**RESEARCH ARTICLE**

# Co-Pilot for Project Managers: Developing a PDF-Driven AI Chatbot for Facilitating Project Management

**KHORSHED ALAM, MAHBUBUL HAQ BHUIYAN, (Graduate Student Member, IEEE), MOHAMMAD SHAFIQUL ISLAM, ABUL HOSSAIN CHOWDHURY, ZAHEED AHMED BHUIYAN, AND SUMAN AHMMED**
Department of Computer Science and Engineering, United International University, Dhaka 1212, Bangladesh

Corresponding author: Suman Ahmmed (suman@cse.uiu.ac.bd)

**ABSTRACT** Our AI-driven PDF Chatbot is specialized for Project Management (PM) Automation and acts as a virtual Project Manager that offers continuous support to global teams. It interprets PDF data like SRS reports and interview transcripts by utilizing Open-Assistant's SFT-1 12B Model. Insights from interviews of 15 project managers have enriched the knowledge base of our chatbot and ultimately enabled informative responses to the stakeholders of the project. Advanced AI techniques ensure efficient text preprocessing, including tokenization, numerical normalization, lowercasing, removing punctuation, removing extra spaces, recursive character text splitter, and lemmatization. It is primarily tailored for e-commerce project and provides precise guidance based on e-commerce data and risk management factors. With an average cosine similarity of 80.80% and semantic similarity score of 85.21%, it consistently aligns with PDF Contents and optimize the project management phases & methodologies. This innovation enhances Human-Robot Interaction, PM Automation, facilitates decision-making, and enables uninterrupted communication. While AI-driven PDF chatbots like ChatPDF and SciSummary exist, our chatbot is uniquely focused on automating project management tasks, providing tailored insights for e-commerce projects and decision-making, thus offering a breakthrough in PM automation. To ensure the chatbot's robustness in context-aware responds, we compare our chatbot with ChatPDF and Sci-summary which are some PDF driven chatbots. Making our work available open-source on https://github.com/codewithkhurshed/SPM-project-repo can enhance its accessibility and promote future research opportunities in PDF driven chatbot development.

**INDEX TERMS** Project management automation, AI Chatbot, PDF data interpretation, open-ASSISTANT SFT-1 12B model, E-commerce projects, text preprocessing, human–robot interaction, semantic similarity, context-aware responses, decision-making support.

## I. INTRODUCTION

Project management is the capacity to lead a team with relevant skills, processes, and experience towards achieving the organizational goal within the given constraints [1]. Effective communication is one of the key components within the team & stakeholders in order to make the project successful. For project managers, the task of overseeing a project often resembles the complexity of solving a Rubik's cube [2]. In the context of remote work, the tasks of the project manager become much more complex, especially because of the different time zones of team members from different locations. While managing a remote team, it is very important to:

- Establish clear communication protocols and working hours.

The associate editor coordinating the review of this manuscript and approving it for publication was Orazio Gambino.

- Use software tools, especially for scheduling the work and to find out overlapping times of meetings.
- Foster a culture of understanding and flexibility in different time zones.
- Documentation of important information and decisions of team members so that other stakeholders may go through the documentation when the team members are offline.
- Encourage asynchronous communication where possible to accommodate different schedules.

Project managers are tasked with the intricate coordination of various project components while simultaneously considering multiple potential outcomes and accurately estimating the project's timeline for completion. Throughout the project's life cycle, common challenges tend to surface. These include the need for a precise definition of the project's scope, meticulous management of the project's budget, the imperative of maintaining effective communication within the project team, and the necessity to address unforeseen obstacles as they arise. A project manager's adeptness in navigating and successfully resolving these challenges plays a pivotal role in ensuring the overall success of the project they are responsible for. However, despite technological advancements, there is a notable absence of specialized AI chatbots to assist project managers in their absence, acting as AI project managers to team members and stakeholders in their project-related duties.

The AI-driven PDF Chatbot for Project Management offers project managers a valuable tool for making informed decisions by providing continuous communication and support to global and remote teams. By automating information retrieval from PDF sources and utilizing advanced AI techniques such as Large Language Models (LLMs), the chatbot streamlines the process of accessing relevant data and insights. Through text data preprocessing techniques and semantic understanding, the chatbot can generate coherent responses to queries, drawing from a dataset that includes e-commerce data, software specifications, and risk management scenarios. Acting as a virtual project manager, the chatbot fills a crucial role in ensuring uninterrupted management, especially in the absence of a human manager, thus enhancing decision-making processes in project management [3], [4].

Our primary goal is to facilitate seamless communication on behalf of the project manager through the integration of advanced AI technology and natural language processing. As there are no such works available that will automate project management by using an AI chatbot to work as a virtual project manager when the actual project manager is offline, we have developed an AI Chatbot designed to function as a project manager to automate project management. This involves constructing a specialized chatbot dataset and training it using the Large Language Model (LLM) provided by OpenAssistant's oasst-sft-1-pythia-12b. In order to make our chatbot more responsive, we used text data preprocessing, including techniques such as tokenization, numerical normal-

ization, lowercasing, removing punctuation, removing extra spaces, recursive character text splitter, and lemmatization. Furthermore, we have developed a web interface to use the chatbot. Our key contribution is an approach for automating information retrieval by designing a PDF-driven chatbot using LLMs in the context of acting as a virtual project manager by answering questions. Thus, the developers can easily communicate information related to the project by simply asking the chatbot. This paper includes the following features and contributions:
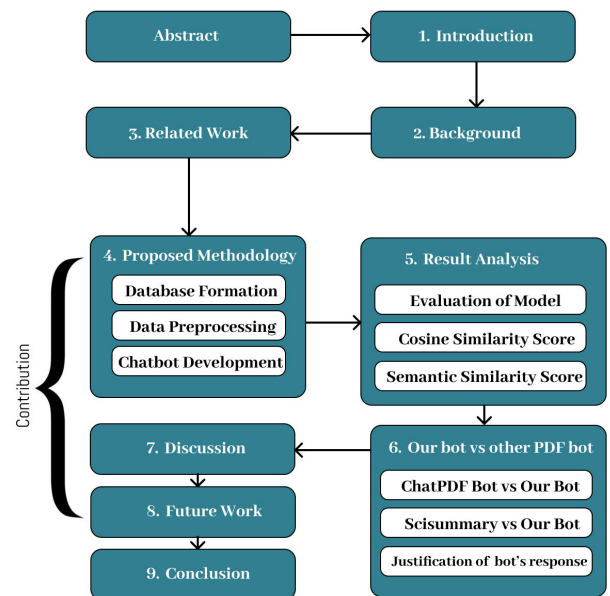


**FIGURE 1.** Paper organization.

- Creation of a PDF-driven AI chatbot: Designed for use in mobile and web applications, it acts as a substitute project manager. This will help to automate project management using an AI chatbot, which is yet unseen in the research domain.
- Dataset Development: Building a tailored dataset focused on e-commerce project development, encompassing data, specifications, and decision-making scenarios common in this field.
- Addressing a Market Gap: Due to the recent innovations in PDF-driven chatbots in 2023, there are only a handful of options available. We've chosen ChatPDF and Sci-summary for benchmarking purposes because they represent the leading edge in handling diverse PDF content. Our own PDF chatbot, however, takes a unique approach. It's a project management-specific PDF-driven AI chatbot, a niche currently unoccupied in the market.
- Enhancing Project Management: Aiming to provide a helpful assistant for project teams, especially during the absence of a human project manager.

The paper is organized into several sections and subsections to provide a comprehensive overview of the

development of a PDF-driven AI chatbot for streamlining project management, as shown in Fig. 1. It begins with an introduction in Section I, a background section which covers the contextual knowledge of this project, a literature review in Section III, and a discussion of existing works in the field. Section IV goes into data and methodology, detailing the dataset creation, preprocessing, and methodology used in the project. Following this, Section V explains the evaluation criteria for the chatbot's performance, while Section VI discusses the comparison between the proposed chatbot and existing PDF chatbots. The discussions of the system are presented in Section VII, offering insights into the proposed chatbot's effectiveness. The paper concludes with a summary of the work, outlining potential future directions for research and development in sections VIII, IX and X.

## II. BACKGROUND
This section provides a contextual overview of the topics pertinent to our research, designed to assist readers from disciplines beyond computer science.

### A. INTRODUCTION TO PROJECT MANAGEMENT (PM) AUTOMATION
Project Management Automation streamlines project workflows by integrating software tools to handle routine tasks and processes. It utilizes task automation to manage scheduling, notifications, and progress tracking, reducing manual effort and error [5]. Workflow automation ensures tasks are assigned, deadlines are met, and dependencies are managed without manual intervention. Resource management tools optimize the allocation of team members and materials [6]. Automated reporting generates real-time project status updates and performance metrics, aiding in decision-making. Integration with other systems ensures seamless data flow and consistency. Risk management automation identifies potential issues and suggests mitigation strategies. Collaboration features enhance team communication and document sharing. Overall, automation improves efficiency, accuracy, and project visibility.

### B. PDF CHATBOT
The PDF Chatbot is an advanced system where PDF file serves as the data source and the chatbot functions as the search engine [7]. It uses a PDF Parsing Module to extract text and structured data from the PDF. This text is indexed by a Search Engine to facilitate quick and accurate searches. The Natural Language Processing (NLP) Engine interprets user queries, recognizing intent and extracting relevant entities. Query Processing translates user input into structured search requests. Results are formatted and presented in the chat interface, with options to navigate directly to relevant sections of the PDF. The system ensures secure handling of data and integrates seamlessly with various platforms [8].

### C. PDF CHATBOT FOR BUSINESS INTELLIGENCE
The PDF Chatbot for Business Intelligence (BI) is a specialized tool designed to extract actionable insights from PDF documents containing business data [9], [10]. It employs a PDF Parsing Module to convert document content into a structured, searchable format. The Indexing Engine organizes this data to enable efficient retrieval. Utilizing Natural Language Processing (NLP), the chatbot interprets user queries related to business metrics, trends, and reports. It applies Data Extraction Techniques to pull relevant information and presents it in a user-friendly interface. The system supports dynamic querying of financial reports, performance summaries, and other BI-related documents. Security features ensure that sensitive business information is protected. The chatbot enhances decision-making by providing quick, contextual answers and summaries from complex PDFs.

