WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Interactive VR Using RAG in Education

Christo Vincent [a], Allen Antony [a], Athul Asok [a], Abin Antony [a], Dr. Anita Brigit Mathew [b]

[a] UG Student, Artificial Intelligence and Data Science, Viswajyothi College of Engineering and Technology, Muvattupuzha, 686661, India
[b] Head of Department, Artificial Intelligence and Data Science, Viswajyothi College of Engineering and Technology, Muvattupuzha, 686661, India

ABSTRACT:

The integration of Artificial Intelligence and Virtual Reality technologies can revolutionize the field of education by introducing innovative teaching methods and enhancing learning experiences. This paper presents AI-VR Tutor, a novel approach to education that leverages the power of virtual reality (VR) and artificial intelligence (AI). Students can enter a VR environment built using Unity Engine, where they can import 3D models representing their learning content. Within this immersive space, a virtual tutor avatar awaits, ready to answer questions and provide clarifications. The key to interaction lies in the Retrieval-Augmented Generation (RAG). A massive subject-specific data corpus is fed into the RAG system, empowering the AI tutor to access and process relevant information. This innovative combination of VR technology and AI-driven tutoring fosters an engaging and effective learning experience, promoting deeper understanding and knowledge retention.

Keywords: *Virtual Reality (VR), Large Language Models(LLM), Retrieval-Augmented Generation(RAG), Hypertext Transfer Protocol(HTTP), Artificial Intelligence(AI)*

1. Introduction

In recent years, the fusion of Artificial Intelligence (AI) and Virtual Reality (VR) technologies has driven significant progress across various fields. This study explores integrating Retrieval-Augmented Generation(RAG) into VR environments to create personalized, immersive tutoring experiences. Specifically, we focus on developing a VR tutor with text-generative capabilities using the Unity game engine.

Conventional teaching methods often struggle to engage learners effectively due to their inability to cater to diverse learning styles. However, RAG combined with VR offers a solution. Educators can tailor learning experiences to individual needs, enhancing comprehension and retention of complex concepts.

Our system provides personalized guidance and feedback for optimized learning outcomes. Key components include natural language processing for interactive communication. Integrated seamlessly within Unity, this approach aims to transcend traditional teaching frameworks.

This paper put forward a system that integrates Retrieval-Augmented Generation in VR for a more interactive learning experience.

2. Literature Survey

2.1 Retrieval-Augmented Generation for Neural Machine Translation

This study by Y. Lu [1] Retrieval-Augmented Generation for Neural Machine Translation (RAG NMT) is a technique that enhances the translation quality of neural machine translation (NMT) models by incorporating retrieval mechanisms. NMT models traditionally translate sentences based solely on the input sentence itself. RAG NMT augments this process by retrieving additional relevant information, such as similar sentences from a vast repository of text (like a document collection or the web) to bolster the translation process. This additional information provides context that the NMT model can leverage to generate more accurate and nuanced translations, particularly for complex or ambiguous sentences.

2.2 Automatic speech recognition: Deep learning approach

The paper by D. Yu and L. Deng [3] likely discusses how deep learning, specifically deep neural networks (DNNs), are revolutionizing automatic speech recognition (ASR) technology. ASR translates spoken language into text. Traditionally, ASR relied on less complex models. This paper argues that DNNs bring significant advantages to ASR due to their ability to learn complex patterns from vast amounts of speech data. This improvement can lead to more accurate speech recognition, even in challenging situations with background noise or variations in speaking styles.

3. Proposed work

In this section we describe the components of our system and their features.

A. *Retrieval-Augmented Generation*

We use Retrieval-Augmented Generation to generate

textual information of the query by the user from the VR front end. The system also allows the user to select links and PDFs to be stored in the vector databases. Teachers can choose what students should refer to and with what information the text should be generated.

