AI-based Audio Analysis of Music and Soundscapes

Research Projects

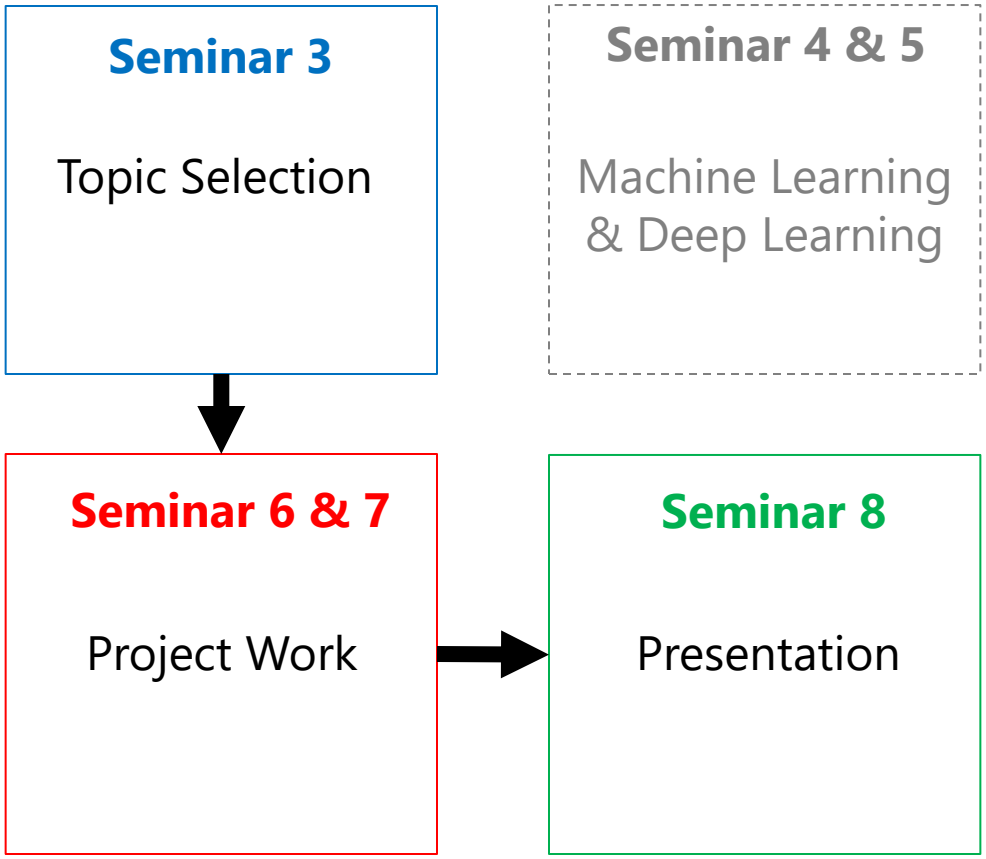
Dr.-Ing. Jakob Abeßer
Fraunhofer IDMT

jakob.abesser@idmt.fraunhofer.de

Outline

- General Comments
- Dataset sources
- Possible Topics

Research Project
Timeline



Research Project
Process

- Form group of 2-3 students & select a research topic
    - Research question?
    - Short literature review
    - How to split the workload?
- Dataset(s)?
- Think about
    - Audio feature representation
    - Modeling approach (machine learning)
    - Evaluation strategy (metrics, dataset split)

Research Project
Audio Datasets

	MIR	Env. Sounds
https://www.audiocontentanalysis.org/datasets.html	✓	
https://ismir.net/resources/datasets/	✓	
https://www.idmt.fraunhofer.de/en/publications/datasets.html	✓	✓
https://zenodo.org	✓	✓
https://homepages.tuni.fi/toni.heittola/datasets		✓
https://towardsdatascience.com/40-open-source-audio-datasets-for-ml-59dc39d48f06	✓	✓

Research Project
Audio

- Get familiar with the audio material (listen to examples)
- Describe the audio
    - What is audible? (isolated sounds / sound mixtures / notes / melodies …)
    - Sample rate, #channels
- How was the audio recorded?
    - Studio vs. field recording
- Under which license was the dataset published?

Research Project
Annotation

- Describe the available annotations
    - Which classes exist?
    - How are they distributed?
    - (Annotate if necessary)
- How many annotators? Which background?
- Does the dataset provide a pre-defined split into training/test sets?
    - If not, how you could create such a split? *(make your research reproducible!)*

Research Project
Task & Modeling

- Which task do you want to solve using audio processing & machine learning?
    - Classification / regression …
- What could be a good (quick to implement) baseline algorithm?
- How can you evaluate the performance of your algorithm?

