Topic Description

The project focuses on creating a Smart Voice Assistant, integrating Automatic Speech Recognition (ASR), Large Language Models (LLMs), and Text-to-Speech (TTS) systems. The project includes:

1. ASR Techniques (e.g., using Whisper or similar models).
2. API integration of LLMs for conversational AI (e.g., OpenAI GPT-4, DeepSeek).
3. TTS Techniques (e.g., using Google Cloud TTS or ElevenLabs).
4. A demonstration of the voice assistant with integrated components in your target domain.

Grading Details

The grading is divided into three categories:

1. **Basic Functionality (60%)**:
    - ASR, LLM, and TTS modules working together.
2. **Extended Functionality (20%)**:
    - Useful features, novel design, better user interface, low latency.
3. **Presentation (20%)**:
    - Clear, fluent presentation with a demonstration of the system.

Report Requirements

- A **final report** may contain:
    - A description of the system architecture.
    - Explanation of how ASR, LLM, and TTS are integrated.
    - Detailed technical approach and challenges faced.
    - The introduction and adaptation of the system to your target domain.
    - Code comments and explanations.
    - Specific contributions of each team member in the project.
    - The length of the report is at most **4 pages**.

Presentation

- Present your system with a clear, structured explanation.
- Demonstrate the running voice assistant during the presentation.
- Prepare a presentation for about **8 mins** per group.
- The presentation will held in the lectures of the last week.

Code Submission

- Submit all **source code** of the project.

- Ensure the code is well-organized, commented, and executable.

Group Requirements

- The project can have **at most 3 people** in the group.
- The final report must also detail the **contributions of each student** to the project.
- Only one of the group members needs to submit the required files.

Please ensure that all components of the project, from coding to documentation, meet these requirements.