The system allows teachers to name game objects that will be passed as context to RAG for generating relevant content.RAG uses FAISS vector databases to conduct fast similarity text to get more context this generated information is passed to the front end. Information in FAISS indexes can be updated in real time by selecting links or pdf data. Retrieval-Augment generation is executed in the backend and generated output is sent to unity VR front end by HTTP response.

The information passed to the RAG contains the questions and context generated from the similarity search. The generated text will be based on the information stored in the FAISS vector database. This will enable students to learn from the context of a syllabus that they are supposed to.

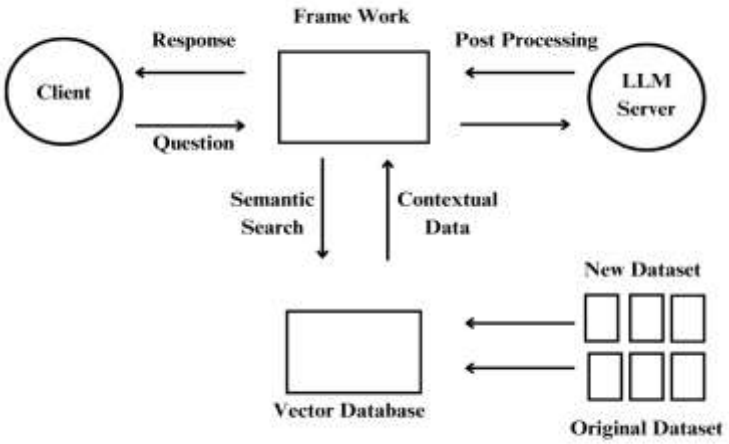

Fig. 1. RAG Architecture

B. Speech recognition

The Python SpeechRecognition library serves as a valuable tool for incorporating speech recognition functionalities into these projects. It acts as a user-friendly interface, enabling users to interact with various speech recognition engines like Google Speech-to-Text or CMU Sphinx. This simplifies the integration of speech recognition tasks without requiring in-depth knowledge of the underlying APIs. The library offers functionalities for audio recording, noise reduction (through adjustments to energy thresholds), and speech-to-text conversion. By leveraging SpeechRecognition, researchers can focus on their core research questions, such as exploring the effectiveness of different dialogue strategies in human-computer interaction or analyzing the impact of noise on speech recognition accuracy. This allows the system to capture audio from the user and convert it into text that is mainly questions based on the interaction in the VR world. This text is passed as a query to generate context to be passed to RAG to generate text and explanation based on information stored in vector databases.

C. Unity

System use unity to generate virtual reality worlds for deep interaction of students with the subject and related contents. Three-dimensional models must be imported and named according to the syllabus so that students get correct feedback. The system is made to work with Oculus but can be tweaked into others using Unity.

Teacher mode enables instructors to arrange and import models and name them which allows some level of control over what students interact with.

Unity enables the creation of interactable objects and import models to the world.