### D. E-COMMERCE PROJECT MANAGEMENT
E-commerce project management presents unique challenges due to its fast-paced and highly competitive nature. The need to manage diverse teams, coordinate supply chains, and adapt to rapidly changing consumer demands requires efficient decision-making and seamless communication. Additionally, handling large volumes of unstructured data from product catalogs, customer reviews, and transaction records complicates the management process. Risk management becomes critical as projects are often influenced by market volatility, logistical uncertainties, and evolving technologies. Traditional approaches often fall short in addressing these complexities due to their inability to provide real-time insights and predictive capabilities.

AI-powered solutions play a pivotal role in addressing these challenges by automating repetitive tasks, extracting actionable insights from unstructured data, and facilitating informed decision-making [11], [12]. AI-driven chatbots, for instance, streamline communication among stakeholders by providing context-aware responses and ensuring uninterrupted project management workflows. Furthermore, AI algorithms analyze historical data to identify trends, predict potential risks, and suggest optimal strategies tailored for e-commerce. These capabilities not only enhance operational efficiency but also enable project managers to focus on high-priority tasks, ultimately driving success in a competitive e-commerce landscape.

## III. RELATED WORK
Recent advancements in the field of AI and chatbots are highlighted through various innovative methodologies. There are limited papers on PDF-driven Chatbot development. For this reason, we considered other types of chatbots for the literature review to gain domain knowledge and also cited papers on PDF-driven chatbots papers. Studies such as shown in Table 1, the AI Decision Assistant & Chatbot for Software Release Planning and Optimized Resource

**TABLE 1.** Previous work.

| Paper | Methodology | Limitations |
|-------|-------------|-------------|
| [13] | AI Decision Assistant & Chatbot, algorithms like XGBoost and K-NN | Lacked ML for task duration prediction |
| [14] | Medical diagnosis chatbot, symptom identification and remedy suggestions | Difficulties in appointment booking implementation |
| [15] | Prisma-based systematic literature review on chatbots in digital business transformation | Research gap due to specific 'chatbot' term focus |
| [16] | AI-based chatbot for hospital management, LSTM neural network | Small sample size, limited generalizability |
| [17] | "Medic agent" hypertension self-management chatbot, Google Cloud's Dialog flow | Small, geographically limited sample, task order biases |
| [18] | AI chatbot for agriculture, Sentence Transformer model; Virtual manager for medical practitioners, Named Entity Extraction model | Less effective for diverse agricultural needs |
| [19] | Virtual manager for medical practitioners, Named Entity Extraction model. | Challenges in dynamic solutions, high maintenance costs |
| [20] | Comprehensive survey on AI chatbots development and potentials | Lack of real-life implementation |
| [21] | Ophthalmology chatbot system, generative model trained on doctor-patient dialogues | No formal evaluation |
| [22] | AI chatbots in educational recommendation, railway technical school context | Limited broader applicability |

Allocation utilized algorithms like XGBoost and K-NN to create a web application, although it lacked machine learning capabilities for task duration prediction [13]. Authors in the paper [14] discussed that the development of a medical diagnosis chatbot marked significant progress in healthcare, offering symptom identification and remedy suggestions. Yet the authors faced difficulties in appointment booking programming. In the digital business transformation, the use of chatbots was examined through a systematic literature review, revealing a research gap due to the focus on papers explicitly using the term 'chatbot' [15].

In a study [16], an AI-based chatbot for hospital management showcased high training accuracy using an LSTM neural network but was limited by its small sample size and generalizability. The authors of the paper [17] developed a hypertension self-management chatbot named "Medic agent" using a user-centered design process and Google Cloud's Dialog flow. Ten participants, with an average age of 60 and a half female, completed usability testing. The study, however, faced limitations like a small, geographically limited sample size and potential biases in task order and duration variability. The field of agriculture also saw the introduction of an AI chatbot using a Sentence Transformer model, adept in addressing farming queries but less effective with complex or diverse agricultural needs [18].

Further explorations in the realm include the creation of a virtual manager for medical practitioners, employing a Named Entity Extraction model but facing challenges in dynamic solutions and high maintenance costs [19]. A comprehensive survey on AI chatbots provided an overview of development and potential but lacked real-life implementation [20]. In ophthalmology, a chatbot system was developed using a generative model trained on doctor-patient conversations, though it did not undergo formal evaluation [21]. Lastly, an intelligent educational recommendation platform utilizing AI chatbots was confirmed for its efficacy

in a railway technical school context, but its specific focus raised questions about its broader applicability [22]. These studies collectively underscore the diverse applications and challenges in the field, offering insights into future research directions and potential improvements in AI and chatbot technologies. However, none of the papers talked about the implantation of an AI chatbot for project managers.

Unklabot 1.0 [25] is an innovative AI-powered chatbot specially designed and proposed for information services at Klabat University with 85% accuracy. It uses LLM (OpenAI GPT 3) with intent recognition and knowledge-based context understanding using semantic searching techniques to retrieve information in external knowledge based on the closest context. The external knowledge is stored in a different chroma vector embedding database. First, the data collection is done, and then it is split into chunks and stored in the chroma database. The information is processed into vector embeddings for a more appropriate answer. The cosine function and the KNN algorithm are used here for measuring the similarity between two sets of documents. The AI language model used here is text-davinci-003 for a more appropriate response.

A system [26] to automate information retrieval from faculty guidelines using a pdf-driven chatbot powered by OpenAI ChatGPT with a Unigram BLEU score of 0.84, a Bigram BLEU score of 0.87, and a loss value of 0.19. This chatbot is aimed at reducing the discomfort of extended reading periods using e-books. The chatbot can generate coherent responses, which are in the PDF. The OpenAI GPT model can be used for translation, summarization, and question-answering systems. The chatbot uses LLM in the context of answering faculty guidelines questions. The LangChain framework, ChatGPT-3.5 and Pinecone are utilized for generating responses.

The paper [27] focuses on the potential risks of using AI tools such as ChatGPT. ChatGPT can be used for

dishonesty in academic examinations. ChatGPT-generated results negatively affect academic integrity. Studies have shown that ChatGPT can be used for cheating and plagiarism, and institutions should take effective steps to prevent this misconduct. Although content generation tools can prevent plagiarism, education and ethics are also very crucial. New formats of assessments that focus on integrity and ethics can prevent this cheating and ensure academic integrity. Authors [28] examine the effectiveness of language-model-powered chatbots in resolving queries related to automotive manuals stored in PDF format. The study evaluates the performance of such chatbots in providing accurate and relevant information to users seeking assistance with understanding or troubleshooting automotive issues. However, the chatbot is unable to reply to the questions where the answers rely on visual elements. Overall, the paper provides valuable insights into the use of advanced NLP technologies for improving user experience and efficiency in accessing information from complex technical documents like automotive manuals stored in PDF format.

Authors [29] offer a comprehensive overview of the methodologies and challenges associated with the implementation of AI-driven conversational chatbots over the period from 1999 to 2022. The paper serves as a valuable resource for researchers, developers, and practitioners interested in understanding the evolution, methodologies, and challenges associated with the implementation of AI-driven conversational chatbots. The authors may offer insights into future trends and directions in AI-driven chatbot development. This could include recommendations for overcoming existing challenges, adopting emerging technologies, and advancing the state-of-the-art in conversational AI. The paper [26] explores the development of a PDF-driven chatbot utilizing OpenAI's ChatGPT to automate information retrieval from faculty guidelines. The paper introduces the problem of manually retrieving information from faculty guidelines, which can be time-consuming and inefficient. It proposes the use of a chatbot powered by OpenAI's ChatGPT to automate this process. The paper provides insights into the design and implementation of a chatbot solution to automate information retrieval from PDF-based faculty guidelines, leveraging the capabilities of OpenAI's ChatGPT model.

The study [30] developed a hybrid model for centralized database access in education, integrating the Transformer Framework and LLMs with a Chatbot system. The methodology included model design, LLM integration, Chatbot implementation, testing with diverse materials, user feedback collection, and optimization. Limitations encompass data availability, scalability, language nuances, security, and user adoption variances.

### A. FILLING THE RESEARCH GAP

Our proposed AI-driven PDF Chatbot for project management automation addresses several limitations observed in previous works in the domain. While various chatbot applications have been explored, particularly in healthcare [14], [16], [17], agriculture [18], education [22], and information services [25], there's been a noticeable gap in the implementation of AI chatbots specifically tailored for project management. Our innovation fills this void by introducing a PDF-driven chatbot designed explicitly to support project management tasks, leveraging advanced AI techniques such as the SFT-1 12B Model [25] and OpenAI's ChatGPT [26]. Unlike existing chatbots, which often cater to specific industries or services, our chatbot is aimed at automating project management processes, offering seamless communication and support to global teams. Moreover, while some studies have explored AI chatbots for information retrieval from PDF documents [26], [28], our chatbot goes beyond mere retrieval by interpreting and responding to PDF data sources such as SRS reports and interview transcripts, thereby providing valuable insights into project management decision-making processes [26]. By focusing on a niche yet crucial area of project management, our work not only extends the application domain of AI-driven chatbots but also contributes to streamlining project execution and enhancing human-robot interaction in the project management landscape.

To the best of our knowledge, no AI chatbot exists that will work as a substitute for the project manager when the project manager is absent. Our proposed chatbot will be able to make decisions based on the e-commerce project's scope and vision.

## IV. METHODOLOGY

Building a chatbot using Natural Language Processing (NLP) is complex because it involves understanding the nuances of human language, which is inherently intricate and varied. Our framework is divided into four parts:

- Dataset Formation: In this part, we create our own custom dataset for training the chatbot model. Data is collected by interviewing 15 project managers based on questionnaires provided by us in the context of e-commerce project development and the Software Requirement Specification (SRS) document.
- Data Preprocessing: In this part, to make the best out of the dataset, we use data preprocessing before model training.
- Chatbot Development: In this part, we have used OpenAssistant's SFT-1 12B Model [13] for our chatbot development, which is shown in Fig 2.
- Maintenance Mechanism: Here, we discussed the maintenance mechanism of our chatbot to help developers.

