

# ChatMyDoc-AI Powered Document Interaction Assistant

Shreyas Chafakande \*<sup>1</sup>, Prof. A.H. Auti\*<sup>2</sup>, Khushi\*<sup>3</sup>, Shreeraj Landge\*<sup>4</sup>

Department of Computer Engineering, Sinhgad Academy of Engineering, Pune , India

---

## ABSTRACT

The exponential growth of digital data has significantly increased the demand for intelligent systems that simplify document comprehension and accelerate information retrieval. Manual reading and analysis of lengthy documents remain cognitively demanding and time-consuming, particularly in academia, research, and professional environments. This review paper examines existing artificial intelligence-based solutions in document processing, including text summarization, natural language querying, annotation systems, and knowledge visualization techniques. It highlights how these technologies aim to improve efficiency but still face limitations in usability, integration, and user interaction.

Based on the gaps identified in current systems, this paper introduces ChatMyDoc, a natural language-driven document interaction assistant designed to enhance user understanding and productivity. Built using modern web technologies and advanced AI models, the system supports PDF querying, abstractive and extractive summarization, annotation storage, and automated flowchart generation. The design focuses on reducing cognitive load, improving accessibility, and enabling intuitive interaction with digital content.

**Keywords:** Artificial Intelligence, Annotation, Document Summarization, Flowchart Generation, Information Retrieval, Knowledge Extraction, Natural Language Processing, PDF Interaction, Review Paper

---

## I. INTRODUCTION

The continuous digitalization of academic and industrial workflows has led to a rapid increase in the volume of textual information. Students, researchers, and professionals rely heavily on digital documents for learning and decision-making. However, traditional methods of interacting with documents require manual reading and analysis, which is time-consuming and cognitively demanding. This often results in reduced efficiency and the possibility of missing important information.

Recent advancements in Natural Language Processing (NLP) have enabled automated text summarization, contextual understanding, and intelligent information retrieval. These technologies have improved the way users process textual data, but most existing solutions focus on single functionalities and lack an integrated approach for better understanding and interaction.

To overcome these limitations, this paper introduces ChatMyDoc, an intelligent document interaction system developed using the MERN stack and powered by the Gemini API. The system supports PDF querying, summarization, annotation, and flowchart generation within a single platform. By combining these features, ChatMyDoc enhances user understanding, reduces cognitive load, and provides an efficient and intuitive way to interact with digital documents.

## II. LITERATURE REVIEW

The increasing shift toward digital learning and content consumption has significantly increased reliance on online documents across academic, healthcare, and corporate domains. Users frequently interact with large PDFs, research articles, and technical manuals, yet traditional document systems provide only static reading capabilities. This limitation forces users to manually extract information, leading to reduced productivity and increased cognitive load. Recent advancements in Artificial Intelligence (AI) and Natural Language Processing (NLP) have enabled intelligent document processing systems that support contextual understanding, automated summarization, and semantic information retrieval.

Early text summarization techniques were primarily based on statistical and rule-based methods, focusing on keyword frequency and sentence ranking. Over time, more advanced approaches such as clustering, graph-based ranking, and optimization techniques were introduced to improve summary quality. The emergence of deep learning and

transformer-based models has significantly enhanced the ability to capture semantic relationships in text. Research by J. Rautaray et al. [1] highlights hybrid optimization techniques for effective summarization, while S. A. Azam et al. [2] discuss current trends and challenges, particularly in handling long documents. Evaluation metrics such as BLEU [3] and ROUGE [4], along with foundational work by Nenkova and McKeown [5]

and Mikolov et al. [6], have further contributed to improving summarization accuracy and semantic understanding. Despite these advancements, existing systems still face limitations in providing interactive and integrated document understanding. Most solutions focus on individual functionalities such as summarization or retrieval, with limited support for dynamic querying, annotation, and visualization. The literature identifies key components including document preprocessing, summarization, semantic retrieval, and annotation; however, these are rarely combined into a unified platform. This gap highlights the need for intelligent systems that integrate multiple capabilities to enhance user interaction and improve overall document comprehension.

