ITA Chatbot

Authors: Arjun Moitra, Ankit Raj, Himanshu Chauhan, Ishaan Chauhan, Param Kumar, Mansi Sharma
Indraprastha Institute of Information Technology

ABSTRACT: This research project addresses the challenges of traditional education systems, focusing on scalability, personalized learning, and administrative efficiency. Recognizing the limitations of conventional schooling in meeting diverse student needs, the project proposes an Intelligent Teaching Assistant (ITA) chatbot integrated into a web-based platform. Motivated by the desire to enhance education quality, the ITA aims to offer timely assistance, personalized learning experiences, and streamlined administrative support. Leveraging advanced AI algorithms and natural language processing, the chatbot analyzes PDF documents uploaded by users, providing tailored responses, quizzes, and summaries. The project's novelty lies in offering dynamic interaction based on document content, filling existing gaps in educational platforms. The methodology involves platform development, data collection, model training, and user testing. Preliminary findings show promising results, with ongoing refinement for real-world effectiveness. In conclusion, the project seeks to revolutionize education through AI-driven personalized learning and administrative support.
KEYWORDS: Education, Artificial Intelligence, Chatbot, Personalized Learning, Administrative Efficiency.

STATEMENT OF ORIGINALITY: In this paper, we present an innovative AI-powered application that integrates multiple advanced technologies to offer users a versatile and user-friendly experience. Our contribution lies in the development of a novel methodology that combines quantized model integration, audio chatting with Whisper AI, image chatting with LLaVA, PDF chatting with a custom ITA model, and SQLite database management for chat history. This approach facilitates seamless interactions with audio, images, and PDFs, enhancing user engagement and accessibility.

1   Introduction

In recent years, the advent of Artificial Intelligence (AI) and Natural Language Processing (NLP) technologies has brought about a new era of innovation across various industries. In the realm of education, the potential of these technologies to revolutionize traditional teaching methodologies and enhance learning outcomes has garnered significant attention. Among the myriad applications of AI and NLP in education, Intelligent Teaching Assistant (ITA) chatbots stand out as a particularly promising innovation. These chatbots leverage advanced algorithms and machine learning techniques to provide personalized assistance, streamline administrative tasks, and foster interactive learning environments.

1.1   Problem Formulation

Despite the promise of ITA chatbots, several challenges persist within the education sector that hinder their widespread adoption and effectiveness. These challenges include scalability issues in accommodating diverse learning needs, the lack of personalized support for students, and the burden of administrative tasks on educators. Moreover, the traditional education system struggles to keep pace with the evolving needs of learners in the digital age, often resulting in outdated teaching methodologies and limited engagement opportunities. As such, there is a pressing need to address these challenges and develop innovative solutions that harness the potential of ITA chatbots to revolutionize education.

1.2   Problem Statement

In light of the aforementioned challenges, the primary focus of this research is to investigate the ability of ITA chatbots to address the shortcomings of the traditional education system. Specifically, the research aims to explore how ITA chatbots can facilitate personalized learning experiences, streamline administrative tasks for educators, and foster a more interactive and engaging learning environment for students. By identifying key areas of improvement and evaluating the impact of ITA chatbots on educational outcomes, this research seeks to provide insights that can inform the design and implementation of future educational technologies.

2   Motivation

The motivation behind the development of an Intelligent Teaching Assistant (ITA) chatbot stems from the pressing need to address the multifaceted challenges entrenched within the traditional education system. These challenges include issues of scalability, personalized learning, and administrative efficiency, which have remained persistent obstacles to the seamless delivery of quality education. In today's educational landscape, educators are faced with the daunting task of catering to the diverse learning needs of a large number of students while grappling with time-consuming administrative tasks. This dichotomy often results in a suboptimal learning experience for students, as individualized assistance becomes

increasingly difficult to provide amidst the administrative burdens faced by educators. Consequently, there exists a critical need for innovative solutions that can bridge these gaps and usher in a new era of educational excellence.

Moreover, the significance of the problem extends beyond the confines of traditional schooling, impacting the overall success and effectiveness of educational systems worldwide. The inability to effectively address the challenges of scalability, personalized learning, and administrative efficiency not only hampers students' academic progress but also undermines the efficacy of educational institutions as a whole. Recognizing the imperative for transformative change, the introduction of an ITA chatbot presents a promising avenue for revolutionizing the educational landscape.

