

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan sistem pembelajaran pelafalan Bahasa Inggris yang efektif menggunakan teknologi Speech Recognition berbasis model Transformer Wav2Vec 2.0. Masalah utama yang diangkat adalah rendahnya kemampuan pelafalan di Indonesia, yang terbukti dari data EF EPI 2024. Kami menggunakan metodologi CRISP-DM untuk memastikan proses penelitian berjalan sistematis. Model Wav2Vec 2.0 yang sudah dilatih sebelumnya (pre-trained) di fine-tuning menggunakan dataset LibriSpeech. Hasil evaluasi menunjukkan model kami memiliki kinerja unggul dengan nilai Word Error Rate (WER) sebesar 0.042778. Angka ini jauh lebih baik dibandingkan Base Model ASR yang hanya mencapai WER 0.070083. Model ini diimplementasikan dalam aplikasi web "Wav2Speech" yang memungkinkan pengguna mendapatkan umpan balik langsung untuk meningkatkan keterampilan pelafalan mereka secara mandiri.

Kata Kunci: Pengenalan Ucapan, Wav2Vec 2.0, Fine-tuning, Pembelajaran Pelafalan Bahasa Inggris.



ABSTRACT

This study aims to develop an effective English pronunciation learning system using Speech Recognition technology based on the Transformer Wav2Vec 2.0 model. The main problem addressed is the low English pronunciation ability in Indonesia, as evidenced by EF EPI 2024 data. We used the CRISP-DM methodology to ensure a systematic research process. The pre-trained Wav2Vec 2.0 model was fine-tuned using the LibriSpeech dataset. The evaluation results show that our model has superior performance with a Word Error Rate (WER) of 0.042778. This figure is much better than the Base ASR Model, which only reached a WER of 0.070083. The model was implemented in a web application "Wav2Speech" that allows users to get direct feedback to improve their English pronunciation skills independently.

Keywords: Speech Recognition, Wav2Vec 2.0, Fine-tuning, English Pronunciation Learning.

