

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. Metodologi CRISP-DM digunakan 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 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 issue addressed is the low pronunciation ability in Indonesia, as evidenced by the EF EPI 2024 data. The CRISP-DM methodology was used to ensure a systematic research process. The pre-trained Wav2Vec 2.0 model was fine-tuned using the LibriSpeech dataset. The evaluation results showed that the model has superior performance with a Word Error Rate (WER) value of 0.042778. This figure is significantly better than the Base ASR Model, which only achieved a WER of 0.070083. This model was implemented in a web application called "Wav2Speech", which allows users to receive instant feedback to improve their pronunciation skills independently.

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

