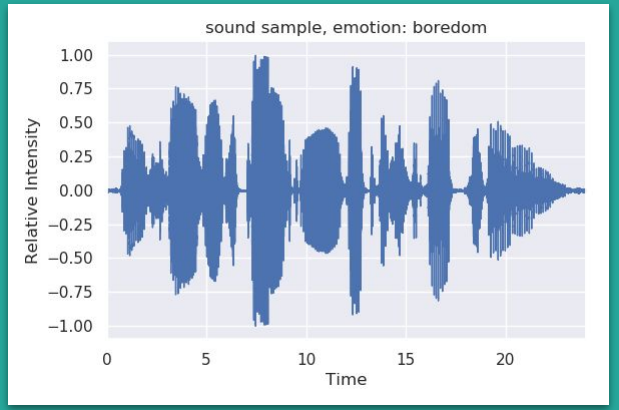
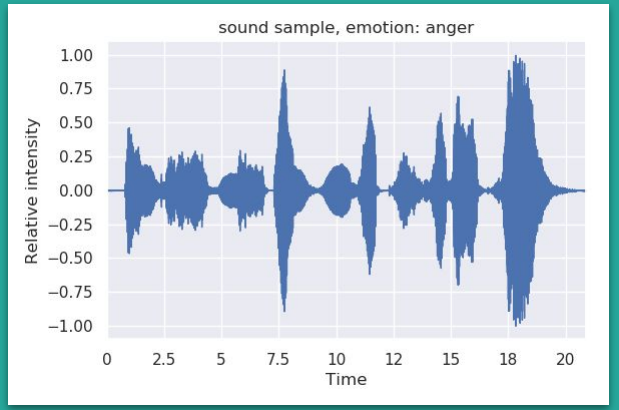
Speech emotion recognition

Charles Dufour

Given an audio recording, can we identify the emotion of the speaker ?



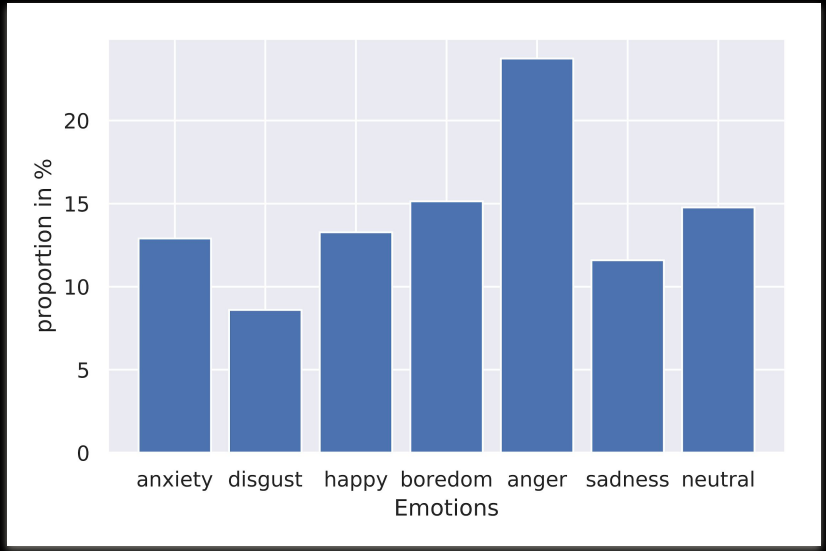


The Data

Source

Emo-DB

- Samples spoken by actors
- 535 audio files, 7 emotions
- Great quality



Building the features

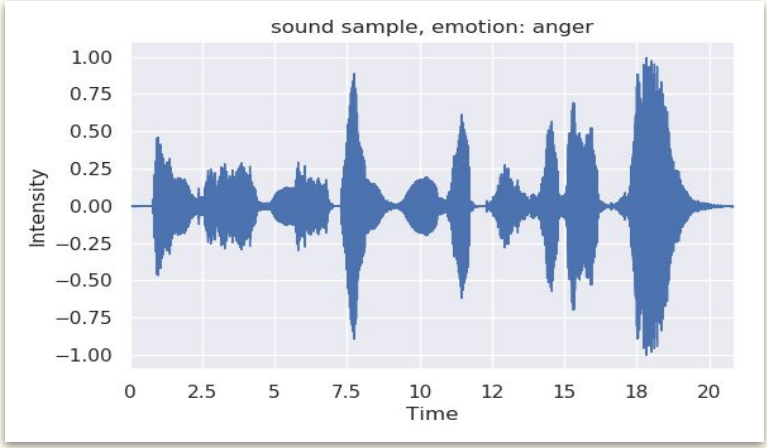
1. Mel-frequency spectrum:

   Mathematical representation of how the human ear perceives the sound.

2. Spectral components (filters, steepness changes,...)

   Only report statistics of these components (mean, standard deviation, minimum, maximum,...)

