used for caching frequently asked queries and their responses, optimizing the system for quicker retrieval of data and enhancing performance.
- FAISS (Facebook AI Similarity Search)
  FAISS is employed for vector similarity searches, a critical component in the system's document processing capabilities. It allows for efficient retrieval of relevant information based on user queries.
- Firebase
  Firebase is integrated for handling two-factor authentication (2FA). This ensures that only authorized personnel have access to the chatbot by verifying their login credentials through email-based OTPs.
- Google Gemini API
  The Google Gemini API integrates large language models (LLMs) for natural language understanding and generation. In the Intelligent Enterprise AI Assistant, it enhances the system's ability to process complex queries and generate accurate, real-time responses. The API supports fine-tuning and custom workflows, making it adaptable to various organizational needs.

Hardware Specifications:

- Processor: Intel i5 or higher with multi-core capability to efficiently manage multiple operations simultaneously.
- RAM: A minimum of 16 GB to support concurrent user requests and ensure smooth real-time document processing.
- Storage: A 512 GB SSD (Solid-State Drive) is recommended for fast data storage and retrieval, enabling quick access to necessary files and information.
- Network: A high-speed internet connection is essential for uninterrupted real-time interactions between the chatbot and users.
- GPU (Optional): For deployments involving complex AI models, a dedicated GPU such as an NVIDIA RTX 2060 or higher is suggested to accelerate machine learning tasks and enhance processing speed.

This combination of software and hardware ensures that the Intelligent Enterprise AI Assistant is scalable, efficient, and adaptable for large organizations, providing real-time query handling and document processing capabilities.

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Experiment and Results for Validation and Verification

Experiments were conducted to test the performance of the system. The Query Response Time showed that the average response time for frequently asked questions (FAQs) cached in Redis was below 200 milliseconds, while queries requiring processing by the large language model (LLM) averaged around 2 to 3 seconds. In terms of Efficiency of Document Processing, the real-time document processing and summarization averaged 5 to 6 seconds per document.

For Security Validation, two-factor authentication was tested, achieving a success rate of 99 percent in OTP generation and validation. The results from these experiments demonstrate that the system effectively meets its desired performance criteria in query handling, document processing, and security measures.

Figure 4.1 depicts the home screen of the Intelligent Enterprise AI Assistant system. The main interface where users can begin interacting with the chatbot or navigate to other features like login and document processing is as shown below.

