

Telephonic call emotion analysis

-Anurag Mishra

audio_emotion_analysis

The objective of this project is to predict the emotion present in any audio file/signal.

Requirements:

1. Anaconda - This installs python along with most popular python libraries including sklearn. If not already installed, install it from <https://www.continuum.io/downloads> .
2. python_speech_features
3. pyaudio

The requirements 2 and 3 can be installed by typing:

```
pip install -r requirements.txt
```

Installation:

Once the requirements are installed, just type `python setup.py install`

Preparing dataset:

- Put all the unlabelled audio files in a folder named `calls`, or any other folder and update the name of folder in `label_dataset.py`.
 - Run the script `label_dataset.py`.
 - It will scan all the audio files, create a set of 30 sec audio chunks.
 - It will play each chunk and then ask for a label (positive/neutral/negative).
 - Enter 1 for negative, 2 for neutral and 3 for positive.
 - Continue until all the chunks are labelled.
 - A new dataset will be prepared in a new folder named `data`, with each audio file of the name in the format `<label>_<counter>.wav`.
-

Scripts description:

feature_extractor.py

Extracts the features from any audio file or signal and returns a feature vector.

emotion_analysis.py

This is the backbone of the project. It contains modules that

- extracts features using the `feature_extractor.py`
- trains the model
- tests the model
- evaluates the model on new dataset

Evaluating a new dataset:

- Put all the new audio files in a folder named `test_calls`, or any other folder and update the name of the folder in `main_script.py`.
- Execute the script `main_script.py` or type

```
python main_script.py
```

- The script will load an existing model or train a new model on the dataset prepared, extract features for each of the new calls, feed it to the trained model.

Results

- For each file, following three things are evaluated:
 - Overall emotion
 - Emotion transition from first half to second half
 - Emotion present in each 20 sec chunk of the file
- These results for each audio file are written to a `.csv` file, on which further analysis can be done.