### A. DATASET FORMATION

For building the PDF-driven chatbot, we've meticulously outlined a series of 50 interview questions for extracting insights from 15 seasoned project managers in terms of their decision-making process on resource allocation, estimation, task breakdown, etc. By focusing on their decision-making
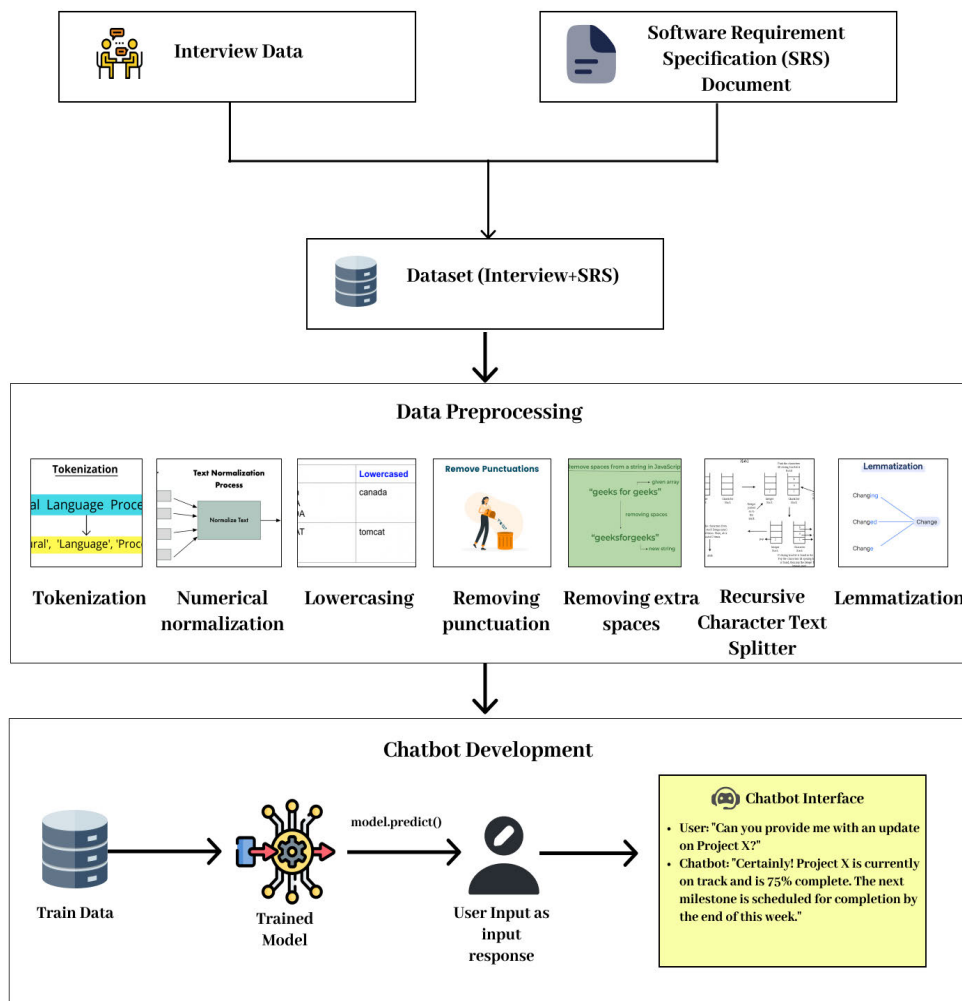
**FIGURE 2.** Methodology diagram.

processes and extracting their responses, we have created a robust dataset. The dataset consists of 25,738 words across 65 pages of the SRS report, specifically tailored to the development of an eCommerce project. This dataset serves as the foundation for our chatbot's knowledge, refining its ability to replicate the decision-making patterns of real project managers within the realm of e-commerce project development. Our approach is centered on capturing the nuanced sequences of responses from these project managers. These sequences serve as training data, enabling the LLM model to adapt and emulate authentic decision-making scenarios encountered in the e-commerce landscape. Furthermore, we have constructed a Software Requirements Specification (SRS) document on our ongoing e-commerce project named 'EkhushBD'. This document consists of a comprehensive understanding of e-commerce project development, serving as a key knowledge base of the project and enriching our

chatbot's knowledge repository. If anyone wants to make a custom dataset like us, then the following information should be considered in a PDF.

1) Scope and vision of the project.
2) Software Specification Fixation Through Appropriate Steps. (Information gathering, Gap analysis, feature list fixation)
3) Functional Requirements.
4) Non-functional requirements. (Performance, Scalability, Availability, Security, Maintainability, and Usability)
5) Hardware requirements.
6) Backup and Recovery Strategy.
7) System Architecture Information.
8) Communication Protocols.
9) Feasibility Analysis.

10) Software Requirements and Technical Description. (Front-end, Back-end, API integration, Payment gateway, and so on)
11) Review Techniques.
12) Budget and team information of the project.

The dataset, enriched through interviews with 15 experienced project managers, extends beyond its e-commerce focus to encompass essential project management criteria applicable across various industries. These criteria include identifying common challenges, addressing risk management factors, and analyzing decision-making processes in real-world project settings. By incorporating these characteristics, the model becomes capable of adapting to a broad spectrum of project types and industries, enabling the chatbot to deliver relevant, context-aware, and insightful guidance. This design ensures that the chatbot offers flexible support, transcending its initial e-commerce scope to address diverse project management scenarios. Whether for structured management or in-depth risk assessment, this adaptability positions the chatbot as a versatile and valuable tool for tackling complex project management needs across industries. You can view and download our custom dataset from section VIII for example.

### B. DATA PREPROCESSING

A chatbot must interpret not just the literal words but also the context, tone, and implied meanings, including idioms and sarcasm. Additionally, it needs to handle a wide range of user inputs, which can include slang, misspellings, and grammatical errors, while also adapting to evolving language trends and regional dialects. This complexity requires sophisticated algorithms and extensive training data to accurately process and respond to user queries in a meaningful way.

We have used data preprocessing techniques like tokenization, Numerical normalization, Lowercasing, Removing punctuation, Removing extra spaces, Recursive Character Text Splitter, and lemmatization for this work. Tokenization is the process of breaking down text into smaller units called tokens. In the case of a chatbot, these tokens are usually words or phrases. This helps in structuring the input text for further processing. The lowecasing step involves converting all characters in the text to lowercase. It is a normalization technique to ensure that the model treats words like "Hello", "hello", and "HELLO" equally, preventing the same word in different cases from being treated as different tokens. By ignoring case distinctions, the model learns better and avoids duplicating representations for words based on their casing. Lemmatization is the process of reducing words to their base or root form, as shown in Fig. 3. For example, "running", "ran", and "runs" would all be lemmatized to "run". This is important in a chatbot as it helps in understanding the semantic meaning of the text more effectively. Numerical Normalization converts the numerical values into a standard form, ensures consistent handling of

numbers, and prevents the model from treating different numeric values as separate entities.

Removing punctuation eliminates punctuation marks that don't carry significant meaning, helps simplify the text, and prevents the model from considering punctuation as separate tokens. It enhances the model's ability to focus on meaningful content while reducing noise in the input data. The technique of removing extra spaces, stripping extra spaces, and normalizing white spaces ensures consistent formatting in text, making it easier for the model to process. It streamlines the text input, preventing the model from interpreting additional spaces as meaningful entities and improving overall text consistency. Recursive Character Text Splitter splits long text sequences into smaller, manageable chunks, helping handle large volumes of data effectively without losing contextual information. In our case, chunk size=500 and chunk overlap=200 are being set for Recursive Character Text splitting. It enables the model to process lengthy texts in smaller segments, aiding in memory efficiency and preserving the context across the text segments.

Each of these preprocessing steps contributes to making the chatbot's natural language understanding more effective and efficient. By cleaning and simplifying the input text, these techniques help reduce complexity and improve the accuracy of the chatbot's responses.

### C. CHATBOT DEVELOPMENT

After preprocessing our data, we trained the data using pre-trained LLM of OpenAssistant's SFT-1 12B [23] Model for our chatbot development. The fine-tuned model, SFT-1 12B, within the OpenAssistant project, represents a milestone in English supervised-fine-tuning (SFT) models. Leveraging the Pythia 12B architecture as its foundation, this iteration has been meticulously fine-tuned using approximately 22,000 instances of human demonstrations showcasing assistant conversations. This extensive training dataset serves as a rich source of diverse conversational patterns, queries, and responses, enabling the model to grasp a wide spectrum of language nuances and context. The Pythia 12B architecture, known for its large-scale capabilities in natural language understanding and generation, forms a robust base that, through the fine-tuning process, has been tailored specifically for chatbot development within the OpenAssistant framework. After training the model with our preprocessed data, we have built a user interface to test our chatbot as shown in Fig. 4. We have used gradio library to build the user interface powered by HTML, CSS and Python.

We are drawn to the potential of OpenAssistant's SFT-1 12B pre-trained LLM for powering our PDF-driven chatbot due to its strengths in several key areas. Firstly, its foundation on the robust Pythia 12B model suggests a strong capability in understanding natural language. This is crucial for our chatbot to accurately interpret user queries and navigate the text content within PDFs. Secondly, SFT-1's focus during
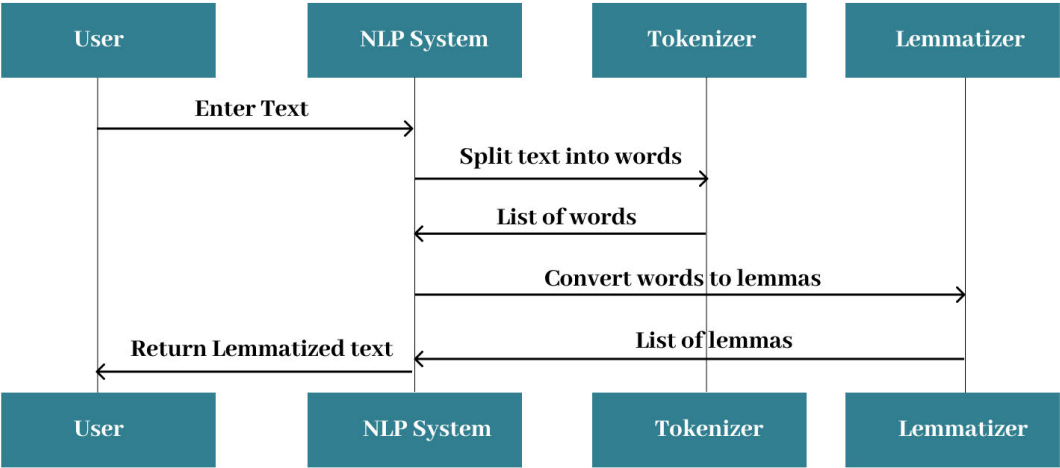
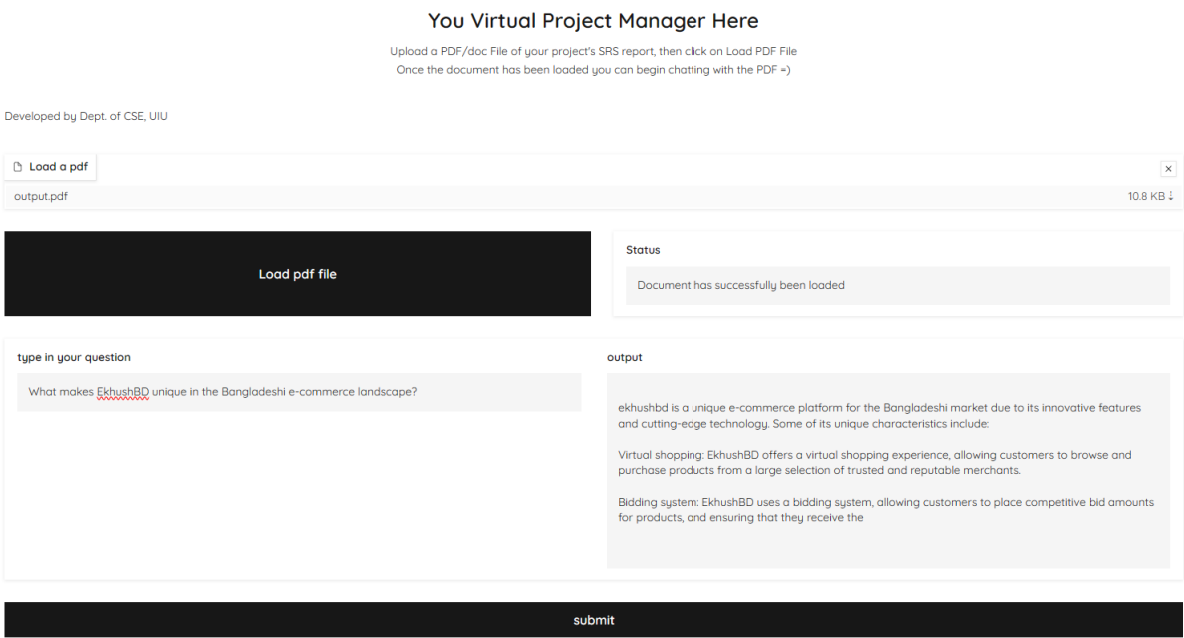**FIGURE 3.** Lemmatization process.