### III. METHODOLOGY

This review adopts a structured methodology to ensure the selection of credible and relevant studies in the field of document summarization and intelligent text retrieval. Research papers were collected from reliable academic databases such as IEEE Xplore, Google Scholar, and open-access repositories. The study focused on publications from 2019 to 2025 related to extractive summarization, semantic information retrieval, and AI-based document assistants. A systematic keyword-based search strategy was applied using terms such as text summarization, NLP-based interaction, and AI document systems.

To maintain research quality, inclusion and exclusion criteria were applied during paper selection. Only peer-reviewed and freely accessible studies with clear methodologies and evaluation metrics such as ROUGE and BLEU scores were included. Irrelevant or outdated studies were excluded to ensure accuracy and relevance. Each selected paper was analyzed based on summarization techniques, semantic retrieval methods, system interactivity, and practical limitations. In total, ten research papers were critically examined to understand existing approaches and their effectiveness.

The analysis was carried out in multiple phases, including literature collection, filtering, comparative evaluation, and gap identification. Based on the findings, key limitations such as lack of integration, limited interactivity, and absence of visualization features were identified. These insights guided the design of the proposed system, ChatMyDoc, which is developed using the MERN stack, integrated with the Gemini API for intelligent processing, and Cloudinary for storage. This methodology ensures a systematic, unbiased, and goal-oriented approach toward developing an efficient document interaction system.

To further strengthen the analysis, a comparative framework was established to evaluate the selected studies based on key performance indicators such as accuracy, efficiency, scalability, and user interaction capabilities. Each system was examined for its ability to handle large documents, support real-time querying, and provide meaningful summaries. Special attention was given to identifying whether the systems supported interactive features such as follow-up queries, contextual understanding, and adaptive responses, which are essential for modern document interaction platforms.

### IV. ARCHITECTURE

The architecture of the proposed system, ChatMyDoc, is based on a web application developed using the MERN stack, enabling seamless interaction between users and intelligent document processing services. On the frontend, users can sign up or log in and access a dashboard where they can upload PDF documents, perform natural language queries, generate summaries, and create flowcharts. The user interface is designed to provide an intuitive experience, allowing easy navigation, result viewing, and annotation management. Upon logout, users are redirected to the home page.

The backend consists of multiple modules, including authentication, document processing, query handling, summarization, annotation management, and visualization. The system integrates the Gemini API for advanced natural language understanding and semantic processing, while Cloudinary is used for efficient document storage. A central MongoDB database manages user data and document information. This modular architecture ensures scalability,

efficient processing, and a clear separation of concerns, resulting in a robust and user-friendly document interaction system.