3 Literature Review

Literature Review: In the context of AI-driven chatbots in education, existing literature reviews provide valuable insights. A comprehensive study by Quiroga Perez et al. (2020) highlights the advantages of incorporating AI chatbots, focusing on aspects such as homework assistance, personalized learning, and skill development. The review also delves into the challenges, limitations, and ethical considerations associated with this technology. To set our proposal apart, we would focus on specific educational domains and address unique challenges not extensively covered in this extensive review. Another significant contribution comes from Chang et al. (2023), which explores educational chatbots across various dimensions like educational field, platform, and design principles. It provides a comprehensive understanding of chatbot interaction styles and their impact on learning outcomes. To differentiate our approach, we would explore additional dimensions or innovative chatbot designs beyond the current literature. Additionally, Hew et al. (2022) investigates the application areas of chatbots in education, emphasizing pedagogical roles, mentoring scenarios, and personalization. The focus on personalized chatbots in related work opens avenues for our proposal to contribute by refining personalization techniques or exploring novel mentoring scenarios. Furthermore, Nkambou et al. (2010) addresses the challenges in constructing intelligent tutoring systems, shedding light on problems, solutions, and the historical context of ITS research. Our proposal would stand out by concentrating on specific features like adaptability

and real-time feedback to enhance the tutoring experience. Unfortunately, accessibility issues prevent a detailed summary of Kasthuri and Balaji (2023), but it likely covers a range of chatbot applications, benefits, and challenges in education. To differentiate our approach, we would explore applications beyond homework assistance and consider integrating chatbots with emerging technologies like AR/VR to address specific educational gapsKasthuri and Balaji (2023); Quiroga Perez et al. (2020); Chang et al. (2023); Hew et al. (2022); Labadze et al. (2023); Nkambou et al. (2010).

4 Novelty

The novelty of the proposed ITA chatbot lies in its innovative approach to addressing longstanding challenges within the education sector. Unlike traditional teaching methods and educational technologies, the ITA chatbot harnesses the power of artificial intelligence and natural language processing to deliver personalized assistance and support to students and educators alike. By incorporating advanced algorithms and machine learning techniques, the chatbot is capable of adapting to the individual learning needs of students, providing tailored recommendations, and streamlining administrative tasks with unprecedented efficiency. Furthermore, the integration of voice recognition capabilities expands the accessibility of the chatbot, enabling users to interact with the system using natural language commands and voice inputs. This fusion of cutting-edge technology and pedagogical principles represents a novel paradigm in education, offering a transformative solution that has the potential to revolutionize the way teaching and learning are conducted in the digital age.

5 Methodology

Our innovative AI-powered application is designed to offer users a versatile and user-friendly experience through the integration of multiple advanced technologies. This section outlines the methodologies employed in the development of our application, highlighting key features and elucidating its functioning.

5.1 Key Features

1. Quantized Model Integration: The application utilizes quantized models, and optimized versions of larger AI models, to ensure efficient performance on consumer-grade hardware. This approach enables broader accessibility without necessitating high-end computing resources.

2. Audio Chatting with Whisper AI: Whisper AI enriches the audio chatting experience by providing robust transcription capabilities. This feature facilitates accurate interpretation and response to voice inputs, fostering smooth and natural conversations.

3. Image Chatting with LLaVA: LLaVA, a fine-tuned LLaMA model, enables image-based chatting within the application. Leveraging image embeddings created using a CLIP model, users can seamlessly discuss and share visual content.

4. PDF Chatting with Custom ITA Model: Our custom ITA Chatbot allows users to upload PDFs and query content. This involves a multi-step process, including data extraction, text cleaning, text vectorization, knowledge base storage, and similarity search, ensuring precise and pertinent responses.

5. SQLite Database for Chat History: A SQLite database is employed to store and manage chat history, enabling seamless access to past interactions. This feature contributes to a continuous and cohesive user experience.

5.2 How It Works

The application integrates these features to provide users with a cohesive and responsive experience. Users interact with the application through various modes, each offering tailored functionalities to enhance the user experience. Advanced AI techniques are utilized to process user interactions and generate appropriate responses. Furthermore, the SQLite database facilitates seamless access to chat history, allowing users to revisit past interactions effortlessly. Upon initialization, the application establishes connections with the necessary resources and loads the quantized models required for audio, image, and PDF processing. Users can then interact with the application through various modes, including audio, images, and PDFs.

In the **audio mode**, users can send voice messages, which are transcribed by Whisper AI in real time. Whisper AI employs advanced natural language processing (NLP) techniques to accurately interpret the voice inputs and generate text-based responses.

For **image-based interactions**, users can share images, which are processed by the LLaVA model. LLaVA understands image embeddings created using a CLIP model, enabling it to recognize and interpret visual content. This allows users to engage in discussions and share visual information seamlessly.

