User Manual

Project Overview

The full name of this project is **Virtual Tutor: Personalized AI Teaching Assistant**

Virtual Tutor is an AI teaching assistant designed for the COMP9331/3331 online course, supporting text and voice interaction, and featuring video feedback with synchronized lip movements. It is based on large language models (LLM), retrieval augmented generation (RAG), automatic speech recognition (ASR), Text To Speech (TTS), and lip-sync technology, providing personalized teaching support anytime and anywhere.

Scope

The project will provide a customizable AI teaching assistant software driven by RAG and multimodal interaction to provide students with personalized and adaptive learning support. This project will start on June 21, 2025, and end on August 7, 2025. The software will include student interface supports interaction with the virtual tutor (AI avatar) via text and voice; admin interface supports uploading course materials and managing virtual tutors; LLM model that integrates RAG functions; interactive module that supports TTS, ASR & 3D avatar; complete backed module with documentation. This project will only include the development of the software.

- Project Duration: June 21, 2025 - August 7, 2025
- The deliverables include: The student client interface supports text and voice interaction and features the display of a virtual teacher (AI avatar).
    - The administrator interface supports uploading course materials, model management, and tutor image management.
    - Integrated RAG-enabled LLM module, supporting Knowledge Base customization and real-time Q&A.
    - The interactive module supports automatic speech recognition (ASR), Text To Speech (TTS), and 3D lip synchronization.

    - Complete backend module and interface documentation.
- Note: This project only develops the software system and does not currently include the deployment and operation phases.

Quick Start

System Requirements

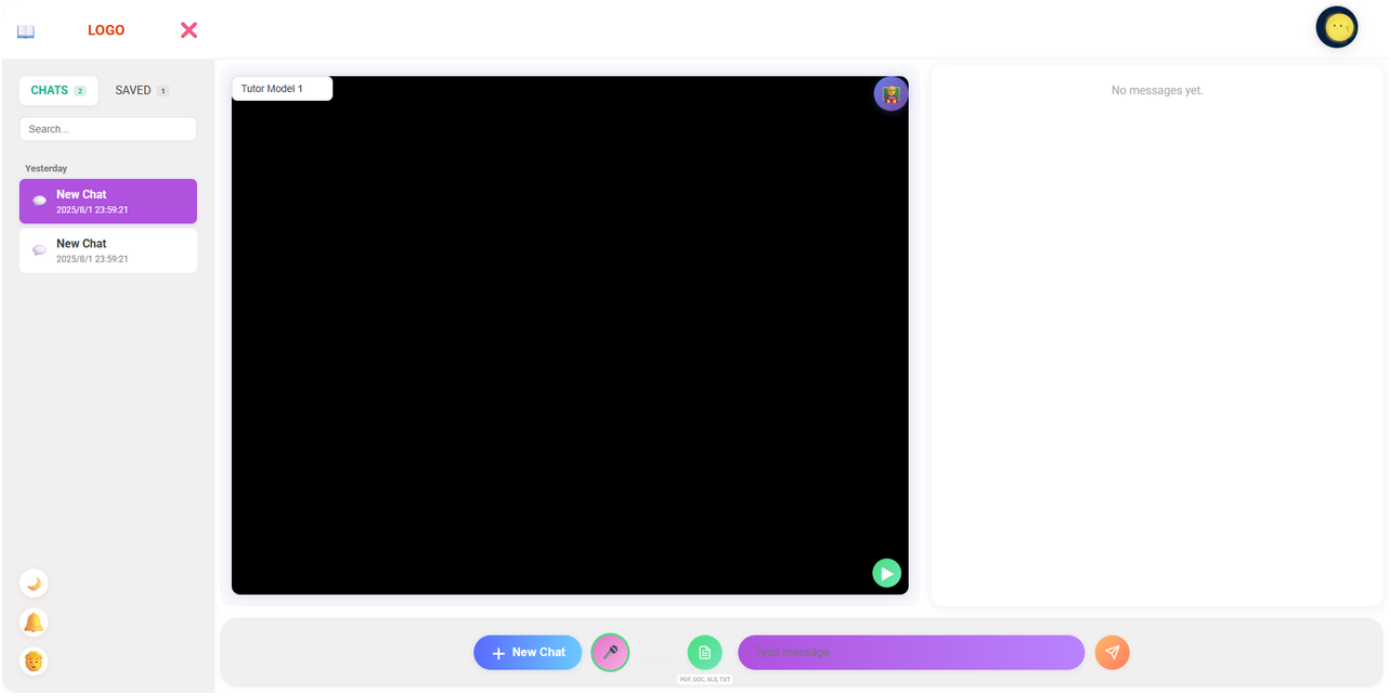
Item	Requirement
Operating System	Any modern browser (Web)
Browser	Chrome / Edge / Firefox (latest version recommended)
Network Connection	Stable network with access to backend services
Others	Microphone (for voice input)