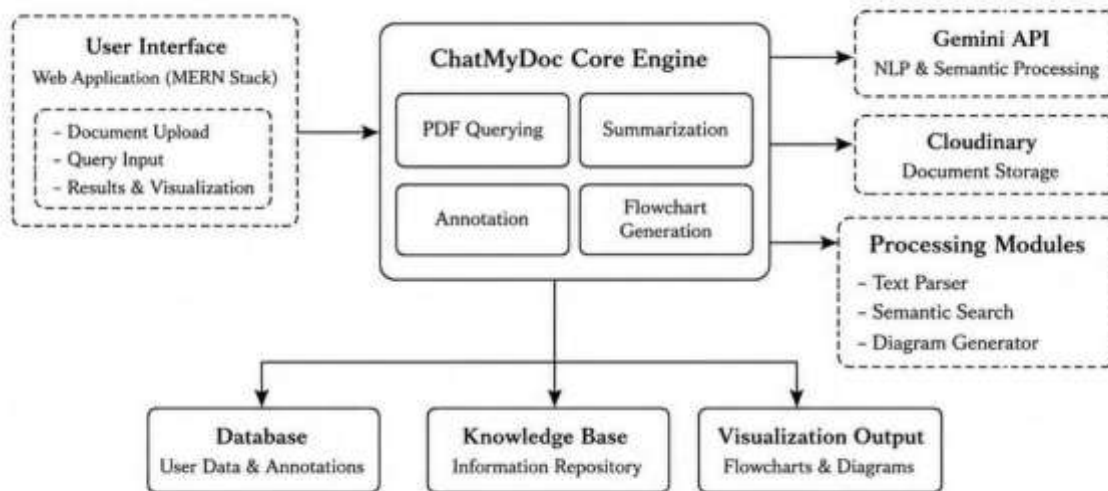


Figure 1: Architecture Diagram

## V. IMPLEMENTATION OVERVIEW

The implemented system, ChatMyDoc, demonstrates a fully functional and user-friendly platform designed for intelligent document interaction. The application integrates multiple modules such as authentication, document management, AI-based querying, and visualization into a unified system. The user interface follows a modern dark theme design, ensuring better readability and an enhanced user experience.

The authentication module provides secure login and registration functionalities. The login interface allows users to enter their credentials, while the registration process collects essential user details such as username, email, and password with confirmation. These details are securely stored and validated through backend services, ensuring safe and controlled access to the system.

The main dashboard serves as the central hub where users can manage their documents efficiently. It displays uploaded PDF files along with details such as file size and upload date, and provides options to open or interact with documents. The dashboard also shows user activity statistics, helping users track their usage. The document upload feature enables users to store files seamlessly, which are then processed for further operations.

The core functionality of ChatMyDoc lies in its ability to interact with documents using natural language queries. Users can ask questions related to their uploaded PDFs, and the system generates context-aware responses using advanced AI processing. In addition to querying, the system supports features such as AI-based summarization, document flow visualization, table extraction, and image extraction, which enhance the overall understanding of document content.

Overall, the implementation ensures smooth integration between frontend and backend components, providing real-time responses and efficient document processing. The system is modular, scalable, and designed with a focus on usability, making it highly suitable for academic, research, and professional applications.