3. Ratio spoken time vs unspoken time

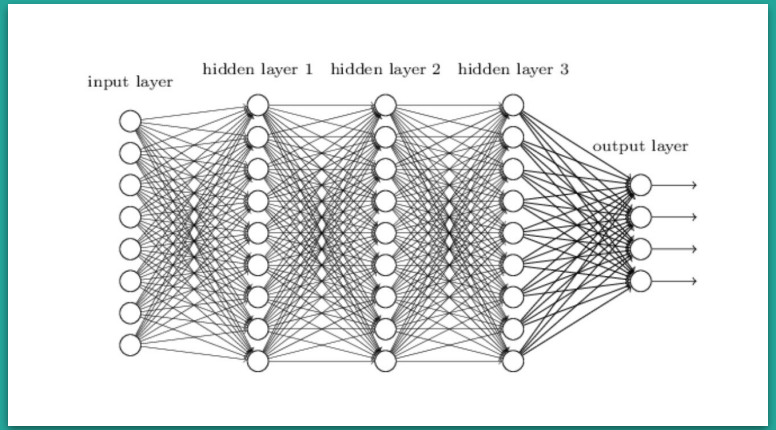


Libraries used:

Librosa, python_speech_features, ...

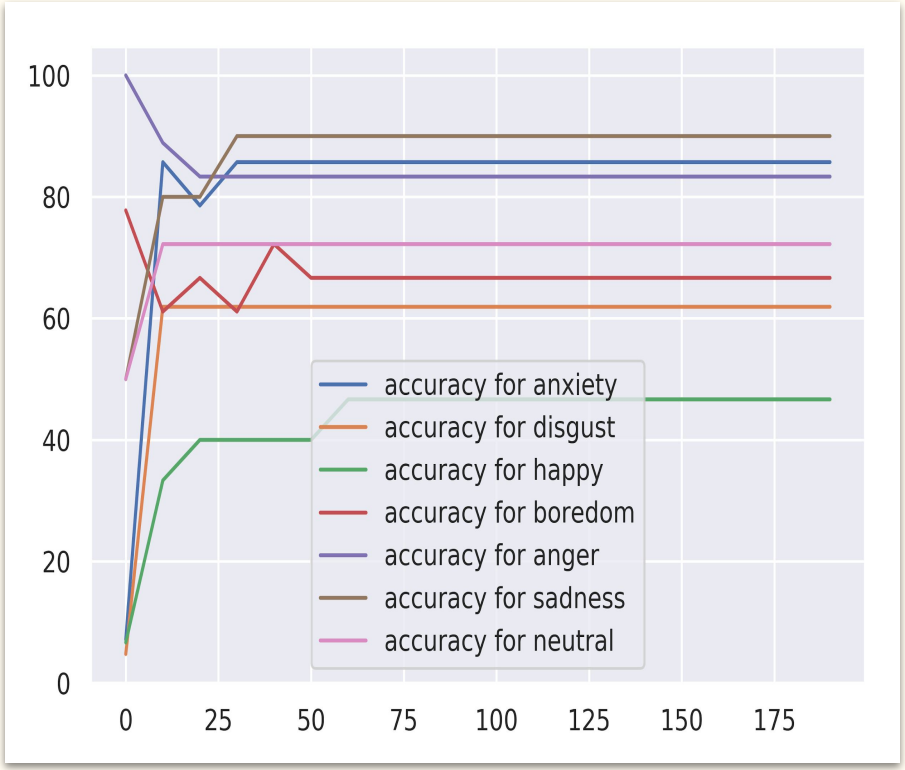
Model

Fully connected neural network with 4 hidden layers and Relu activation



Results

global accuracy vs epochs



accuracy per class vs epochs

Confusion matrix (in %)

Predicted label

True label

	anxiety	disgust	happy	boredom	anger	sadness	neutral
anxiety	84.62	0.0	7.69	0.0	0.0	0.0	7.69
disgust	10.0	60.0	0.0	5.0	10.0	10.0	5.0
happy	14.29	7.14	42.86	0.0	35.71	0.0	0.0
boredom	11.76	0.0	0.0	64.71	0.0	11.76	11.76
anger	5.88	5.88	5.88	0.0	82.35	0.0	0.0
sadness	0.0	0.0	0.0	11.11	0.0	88.89	0.0
neutral	11.76	0.0	0.0	17.65	0.0	0.0	70.59

Sources

- **Emotion recognition from the human voice**. Parlak, Cevahir & Diri, Banu. (2013). *21st Signal Processing and Communications Applications Conference, SIU 2013. 1-4. 0.1109/SIU.2013.6531196*
- **Speech Emotion Recognition: Methods and Cases Study**. Kerkeni, Leila & Serrestou, Youssef & Mbarki, Mohamed & Raoof, Kosai & Mahjoub, Mohamed. (2018). *175-182. 10.5220/0006611601750182*
- **Emotion Recognition from Speech using Discriminative Features**. Chandrasekar, Purnima & Chapaneri, Santosh & Jayaswal, Deepak. (2014). *International Journal of Computer Applications. 101. 31-36. 10.5120/17775-8913*
- Berlin emotional speech database
- Librosa: Audio and music signal analysis in python.McFee, Brian, Colin Raffel, Dawen Liang, Daniel PW Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto
- Hastie, T.; Tibshirani, R. & Friedman, J. (2001), *The Elements of Statistical Learning* , Springer New York Inc. , New York, NY, USA .