Research Project
Baseline System

- Baseline system / Processing pipeline
  - Import audio data
  - Import annotations
  - Normalize data
  - Data split (training set / test set)
  - Audio feature extraction
  - Setup modeling algorithm (classifier)
  - Train classifier
  - Evaluate classifier with test set
  - Error analysis

Research Project
Improvements & Documentation

- Improve baseline system
  - Try different feature representations & modeling algorithms
  - Repeat evaluation
- Documentation
  - Short presentation (4-5 slides)
    - Include audio examples & plots
  - (online documentation?)

Research Project
Topic #1: Sound Event Classification

- Dataset
  - ESC-50 datasets (https://github.com/karolpiczak/ESC-50)
- Task
  - Classify isolated sound recordings (5s) into 50 sound classes
- Aspects to look deeper into
  - Compare different spectrogram representations (STFT, Mel Spectrogram etc.)
  - Data augmentation (https://github.com/iver56/audiomentations)

Research Project
Topic #2: Vehicle Type Classification

- Dataset
  - IDMT-TRAFFIC (https://www.idmt.fraunhofer.de/en/publications/datasets/traffic.html)
- Task
  - Vehicle type classification (bus, car, motorcycle, truck)
  - Movement direction estimation (left > right, right > left)
- Aspects to look deeper into
  - Classify between noisy sound classes
  - How to analyze stereo signals (time-of-arrival differences)

Research Project
Topic #3: Bird Activity Detection

- Dataset
  - warblrb10k dataset (https://dcase.community/challenge2018/task-bird-audio-detection) - 2,000 smartphone recordings
- Task
  - Classify a 10s audio recording for bird activity (active / not active)
- Aspects to look deeper into
  - How to deal with large variety of background sounds?
  - Convolutional Neural Networks to learn spectro-temporal patterns (bird vocalizations)

Research Project
Topic #4: Acoustic Scene Classification

- Dataset
  - DCASE-2013-Task1 (https://dcase.community/challenge2013/task-acoustic-scene-classification)
- Task
  - Classify the acoustic scene (10 classes) given a 30s binaural audio recording
- Aspects to look deeper into
  - How to summarize long-term characteristics of audio signals?
  - Convolutional Neural Networks

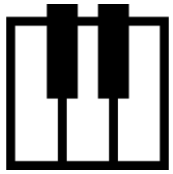
Research Project
Topic #5: Music Genre Classification

- Dataset
  - FMA-small (https://github.com/mdeff/fma) – 8000 30s tracks, 8 genres
- Task
  - Classify the music genre
- Aspects to look deeper into
  - Compare different audio features (rhythm, harmony, timbre)

Research Project
Topic #6: Music Instrument Classification

- Dataset
  - MedleyDB (https://medleydb.weebly.com/) – 196 multitracks
- Task
  - Instrument recognition in multitimbral mixtures or classifying individual stems (one instrument active per stem)
- Aspects to look deeper into
  - How robust is instrument recognition vs. #overlapping instruments?
  - How does instrumentation relate to music genre (also annotated)?
    - Co-occurrence matrix

Research Project
Topic #7: Chord Recognition



- Dataset
  - IDMT-SMT-CHORDS (https://www.idmt.fraunhofer.de/en/publications/datasets/chords.html)
- Task
  - Estimate chord type (3-voiced / 4-voiced chords) from keyboard instruments / guitars
- Aspects to look deeper into
  - Compare classical approach (template matching on chroma vectors) with deep learning based approach (CNN)

Research Project
Topic #8: Record-Your-Own-Soundscapes

- Dataset
  - Soundscape recordings
- Task
  - Sound Event Detection
  - Annotation using Sonic Visualiser
- Aspects to look deeper into
  - Annotator Agreement
  - Background Noises
  - Temporal long-term structure of audio recordings

Fig. 1

Tools

- Python programming
  - Jupyter notebook (https://jupyter.org/)
  - Google Colab (https://colab.research.google.com)
- Audio Editing/Processing
  - Audacity (https://www.audacityteam.org/)
- Annotation
  - Sonic Visualiser (https://www.sonicvisualiser.org/)
- Presentation
  - Powerpoint / Google Slides
- Data Sharing
  - Dropbox / Google Drive

Images

Fig. 1: https://wra-ca.com/wp-content/uploads/2021/02/AudioMoth-photo.jpg