

Data-Driven Pronunciation Generation for ASR

Maria Obedkova

In ASR systems, dictionaries are usually used to describe pronunciations of words in a language. These dictionaries are typically hand-crafted by linguists. One of the most significant drawbacks of dictionaries created this way is that linguistically motivated pronunciations are not necessarily the optimal ones for ASR. The goal of this research was to explore approaches of data-driven pronunciation generation for ASR. We investigated several approaches of lexicon generation and implemented the completely new data-driven solution based on the pronunciation clustering. We proposed an approach for feature extraction and researched different unsupervised methods for pronunciation clustering. We evaluated the proposed approach and compared it with the current hand-crafted dictionary. The proposed data-driven approach could beat the established baselines but underperformed in comparison to the hand-crafted dictionary which could be due to unsatisfactory features extracted from data or insufficient fine tuning.