Public Abstract

First Name:Zeshan
Middle Name:
Last Name:Peng
Adviser's First Name:Yi
Adviser's Last Name:Shang
Co-Adviser's First Name:
Co-Adviser's Last Name:
Graduation Term:FS 2017
Department:Computer Science
Degree:MS
Title:ACOUSTIC FEATURE-BASED SENTIMENT ANALYSIS OF CALL CENTER DATA

With the advancement of machine learning methods, audio sentiment analysis has become an active research area in recent years. For example, business organizations are interested in persuasion tactics from vocal cues and acoustic measures in speech. A typical approach is to find a set of acoustic features from audio data that can indicate or predict a customer’s attitude, opinion, or emotion state. For audio signals, acoustic features have been widely used in many machine learning applications, such as music classification, language recognition, emotion recognition, and so on. For emotion recognition, previous work shows that pitch and speech rate features are important features. This thesis work focuses on determining sentiment from call center audio records, each containing a conversation between a sales representative and a customer. The sentiment of an audio record is considered positive if the conversation ended with an appointment being made, and is negative otherwise. In this project, a data processing and machine learning pipeline for this problem has been developed. It consists of three major steps: 1) an audio record is split into segments by speaker turns; 2) acoustic features are extracted from each segment; and 3) classification models are trained on the acoustic features to predict sentiment. Different set of features have been used and different machine learning methods, including classical machine learning algorithms and deep neural networks, have been implemented in the pipeline. In our deep neural network method, the feature vectors of audio segments are stacked in temporal order into a feature matrix, which is fed into deep convolution neural networks as input. Experimental results based on real data shows that acoustic features, such as Mel frequency cepstral coefficients, timbre and Chroma features, are good indicators for sentiment. Temporal information in an audio record can be captured by deep convolutional neural networks for improved prediction accuracy.