School of Computing

FACULTY OF ENGINEERING AND
PHYSICAL SCIENCES

UNIVERSITY OF LEEDS

Final Report

Enhancing University Learning with Retrieval-Augmented Generation and GPT-3.5 Fine-Tuning

Abdul Karim Abbas

Submitted in accordance with the requirements for the degree of
BSc, MEng Computer Science with Artificial Intelligence

2023/24

COMP3931 Individual Project

The candidate confirms that the following have been submitted:

Items	Format	Recipient(s) and Date
Final Report	*PDF file*	*Uploaded to Minerva (30/04/2024)*
Scanned participant consent forms	*PDF file*	*Uploaded to Minerva (25/04/2024)*
Link to online code repository	*URL*	*Sent to supervisor and assessor (30/04/2024)*

The candidate confirms that the work submitted is their own and the appropriate credit has been given where reference has been made to the work of others.

I understand that failure to attribute material which is obtained from another source may be considered as plagiarism.

(Signature of student) ______________________

Summary

This dissertation explores the integration of Retrieval-Augmented Generation (RAG) and fine-tuning techniques with GPT-3.5 to enhance digital learning in higher education. Given the rapid adoption of digital platforms in educational settings, there is a developing need to enhance the interactive capabilities of educational chatbots to support students effectively. Current AI models, while robust, often fall short in responding to domain-specific queries that require up-to-date knowledge and often need continuous specialised training.

This research aimed to refine GPT-3.5s responsiveness and adaptability to the dynamic nature of educational content by implementing retrieval-augmented generation to enhance the chatbot's response accuracy and relevance; ensuring that the chatbot could handle a wide array of student inquiries on a specialised domain with precision.

The methodology involved the integration of RAG to supplement GPT-3.5's knowledge base, enabling it to retrieve and utilise the most relevant information from a curated dataset of university lecture materials. Additionally, the project explored fine-tuning techniques to tailor GPT-3.5's responses to the specific language and needs of university-level material.

The project achieved the development of an enhanced chatbot that successfully leverages retrieval-augmented generation to access and integrate targeted information into its responses, thereby significantly improving the specificity and accuracy of its assistance. Fine-tuning was applied to better understand and process complex academic queries specific to university courses, achieving a high degree of contextual alignment to the source lecture and response relevance. The model was rigorously tested and validated within an educational setting, demonstrating improved performance in generating accurate responses compared to the standard GPT model. Feedback from real university students confirmed that the enhancements could substantially improve their learning experience by providing precise and informative responses.

Acknowledgements

I want to express my gratitude to my project supervisor, Professor Eric Atwell, for his wisdom and guidance throughout this project. His expertise in this field was truly invaluable.

I am also grateful to Dr. Nishant Ravikumar for his essential intermediate feedback as project assessor. Their insight always helped ensure I was on the right path.

Finally, I thank my mom, dad, and sister for their unconditional love and support throughout this project and university as a whole.

Table of Contents

Chapter 1
Introduction and Background Research

1.1 Introduction

In education, particularly within the context of digital learning, chatbots represent a promising opportunity for supplementary student support. Large language models such as GPT-4 (Open AI, 2023) and BARD (Google) are already emerging as valuable tools for student support. These advanced AI models are increasingly recognised for their ability to understand and generate human-like text, making them ideal for interactive learning environments. They offer real-time responses to inquiries, can guide through complex concepts, and provide easy access to a vast array of information. By leveraging these technologies, educational institutions can deliver more responsive and adaptive learning experiences, catering to the diverse and unique needs of students across different learning settings. This integration not only improves accessibility and engagement but also allows educators to focus on other more demanding tasks, enhancing overall educational quality.

However, a significant challenge remains in the domain-specific limitations of chatbots such as GPT-3.5. When tasked with answering specific questions—such as those about new developments beyond its last training data or detailed queries about a particular university course—GPT-3.5 frequently encounters difficulties in providing accurate responses and often resorts to generating incorrect information that mimics what a correct response may look like. This issue highlights the necessity for ongoing updates and specialised training for AI models to maintain their relevance and effectiveness.

This project proposes an intuitive approach to overcome these limitations by integrating the capabilities of GPT technology, Retrieval-Augmented Generation (RAG), and fine-tuning techniques to enhance the chatbot's knowledge base with a tailored, supplementary dataset; thereby allowing the chatbot to gain the ability to access the most pertinent and up-to-date information across various domains without the need for constant retraining.

Additionally, the project will undertake a series of experiments to both quantitatively and qualitatively assess the effects of this framework on the chatbot's performance, focusing particularly on enhancing the coherence, relevance, and accuracy of its responses within educational settings.

Ultimately, the success of this project will be determined by its ability to effectively integrate advanced AI functionalities, making educational chatbots more responsive and adept at handling the specificities of academic discourse.

1.2 Digital Learning and AI

The advancement of digital learning is evident through the growing interaction with online technological platforms. Massive Open Online Courses (MOOCs) such as Coursera, introduced in 2018, have played a pivotal role in access to digital education, allowing learners worldwide to access courses from leading universities at minimal or no cost. Similarly, Learning Management Systems (LMS) such as Blackboard, launched with new features in 2023, have become integral to organising and delivering educational content, facilitating communication between educators and students, and tracking student progress in both traditional and online learning environments. Additionally, the rise of educational apps, for instance, Duolingo's language learning platform in 2019, highlights the shift towards accessible and interactive learning experiences, enabling users to study new languages on-the-go with engaging, gamified content.

Digital learning's evolution continues as it embraces increasingly complex data analytics and feedback mechanisms. By leveraging big data, educators and institutions can uncover insights into learning patterns, predict student performance, and offer targeted interventions to support learners at risk of falling behind. The potential for data-driven customisation in digital education paves the way for more nuanced and effective learning pathways, marking a significant leap forward in educational technology's ability to cater to diverse learning styles and needs.

AI chatbots can significantly contribute to the scalability of digital learning, enabling the management of large volumes of student interactions without the need for extensive human resources. This is particularly beneficial in scenarios where an instructor's attention is spread across many students. Additionally, according to Kooli (2023), chatbots have the potential to personalise the learning pathway by analysing the students' learning abilities, suggesting additional learning resources, and providing revision on challenging topics, thereby enhancing overall learning outcomes.

To further understand how these AI systems can be optimised for educational purposes, the following chapter delves into the GPT transformer architecture, which underpins models like GPT-3.5. Outlining the basic mechanisms of how GPT models process and generate language will provide a foundation for their improvement.