In the PDF mode, users can upload PDF documents and ask questions about the content. The custom ITA Chatbot handles these queries by extracting text from the PDF, cleaning it to remove unwanted characters, and converting it into numerical vectors through tokenization and word embeddings. These vectors are then stored in a knowledge base with relevant metadata for efficient retrieval. When a user asks a question, the ITA Chatbot performs a similarity search to find the most relevant information in the knowledge base and generates a response accordingly.

Throughout these interactions, the application maintains a SQLite database to store and manage chat history. This database ensures that users can access their past interactions easily, providing a seamless and continuous user experience.

6 Database

In our project, we utilized a variety of datasets to train, validate, and evaluate the performance of the ITA chatbot. These datasets include:

- **PaperChecking.csv**: This dataset comprises questions typically used for paper checking or quiz grading. It includes columns such as question_id, question, options, and correct_option. The questions are in multiple-choice format, allowing the chatbot to generate appropriate responses based on the user's selection.

- **MCQ.csv**: The MCQ dataset contains multiple-choice questions related to Information Retrieval (IR) topics. It consists of columns such as question_id, question, options (option1, option2, option3, option4), and correct_answer. These questions cover a range of IR concepts, providing users with opportunities to test their knowledge and receive informative responses from the chatbot.

- **SingleQA.csv**: This dataset contains single-answer questions along with their corresponding answers. Each entry in the dataset consists of a question_id, question, and answer. These questions are designed to elicit specific responses from the chatbot, allowing users to obtain precise information or clarification on various topics related to IR.

- **CourseMaterial.csv**: The CourseMaterial dataset includes educational materials such as lecture notes, presentations, or supplementary resources related to IR. It contains columns such as material_id, title, description, and link. The chatbot can utilize this dataset to provide users with relevant course materials, aiding in their learning process and understanding of IR concepts.

7 Code

The implementation of our solution is available on GitHub at [link], along with documentation and instructions for usage.

8 Evaluation

To evaluate the effectiveness of the Intelligent Teaching Assistant (ITA) chatbot, we conducted comprehensive assessments focusing on two key aspects:

- Comparison with baselines and system performance on existing data

- State-of-the-art (SOTA) performance on different evaluation metrics and its ability to handle new data and different cases.

8.1 Comparison with baseline results

- The ITA chatbot's performance is compared with baseline systems, which represent simplistic or rule-based approaches to chatbot development. Through this comparison, we ascertain the chatbot's superiority and effectiveness in handling complex user queries and providing accurate responses.

- Using existing datasets commonly used for training and testing educational chatbots, we evaluate the chatbot's performance on diverse user queries, responses, and interactions. This assessment helps us gauge the chatbot's robustness and generalization capabilities in real-world scenarios.

```
Metrics for CourseMaterial.csv:

Precision  :      0.12698412698412698
Recall     :      0.7272727272727273
F1-score   :      0.2162162162162162
```

```
Metrics for PaperChecking.csv:

Precision  : 0.2
Recall     : 0.9090909090909091
F1-score   : 0.32786885245901637
```

```
Metrics for MCQ.csv:

Precision  : 0.45
Recall     : 0.8181818181818182
F1-score   : 0.5806451612903226
```

```
Metrics for singleQA.csv:

Precision  : 0.32
Recall     : 0.48484848484848486
F1-score   : 0.38554216867469876
```

```
Updated Metrics for CourseMaterial.csv:

Precision  :      0.8571428571428571
Recall     :      0.675
F1-score   :      0.7552447552447553
```

```
Updated Metrics for PaperChecking.csv:

Precision  : 0.8333333333333334
Recall     : 0.32894736842105265
F1-score   : 0.4716981132075471
```

```
Updated Metrics for MCQ.csv:

Precision  : 0.95
Recall     : 0.2235294117647059
F1-score   : 0.3619047619047619
```

```
Updated Metrics for singleQA.csv:

Precision  : 0.9130434782608695
Recall     : 0.3088235294117647
F1-score   : 0.46153846153846156
```

8.2 SOTA - State-of-the-Art

- The ITA chatbot is benchmarked against state-of-the-art (SOTA) models and techniques in the field of educational chatbots. This benchmarking enables us to assess the chatbot's performance relative to cutting-edge models and technologies.

- We evaluate the chatbot's performance on different evaluation metrics such as F1 score, precision, recall, and accuracy. By comparing these metrics with SOTA benchmarks, we determine the chatbot's level of advancement and innovation in educational assistance.

