

BAB 1 PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi pengenalan suara (*Speech Recognition*) telah mendorong lahirnya berbagai penelitian yang memanfaatkan teknologi ini untuk otomasi dan efisiensi kerja. Setia Buana (2020) mengembangkan aplikasi *speech-to-text* berbasis Python untuk membantu wartawan mentranskripsikan wawancara secara otomatis, dan berhasil mencapai akurasi 94,75% menggunakan modul *speech recognition* dan algoritma FastICA [7]. Penelitian ini membuktikan bahwa teknologi STT mampu bekerja dengan baik pada kondisi audio yang bervariasi, namun masih terbatas pada penggunaan desktop, tidak mendukung interaksi real-time, dan belum terintegrasi dengan platform interaktif.

Perkembangan selanjutnya dilakukan oleh Rifaldi & Cahyono (2024) yang mengintegrasikan model Whisper AI dari OpenAI ke dalam aplikasi notulensi otomatis bernama SPERCO, dilengkapi fitur *Speaker Diarization* menggunakan pustaka Pyannote. Penelitian ini menunjukkan bahwa Whisper AI memiliki performa transkripsi multibahasa yang tinggi bahkan dalam lingkungan berkebisingan [8]. Meskipun demikian, penerapan Whisper AI dalam penelitian tersebut masih terbatas pada konteks pencatatan rapat dan belum dikembangkan untuk kebutuhan layanan informasi akademik berbasis interaksi suara secara real-time di lingkungan kampus.

Dari kedua penelitian tersebut teridentifikasi celah yang belum dijawab, yaitu belum adanya sistem yang mengintegrasikan model Whisper AI ke dalam platform interaktif berbasis Unity untuk kebutuhan layanan informasi akademik mahasiswa. Di sisi lain, kebutuhan akan akses informasi akademik yang cepat dan mudah di Program Studi Informatika Universitas Muhammadiyah Malang (UMM) semakin mendesak. Informasi seperti jadwal kuliah, lokasi ruang kelas, dosen pengampu, hingga prosedur administrasi akademik terkadang tidak mudah diakses secara cepat oleh mahasiswa [3]. Meskipun pengembangan virtual assistant di perguruan tinggi Indonesia telah banyak dilakukan, implementasinya masih didominasi pendekatan berbasis teks seperti chatbot dan belum mengintegrasikan teknologi *Speech-to-Text* sebagai antarmuka interaksi utama [35][36].

Berdasarkan hal tersebut, penelitian ini hadir sebagai upaya penyempurnaan dari penelitian-penelitian sebelumnya dengan mengintegrasikan model Whisper AI ke dalam asisten virtual berbasis Unity yang dirancang khusus untuk memberikan layanan informasi akademik. Model ini mampu mentranskripsikan ucapan dalam berbagai bahasa dan dialek termasuk Bahasa Indonesia, sehingga mahasiswa cukup menyampaikan pertanyaan secara lisan dan sistem akan memberikan jawaban berdasarkan data akademik yang telah ditentukan [1][4]. Penggunaan Unity sebagai platform pengembangan memungkinkan pembuatan antarmuka yang interaktif dan dinamis, sangat cocok untuk integrasi dengan AI dan pengenalan suara [5]. Selain itu, pendekatan pengembangan yang digunakan adalah model Spiral karena mendukung pengembangan sistem secara bertahap dan iteratif dengan penekanan pada analisis dan mitigasi risiko pada setiap siklus [6].

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah diatas, penulis merumuskan masalah sebagai berikut:

- a. Bagaimana cara mengintegrasikan model Whisper AI sebagai sistem pengenalan suara (Speech-to-Text) dalam aplikasi Unity?
- b. Bagaimana proses integrasi Whisper AI dapat mendukung asisten virtual dalam menjawab pertanyaan seputar informasi akademik Prodi Informatika UMM?
- c. Seberapa efektif sistem ini dalam mengenali dan mentranskripsi ucapan pengguna?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dijabarkan diatas, penelitian ini bertujuan untuk:

- a. Mengembangkan aplikasi asisten virtual berbasis Unity yang mampu menerima input suara dari pengguna.
- b. Mengintegrasikan model Whisper AI sebagai komponen Speech-to-Text dalam aplikasi virtual assistant berbasis Unity.
- c. Memberikan kemudahan akses informasi akademik kepada mahasiswa melalui interaksi berbasis suara.

1.4 Batasan Masalah

- a. Aplikasi dibangun menggunakan Unity dan dijalankan pada platform desktop (Windows).
- b. Model Speech-to-Text yang digunakan adalah Whisper AI dari OpenAI dengan versi open-source.
- c. Dataset percakapan terbatas pada informasi akademik Prodi Informatika UMM, seperti jadwal kuliah, dosen pengampu, dan lokasi ruangan.
- d. Sistem hanya mendukung bahasa Indonesia sebagai input suara.