

ABSTRAK

Kesiapan menghadapi wawancara kerja menjadi salah satu tantangan yang masih banyak dialami oleh pencari kerja, terutama dalam menyampaikan jawaban secara jelas, percaya diri, dan sesuai dengan konteks pertanyaan. Penelitian ini mengembangkan platform pelatihan wawancara berbasis *website* yang dapat digunakan pengguna untuk berlatih secara mandiri dan memperoleh umpan balik secara langsung. Sistem ini menerapkan Large Language Model (LLM), Facial Expression Model (FEM), dan OpenAI Whisper untuk menghasilkan umpan balik multimodal yang mencakup aspek verbal dan nonverbal. OpenAI Whisper digunakan untuk mengubah jawaban lisan pengguna menjadi teks, FEM digunakan untuk menganalisis ekspresi wajah selama sesi latihan, sedangkan LLM digunakan untuk menghasilkan pertanyaan wawancara serta mengevaluasi kualitas jawaban pengguna. Platform yang dikembangkan memiliki fitur utama berupa autentikasi pengguna, pengelolaan profil, pembuatan sesi latihan, perekaman jawaban, tampilan hasil evaluasi, dan riwayat sesi. Hasil pengujian menunjukkan bahwa seluruh 15 skenario Black-Box berhasil dijalankan. Model ASR mampu memproses 687 data audio dengan nilai WER cleaned sebesar 0,0865, CER cleaned sebesar 0,0306, dan rata-rata latency 0,4262 detik per file. Model FEM memperoleh akurasi sebesar 93,2%, sedangkan pengujian LLM menunjukkan bahwa DeepSeek V4 Flash lebih sesuai untuk pembuatan pertanyaan dan Gemini 2.5 Flash lebih sesuai untuk evaluasi jawaban. Hasil User Acceptance Testing juga menunjukkan bahwa sistem berada pada kategori sangat baik, sehingga layak digunakan sebagai media latihan wawancara berbasis *website*.

Kata Kunci: Pelatihan Wawancara, Umpan Balik Multimodal, Large Language Model, OpenAI Whisper, Facial Expression Model

ABSTRACT

Interview preparation remains a common challenge for job seekers, particularly in delivering answers clearly, confidently, and in accordance with the given questions. This study develops a web-based interview training platform that allows users to practice independently and receive direct feedback after each session. The system applies a Large Language Model (LLM), Facial Expression Model (FEM), and OpenAI Whisper to provide multimodal feedback covering both verbal and nonverbal aspects. OpenAI Whisper is used to transcribe users' spoken answers into text, FEM is used to analyze facial expressions during the interview session, while the LLM is used to generate interview questions and evaluate the quality of users' answers. The developed platform includes several main features, such as user authentication, profile management, interview session creation, answer recording, evaluation result display, and session history. The testing results show that all 15 Black-Box testing scenarios were successfully executed. The ASR model processed 687 audio data with a cleaned WER of 0.0865, a cleaned CER of 0.0306, and an average latency of 0.4262 seconds per file. The FEM model achieved an accuracy of 93.2%, while the LLM evaluation showed that DeepSeek V4 Flash was more suitable for question generation and Gemini 2.5 Flash was more suitable for answer evaluation. The User Acceptance Testing results also indicate that the system falls into the very good category, making it suitable to be used as a web-based interview training medium.

Keywords: Interview Training, Multimodal Feedback, Large Language Model, OpenAI Whisper, Facial Expression Model