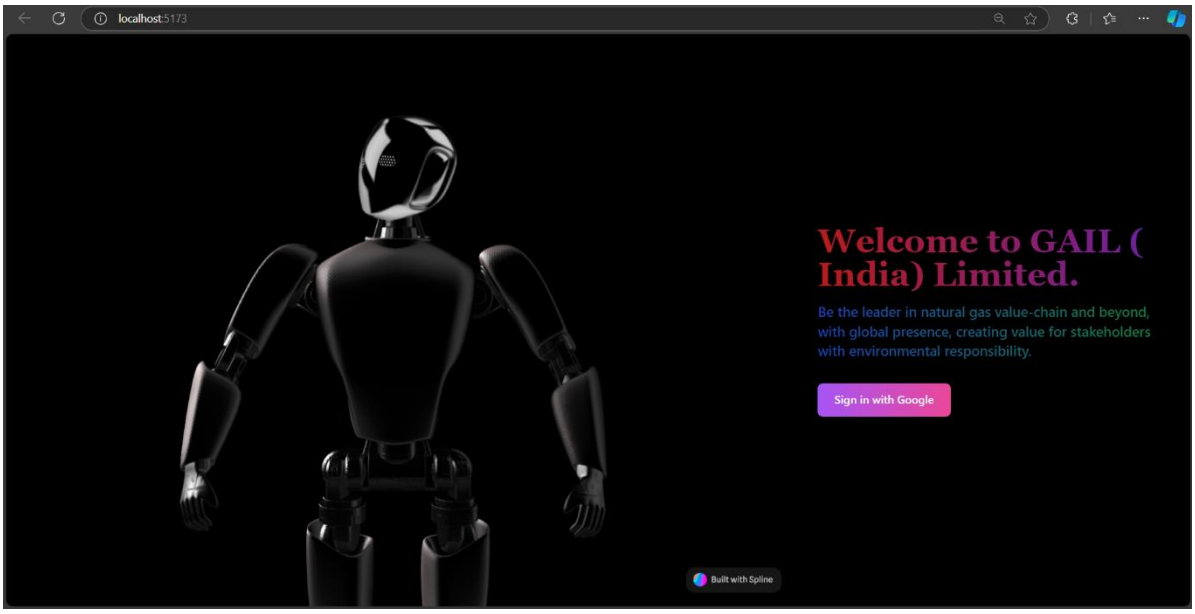

Fig 4.1: Home page

Figure 4.2 shows the login interface where users are required to enter their credentials. Two-factor authentication (2FA) via email-based OTP is highlighted as a security feature on this screen.

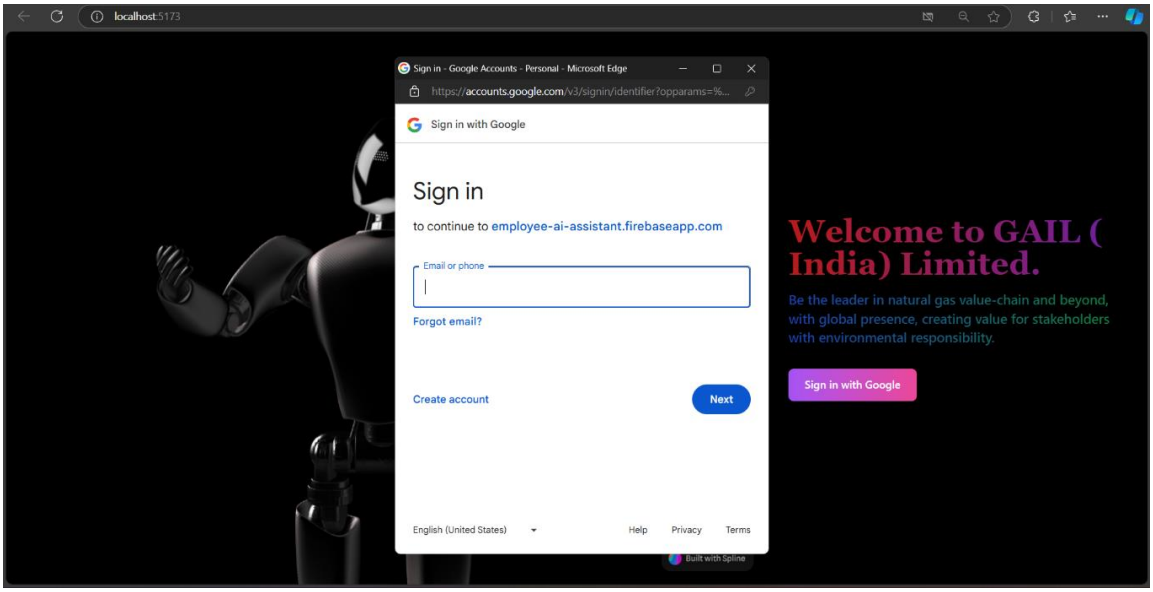

Fig 4.2: Login screen

Figure 4.3 showcases a conversation with the GAIL chatbot, where users can ask questions related to HR or IT issues, and the system responds with relevant answers based on AI and NLP processing.