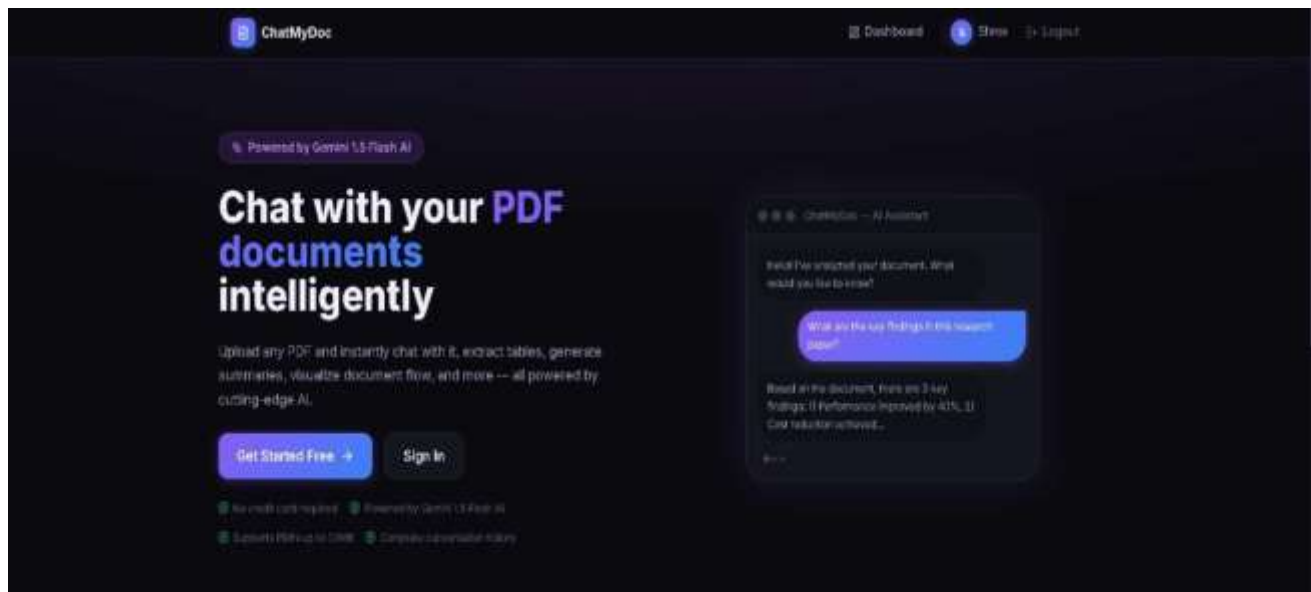


Figure: 2

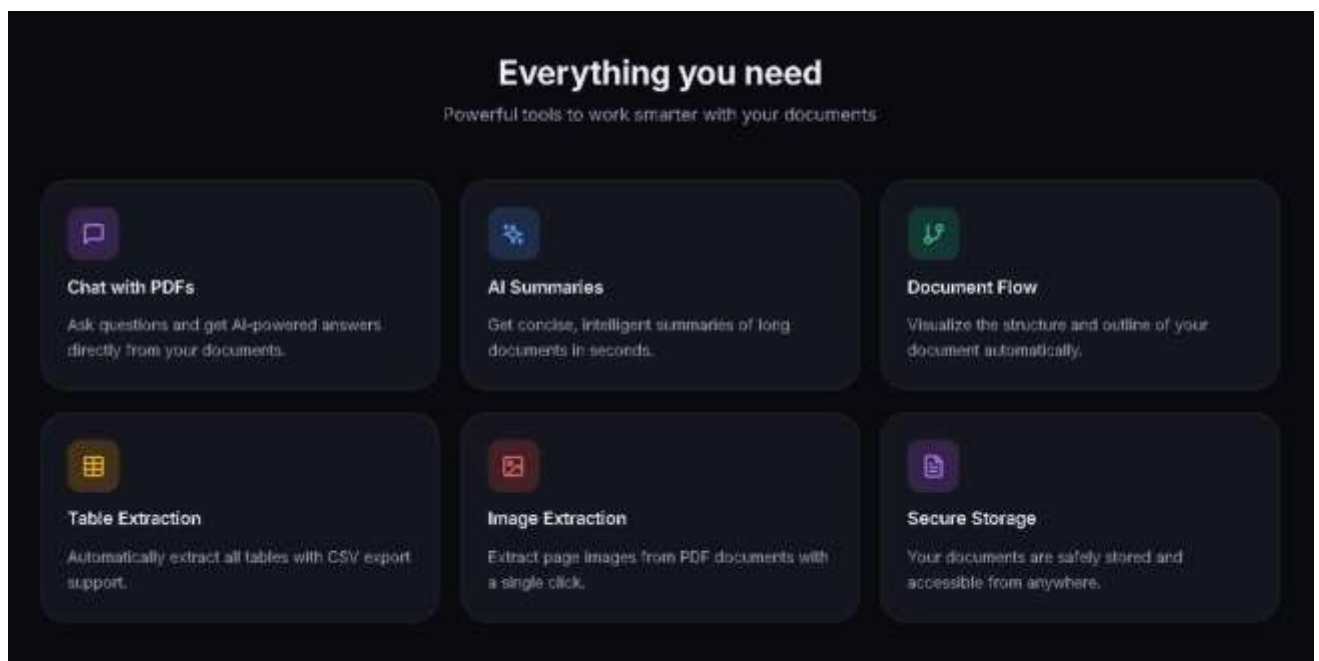


Figure: 3

#### Features:

The Chat with PDFs feature allows users to interact with their uploaded documents using natural language queries. Users can ask questions directly related to the content, and the system provides intelligent, context-aware answers. The AI Summaries feature generates concise and meaningful summaries of long documents, helping users quickly grasp important information without reading the entire content. Additionally, the Document Flow feature visualizes the structure of the document, enabling users to understand the logical flow and relationships between different sections.