**FIGURE 4.** Chatbot web interface.

fine-tuning aligns perfectly with our chatbot's core function: responding to user prompts and questions through text. Finally, the inherent adaptability of LLMs like SFT-1 is appealing. As our chatbot interacts with users and processes more PDFs, the model has the potential to become more adept at handling tasks specific to the information our PDFs contain. However, we must acknowledge that additional training might be necessary for SFT-1 to effectively process and extract information directly from PDFs. For that reason, we preprocessed the data with described data preprocessing techniques in section 3.2. This crucial step made our chatbot more robust and context aware as it can process the pdf's data to a large degree.

### D. MAINTENANCE MECHANISM

**Figure 5** provides a comprehensive overview of the maintenance mechanism of our chatbot, highlighting the iterative process that ensures the chatbot evolves in response to user interactions. This mechanism is pivotal for maintaining the

**TABLE 2.** Response time of each query from chatbot.

| Question | Bot Response | Response Time |
|---|---|---|
| Who are the target users for EkhushBD? | The target users for EkhushBD are individuals, companies, and organizations that engage in e-commerce in Bangladesh. The platform is designed to provide a range of features and functionality to these users, including augmented reality, virtual voice search, and more, to enhance their shopping experience and increase sales. | 5 secs |
| What is the role of ARCore technology in EkhushBD? | Augmented RealityCore (ARCore) is a framework developed by Google that provides a set of tools and libraries for developing AR applications on Android devices. It provides an efficient way to display augmented reality content on a device, using the camera and motion sensors to track the environment and translate virtual objects into interactive, real-world objects. In EkhushBD, ARCore can be used to enhance the user experience by providing more engaging and interactive content, such as virtual objects that can be placed. | 5 secs |
| What technology stack will be used to develop EkhushBD? | The technology stack will likely include JavaScript for the front-end, PHP for the back-end, and a MySQL database for storing user data. The EkhushBD platform will likely use a combination of APIs and other tools to connect with various third-party systems and data sources. | 4 secs |
| How does the product bidding system work on EkhushBD? | Bidding is a feature on EkhushBD that allows users to place a bid on a product. This bid can then be accepted or rejected by the seller. The system then calculates the final selling price based on the bid and the current market prices. | 4 secs |
| What hosting and domain provider will be used for EkhushBD? | EkhushBD will be hosted on the SNBD HOST cloud server platform. The domain name for the website will be registered with a domain name provider. | 3 secs |
| How will EkhushBD handle potential legal and regulatory challenges? | Potential legal and regulatory challenges in the e-commerce space in Bangladesh are a major concern for e-commerce businesses. With a young and rapidly growing digital economy, there is a need to ensure compliance with all relevant laws and regulations. EkhushBD will work closely with legal and regulatory experts to ensure that the platform operates within the law and is compliant with all relevant regulations. The team will also implement measures to ensure data protection and privacy and will regularly monitor the legal and regulatory landscape to ensure that. | 7 secs |

relevance and effectiveness of the chatbot, as it adapts to new user behaviors and preferences over time.

The process starts when a user interacts with the chatbot. Each interaction, whether it's a question, command, or conversational exchange, is logged by the chatbot system. This logging is critical as it serves as the raw data from which the chatbot can learn and improve. Over time, these interactions accumulate in a dataset dedicated to capturing user communications. This dataset becomes a vital resource for understanding user needs and identifying areas where the chatbot may not be performing optimally.

Periodically, a data scientist accesses this dataset to perform a thorough analysis. The goal of this analysis is to sift through the data to identify patterns, anomalies, or any new types of queries that the chatbot is struggling with. This step is crucial as it informs the areas of improvement needed in the chatbot's responses and functionalities.

Following the analysis, the data scientist uses the insights gained to update the model that powers the chatbot. This ML model is what enables the chatbot to understand and respond to user queries. The update process might involve retraining the model with new data, tweaking algorithms, or introducing new parameters that enhance the chatbot's accuracy and efficiency. The updated model is then thoroughly tested to ensure that it performs better than its predecessor in handling real-world interactions.

Once the updated model meets the performance criteria, it is deployed back into the chatbot system. This deployment is seamlessly integrated so that users can immediately benefit from the enhanced capabilities of the chatbot. With the new model in place, the chatbot can respond to queries more accurately, manage more complex conversations, and provide a more engaging user experience.

This entire cycle, from data collection, analysis, model updating, and deployment, is crucial for the chatbot's ability to adapt and improve continuously. It ensures that the chatbot can handle evolving user expectations and emerging trends in user behavior. Furthermore, this ongoing maintenance helps in mitigating issues like response inaccuracies or outdated information, which are critical for maintaining user trust and satisfaction. This cyclical maintenance ensures that the chatbot remains a helpful, reliable, and adaptive tool for users, consistently improving in response to their evolving needs.

## V. RESULT ANALYSIS

The chatbot retrieves the information automatically from PDF-Document as a response to the questions from the user about the project. In this research, we use LangChain Framework and OpenAssistant's SFT-1 12B Model to get the response, with some preprocessing mentioned in the previous section. We achieve an average cosine similarity score of 0.8080 or 80.80% and an average Semantic Similarity score of 0.8521 or 85.21%.

**TABLE 3.** Expected answer vs Bot response.

| Question | Expected Answer | Bot Response |
|---|---|---|
| Who are the target users for EkhushBD? | EkhushBD primarily targets consumers within Bangladesh seeking a diverse and dynamic online shopping experience, particularly those interested in trying on clothes virtually and participating in product auctions. | The target users for EkhushBD are individuals, companies, and organizations that engage in e-commerce in Bangladesh. The platform is designed to provide a range of features and functionality to these users, including augmented reality, virtual voice search, and more, to enhance their shopping experience and increase sales. |
| What is the role of ARCore technology in EkhushBD? | EkhushBD integrates ARCore to allow users to virtually try on garments and accessories through their smartphones or tablets, enhancing product visualization and decision-making before purchase. | ARCore is a technology developed by Google that allows developers to create augmented reality experiences for mobile devices. In the context of EkhushBD, ARCore technology could be used to create an augmented reality app that helps customers in virtually try on garments and accessories through their smartphones on the products they are interested in buying. This app could use ARCore's advanced capabilities, such as realistic rendering and motion tracking, to enhance the user experience and help customers make informed purchasing decisions. |
| What technology stack will be used to develop EkhushBD? | The technology stack of EkhushBD involves HTML, CSS, Javascript, and PHP for the back-end and MySQL for Database Development. Few third-party APIs will be there for payment gateway, authentication. | The technology stack will likely include JavaScript for the front-end, PHP for the back-end, and a MySQL database for storing user data. The EkhushBD platform will likely use a combination of APIs and other tools to connect with various third-party systems and data sources. |
| How does the product bidding system work on EkhushBD? | Users can participate in auctions for specific products, submitting bids and competing with others to win the item at a potentially lower price. The system will have clear rules, notifications, and winner management mechanisms. | Bidding is a feature on EkhushBD that allows users to place a bid on a product. This bid can then be accepted or rejected by the seller. The system then calculates the final selling price based on the bid and the current market prices. |
| What hosting and domain provider will be used for EkhushBD? | SNBD Host Ltd. will manage the domain and host the website on their VPS Server. | EkhushBD will be hosted on SNBD HOST cloud server platform. The domain name for the website will be registered with a domain name provider. |
| How will EkhushBD handle potential legal and regulatory challenges? | The development team will ensure compliance with all relevant Bangladeshi laws and regulations related to e-commerce, data privacy, and consumer protection. | Potential legal and regulatory challenges in the e-commerce space in Bangladesh are a major concern for e-commerce businesses. With a young and rapidly growing digital economy, there is a need to ensure compliance with all relevant laws and regulations. EkhushBD will work closely with legal and regulatory experts to ensure that the platform operates within the law and is compliant with all relevant regulations. The team will also implement measures to ensure data protection and privacy, and will regularly monitor the legal and regulatory landscape to ensure that. |
| Does EkhushBD accept cryptocurrency payments at this time? | No. EkhushBD is not accepting cryptocurrency as a form of payment at this time. | No. As of my knowledge EkhushBD is not accepting cryptocurrency as form of payment at this time. |

Measuring the performance of an AI chatbot can involve several metrics, such as cosine similarity score, semantic similarity score, response time, and user satisfaction. We are currently conducting a user satisfaction survey to evaluate the performance of this chatbot. The chatbot's efficiency, as shown in Table 2, appears to be quite high, with response times ranging from 4 to 7 seconds for a variety of complex questions pertaining to the project's platform. The ability to deliver concise and relevant answers within this timeframe indicates a well-optimized system capable of handling user queries effectively. Overall, the chatbot's response time is indicative of a system that is both responsive and adept at managing a diverse set of inquiries. This level of efficiency can significantly enhance user experience by providing quick and accurate information, which is especially important in the fast-paced environment of e-commerce. Based on the comparison between the expected answers shown in Table 3 and the bot's responses, the chatbot appears to be performing well. It provides comprehensive and relevant

information that aligns closely with the expected answers. The chatbot's answers are sufficiently detailed, indicating a well-programmed system capable of understanding and responding to various topics related to the project's platform.

