I. Music Genre Classification

Goal
Classify music signal into one of the genres.

Reference
Tzanetakis, G., & Cook, P. (2002). Musical genre classification of audio signals. IEEE Transactions on speech and audio processing, 10(5), 293-302.

Data
— Potential Data Source (1000 tracks with 10 genres and 100 tracks per genre)
http://opihi.cs.uvic.ca/sound/genres.tar.gz

— Million song dataset
http://labrosa.ee.columbia.edu/millionsong/

II. Music Speech Discrimination From Short Audio Snippets

Goal
Characterize a given signal as speech or music (can also add other signal types like environmental sounds)

Reference
Chen, Lei, Sule Gunduz, and M. Tamer Ozsu. "Mixed type audio classification with support vector machine." 2006 IEEE International Conference on Multimedia and Expo. IEEE, 2006.

Data
— Potential Data Source (120 tracks of speech and music (each with 30sec))
http://opihi.cs.uvic.ca/sound/music_speech.tar.gz

III. ECG Signal Quality Assessment

Goal
To predict the quality of the measured ECG signal

Reference
Behar, Joachim, et al. "ECG signal quality during arrhythmia and its application to false alarm reduction." IEEE Transactions on Biomedical Engineering 60.6 (2013): 1660-1666.

Data
MIMIC-II Dataset
https://physionet.org/mimic2/

MIT-DIH Arrhythmia Database
https://physionet.org/physiobank/database/mitdb/

IV. EEG Signal Analysis

Goal

Identifying the presence of alcohol in the collected EEG signal - AR modeling/feature selection.

Reference

Faust, Oliver, et al. "Analysis of EEG signals during epileptic and alcoholic states using AR modeling techniques." IRBM 29.1 (2008): 44-52.

Data
Collected from 10 alcoholic and 10 normal subjects.
https://archive.ics.uci.edu/ml/datasets/EEG+Database

V. Dimensionality Reduction of fMRI data

Goal
Reducing the dimensionality of fMRI data for visualization and classification.

Reference
Van Der Maaten, Laurens, Eric Postma, and Jaap Van den Herik. "Dimensionality reduction: a comparative." J Mach Learn Res 10 (2009): 66-71.

Data
6 subjects with manual annotation
http://www.cs.cmu.edu/afs/cs.cmu.edu/project/theo-81/www/

VI. Speaker Diarization

Goal
Segment speech based on speakers

Reference
Vijayasenan, Deepu, Fabio Valente, and Hervé Bourlard. "An information theoretic approach to speaker diarization of meeting data." IEEE transactions on audio, speech, and language processing 17.7 (2009): 1382-1393.

Data
40 speakers with orthographic annotations
http://buckeyecorpus.osu.edu/

VII. Natural and Synthetic Speech Classification

Goal
Classify speech as natural (genuine) or synthetic (machine generated)

Reference
Alam, Md Jahangir, et al. "Spoofing Detection on the ASVspoof2015 Challenge Corpus Employing Deep Neural Networks.

Data
http://datashare.is.ed.ac.uk/handle/10283/782

VIII. Whole Word Detection From Speech Using Recurrent Networks

Goal
Detect the presence of a keyword in speech using recurrent networks

Reference
Wöllmer, Martin, et al. "Bidirectional LSTM networks for context-sensitive keyword detection in a cognitive virtual agent framework." *Cognitive Computation* 2.3 (2010): 180-190.

Data
40 speakers with orthographic annotations
http://buckeyecorpus.osu.edu/

IX. Emotion Detection from Facial Images

Goal
Classifying the facial image as one of 15 emotions

Reference
Azcarate, A., Hageloh, F., van De Sande, K., & Valenti, R. (2005). Automatic facial emotion recognition. University of Amsterdam Technical Report.

Data
11 subjects with 15 emotions
http://vision.ucsd.edu/content/yale-face-database

X. Pose Estimation from Images

Goal
Estimating the pose from image

Reference
Murphy-Chutorian, Erik, and Mohan Manubhai Trivedi. "Head pose estimation in computer vision: A survey." IEEE transactions on pattern analysis and machine intelligence 31.4 (2009): 607-626.

Data
http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html

XI. Recognizing Imagined Speech from Multi-channel EEG.

Goal
Recognizing speech class categories with multi-channel EEG recordings.

Data
http://www.cs.toronto.edu/~complingweb/data/karaOne/karaOne.html

Reference
Shunan Zhao and Frank Rudzicz (2015) Classifying phonological categories in imagined and articulated speech. In Proceedings of ICASSP 2015, Brisbane Australia.

XII. Face Recognition from Limited Data

Goal
Identifying faces from very limited training data

Reference
Lu, Jiwen, Yap-Peng Tan, and Gang Wang. "Discriminative multimanifold analysis for face recognition from a single training sample per person." IEEE transactions on pattern analysis and machine intelligence 35.1 (2013): 39-51.

Data
http://vision.ucsd.edu/content/yale-face-database
http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html