Project Link

Open the browser and visit: Open

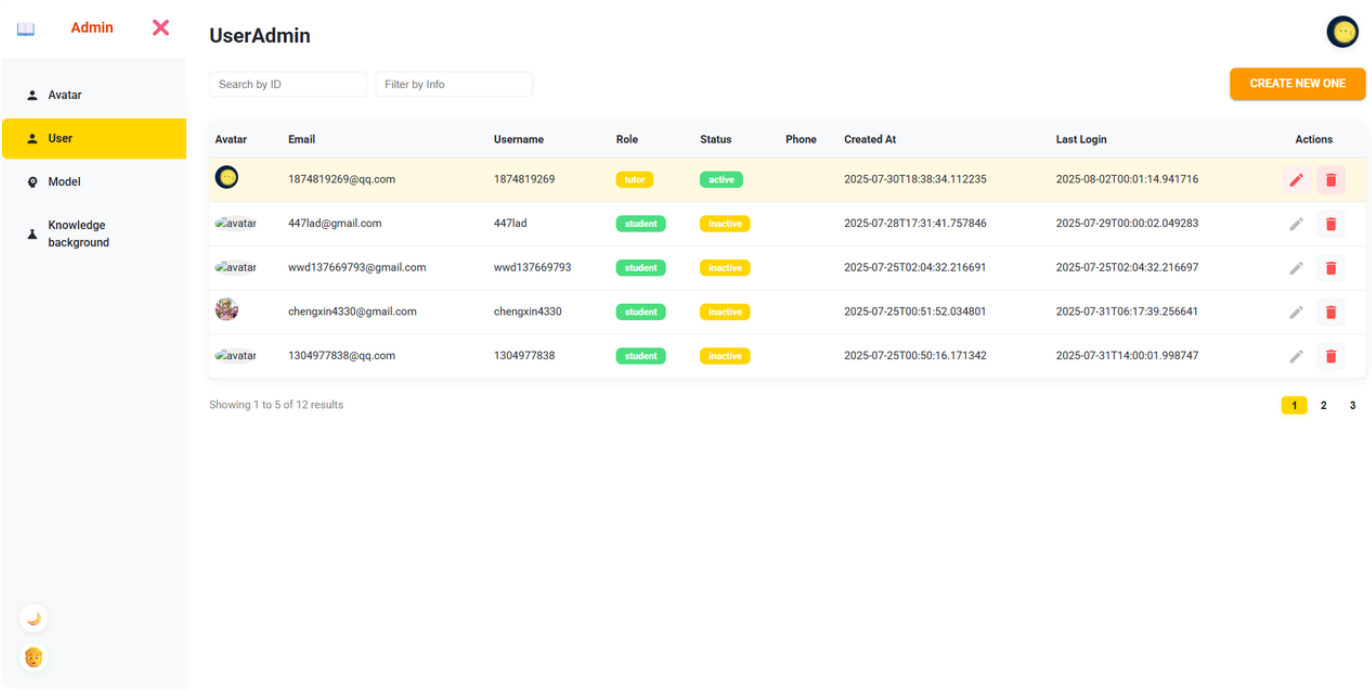
UI Overview

1. User Homepage



Area	Function Description
Input Box	Supports entering course-related questions, commands, or requesting suggestions
Chat Area	Displays the current conversation with the Virtual Tutor
Voice Button	Click to enable speech recognition, convert to text, and auto-submit
Video Area	Shows AI virtual tutor's synchronized lip-sync video feedback; supports tutor selection
Chat History Window	Browse past conversations; pin important questions for future reference
Settings Menu	Language switch, mode switch (voice/text), user info modification, etc.

2. Administrator Homepage



Area	Function Description
User Management	Manage all user data
Knowledge Base Management	Edit the RAG background knowledge base
Model Management	Manage or switch TTS/ASR/Lip-sync/LLMs models
Tutor Avatar Management	View, create, or delete currently supported tutor avatars

How to Use

Registration Process

1. Enter the registration/login interface and click "Register a New Account"
2. Enter the email address that can receive emails, and click Send Verification Code
3. Check your email for the verification code and fill it in
4. Set a password (must include uppercase and lowercase letters, numbers, symbols, and be at least eight characters long)
5. Click the "Register" button to complete registration and automatically jump to the login page

Login Process

- Ordinary User:
  - Enter your email and password
  - Log in and then enter the student homepage
- Administrator:
  - Enter your email and password
  - Log in and then enter the management interface

Text Interaction

- Enter a question in the text box, e.g., "Explain TCP congestion control"
- Click the send button or press Enter
- The system returns a text response and simultaneously displays the lip-sync video (if enabled)

Voice interaction (optional)

- Click the microphone icon and start speaking
- Voice is automatically converted to text, and the system responds
- Answers are presented in the form of text + video (AI lip-sync)

Video Output (Optional)

- After enabling video mode, Tutor's responses will be accompanied by lip-sync animation videos