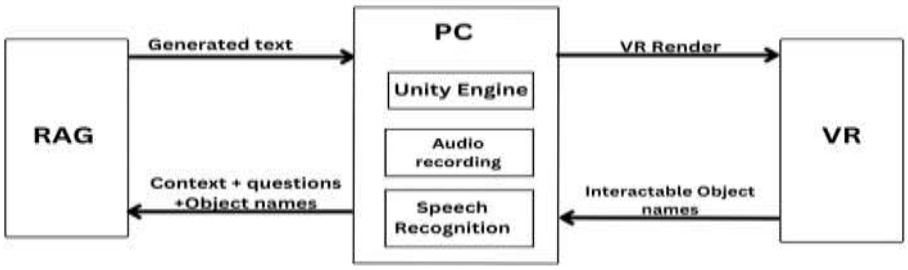

Fig. 2. Overall Architecture

4. Results and Discussion

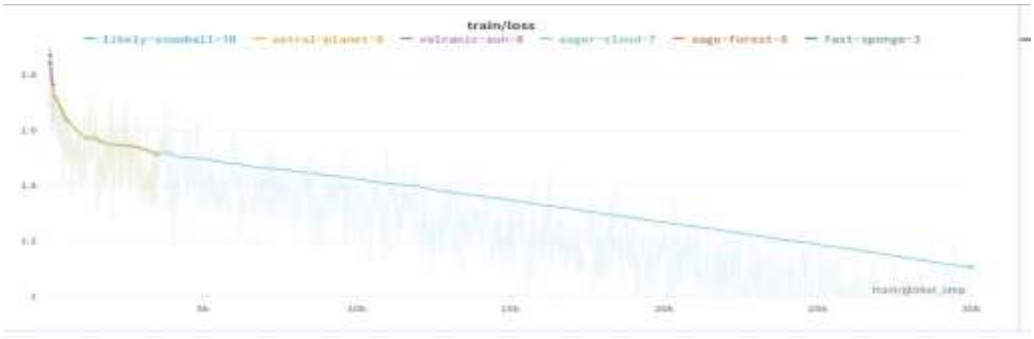

fig. 3 Falcon 7b performance graph.

5. Conclusion and Future Scope

In conclusion, the integration of Retrieval-Augmented Generation (RAG) within Virtual Reality (VR) environments holds immense potential to revolutionize education by offering personalized, immersive, and interactive learning experiences. Our study introduces the AI VR Tutor system, which harnesses the power of Artificial Intelligence (AI) to enhance educational practices within the Unity game engine.

Traditional teaching methods often face challenges in engaging learners effectively and catering to diverse learning styles. By leveraging RAG within VR, educators can customize learning experiences to individual needs, thereby improving comprehension and retention of complex concepts.

The AI VR Tutor system facilitates interactive communication through natural language processing, providing personalized guidance and feedback to learners. Integrated seamlessly within Unity, this approach transcends conventional teaching frameworks, offering a dynamic and engaging learning environment.

Overall, our research demonstrates the potential of integrating RAG with VR to create innovative educational solutions that promote effective learning outcomes. As technology continues to advance, further exploration and refinement of such systems have the potential to reshape the landscape of education, making learning more accessible, engaging, and impactful for learners worldwide.

Acknowledgment

We would like to express our sincere gratitude toViswajyothi College of Engineering and Technology, Vazhakulam for their generous support. We would like to express our heartfelt thanks to our Project Coordinators Mrs.Mary Nirmala George and Mrs Anupriya Mohan for their motivation, assistance and help. We also thank all the faculty members of the Artificial Intelligence and Data Science Department for their constant support and encouragement.

REFERENCES

[1] Y. Lu et al., "Retrieval-Augmented Generation for Neural Machine Translation," in Transactions on Audio, Speech, and Language Processing, vol. 28, no. 5, pp. 1144-1154, 2020.

[2] H. Hermansky, "Neural networks for speech recognition," in Speech Communication, vol. 16, no. 1, pp. 5-16, 1995.

[3] D. Yu and L. Deng, "Automatic speech recognition: Deep learning approach," in Speech Processing Magazine, IEEE, vol. 28, no. 1, pp. 14–22, 2017.

[4] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, S. Riedel, and D. Kiela (2020), "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks" in arXiv preprint arXiv:2005.11401.

[5] Buyuk Salih, I., Bayburt, S., Buyuksalih, G., Baskaraca, A. P., Karim, H., and Rahman, A. A.: 3D "Modeling and Visualization based on the Unity Game Engine – Advantages and Challenges*",* ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., IV-4/W4, 161–166 https://doi.org/10.5194/isprs-annals-IV-4-W4-161-2017, 2017.

[6] Chen, J., Lin, H., Han, X., & Sun, L. (2024). "Benchmarking Large Language Models in Retrieval-Augmented Generation. Proceedings of the AAAI Conference on Artificial Intelligence", 38(16), 17754-17762. https://doi.org/10.1609/aaai.v38i16.29728

.