

Abstract

Speech recognition is one of identification processes that are being developed. Gender classification based on speech affects the performance of speech recognition. Thus, in this research, we build a voice-based gender classification system. We use 2264 voices as dataset, which are split into two classes, i.e. "male" and "female". For feature extraction, we use Mel-frequency Cepstral (MFCC) to generate feature vector from the voice signal. In the classification process, we use Support Vector Machine (SVM) method. The highest SVM accuracy for gender-based voice classification is 100% using Polynomial kernel of degree 1.

Keywords: Gender classification, speech recognition, MFCC, SVM