- Performance metrics such as F1 score, precision, recall, and accuracy are used to quantify the chatbot's performance relative to SOTA benchmarks. Achieving superior results in these metrics signifies the chatbot's ability to outperform existing models and technologies in educational chatbot applications.

8.3 Performance on New Data and Handling Different Cases:

- In addition to evaluating the chatbot's performance on existing data, we assess its ability to handle new data and diverse cases. This includes scenarios where user queries vary in complexity, ambiguity, and specificity.

- By analyzing the chatbot's performance on new data, we gain insights into its adaptability, versatility, and effectiveness in addressing a wide range of user needs and requirements. Furthermore, we examine how the chatbot responds to previously unseen queries and whether it maintains high accuracy and relevance in such cases.

9 APPENDIX I

9.1 Interview Questions for Professors:

Q1. What are your thoughts on integrating AI-driven chatbots as teaching assistants in your classroom?

Q2. Do you see them as beneficial tools for enhancing the learning experience?

Q3. Have you previously used any AI-driven tools or chatbots in your teaching practices? If so, what was your experience?

Q4. What specific challenges do you face in managing administrative tasks and providing personalised assistance to students?

Q5. What specific tasks or areas do you believe they could assist with effectively?

Q6. How do you see AI chatbots to address these challenges and enhance the educational experience for both you and your students?

Q7. What specific tasks or areas do you believe they could assist with effectively?

Q8. In your opinion, what are the key features or functionalities that an ideal AI teaching assistant chatbot should possess?

Q9. Are there any particular capabilities that you consider essential for their successful integration into classroom activities?

Q10. Are there any concerns or reservations you have regarding the implementation of AI chatbots in education, particularly as teaching assistants?

Q11. What potential risks or drawbacks do you foresee, if any?

Q12. How do you see AI chatbots impacting the dynamics of teacher-student interactions and the overall classroom environment?

Q13. Do you see them as potentially altering traditional teaching roles or fostering new forms of engagement?

9.2 Interview Questions for Students:

Q1. Have you ever interacted with an AI-driven chatbot in an educational setting before? If yes, could you share your experience?

Q2. What are some challenges you encounter while studying or seeking assistance from your instructors or teaching assistants?

Q3. How do you think AI chatbots could help address these challenges and improve your learning experience?

Q4. What specific features or functionalities would you like to see in an AI teaching assistant chatbot to support your academic needs?

Q5. Are there any concerns or apprehensions about interacting with AI chatbots for academic purposes?

Q6. How do you perceive the role of AI chatbots in supplementing traditional teaching methods and personalized learning experiences?

Q7. Do you believe AI chatbots have the potential to positively impact your academic performance and overall engagement with course materials?

9.3 Interview Questions for Teaching Assistants:

Q1. What are some of the primary responsibilities and tasks you perform as a teaching assistant?

Q2. How do you currently address the challenges of managing administrative duties and providing personalised assistance to students?

Q3. In what ways do you think AI chatbots could assist you in your role as a teaching assistant?

Q4. What features or functionalities do you believe are essential for an AI teaching assistant chatbot to effectively support you in your tasks?

Q5. Do you have any concerns or reservations about integrating AI chatbots into your workflow as a teaching assistant?

Q6. How do you envision collaborating with an AI chatbot in facilitating student learning and engagement?

Q7. What opportunities do you see for AI chatbots to enhance the overall effectiveness and efficiency of teaching assistant roles in educational settings?

References

Chang, D., Lin, M., Hajian, S., and Wang, Q. (2023). Educational design principles of using ai chatbot that supports self-regulated learning in education: Goal setting, feedback, and personalization. *Sustainability*, 15:12921.

Hew, K., Huang, W., Du, J., and Jia, C. (2022). Using chatbots to support student goal setting and social presence in fully online activities: learner engagement and perceptions. *Journal of Computing in Higher Education*, 35:1–29.

Kasthuri, E. and Balaji, S. (2023). Natural language processing and deep learning chatbot using long short term memory algorithm. *Materials Today: Proceedings*, 81:690–693. International Virtual Conference on Sustainable Materials (IVCSM-2k20).

Labadze, L., Grigolia, M., and Machaidze, L. (2023). Role of ai chatbots in education: systematic literature review. *International Journal of Educational Technology in Higher Education*, 20.

Nkambou, R., Bourdeau, J., and Psyché, V. (2010). *Building Intelligent Tutoring Systems: An Overview*, volume 308, pages 361–375.

Quiroga Perez, J., Daradoumis, T., and Puig, J. (2020). Rediscovering the use of chatbots in education: A systematic literature review. *Computer Applications in Engineering Education*, 28.