Improving Mispronunciation Detection and Diagnosis for Non-native Learners of the Arabic Language

Nina Hosseini-Kivanani, Norah Alrashoudi, & Aswin Giridhar Subramanian
Mentor: Stan Salvador

The mispronunciation detection and diagnosis (MDD) plays a vital role in computer-assisted pronunciation training (CAPT), offering second language learners (L2) an opportunity to enhance their speaking abilities. Arabic, a language spoken by over 422 million people, holds immense significance as the language of The Holy Quran. While most current Arabic MDD systems utilize learning-based techniques, this study proposes the development of a deep learning-based MDD model for non-native Arabic speakers. This will help improve their language learning and pronunciation skills.

The proposed model integrates both acoustic features extracted from speech data and linguistic features to identify mispronounced phonemes. The study focuses on specific Arabic phonemes that non-native speakers often mispronounce, such as /x/, /ɣ/, /q/, /ʕ/, and /ħ/. A dataset was created, containing speech data from native and non-native Arabic speakers, each reading ten sentences containing one or more of the selected phonemes.

The performance of the proposed model was evaluated using the word error rate (WER) and character error rate (CER). The model was trained using Wav2Vec from HuggingFace and achieved a 80% CER. However, the model struggled to recognize phonemes accurately based on the obtained WER, possibly due to the small size of the dataset or the different transcription used.

Table 1: Dataset Description

Information	Used part
Language	Arabic
Speakers	10 African L1 (learn Arabic)
Level	Intermediate learner
Data	Read sentences
Hours	150 utterances, 19 minutes
Age	Adult
Annotation	Word & phoneme level
Mispronunciation	Deletion/addition/substitution
Labeling	Al-Tamimi Romanization of Arabic Orthography (ATR)