**Cosine Similarity Score:** Cosine similarity measures the cosine of the angle between two non-zero vectors in an n-dimensional space [34]. In NLP, these vectors often represent the term frequency-inverse document frequency (TF-IDF) or embedding of words, sentences, or documents. By converting text into numerical vectors, it efficiently compares the query and PDF content, providing a score that reflects their similarity, as shown in Fig. 6. This metric is scalable for large PDFs and helps gauge how well the chatbot retrieves relevant information. It helps in representing documents as vectors in a high-dimensional space, allowing comparison and measuring similarity between different documents efficiently. Our chatbot achieves an Average Cosine Similarity Score of 0.8080. We calculate our cosine similarity score by following
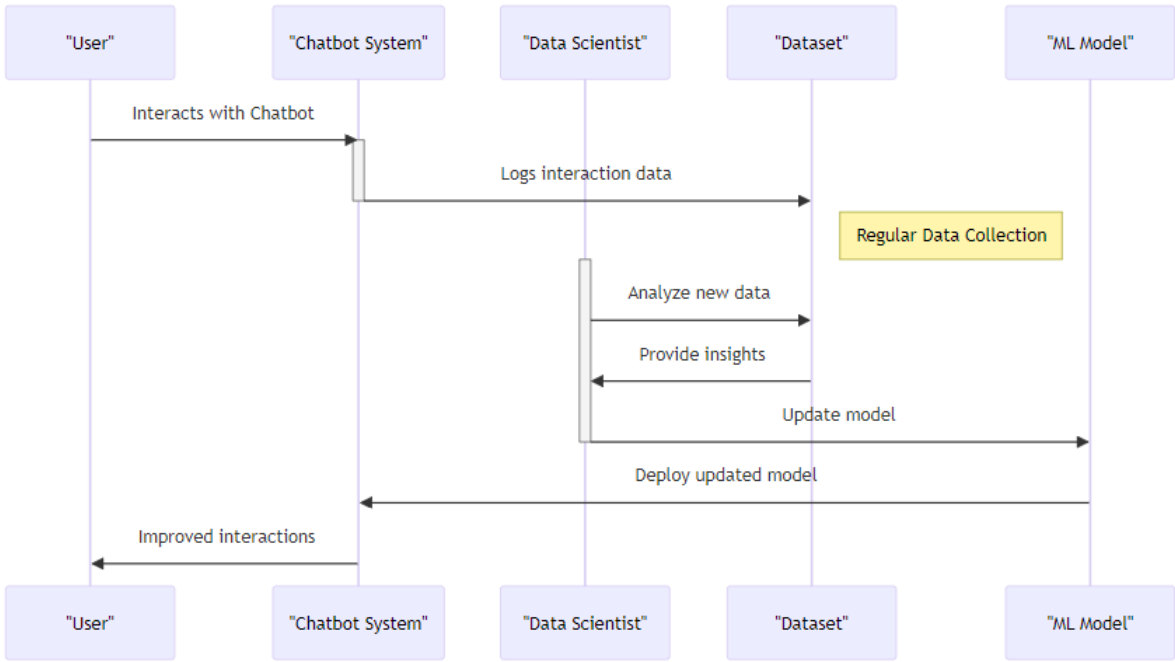
**FIGURE 5.** Maintenance mechanism diagram.

equation (1).

$$CS = \cos(\theta) = \frac{A \cdot B}{||A|| \cdot ||B||} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}} \quad (1)$$

Considering cosine similarity as a performance indicator for PDF chatbot development is crucial for several reasons. Firstly, cosine similarity measures the likeness between two documents (in our case, the expected response document and the bot response document), enabling the chatbot to assess the relevance of its responses to user queries against the content stored in the PDF repository. This ensures that the chatbot returns documents closely matching the user's needs, thereby enhancing the user experience. Secondly, by optimizing cosine similarity scores, developers can improve the precision and accuracy of document retrieval, leading to more reliable results. Thirdly, user satisfaction is significantly enhanced when the chatbot delivers relevant documents promptly. Cosine similarity facilitates this by efficiently handling large document volumes and scaling effectively to varying user demands. Additionally, cosine similarity serves as an essential metric for evaluating and improving the chatbot's performance over time, guiding iterative enhancements to algorithms and techniques used in document retrieval.

**Semantic Similarity Score:** Semantic similarity assesses the degree of likeness between two pieces of text based on their meaning and context rather than just their lexical similarity [33]. It measures how closely related two sentences or documents are in terms of their underlying meaning or semantics. In the context of a PDF chatbot, which deals with

**TABLE 4.** Performance analysis.

| Performance Matrix | Score |
|---|---|
| Cosine Similarity | 0.8080 |
| Semantic Similarity | 0.8521 |
| Response Time | 4-7 secs |

a wide range of queries and documents, understanding the context is crucial. Semantic similarity helps gauge how well the chatbot comprehends the meaning and intent behind user queries or documents. By evaluating semantic similarity, the chatbot can provide more accurate and contextually relevant responses. It helps in better understanding user queries and matching them with relevant content from PDF documents. Our chatbot achieves an Average Semantic Similarity of 0.8521.

Considering the Semantic Similarity Score as a performance indicator for PDF chatbot development, it holds significant importance for various reasons. Unlike traditional metrics like cosine similarity, semantic similarity reaches deeper into the meaning and context of the documents and user queries rather than just their lexical or syntactic similarity. By incorporating semantic similarity, chatbots can better understand the underlying intent behind user queries and match them with relevant documents more accurately. This ensures that the chatbot not only retrieves documents that contain similar keywords but also those that share similar concepts and themes, thus improving the overall quality of responses. Semantic similarity also
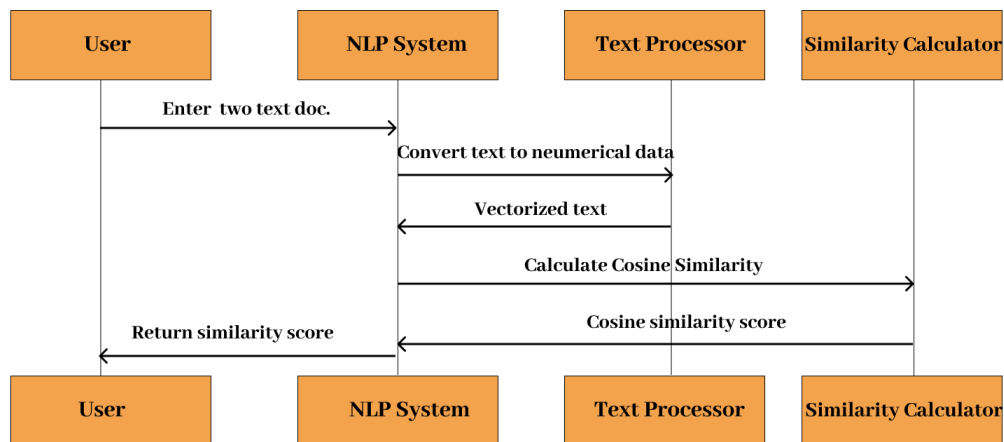
**FIGURE 6.** Cosine similarity score calculation process.

enables chatbots to handle queries that might use different phrasings or synonyms but convey the same meaning, enhancing the chatbot's versatility and effectiveness in serving users. Moreover, by using semantic similarity scores, developers can continually refine and optimize the chatbot's understanding and retrieval capabilities, ultimately leading to a more robust and user-friendly system. Thus, considering semantic similarity as a performance indicator not only enhances the relevance and accuracy of document retrieval but also contributes to the overall sophistication and efficacy of PDF chatbot development.

The cosine similarity score (80.80%) and semantic similarity score (85.21%) serve as essential benchmarks for evaluating the chatbot's ability to align its responses with the content of the PDF documents it processes. The cosine similarity score quantifies how closely the chatbot's generated responses match the vectorized representation of the original document, ensuring that the factual and contextual information from the source is retained. Meanwhile, the semantic similarity score goes beyond surface-level text matching, assessing how well the chatbot captures the underlying meaning and intent of the content. This is crucial in project management automation, where accurate and context-aware responses are required for effective decision-making. Together, these scores indicate that the chatbot not only reproduces content faithfully but also delivers semantically meaningful insights, making it a reliable tool for automating complex tasks while maintaining the integrity of the original project data.

**Response Time:** Measuring the time it takes for the bot to respond to a query. Python's time module is used to record the start and end times of each query-response cycle [35].

## VI. PERFORMANCE COMPARISON WITH OTHER EXISTING PDF CHATBOTS

ChatPDF [31] is a tool made to make reading and understanding PDF files easier, especially useful for academic papers. Users can ask questions about anything in the PDF, and the bot will directly answer with information extracted from the PDF. This helps users get key details and comprehend complex information without reading the whole document.

SciSummary [32] is also an automatic summarizing tool made specifically for scientific texts. It uses advanced natural language processing techniques and allows researchers, students, and academicians to quickly understand the main findings and importance of a study, saving time and helping to get a broader understanding of the topic.

Due to the recent innovations in PDF-driven chatbots in 2023, there are only a handful of options available. We've chosen ChatPDF and Sci-summary for benchmarking purposes because they represent the leading edge in handling diverse PDF content. Our own PDF chatbot, however, takes a unique approach. It's a project management-specific PDF-driven AI chatbot, a niche currently unoccupied in the market.

Table 5 shows how our bot has a deeper understanding of who EkhushBD is for, compared to ChatPDF. While ChatPDF focuses mainly on Bangladeshi shoppers looking for exciting online shopping features like virtual try-ons and auctions, our bot sees a bigger picture. It recognizes that EkhushBD can serve not just shoppers but also business leaders, researchers, and government officials, highlighting its significant role in boosting the Bangladeshi economy. When it comes to ARCore technology, our bot goes beyond ChatPDF's simple view that it's just for trying on clothes virtually. It introduces ARCore as a cutting-edge technology by Google that can make online shopping more engaging and realistic, offering a fuller picture for those new to the concept.