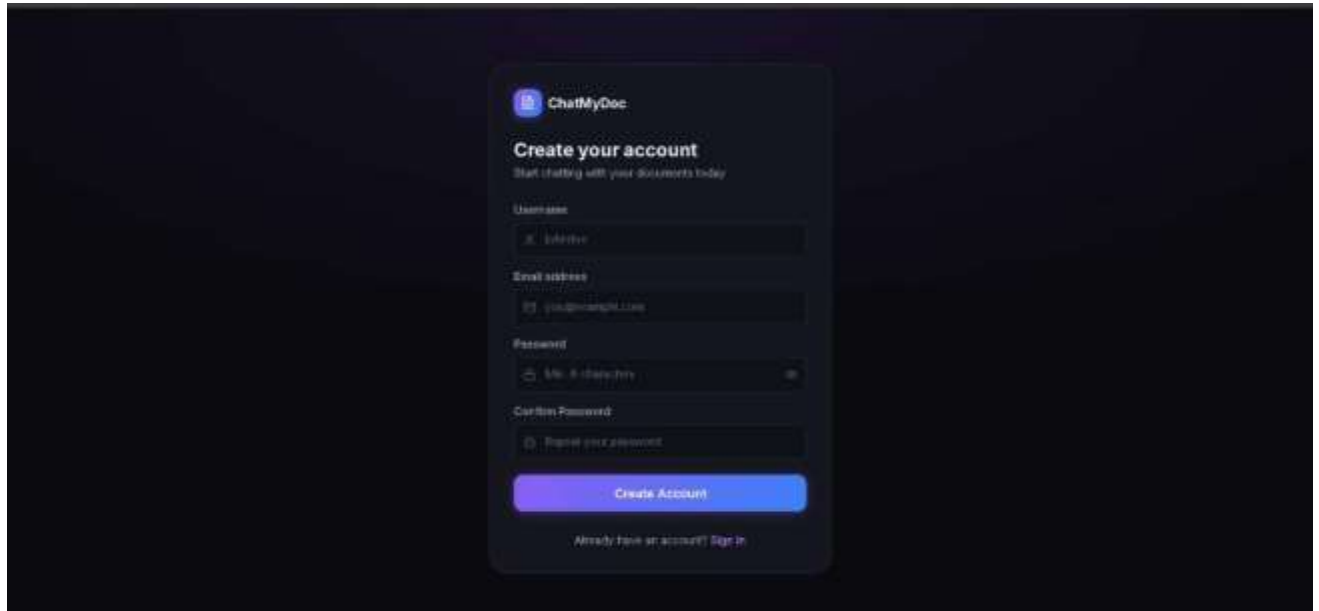


Figure : 4

#### Registration Page:

The signup page of ChatMyDoc provides a simple and secure interface for new users to create an account and access the system. The design follows a modern dark-themed layout with a clean and user-friendly structure, ensuring ease of use and better visual experience.

The interface includes input fields for username, email address, password, and confirm password, allowing users to enter their credentials accurately. Password validation ensures basic security requirements, such as minimum character length. The confirmation field helps prevent errors during account creation. All user data entered through this form is securely transmitted to the backend and stored in the database for authentication purposes.

The page also includes a "Create Account" button that initiates the registration process. Additionally, users who already have an account are provided with a direct navigation link to the login page, ensuring smooth user flow. Overall, this interface ensures a secure, intuitive, and efficient onboarding experience for new users of the ChatMyDoc system.

#### Login Page:

The login page of ChatMyDoc provides a secure and user-friendly interface for existing users to access the system. It features a clean, modern dark-themed design with input fields for email and password, ensuring easy and efficient credential entry. Password masking is implemented to maintain security and privacy. The "Sign In" button initiates the authentication process, and upon successful login, users are redirected to their personalized dashboard. Additionally, a navigation link is provided for new users to create an account, ensuring a smooth and seamless user experience.

