**Spot Emotions in human speech using SVM, LSVM and RF method**

Yuan Tian, Xin Li, Lin Qiao Shang, Wang Ming Situ

***Abstract***

1. **Objective of Project:**

In this project, our aim is to use support vector machine(SVM), Least Square Support Vector Machine Classifier, Random Forest to classify different speech samples by emotional status.

1. **Tools and Approach**
   1. **Language and platform**:

MATLAB, Python.

* 1. **Training Data Source:**

The Interactive Emotional Dyadic Motion Capture (IEMOCAP) Database from University of Southern California. (http://sail.usc.edu/iemocap/)1

* 1. **Software Tools and Libraries:**

MATLAB machine learning tool box.

Tensor Flow

* 1. **Approach:**

We designed a two-step experiment:

Firstly, we use basic SVM and Least Squares SVM classifier, and compared their performance. In this step, our reference papers are: *A Tutorial on Support Vector Machines for Patten recognition.*2 *Least Square Support Vector Machine Classifier.*3

Secondly, we use Random Forest model. Our reference papers are: *Divide-and-Conquer based Ensemble to Spot Emotions in Speech using MFCC and Random Forest.*4

We plot the output data by MATLAB2015b, and analyze the performance for each model.

1. **Data Collection and Processing**

Original datasets are voices of five different emotions including angry, sad, happy, neutral and fear. We collected four emotional voices performed by professional actors/actress from The Interactive Emotional Dyadic Motion Capture database(IEMOCAP) and recorded neutral voices from the Voice of America (VOA) announcers. voices for each emotion are about 10 (.mp3) files, proportion of male and female is about 1:1 to avoid the error cause by gender. Language is English.

Using goldware to convert voice samples (.mp3) into digital signal(wave) and then save it into (.txt) file.

Data cleaning include getting rid of noises and blankness (when it’s quiet, voice value = 0) from our voice samples.

We got 50 sets Training samples and 10 sets test samples based on this procedure.

1. **Model design**

**4.1 SVM**

**4.2 LSVM**

**4.3 RF**

1. **Coding and compile**
2. **Output and screenshot**
3. **Performance Comparison and Analysis**

***Reference:***

[1] The Interactive Emotional Dyadic Motion Capture (IEMOCAP) Database [Advertisement]. (2004). Retrieved December 10, 2017, from http://sail.usc.edu/iemocap/

[2] A Tutorial on Support Vector Machines for Patten recognition.

[3] Least Square Support Vector Machine Classifier

[4] Badshah, A. M., Ahmad, J., Lee, M. Y., & Baik, S. W. (2016). Divide-and-Conquer based Ensemble to Spot Emotions in Speech using MFCC and Random Forest. Retrieved November 16, 2017, from https://arxiv.org/abs/1610.01382.

***Contributor:***

[1] Yuan Tian (10422134)

Paper reading, find public database and apply for access, model design and implementation, coding and report writing.

[2] Xin Li (10421856)

Topic come up, related paper finding and reading, data collection including training samples and test samples, data cleaning coding, report writing

[3] Lin-qiao Shang (10421310)

[4] Wang Ming Situ