Looking at future trends, ChatPDF is optimistic about EkhushBD embracing new tech like cryptocurrency payments and social shopping. However, our bot takes a more realistic approach. It points out the uncertain future of digital

marketing and e-commerce, suggesting EkhushBD is ready for ongoing innovation without making uncertain promises. This practical viewpoint is also seen in how our bot talks about legal challenges. Unlike ChatPDF, which only gives a basic assurance about following laws, our bot explains how the development team is actively working with legal experts to foresee and tackle potential issues.

Both our bot and ChatPDF agree on where EkhushBD stands with cryptocurrency payments right now. In summary, by offering deeper insights, clearer explanations, and a well-rounded view, our bot significantly improves how users understand and interact with EkhushBD, setting a new standard for engaging and informative content.

Table 6 showcases a comparative analysis of the responses between our bot and SciSummary regarding the EkhushBD e-commerce platform. Notably, our bot presents a richer contextual understanding across five key inquiries. Firstly, while SciSummary narrowly categorizes EkhushBD's target audience as consumers within Bangladesh eager for a versatile online shopping realm, our bot broadens this horizon, encapsulating not only potential buyers but also company insiders, academics, and policy formulators, thereby appreciating the platform's extensive influence on the national economy. In discussing ARCore technology, our bot goes beyond the basic application of virtual try-ons, enlightening users about ARCore's origins with Google and its broader implications for enhancing digital interactions, a stark contrast to SciSummary's limited explanation. On the topic of future innovations, our bot adopts a pragmatic stance, acknowledging the unpredictable pace of e-commerce advancements while highlighting EkhushBD's commitment to continual technological exploration without resorting to speculative claims. Regarding legal and regulatory hurdles, our bot underscores a proactive approach by the development team, who prioritize collaboration with experts to navigate compliance complexities, showing a deeper level of engagement compared to SciSummary's general compliance statement. Both platforms concur on the current non-utilization of cryptocurrency payments. Collectively, these detailed and nuanced responses from our bot underscore its superior capability to synthesize and convey information, offering users a more comprehensive and insightful understanding of EkhushBD, thus enhancing the decision-making process for a varied audience.

Table 7 illustrates how the bot's responses outshine those of ChatPDF and SciSummary by providing more detailed and contextual information. The bot's strength lies in its ability to offer comprehensive explanations that delve deeper into the topics related to the EkhushBD e-commerce platform, surpassing the surface-level descriptions provided by the other tools. Moreover, the bot caters to a wider range of user needs by considering various perspectives and stakeholders, including customers, company officials, researchers, and policymakers. This inclusive approach ensures that the bot's responses are relevant and beneficial to a diverse audience. By furnishing context-specific details

and insights into technologies like ARCore within the context of EkhushBD, the bot enriches the user experience and facilitates a more thorough understanding of the platform and its functionalities. Overall, the bot's informative and valuable interaction sets it apart by empowering users to make informed decisions and grasp the complexities of the EkhushBD e-commerce platform effectively.

## VII. DISCUSSION

This paper addresses a unique problem in project management by introducing an AI chatbot that leverages PDF data to enhance communication and streamline project processes. This innovative approach aims to overcome challenges faced by project managers in global, diversified, and remote environments where effective communication is crucial for success. One key aspect highlighted in the paper is how AI chatbots can significantly improve communication for project managers in remote teams. By acting as a virtual project manager, the AI chatbot ensures round-the-clock communication with team members and stakeholders, irrespective of their locations and time zones. This continuous availability facilitates seamless information sharing, task coordination, and decision-making, thereby enhancing team collaboration and productivity in remote work settings. Furthermore, the benefits of using a PDF-Driven AI Chatbot for project management are manifold. Firstly, the AI chatbot offers data-driven insights for decision-making and risk management, enabling project managers to make informed choices based on real-time information extracted from PDF documents. This data-driven approach enhances the accuracy and efficiency of project management processes.

Secondly, the integration of advanced AI technology and natural language processing in the AI chatbot streamlines project management by automating repetitive tasks, such as data preprocessing and information retrieval from PDF files. This automation not only saves time but also reduces the likelihood of human errors, ensuring consistency and reliability in project operations. The AI chatbot's ability to analyze and interpret PDF data using techniques like tokenization, numerical normalization, and lemmatization enhances the quality of responses and recommendations provided to project managers. This intelligent processing of information enables the chatbot to generate coherent and contextually relevant responses aligned with the content of the PDF documents, thereby improving the overall decision-making process.

The scalability of the AI-driven PDF Chatbot for Project Management extends beyond its current focus on e-commerce projects to various other project scenarios. The chatbot's scalability lies in its adaptability to different industries, project types, and organizational contexts, making it a versatile tool for project managers across diverse domains. By utilizing advanced AI techniques and natural language processing, the chatbot can be customized and trained to understand and respond to specific project

**TABLE 5.** ChatPDF response vs Our response.

| Question | ChatPDF Answer | Our Bot Response |
|---|---|---|
| Who are the target users for EkhushBD? | The target users for EkhushBD primarily include consumers within Bangladesh seeking a diverse and dynamic online shopping experience, particularly those interested in trying on clothes virtually and participating in product auctions. | The target users for EkhushBD are anyone who would benefit from an enhanced digital experience. This could include customers, company officials, researchers, and policymakers. By providing a more sophisticated system, EkhushBD hopes to help Bangladesh become a more digitally advanced country, drive economic growth, create jobs, and improve the lives of its citizens. |
| What is the role of ARCore technology in EkhushBD? | ARCore technology in EkhushBD allows users to virtually try on garments and accessories through their smartphones or tablets, enhancing product visualization and decision-making before purchase. This feature aims to provide a more interactive and engaging shopping experience compared to traditional e-commerce platforms. | ARCore is a technology developed by Google that allows developers to create augmented reality experiences for mobile devices. In the context of EkhushBD, ARCore technology could be used to create an augmented reality app that helps customers in virtually try on garments and accessories through their smartphones on the products they are interested in buying. This app could use ARCore's advanced capabilities, such as realistic rendering and motion tracking, to enhance the user experience and help customers make informed purchasing decisions. |
| Will EkhushBD explore incorporating additional innovative technologies like cryptocurrency payments or social commerce features in the future? | Yes, EkhushBD is open to exploring the incorporation of additional innovative technologies like cryptocurrency payments or social commerce features in the future. | The evolution of e-commerce and digital marketing is rapid and ever-changing. The industry is constantly evolving and it is likely that new technologies and services will continue to emerge. It is impossible to predict the future, but it is likely that EkhushBD will continue to innovate and explore new technologies to improve the user experience and increase sales. |
| How will EkhushBD handle potential legal and regulatory challenges? | EkhushBD will ensure compliance with all relevant Bangladeshi laws and regulations related to e-commerce, data privacy, and consumer protection. The development team will be responsible for addressing potential legal and regulatory challenges to ensure the platform's adherence to the applicable legal framework. | EkhushBD will handle potential legal and regulatory challenges by following all relevant laws and regulations, ensuring that all of their activities are carried out in a responsible and ethical manner. They will also work closely with legal and regulatory experts to ensure that they are fully informed of any potential issues and have the necessary resources and expertise to address them effectively. |
| Does EkhushBD accept cryptocurrency payments at this time? | No, EkhushBD does not currently accept cryptocurrency payments. | At this time, EkhushBD does not accept cryptocurrency payments. |

management requirements, regardless of the industry or sector. In terms of scalability across different project scenarios, the chatbot's dataset development approach can be tailored to cover a wide range of data, specifications, and decision-making scenarios relevant to specific project domains. This flexibility allows the chatbot to provide precise and relevant guidance in various project contexts, ensuring its scalability and applicability beyond e-commerce projects.

Moreover, the chatbot's ability to interpret and respond to PDF data from sources like SRS reports can be used in different project scenarios to extract valuable insights and support decision-making processes. This adaptability enables the chatbot to scale effectively across diverse project management landscapes, offering support and automation capabilities to project managers in various industries and settings. The chatbot's context-aware responses play a pivotal role in enhancing stakeholder engagement by fostering effective communication, improving decision-making, and aligning actions with project goals. Its ability to interpret user queries and respond with contextual relevance ensures a seamless interaction experience, reducing ambiguities and misunderstandings. This allows stakeholders to focus

on critical project management tasks without the need to repeatedly clarify or rephrase their concerns. By providing personalized and context-specific insights, the chatbot ensures stakeholders feel heard and understood, thereby building trust and confidence in the system. For instance, when stakeholders raise concerns about resource allocation or potential risks, the chatbot leverages its knowledge base to offer targeted recommendations and actionable steps, directly addressing the issues at hand. This immediate and precise assistance not only reduces response times but also empowers stakeholders to make informed decisions promptly. The chatbot's 24/7 availability further enhances its value by facilitating continuous engagement, especially for global teams operating across multiple time zones. Stakeholders can access information and support whenever needed, ensuring uninterrupted collaboration. This capability strengthens their involvement and commitment to project outcomes.