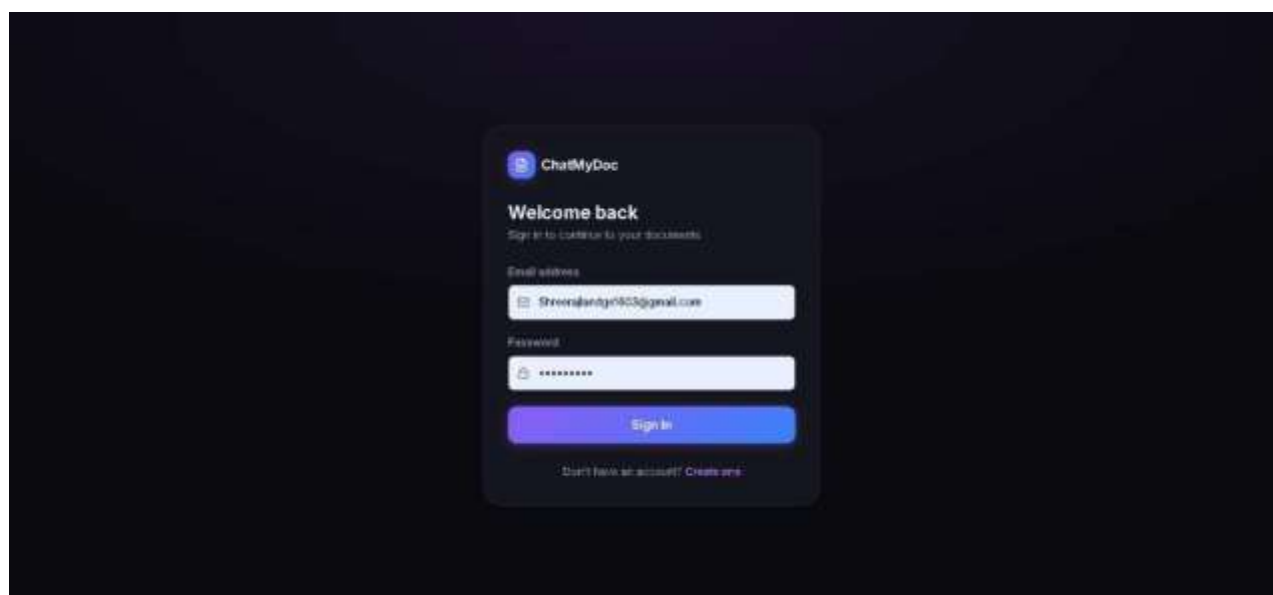


Figure: 5

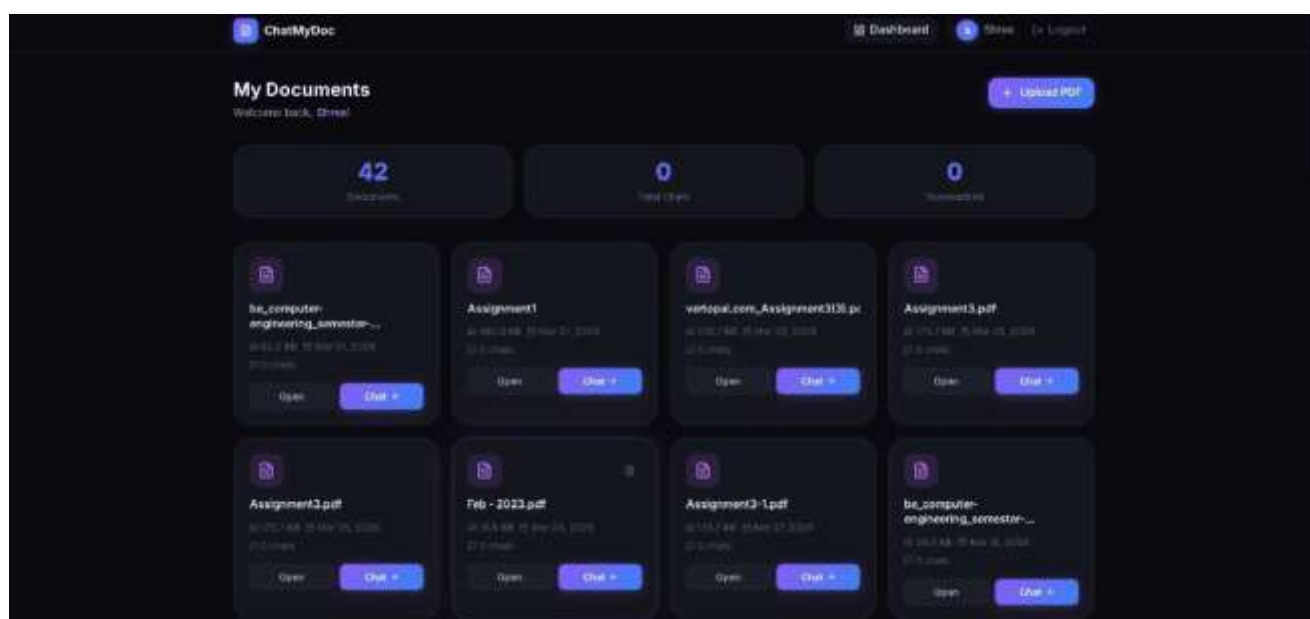


Figure: 6

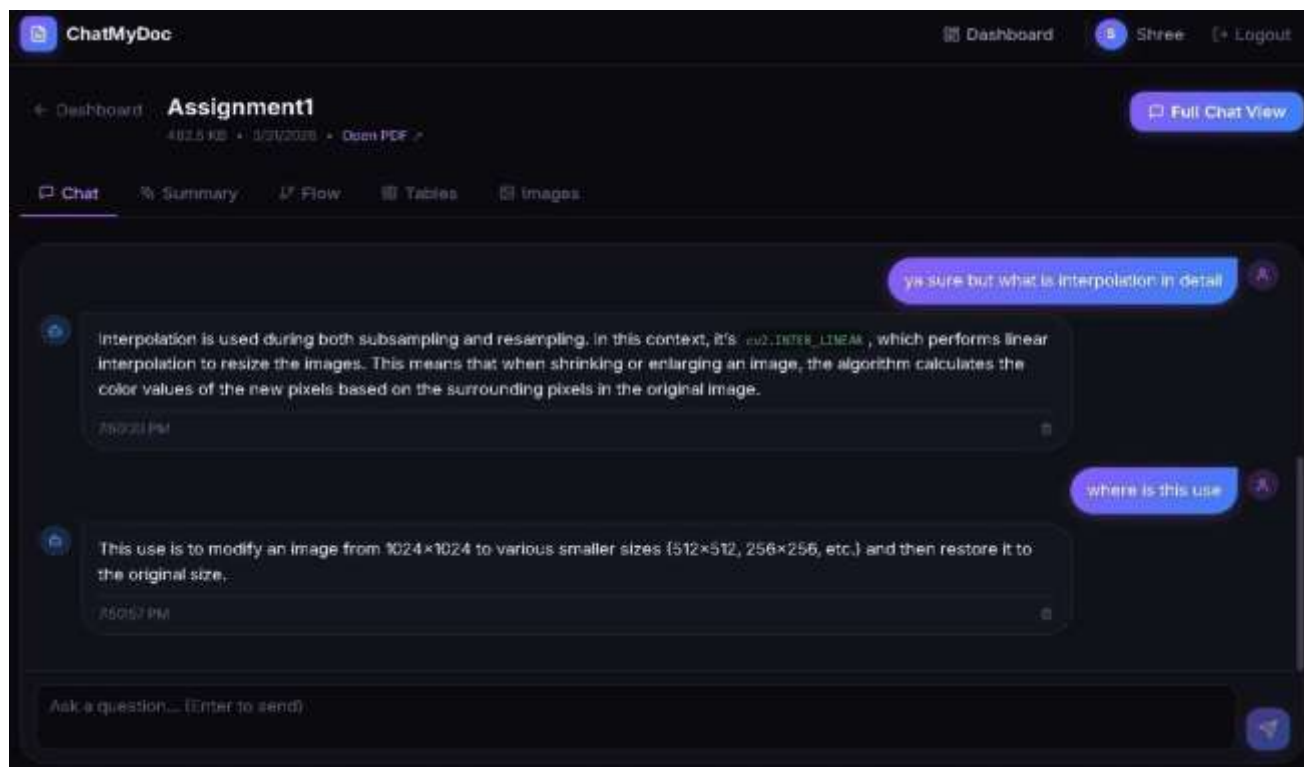


Figure : 7

## VI. RESULTS AND DISCUSSION

### 1. Key Findings of the Survey

This study highlights several important observations regarding the use of Artificial Intelligence and Natural Language Processing in document interaction systems. Based on the reviewed literature and system implementation, it is observed that AI-based document assistants significantly improve the efficiency of information retrieval and understanding. Text summarization techniques help users quickly grasp key points from large documents, while semantic query systems enable users to interact with documents using natural language. Additionally, integrating multiple functionalities such as querying, summarization, annotation, and visualization into a single platform enhances overall usability and user experience, which is effectively demonstrated in the ChatMyDoc system.

### 2. Challenges in the Field

Despite advancements in document processing technologies, several challenges still exist. One major issue is handling large and complex documents while maintaining contextual accuracy in summarization and query responses. Another challenge is ensuring efficient processing speed, especially when dealing with real-time user queries. Additionally, understanding diverse user queries and