

ABSTRACT

Gender classification based on voice is crucial for speech recognition, which can be applied to various applications. It is generally developed using conventional machine learning and deep learning approaches. In this research, a gender classification model based on speech is developed using Bidirectional Long Short-Term Memory (BLSTM). The Mel Frequency Cepstral Coefficient (MFCC) is exploited to extract the features to train the BLSTM. Evaluation using a low dataset of 1,000 utterances, 500 males and 500 females, for five runs shows that the model is accurately capable of classifying the gender of the speakers. With a train-test split portion of 80:20, the model obtains an average accuracy of 86.7%, where the highest and the lowest accuracy are 90.5% and 81.0%, respectively. Reducing the portion decreases its performance. It is still stable for the 50:50 train-test split.

Keywords: bidirectional long short term memory, deep learning, gender classification, Mel frequency cepstral coefficient, recurrent neural network.