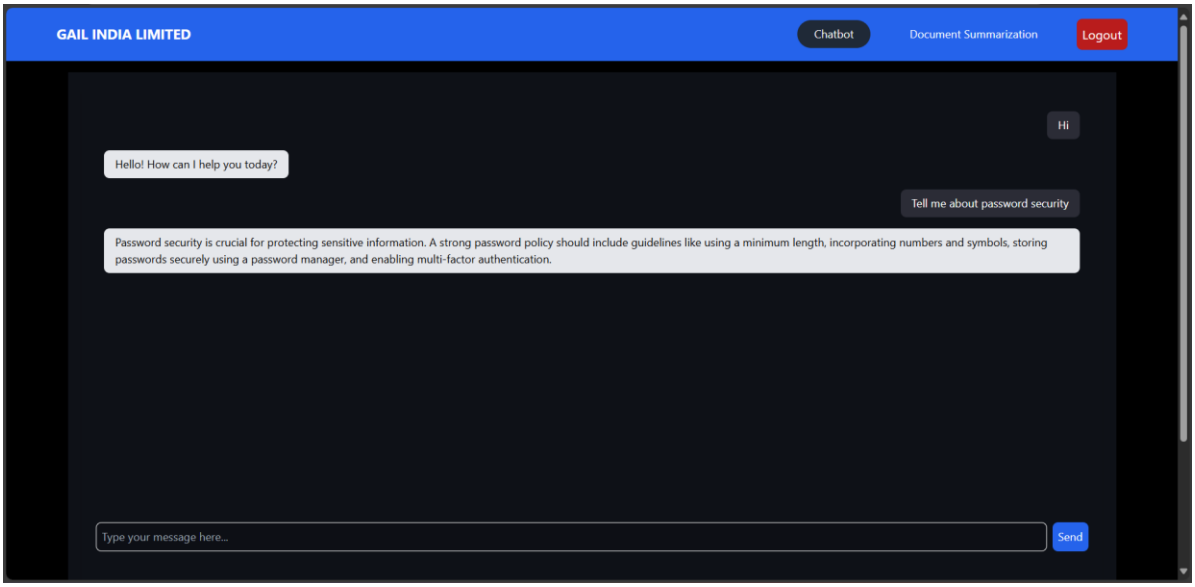

Fig 4.3: GAIL Chatbot

Figure 4.4 illustrates the system's interface where users interact with the AI chatbot by uploading Documents. The chatbot processes these documents in real time, providing responses based on the content of the uploaded file. This demonstrates the document processing capability integrated into the Intelligent Enterprise AI Assistant, enabling efficient query resolution using AI-based document summarization.