providing accurate, context-aware responses remains a complex task. Integration of multiple features into a single system without affecting performance and scalability is also a significant challenge in developing intelligent document interaction platforms.

### 3. Limitations of the Survey

While this study provides a comprehensive overview of AI-based document interaction systems, certain limitations are present. The scope of the review mainly focuses on summarization and semantic retrieval techniques, while other advanced areas such as multimodal AI and deep contextual learning are not explored in detail. Furthermore, the analysis is primarily based on existing literature and system design, with limited real-world user testing and empirical validation. Additionally, the system performance may vary depending on document complexity and size, which may require further optimization in future improvements.

## VII. CONCLUSION

This paper presented ChatMyDoc, an intelligent document interaction system designed to simplify document understanding and improve information retrieval. By integrating features such as PDF querying, summarization, annotation, and flowchart generation, the system provides a unified and efficient platform for users to interact with digital documents. The use of modern technologies like the MERN stack and AI-based processing enhances performance and usability. Overall, ChatMyDoc effectively reduces cognitive effort and offers a practical solution for academic, research, and professional applications.

## VIII. ACKNOWLEDGEMENTS

We would like to express our sincere gratitude to all those who supported and guided us throughout the development of this project, ChatMyDoc. We are especially thankful to our project guide for their valuable suggestions, continuous encouragement, and technical guidance, which helped us in successfully completing this work.

We also extend our gratitude to our department and institution for providing the necessary resources and environment to carry out this project. Finally, we would like to thank our friends and peers for their support and cooperation during the development process.

## IX. REFERENCES

- **-Based Feature Extraction Technique for Single Document** *IEEE Access*, vol. 13, pp. 24515–24521, Feb. 2025.
- **S. A. Azam et al., “Current Comprehensive Review,”**
- **K. Papineni et al., “BLEU: A Method for Automatic Evaluation of Machine Translation,”** in *Proc. ACL*, pp. 311–318, 2002.
- Trends and Advances in Extractive Text Summarization: A** *IEEE Access*, vol. 12, pp. 114586–114612, 2024.
- **C. Lin, “ROUGE: A Package for Automatic Evaluation of Summaries,”** in *Proc. ACL Workshop*, pp. 74–81, 2004.
- **Found. Trends Inf. Retrieval**, vol. 5, no. 2–3, pp. 103–233, 2012.
- **“Distributed Representations of Words and Phrases and their”** *Proc. NeurIPS*, 2013.