



Evaluasi Performansi Transformer Whisper pada *Dataset* Bahasa Indonesia

Ratna Atika,^{*a,b} Suci Dwijayanti,^c Bhakti Yudho Suprpto^c

^a *Jurusan Teknik Elektro Program Studi Sarjana Terapan Teknik Elektro, Politeknik Negeri Sriwijaya, Palembang, Sumatera Selatan, Indonesia;* ^b *Program Doktor Ilmu Teknik, Fakultas Teknik, Universitas Sriwijaya, Palembang, Sumatera Selatan, Indonesia;* ^c *Jurusan Teknik Elektro, Fakultas Teknik, Universitas Sriwijaya, Palembang, Sumatera Selatan, Indonesia.*

*E-mail: ratna.atika@polsri.ac.id

ABSTRAK

Untuk mendukung interaksi manusia-robot secara alami, pengembangan teknologi robot memerlukan kemampuan pemrosesan suara yang handal. Salah satu model mutakhir dalam *automatic speech recognition* (ASR) adalah Whisper, sebuah arsitektur Transformer yang dikembangkan oleh OpenAI. Namun, performansi Whisper untuk bahasa Indonesia, khususnya dalam konteks aplikasi robotika, masih terbatas untuk dievaluasi secara mendalam. Penelitian ini bertujuan untuk mengevaluasi performansi model Whisper dalam mengenali ucapan berbahasa Indonesia dengan memanfaatkan *dataset* Mozilla model Common Voice Corpus 20.0. Metode yang digunakan meliputi pengujian model Whisper terhadap data audio berbahasa Indonesia dengan variasi panjang rekaman, aksent, serta kebisingan latar. Evaluasi dilakukan dengan mengukur *word error rate* (WER) dan *character error rate* (CER) sebagai indikator akurasi. Hasil pengujian menunjukkan bahwa Whisper mampu mencapai rata-rata WER sebesar 14,2% dan CER sebesar 13,9% pada kondisi audio bersih. Hasil ini menunjukkan bahwa whisper memiliki potensi tinggi untuk diimplementasikan pada sistem robot berbahasa Indonesia, meskipun perlu dilakukan adaptasi lebih lanjut terhadap kondisi *realtime* yang bervariasi. Hasil ini memberikan kontribusi penting dalam pengembangan interaksi manusia dengan robot yang mendukung komunikasi berbasis suara dalam bahasa Indonesia.

Kata Kunci: Whisper, Transformer, *Speech Recognition*, Bahasa Indonesia, Robot