SCHOOL OF ENGINEERING & COMPUTING

Harnessing Large Language Models for Multitasking AI Chatbot

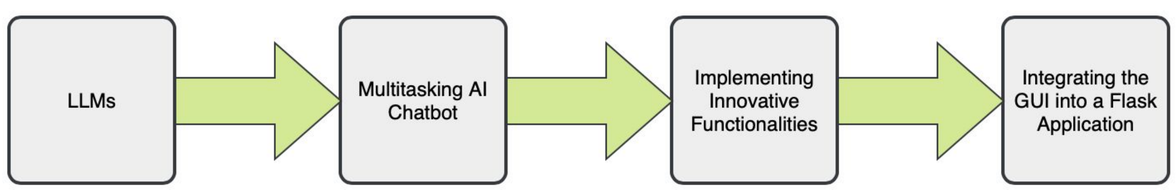
Connor Hehn (Software Engineering), Reyes Huerta (Cybersecurity), David McNulty (Software Engineering), Andrew Visceglia (Data Science)

Advisor: Dr. Sidike Paheding

INTRODUCTION & BACKGROUND

- In recent years, the growth of AI has been fueled by the development of generative AI applications from OpenAI, Google, and Microsoft.
- IBM defines chatbots as AI-driven programs simulating human conversation through NLP.
- Transformer models (e.g., GPT-3) are essential for modern AI chatbot progression, with the attention mechanism enhancing contextual understanding.
- LLMs → multitasking AI chatbot (voice and text commands in English and Spanish) → implementing innovative functionalities (mapping and music)
- Running the model on a computer → integrating the GUI into a Flask application

Proposed Workflow



OBJECTIVE

The overarching objective of this project is to deliver an accessible and multitasking AI chatbot application that is tailored to meet the needs of prospective Fairfield University students through the integration of speech recognition and text-to-speech capabilities.

METHODOLOGY

- Our AI-powered chatbot utilizes several different modules to achieve its accessible functionality.
- Below are all of the modules used along with their purpose in the project.

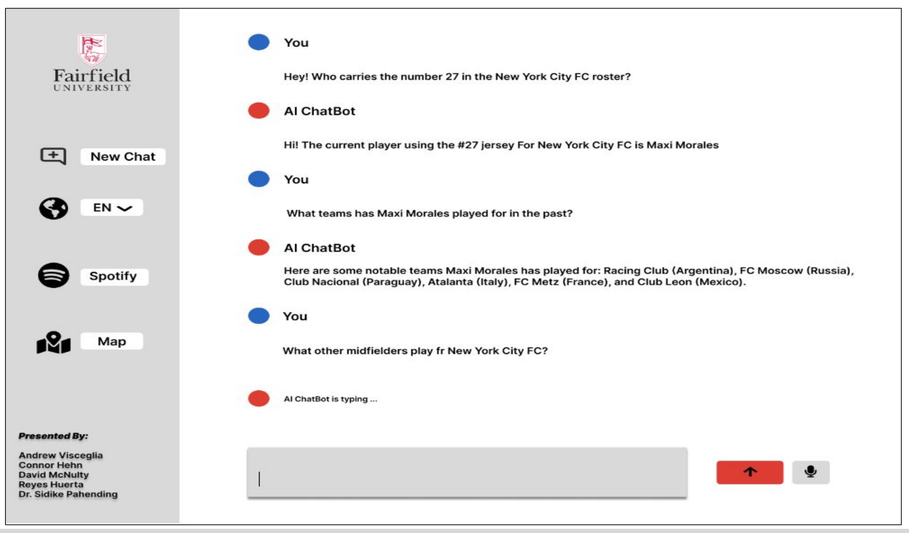
Module Name	Functionality
Hugging Face	Allows interaction with Hugging Face Hub, specifically the LLM: Mixtral-8x7B.
Flask	Python web framework used for handling routing, templating, and web related tasks.
json	Syntax for json data which is the information accessed through our API calls.
gTTS	Provides us with an interface to Google's Text-to-Speech API.
speech_recognition	Library which performs speech recognition with support for several APIs and engines. Capable of both online and offline functionality.
pygame	Set of Python modules that are designed for creating fully featured multimedia programs.
io	Module for accessing Python interfaces to stream handling.

MODEL FORMULATION

- Large Language Model (LLM): Mixtral-8x7B
  - Generative model based on Sparse Mixture of Experts (SMoE) Architecture.
  - Several "experts" handle separate tasks.
  - Outperforms Llama 2 70B on benchmark tests.
  - Implemented via Hugging Face Hub.
- Spotify API:
  - Allows users to request and play music.
- Mapping API: GraphHopper
  - Current plan is FREE and allows for 500 credits/day, up to 5 locations/request, and up to 1 vehicles/request.
  - GraphHopper Geocoding and Routing APIs
  - Provides step-by-step walking directions, the total distance and time for traveling from the start to end locations, and a map tracing the path for the user to follow.
  - Practical Use: Navigating Fairfield University's campus

RESULTS

- Our ongoing development encompasses an immersive chatbot experience, boasting a plethora of dynamic features:
  - Seamless speech interaction, enabling effortless communication
  - Engaging text-to-speech functionality for an immersive conversational experience
  - Intuitive language selection, allowing users to effortlessly switch between five languages (English, Spanish, French, German, Italian)
  - Interactive mapping capabilities enhancing user navigation
  - Spotify integration, providing users with personalized music listening



CONCLUSIONS & FUTURE PLANS

- Future plans include:
  - iOS development
  - Upload application to Fairfield University AI Laboratory

ACKNOWLEDGMENT / REFERENCES

[1] Kasneci, Enkelejda, et al. "ChatGPT for good? On opportunities and challenges of large language models for education." *Learning and Individual Differences*, vol. 103, 2023, p. 102274.
[2] Radford, Alec, et al. "Language models are unsupervised multitask learners." *OpenAI Blog*, vol. 1, no. 8, 2019, p. 9.
[3] Samant, Rahul Manohar, et al. "Framework for deep learning-based language models using multi-task learning in natural language understanding: A systematic literature review and future directions." *IEEE Access*, vol. 10, 2022, pp. 17078-17097.
[4] Wei, Jason, et al. "Emergent abilities of large language models." *arXiv preprint*, arXiv:2206.07682, 2022.
[5] Kasirzadeh, Atoosa, and Iason Gabriel. "In conversation with artificial intelligence: aligning language models with human values." *Philosophy & Technology*, vol. 36, no. 2, 2023, pp. 27.
[6] Bouschery, Sebastian G., Vera Blazevic, and Frank T. Piller. "Augmenting human innovation teams with artificial intelligence: Exploring transformer-based language models." *Journal of Product Innovation Management*, vol. 40, no. 2, 2023, pp. 139-153.
[7] Al Smadi, Takialddin, et al. "Artificial intelligence for speech recognition based on neural networks." *Journal of Signal and Information Processing*, vol. 6, no. 02, 2015, pp. 66.
[8] Chen, Yi-Chen, et al. "Aipnet: Generative adversarial pre-training of accent-invariant networks for end-to-end speech recognition." *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2020.
[9] "mistralai/Mixtral-8x7B-v0.1 · Hugging Face," Huggingface.co, 2024. https://huggingface.co/mistralai/Mixtral-8x7B-v0.1.

ASEE AMERICAN SOCIETY FOR ENGINEERING EDUCATION
ASEE NE Conference, Fairfield CT, April 19-20, 2024

Fairfield University
School of Engineering & Computing

INNOVATIVE RESEARCH SYMPOSIUM
Fairfield CT, April 25th, 2024