Overall, the AI-driven PDF Chatbot for Project Management demonstrates scalability in its potential to be customized, trained, and applied to different project scenarios, making it a valuable asset for project managers seeking automated support and enhanced decision-making

**TABLE 6.** Scisummary response vs Our response.

| Question | SciSummary Answer | Our Bot Response |
|---|---|---|
| Who are the target users for EkhushBD? | EkhushBD strategically targets Bangladeshi consumers seeking a dynamic online shopping experience, as well as virtual shoppers intrigued by the prospect of virtually trying on products and engaging in product auctions. | The target users for EkhushBD are anyone who would benefit from an enhanced digital experience. This could include customers, company officials, researchers, and policymakers. By providing a more sophisticated system, EkhushBD hopes to help Bangladesh become a more digitally advanced country, drive economic growth, create jobs, and improve the lives of its citizens. |
| What is the role of ARCore technology in EkhushBD? | The ARCore technology in EkhushBD allows users to virtually try on garments and accessories through their smartphones or tablets, enhancing product visualization and decision-making before purchase. | ARCore is a technology developed by Google that allows developers to create augmented reality experiences for mobile devices. In the context of EkhushBD, ARCore technology could be used to create an augmented reality app that helps customers in virtually try on garments and accessories through their smartphones on the products they are interested in buying. This app could use ARCore's advanced capabilities, such as realistic rendering and motion tracking, to enhance the user experience and help customers make informed purchasing decisions. |
| Will EkhushBD explore incorporating additional innovative technologies like cryptocurrency payments or social commerce features in the future? | Staying open to future technological advancements and user preferences will allow EkhushBD to potentially explore incorporating innovative technologies like cryptocurrency payments and social commerce features to adapt and evolve over time. | The evolution of e-commerce and digital marketing is rapid and ever-changing. The industry is constantly evolving and it is likely that new technologies and services will continue to emerge. It is impossible to predict the future, but it is likely that EkhushBD will continue to innovate and explore new technologies to improve the user experience and increase sales. |
| How will EkhushBD handle potential legal and regulatory challenges? | The development team will ensure compliance with all relevant Bangladeshi laws and regulations related to e-commerce, data privacy, and consumer protection to handle potential legal and regulatory challenges. | EkhushBD will handle potential legal and regulatory challenges by following all relevant laws and regulations, ensuring that all of their activities are carried out in a responsible and ethical manner. They will also work closely with legal and regulatory experts to ensure that they are fully informed of any potential issues and have the necessary resources and expertise to address them effectively. |
| Does EkhushBD accept cryptocurrency payments at this time? | As of now, the document does not mention whether EkhushBD accepts cryptocurrency payments. | At this time, EkhushBD does not accept cryptocurrency payments. |

**TABLE 7.** Justification of each bot's response.

| Topic | ChatPDF's Response | SciSummary's Response | Our Bot Response |
|---|---|---|---|
| Target Users | ChatPDF identifies consumers in Bangladesh as the target users who will be interested in virtual tryouts and auctions. | Same as ChatPDF | Our bot goes far beyond the generic responses from the previous two bots and identifies target users as anyone who will benefit from an enhanced digital experience, including customers, company officials. researchers and policymakers, who will impact the Bangladeshi economy as a whole. |
| ARCore Technology | ChatPDF States the basic functionality of ARCore, which allows users to virtually try on clothes. But EkushBD was not only about clothing items | Same as ChatPDF | Our bot states that ARCore not only helps with virtual try-outs clothes but other products as well, which was missing in ChatPDF and SciSummary's responses. Our bot's more contextualized responses help someone who's not aware of context. Our bot tries to provide more context than ChatPDF and SciSummary. |
| Legal and Regulatory Challenges | ChatPDF states that EkushBD will comply with Bangladeshi laws and regulations. | Same as ChatPDF | Our bot states that EkushBD will actively engage with experts to navigate and adhere legal and regulatory frameworks of Bangladesh. |

capabilities across a wide range of projects and industries. The paper's innovative approach to developing a PDF-Driven AI Chatbot for project management offers a unique solution to the challenges faced by project managers in remote and diverse work environments. By harnessing the power of AI technology and PDF data, the chatbot not only improves communication and collaboration within remote teams but also enhances the efficiency, and effectiveness of project management processes.

## VIII. FUTURE WORK
Integration of image-based outputs in our chatbot in the future. Incorporating image-based outputs in the AI chatbot developed for project management can significantly enhance

its capabilities and user experience. By integrating image recognition and processing technologies, the chatbot can analyze and interpret visual information from images, diagrams, or charts shared by project team members. This integration can enable the chatbot to provide contextual responses, insights, and recommendations based on visual data, thereby enhancing communication and decision-making within project teams.

## IX. CUSTOM DATASET
Here, you can download our custom dataset to get an idea about what types of data we are giving to our chatbot, which can substitute project management tasks without a real project manager.

Custom Dataset: https://github.com/codewithkhurshed/SPM-project-repo/

## X. CONCLUSION
In conclusion, this paper revolves around the development of a PDF-Driven AI Chatbot designed to act as a virtual project manager for e-commerce projects. The study emphasizes the importance of leveraging advanced AI technology and natural language processing to automate routine tasks, provide data-driven insights, and fill the current void of AI chatbots capable of effectively assisting project managers. The research highlights the significance of tailored dataset development, advanced data preprocessing techniques, and the utilization of the Large Language Model (LLM) to train the chatbot. This chatbot achieves an average cosine similarity score of 0.8080 or 80.80% and an average Semantic Similarity score of 0.8521 or 85.21%. Furthermore, the study fills the market gap in AI chatbots for project management and aims to enhance project management practices by providing a helpful assistant for project teams, especially in the absence of a human project manager. Overall, the work demonstrates the potential of AI chatbots in revolutionizing project management and contributing to the advancement of AI and natural language processing in the context of e-commerce project development.

## REFERENCES

[1] S. Onu, "Perceptual differences between project managers and sponsors in the initiation phase of a project," in *Proc. IEEE Int. Conf. Ind. Eng. Eng. Manage.*, Hong Kong, Dec. 2012, pp. 2003–2007, doi: 10.1109/IEEM.2012.6838097.

[2] D. Tawfik. *10 Common Project Management Challenges (and How to Overcome Them)*. Accessed: Oct. 21, 2023. [Online]. Available: https://monday.com/blog/project-management/project-management-challenges/

[3] J.-S. Chou, P.-L. Chong, and C.-Y. Liu, "Deep learning-based chatbot by natural language processing for supportive risk management in river dredging projects," *Eng. Appl. Artif. Intell.*, vol. 131, May 2024, Art. no. 107744, doi: 10.1016/j.engappai.2023.107744.

[4] A. M. Felicetti, A. Cimino, A. Mazzoleni, and S. Ammirato, "Artificial intelligence and project management: An empirical investigation on the appropriation of generative chatbots by project managers," *J. Innov. Knowl.*, vol. 9, no. 3, Jul. 2024, Art. no. 100545, doi: 10.1016/j.jik.2024.100545.

[5] T. Wang and H.-M. Chen, "Integration of building information modeling and project management in construction project life cycle," *Autom. Construction*, vol. 150, Jun. 2023, Art. no. 104832, doi: 10.1016/j.autcon.2023.104832.

[6] V. Azamfirei, F. Psarommatis, and Y. Lagrosen, "Application of automation for in-line quality inspection, a zero-defect manufacturing approach," *J. Manuf. Syst.*, vol. 67, pp. 1–22, Apr. 2023, doi: 10.1016/j.jmsy.2022.12.010.

[7] Z. Yan, N. Duan, J. Bao, P. Chen, M. Zhou, Z. Li, and J. Zhou, "DocChat: An information retrieval approach for chatbot engines using unstructured documents," in *Proc. 54th Annu. Meeting Assoc. Comput. Linguistics (Volume 1: Long Papers)*, 2016, pp. 1–12, doi: 10.18653/v1/p16-1049.

[8] D. Li, X. Hu, Z. Sun, B. Hu, S. Ye, Z. Shan, Q. Chen, and M. Zhang, "TruthReader: Towards trustworthy document assistant chatbot with reliable attribution," in *Proc. Conf. Empirical Methods Natural Lang. Processing: Syst. Demonstrations*, Miami, FL, USA, 2024, pp. 89–100.

[9] S. P. Nayak, A. Rai, K. Vankataramanappa, J. A. Pansuriya, and J. Singler, "Adaptive MoD chatbot: Toward providing contextual corporate summarized document as suggestions and reported issue ticket routing," in *Advances in Artificial Intelligence and Data Engineering* (Advances in Intelligent Systems and Computing). Singapore: Springer, Aug. 14, 2020, pp. 659–674, doi: 10.1007/978-981-15-3514-7_50.

[10] A. C. Sari, N. Virnilia, J. T. Susanto, K. A. Phiedono, and T. K. Hartono, "Chatbot developments in the business world," *Adv. Sci., Technol. Eng. Syst. J.*, vol. 5, no. 6, pp. 627–635, Nov. 2020, doi: 10.25046/aj050676.

[11] J. Cho and G. Lee, "A chatbot system for construction daily report information management," in *Proc. Int. Symp. Autom. Robot. Construct. (IAARC)*, May 24, 2019, pp. 429–437, doi: 10.22260/isarc2019/0058.

[12] C.-N. Bodea, M.-I. Dascalu, and A. Hang, "Chatbot-based training for project management: Another way of corporate training or a must-have tool for sustainable education?" in *Research on Project, Programme and Portfolio Management* (Lecture Notes in Management and Industrial Engineering). Cham, Switzerland: Springer, 2020, pp. 249–259, doi: 10.1007/978-3-030-60139-3_17.

[13] E. Ebrahim, M. Sayed, M. Youssef, H. Essam, S. A. El-Fattah, D. Ashraf, O. Magdy, and R. ElAdawi, "AI decision assistant ChatBot for software release planning and optimized resource allocation," in *Proc. IEEE Int. Conf. Artif. Intell. Test. (AITest)*, Athens, Greece, Jul. 2023, pp. 55–60, doi: 10.1109/AITest58265.2023.00018.

[14] M. L. Prasad, C. Srinivasulu, M. Meenakshi, and P. Rohitha, "Chatbots in healthcare," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 11, no. 8, pp. 1087–1091, Aug. 2023, doi: 10.22214/ijraset.2023.55295.

[15] A. Miklosik, N. Evans, and A. M. A. Qureshi, "The use of chatbots in digital business transformation: A systematic literature review," *IEEE Access*, vol. 9, pp. 106530–106539, 2021, doi: 10.1109/ACCESS.2021.3100885.

[16] S. R. Dammavalam, N. Chandana, T. R. Rao, A. Lahari, and B. Aparna, "AI based chatbot for hospital management system," in *Proc. 3rd Int. Conf. Comput., Anal. Netw. (ICAN)*, Punjab, India, Nov. 2022, pp. 1–5, doi: 10.1109/ICAN56228.2022.10007105.

[17] A. C. Griffin, S. Khairat, S. C. Bailey, and A. E. Chung, "A chatbot for hypertension self-management support: User-centered design, development, and usability testing," *JAMIA Open*, vol. 6, no. 3, pp. 1–11, Jul. 4, 2023, doi: 10.1093/jamiaopen/ooad073.

[18] R. Biswas and N. Goel, "Intelligent chatbot assistant in agriculture domain," in *Agriculture-Centric Computation*. Cham, Switzerland: Springer, 2023, pp. 180–194, doi: 10.1007/978-3-031-43605-5_14.

[19] V. More, S. Khose, S. Sarode, and S. Rukhande, "Virtual manager for medical practitioners," in *Proc. 4th Biennial Int. Conf. Nascent Technol. Eng. (ICNTE)*, Navi Mumbai, India, Jan. 2021, pp. 1–4, doi: 10.1109/ICNTE51185.2021.9487766.

