Interactive English Learning Chatbot: An LLM-Based Teacher For Young Learners

Presented by Cheung Tuen King
Advised by Dr. Cecia KI CHAN

THE DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

OVERVIEW

Foreign language learning has been increasingly recognized as a vital element for children cognitive and social growth, yet non-native learners face challenges and traditional education often prove inefficient.

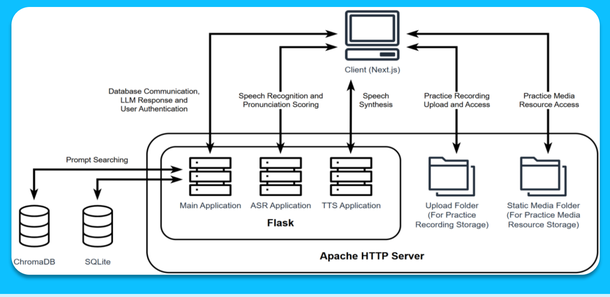
As such, this project, supported by AiTalk IT Limited the tech startup, aims to develop an AI-driven educational chatbot to specifically aid the English learning of the Hong Kong children.

OBJECTIVES

- Build a chatbot to recognize Cantonese and English speech, output dynamic, simple and age-appropriate content, and deliver natural speech output within acceptable time range
- Develop a pronunciation and reading and comprehension practice module
- Integrate free services and a cheap deployment device for cost-efficiency
- Design child-oriented user interface

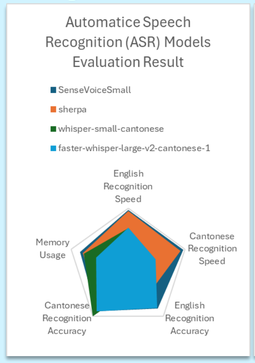
SYSTEM ARCHITECTURE

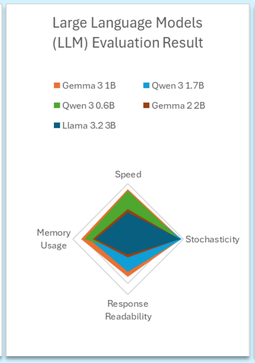
The application was developed using **Flask**, **Next.js, ChromaDB, SQLite & Apache HTTP Server.**

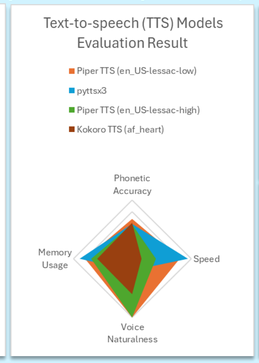


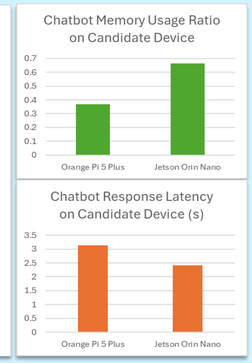
MODEL EVALUATION

Initially, models were researched and evaluated on various attributes to suit our objectives. Based on the results, **SenseVoiceSmall, Gemma 3 1B & Piper TTS** were considered. A chatbot prototype was built to assess their combined runtime performance on candidate device. Finally, **Orange Pi 5 Plus** was selected due to its relatively low memory consumption ratio.

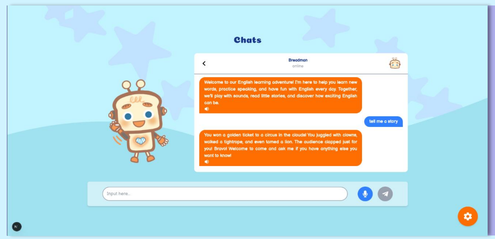




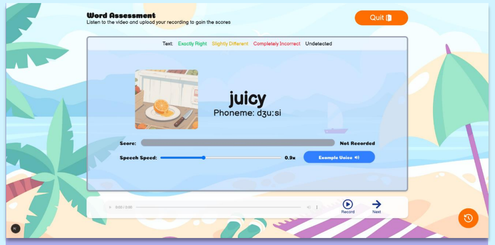




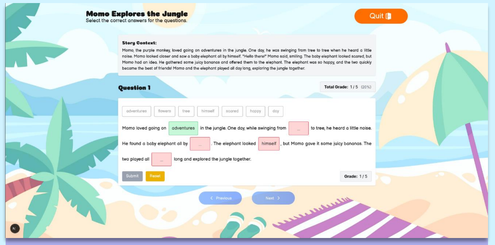
KEY FEATURES



Chatbot Interaction
Users could provide prompts via text input or audio recording. For recording input, speech is detected by VAD, which then pass to ASR-LLM-TTS pipeline for responding. The returned LLM response and synthesized speech are immediately rendered on the UI and schedule for playback, respectively.

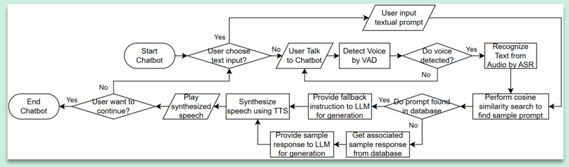


Pronunciation Practice
Word-, sentence- and story-level practice are built for progressive learning. A sample pronunciation was given for user reference. During recording, word color and score are updated continuously.

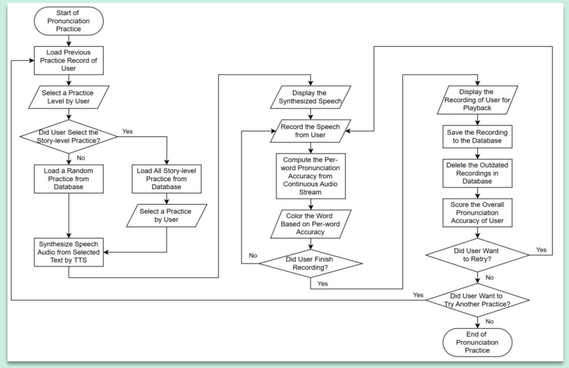


Reading and Comprehension Practice
MC, fill-in-the-blanks and short questions are developed. For short questions, answer is graded by LLM, while others are scored by direct comparison. Users should select the story for reading and training in prior.

WORKFLOW DESIGN



Chatbot Interaction Flowchart



Pronunciation Practice Flowchart

CONCLUSION

We have achieved most of the objectives by:

- Systematically evaluating the candidate models using comprehensive metrics at the early stage.
- Integrating cost-efficient devices and free models and libraries with favorable runtime performance.
- Applying child-oriented design principles derived from established academic literature, like scrolling prevention and immediate feedback.

However, certain issue was also found during testing, requiring following improvement in the future:

- Expand the RAG database to address the irrelevant chatbot responses due to insufficient samples.
- Refine the pronunciation assessment algorithm to resolve the potential evaluation inaccuracies.