Lyrebird AI Engineer Skill Assessment Challenge

Challenge Description

Create a simplified consultation note-taking web application inspired by Lyrebird, focusing primarily on Large Language Model (LLM) integration and experimentation. We encourage you to try our trial at Lyrebirdhealth.com for context.

Objectives

Build a backend-focused solution that enables:

- Audio recording of a consultation session (you may simulate or stub this part).
- Generating a consultation note after the recording finishes using a real LLM service.

Important: You **must** integrate a real LLM service (e.g., OpenAI GPT, Anthropic) for note generation and implement an **LLM experimentation framework** to support prompt tuning, evaluation, and iterative improvement.

Focus Areas

- **LLM Service Integration:** Connect to an external LLM API to generate consultation notes from input data.
- **Prompt Engineering & Experimentation:** Design and refine prompts using an experimentation framework or methodology to improve output quality.
- **User Feedback Loop:** Implement a backend mechanism to collect feedback on generated notes and use it to guide prompt/model refinements.
- **Backend Reliability & Scalability:** Focus on data processing, API reliability, and system robustness.

Note: Front-end implementation can be minimal—basic UI or APII endpoints are sufficient.

Submission Requirements

Before the interview, please provide:

- **Source Code:** Complete backend solution with LLM integration and experimentation capabilities.
- **Setup Instructions:** Clear steps to run the project locally, including environment variables and API key setup.
- **Solution Overview:** Explanation of your architecture, prompt design, experimentation strategy, and feedback loop.

Interview Agenda

During the interview, you will:

- **Present Your Solution:** Focus on the LLM integration, experimentation framework, and backend design.
- **Discuss Decisions & Assumptions:** Explain your prompt engineering and feedback loop choices.
- **Explore Edge Cases:** Address LLM limitations, failure handling, and ambiguous outputs.
- **Production Strategy:** Describe deployment, monitoring, and iterative improvements in production.

Key Outcomes

This challenge assesses your ability to:

- Integrate and optimize real LLM services in backend workflows.
- Apply systematic prompt experimentation and refinement.
- Design feedback-driven improvements for AI-generated content.
- Build reliable, maintainable backend systems.
- Document any assumptions clearly.
- Contact your POC if you have questions during the challenge.