



Educational Resources Advisor Chatbot for Diverse Audience

GLC RESEARCH GROUP 2: Responsible access and availability of useful AI materials to learners

[Akuadasuo Ezenyilimba](#)

Postdoctoral Research Scholar
School for the Future of Innovation
in Society Arizona State University
USA

[Gokul S KRISHNAN \(Ph.D.\)](#)

Research Scientist
Centre for Responsible AI
IIT Madras
India

Recent advancements in the field of AI and Natural Language Processing have seen applications of Large Language Models across sectors. However, they do have concerns in terms of various aspects such as Hallucinations (generating undesirable content), causing bias issues, and cannot be trusted in certain crucial sectors. However, users across the globe have started consuming these LLM based chatbot applications for various purposes including assistance in finding educational resources and career guidance. Therefore, it is important to address these issues and ensure that such applications adhere to various aspects of Responsible AI making these models suitable and safe for usage. In this work, we intend to look at existing LLM based chatbots as well as based LLMs that can be finetuned to ensure accurate content that is free from fairness/bias issues, making it safer to use for diverse audiences across the globe.

Aim: To design and develop a language model based chatbot application that can provide useful educational resources for diverse audiences across the globe.

Objectives:

1. To finetune a LLM for conversational use cases in educational fields and information.
2. To have an effective mechanism for detecting and weeding out false information (due to hallucination of LLMs) and correct them while providing responses to users
3. To design an effective approach to detect and mitigate geographical/racial/gender biases from the responses using Responsible AI principles and techniques
4. To develop a Web/Mobile app that can take user inputs and provide responses.

Pre-requisites: Basics of NLP, Deep Learning, Large Language Models, Responsible AI, Web App development

Research Questions:

1. Can we finetune a LLM to give us accurate information on educational resources to act as a conversational agent?
2. Can we ensure that the outputs for the proposed conversational agent are fair and inclusive?

Tentative Methodology:

1. Dataset – Find datasets that can be used to finetune the LLM for the proposed conversational agent.
2. Model – LLMs like GPT, Llama and smaller LMs like Phi, T5, etc. can be considered to build the conversational agent
3. Fairness Mechanism – We can have post-hoc approaches to detect/weed out biases from the outputs of the trained LLMs
4. Building UI and Web/Mobile app using the trained model.



Milestones

Feature name	PRD	PM	Start date	End date
Project Ideation	https://drive.google.com/drive/u/1/folders/1582FPQ2oxbBXKPk_k2_BE5FNVfcXKrlq5	Akuadasuo Ezenyili mba	Feb 15	mars. 15
Literature Review		Akuadasuo Ezenyili mba	Feb 23	mars. 15
Project Proposal		Akuadasuo Ezenyili mba	mars. 15	Apr 12
Project Prototype		Akuadasuo Ezenyili mba	Apr 12	juin.28
Project Pilot		Akuadasuo Ezenyili mba	juin.28	Aug 16