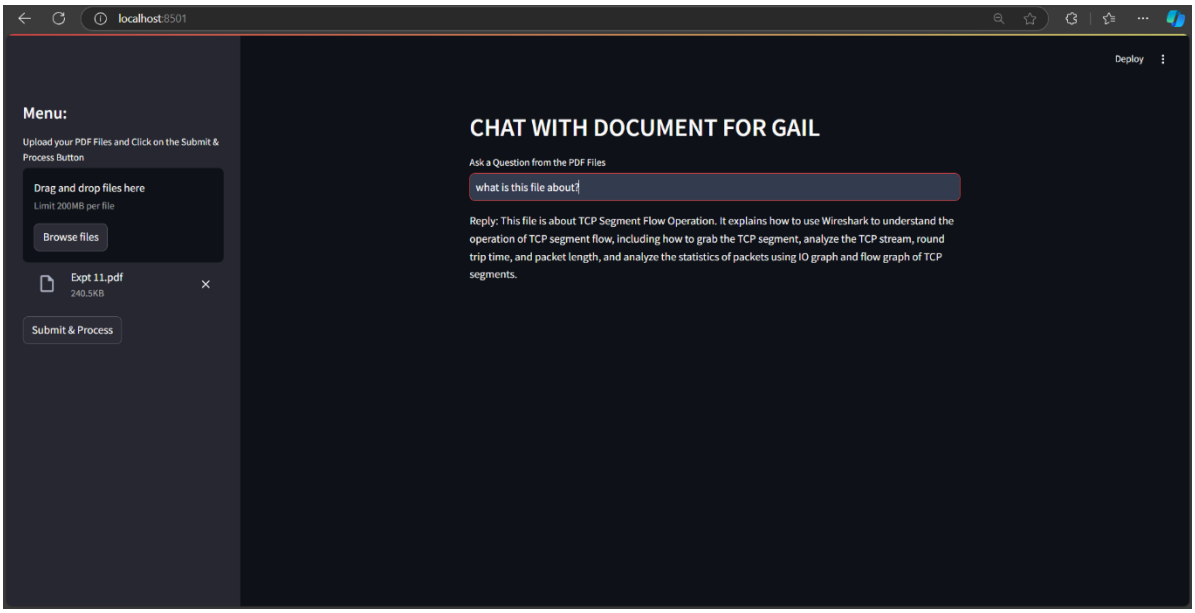

Fig 4.4: Chat with pdf

4.2 Analysis

The development of the Intelligent Enterprise AI Assistant has positively impacted organizational efficiency by automating routine tasks, resulting in faster and more accurate query handling. This reduces reliance on manual workflows, allowing employees to focus on more value-generating activities. The system effectively manages concurrent queries without performance degradation, making it ideal for large enterprises. As Lao Tzu noted, "A journey of a thousand miles begins with a single step," and this system marks a significant step toward fully automated enterprise operations, paving the way for more intelligent integrations

CHAPTER 5
CONCLUSION AND FUTURE WORK

In summary, the Intelligent Enterprise AI Assistant significantly streamlines organizational processes by leveraging advanced AI and NLP techniques. This system not only automates routine tasks but also ensures that complex queries are answered with contextual accuracy, enhancing overall responsiveness. Its ability to process documents efficiently, coupled with robust security features, makes it exceptionally well-suited for large-scale public sector organizations where operational efficiency and data protection are paramount. By freeing up human resources from mundane tasks, the AI Assistant empowers employees to focus on more strategic and value-driven activities, thereby fostering innovation and productivity. As organizations increasingly embrace digital transformation, this solution represents a critical step toward fully automated operations, paving the way for enhanced service delivery and organizational growth.

Future versions of the system could include voice recognition for more natural interactions and expand multilingual support to meet diverse organizational needs. Additionally, integrating advanced security measures, such as biometric authentication, and incorporating predictive analytics to anticipate user requirements would further enhance the system's capabilities.

REFERENCES

[1] Przegalinska, A., Ciechanowski, L., Stroz, A., Gloor, P., Mazurek, G. "A Survey on Chatbot Implementation in Customer Service," Proceedings of the 2019 International Conference on Chatbots and Conversational Agents, pp. 45-52, 2019.

[2] Gao, J., Galley, M., Li, L. "Deep Learning for Chatbots: A Review," Proceedings of the 2020 IEEE International Conference on Big Data, pp. 3321-3326, 2020.

[3] Joulin, A., Mikolov, T., Grave, E., Bojanowski, P., Maziarz, K. "Chatbot Design for Conversational Agents," Proceedings of the 2020 International Conference on Natural Language Processing and AI, pp. 110-116, 2020.

[4] Akkiraju, R., Sinha, V., Hu, T., Xu, A., Liu, Z., You, Q., Guo, Y. "Building Conversational Agents with Reinforcement Learning," Proceedings of the 2019 International Conference on Artificial Intelligence, pp. 58-65, 2019.

[5] Tewari, A., Panigrahi, B.K., Kalra, K. "A Hybrid Approach to Chatbot Development: Combining Rule-Based and Machine Learning Techniques," Proceedings of the 2020 International Conference on Intelligent Systems and Computing, pp. 99-105, 2020.

[6] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I. "Natural Language Processing for Chatbots," Proceedings of the 2020 International Conference on Machine Learning, pp. 90-100, 2020.

[7] Sutskever, I., Vinyals, O., Le, Q.V. "Sequence-to-Sequence Models for Conversational AI," Proceedings of the 2017 International Conference on Neural Information Processing Systems, pp. 3104-3112, 2017.

[8] Følstad, A., Nordheim, C.B., Bjørkli, C.A. "Evaluating the Effectiveness of Conversational Agents in Higher Education," Proceedings of the 2017 International Conference on Learning Technologies, pp. 150-157, 2017.

[9] Kenny, R., Grimes, J., Firth, A., et al. "Conversational Agents and Natural Language Processing: Trends and Advances," Proceedings of the 2021 International Conference on Human-Computer Interaction, pp. 45-52, 2021.

[10] Canhoto, A.I., Clear, F. "Exploring Chatbot Design: A User-Centric Approach," Proceedings of the 2020 International Conference on Human Factors in Computing Systems, pp. 205-212, 2020.