- The video will read aloud the response content in the right chat in real time
- Click the drop-down menu in the upper left corner of the page to change the image of different AI teachers (the first load may be slow)

Example Cases

Example 1: Text Q&A

User input: Explain TCP 3-way handshake.

System response: The TCP 3-way handshake involves SYN, SYN-ACK, and ACK... (with synchronized video)

Example 2: Voice Q&A

User voice: What is packet loss?

System response: Packet loss occurs when packets are dropped in transit... (Lip-sync video)

Example 3: Contextual Continuous Questioning

User: What is the purpose of sliding window in TCP?

System response: The sliding window allows multiple packets to be sent...

Frequently Asked Questions (FAQ)

Q: The microphone is not working?

- Please ensure that the browser authorizes microphone access and check the system settings.

Q: Is the answer inaccurate or irrelevant?

- This system is only based on the imported course Knowledge Base, and if relevant content is not covered, it will affect accuracy.

Q: Why is the video out of sync with the audio?

- Sometimes voice generation and animation rendering are out of sync. Refreshing the page or waiting a few seconds can resolve this issue.

Q: How to switch between voice and text response?

- Click the play button ▶️ at the bottom right of the virtual tutor to start or stop the video response.

Limitations and Precautions (Known Limitations)

1. This system only covers content related to the COMP9331/3331 course and is not applicable to general Q&A;
2. Automatic speech recognition (ASR) may be affected by environmental noise, so it is recommended to use it in a quiet environment;
3. The response time of the first lip-sync video is slightly longer, which is a normal phenomenon during preheating and loading;
4. Disabling video output does not affect basic learning and Q&A functions;
5. The current version does not support the mobile end App (only available for use in web browsers).

Feedback & Support

- Project Team Email:
- GitHub Issues (e.g., open source)
- Submit the feedback form:

Appendix

Demo Video

- See the demonstration video of the final presentation for details.