[20] R. Kulkarni and D. R. C. Jaiswal, "A survey on AI chatbots," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 11, no. 9, pp. 1738–1744, Sep. 30, 2023, doi: 10.22214/ijraset.2023.55911.

[21] J. H. Lee, M.-S. Jeong, J.-U. Cho, H.-K. Jeon, J.-H. Park, K.-D. Shin, S.-J. Song, and Y.-G. Cheong, "Developing a ophthalmic chatbot system," in *Proc. 15th Int. Conf. Ubiquitous Inf. Manage. Commun. (IMCOM)*, Jan. 2021, pp. 1–7, doi: 10.1109/IMCOM51814.2021.9377398.

[22] T. Kingchang, P. Chatwattana, and P. Wannapiroon, "Intelligent educational recommendation platform with AI chatbots," *Int. Educ. Stud.*, vol. 16, no. 5, p. 19, Sep. 24, 2023, doi: 10.5539/ies.v16n5p19.

[23] S. Sentdex. *Sentdex/OpenAssistant_API_Pythia_12B: Creating and Using an Open Assistant API Locally (Pythia 12B GPT Model)*. GitHub. [Online]. Available: https://github.com/Sentdex/OpenAssistant_API_Pythia_12B/

[24] Q. L. Tran, A.-C. Le, and V.-N. Huynh, "Towards a human-like chatbot using deep adversarial learning," in *Proc. 14th Int. Conf. Knowl. Syst. Eng. (KSE)*, Oct. 2022, pp. 1–5, doi: 10.1109/KSE56063.2022.9953771.

[25] S. W. Taju, G. F. Mandias, J. H. Moedjahedy, A. K. Wahyudi, R. Rotikan, and E. Y. Putra, "AI-powered chatbot for information service at klabat university by integrating OpenAI GPT-3 with intent recognition and semantic search," in *Proc. 5th Int. Conf. Cybern. Intell. Syst. (ICORIS)*, Pangkalpinang, Indonesia, Oct. 2023, pp. 1–6, doi: 10.1109/icoris60118.2023.10352259.

[26] M. A. Khadija, A. Aziz, and W. Nurharjadmo, "Automating information retrieval from faculty guidelines: Designing a PDF-driven chatbot powered by OpenAI ChatGPT," in *Proc. Int. Conf. Comput., Control, Informat. Appl. (IC3INA)*, Bandung, Indonesia, Oct. 2023, pp. 394–399, doi: 10.1109/IC3INA60834.2023.10285808.

[27] R. J. M. Ventayen, "OpenAI ChatGPT generated results: Similarity index of artificial intelligence-based contents," *SSRN Electron. J.*, 2023, doi: 10.2139/ssrn.4332664.

[28] T. Medeiros, M. Medeiros, M. Azevedo, M. Silva, I. Silva, and D. G. Costa, "Analysis of language-model-powered chatbots for query resolution in PDF-based automotive manuals," *Vehicles*, vol. 5, no. 4, pp. 1384–1399, Oct. 2023, doi: 10.3390/vehicles5040076.

[29] C.-C. Lin, A. Y. Q. Huang, and S. J. H. Yang, "A review of AI-driven conversational chatbots implementation methodologies and challenges (1999–2022)," *Sustainability*, vol. 15, no. 5, p. 4012, Feb. 22, 2023, doi: 10.3390/su15054012.

[30] D. Bratić, M. Šapina, D. Jurečić, and J. Ž. Gršić, "Centralized database access: Transformer framework and LLM/Chatbot integration-based hybrid model," *Appl. Syst. Innov.*, vol. 7, no. 1, p. 17, Feb. 15, 2024, doi: 10.3390/asi7010017.

[31] ChatPDF. (Nov. 13, 2023). *Chat With Any PDF!*. [Online]. Available: https://www.chatpdf.com/

[32] SciSummary. (Feb. 2, 2023). *Use AI to Summarize Scientific Articles*. [Online]. Available: https://scisummary.com/

[33] P. Sunilkumar and A. P. Shaji, "A survey on semantic similarity," in *Proc. Int. Conf. Adv. Comput., Commun. Control (ICAC3)*, Dec. 2019, p. 2019, doi: 10.1109/ICAC347590.2019.9036843.

[34] A. R. Lahitani, A. E. Permanasari, and N. A. Setiawan, "Cosine similarity to determine similarity measure: Study case in online essay assessment," in *Proc. 4th Int. Conf. Cyber IT Service Manage.*, Apr. 2016, pp. 1–6, doi: 10.1109/CITSM.2016.7577578.

[35] U. Gnewuch, S. Morana, M. T. P. Adam, and A. Maedche, "Opposing effects of response time in human–chatbot interaction," *Bus. Inf. Syst. Eng.*, vol. 64, no. 6, pp. 773–791, Dec. 2022, doi: 10.1007/s12599-022-00755-x.

**MAHBUBUL HAQ BHUIYAN** (Graduate Student Member, IEEE) received the Bachelor of Science degree in computer science and engineering from Independent University, Bangladesh, and the Master of Science degree in computer science and engineering from United International University. He is currently an Emerging Scholar in the field of AI and ML, with a keen focus on signal processing and brain development projects. As a Lecturer in computer science and engineering (CSE) with Southeast University, he is known for his dedication to foster research skills among undergraduate students. He actively encourages young minds to delve into the world of research, highlighting its importance in the evolving tech landscape. He is deeply involved in mentoring and supervising numerous student-led research projects, facilitating their exploration of innovative ideas in AI, ML, and ed-tech. His approach in the classroom is characterized by a blend of practical insights, a question-based approach alongside theoretical knowledge, aimed at preparing students for the challenges of modern technological advancements.

**MOHAMMAD SHAFIQUL ISLAM** received the master's degree in computer science and engineering from United International University and the master's degree in business studies from the National University of Bangladesh. He is currently a Senior Technical Project Manager with over 23 years of experience in the IT industry, working across diverse global markets, including Bangladesh, Saudi Arabia, the Philippines, Uganda, Bahrain, United Arab Emirates, and the USA. A Prince2-Certified Practitioner, he is highly skilled in web development, database management, and various programming languages, with extensive experience leading ERP implementations across multiple domains. Currently, based in Bangladesh, he manages large-scale software development projects for a Dubai-based company. Beyond his project management expertise, he has a strong academic background, having taught and conducted corporate training for prestigious organizations, such as Canon, Oracle, Microsoft, American Express, and the United Nations. He is a certified trainer in 16 different skills of Aptech Worldwide and was honored with the International Best Faculty Award, in 2007 and 2008. With a keen interest in AI, ML, the IoT, cloud computing, and digital transformation; he is dedicated to advancing both the technical and managerial aspects of platform development, automation, and innovation in today's rapidly evolving tech landscape.

**KHORSHED ALAM** received the B.Sc. degree in computer science and engineering (CSE) from East Delta University (EDU) and the M.Sc. degree in CSE from United International University (UIU). He has two years of professional experience as a Software Engineer specializing in android development. He is currently a Research Assistant (RA) with Independent University, Bangladesh (IUB). His academic excellence has been recognized through several prestigious awards, including the Emerging Leader Award, in 2024, the Dean's List Award, research grants, and multiple Merit Scholarships. His research interests include artificial intelligence, machine learning, signal processing, human–computer interaction, med-tech, computer networks, and software engineering. He is passionate about utilizing technology for innovative solutions and impactful research. He is actively seeking Ph.D. opportunities in the domains of artificial intelligence, machine learning, human–computer interaction, computer networks, and software engineering.

**ABUL HOSSAIN CHOWDHURY** received the B.Sc. degree in computer science and engineering from BRAC University. He is currently pursuing the M.Sc. degree in software engineering with United International University. He is currently a Software Quality Assurance professional with specialties in software testing methodologies, automation, and process optimization. His research interests include enhancing software quality through innovative testing techniques, with a focus on automation testing techniques, performance testing, and user-centric design for better user experience and overall performance of the software. His professional experience in domains like enterprise resource planner, database management, e-commerce, affiliate networks, and static application security testing solutions has provided him with unique insights into integrating technical solutions with business needs. He is proficient in popular testing tools, programming languages like Java and Python and in SQL, which is suitable for advanced testing and data validation. He is committed to academic and industry collaboration and has dedication to explore novel approaches in software quality assurance and contribute to the advancement in software engineering process.

**ZAHEED AHMED BHUIYAN** received the master's degree in computer science and engineering (CSE) from United International University (UIU), Bangladesh. He is currently pursuing the Ph.D. degree. Currently, he is a Graduate Assistant (GA) with the MSCSE Program and a Research Assistant (RA) under Dr. Md Motaharul Islam with UIU. He majored in cybersecurity in his master's, and he is interested in the Internet of Things (IoT), cloud computing, cloud security, networking, software defined networking, healthcare technologies, green computing, artificial intelligence, satellite internet communication, and grid-level energy storage systems. His research interests include the fourth industrial revolution (4IR) because of its fusion of AI, robotics, the IoT, and quantum computing advances. He has published a few research articles.

**SUMAN AHMMED** received the B.Sc.-Eng. and M.Sc. degrees from the Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology (BUET), and the Doctorate degree in ICT-based pedagogy system design, development, and implementation from PPA, Paris. He is currently an Associate Professor with the Department of Computer Science and Engineering, United International University (UIU). He is also a Founder Trustee of United International University, a renowned private university in Dhaka, Bangladesh, and also a Founder of Bangladesh Digital Education Research Ltd. (bdeducation), an ICT-enable education research-oriented company. He is also an Academic and IT Consultant and a member of managing committee for the different academic institutions in Bangladesh. He is also the Director of the Center for Development of IT Professionals (CDIP), UIU, with a vision migrating from academia to industry. His major research interests include ICT-enabled education, technology enhanced learning (TEL), innovative software/system design and development, artificial intelligence, software project management, HCI, machine learning, and artificial neural networks. He used to teach system analysis and design, software project management, system specification and analysis, IT audit, and assessment courses. He has published one book on project management on ICT-based pedagogy developed and implementation. He has published around 31 research papers into reputed journals and conferences project management. He has earned BASIS ICT Award (Education), in 2022, awarded Innovation fund from the government of Bangladesh, Research Grant from Institute of Advanced Research (IAR), UIU, and attained some research project in education from the United Trust, a Welfare Wing of United Group.

• • •