**Deep Learning for Audio**

1. Textbooks
   1. Goodfellow et al: “Deep Learning” deeplearningbook.org
   2. “Computational Analysis of Sound Scenes and Events” cassebook.github.io
   3. “Fundamentals of Music Processing” https://link.springer.com/book/10.1007/978-3-319-21945-5
   4. Zhang et al: “Dive into Deep Learning” <https://d2l.ai/>
   5. Murphy: “Machine Learning” www.cs.ubc.ca/ murphyk/MLbook/
   6. QM library: Hard copies AND ebooks
2. Our programming framework:
   1. Python 3
   2. PyTorch
   3. Plus numpy, matplotlib, librosa
   4. QMUL compute servers
      1. https://jhub.eecs.qmul.ac.uk/: Login with your EECS credentials
3. Paper presentations
   1. 15 minute each, 10% of module mark
   2. Read a research paper
   3. Compose some simple slides on it
   4. Present 10–12 minutes (plus questions)
   5. Also submit slides via QMplus before the lecture time
   6. Suggested slide organisation:
      1. Paper title, authors, and general aim
      2. Method – what they did
      3. Novelty – what was new/interesting
      4. Limitations – in your opinion
      5. Implications – what do we learn from this?
      6. Approx. 12 slides You MAY reuse figures/tables etc from the paper.
   7. Sample papers:
      1. Fairbrass et al (2018): CityNet—Deep learning tools for urban ecoacoustic assessment
      2. Davies and Boeck (2019): Temporal convolutional networks for musical audio beat tracking
      3. Pandeya et al (2018): Domestic cat sound classification using learned features from deep n.n.s
      4. Salamon et al (2017): Fusing shallow and deep learning for bioacoustic bird species classification
      5. Sethi et al (2019): Combining machine learning and a universal acoustic feature-set yields efficient automated monitoring of ecosystems
      6. Kim et al (2018): CREPE: A Convolutional Representation for Pitch Estimation Cakir et al (2017) Convolutional recurrent n.n.s for polyphonic sound event detection
      7. Dieleman et al (2014): End-to-end learning for music audio
      8. Pons et al. (2019): Randomly weighted CNNs for (music) audio classification
   8. You can choose:
      1. Present in class: upload your presentation to QM+ in advance
      2. Record video:
         1. upload your presentation to the video QM+ in advance,
         2. upload the video (e.g. into SharePoint) in advance,
         3. share the link of the video via Student Forum on QM+,
         4. response to possible questions.
4. Coursework
   1. Specification
      1. Apply two different deep learning methods to infer two different aspects of an audio/music task
      2. Explore empirically how to combine the two different inferences
      3. Pick an audio file that will be the case study to inspect the performance of the models in (1) and (2)
   2. Coursework proposal submission:
      1. Deadline: 4th March
      2. Proposals will not be accessed and your proposal can be adjusted
   3. Coursework submission:
      1. Deadline: 14th April
      2. What to submit: 1500-word report PDF.
5. Open science
   1. We like:
      1. Open source — Reuse/study code of others; publish your own (e.g. Github)
      2. Open data — We’ll use research datasets
      3. Open access — Research papers freely online
      4. Be aware of licences: Creative Commons, GPL, MIT...
6. Machine Learning:
   1. y=f\_theta(x): Parameters theta to be chosen with the aid of data
7. Deep learning:
   1. y=f\_theta’’’’(f\_theta’’’(f\_theta’’(f\_theta’(f\_theta(x)))))