Lecture WS 2018/2019
Deep learning
for speech & language processing

Thang Vu, Heike Adel, Daniel Ortega, Sabrina Jenne

Date: 20.01.2018

Lecture organizations

Date	Topic
21.01.2019	Projects – No Lecture
24.01.2019	Projects – No Lecture
28.01.2019	Projects – No Lecture
01.02.2019	Optional Test Project Submission
04.02.2019	Projects – No Lecture
08.02.2019	Final Project Submission
14.02.2019	Written Exam

Policy

- **Projects:**
  - You have to submit a solution to be able to write the final exam
  - BONUS points for written exam
    - Max. 20 points
- Written exam:
  - 100 points
  - You are allowed to write only if you meet the requirements of the exercises and submit a project solution
  - 60 points needed to pass

Projects - Policy

- **Individual submissions (no group work)**
- We offer two topics – you choose one of them
- We provide you with
  - Labeled training data
  - Test data (without labels)
- You submit the results (labels) of your network on the test data as well as your final running code
- For each topic, we will create a ranking of results

Topics (I)

- Topic 1: Dialog Act Classification from Text
- Topic 2: Emotion Recognition from Speech

Topic 1: Dialog Act Classification

- Develop a classifier which predicts the dialog acts (e.g. question/statement/backchannel) given an utterance and its context using lexical cues
- Example:
  - “what is that” → question
  - “anyway it 's the only car that says you know sporty and class uh” → opinion

Topic 1: Dialog Act Classification

- Two files are given for the training:
  - utterances.train
  - utterances.valid
- with following format:
  - utt_ID dialog_act(utt_t) utt_t-3 ; utt_t-2 ; utt_t-1 ; utt_t

- One test file without label is given: utterances.test
- with the following format:
  - utt_ID utt_t-3 ; utt_t-2 ; utt_t-1 ; utt_t

Topic 1: Dialog Act Classification

- Some statistics:
  - Training: 196,502
  - Validation: 20,000
  - Test: 20,000

Topic 1: Dialog Act Classification

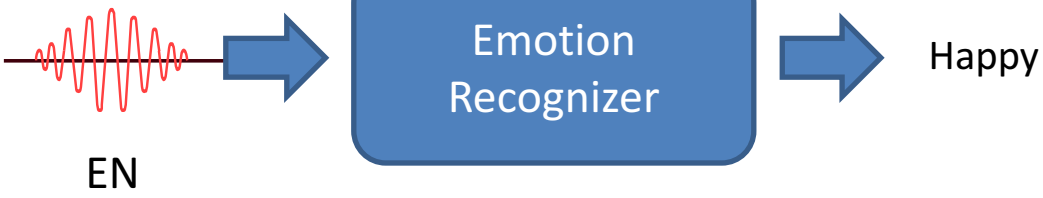
- Dialog act set: (see README file for more details)
  - % : Indecipherable
  - %-- : Abandoned
  - 2 : Collaborative-Completion
  - aa : Accept, Yes Answers
  - aap : Partial Accept
  - ar : Reject, No Answers
  - b : Backchannel
  - ba : Appreciation
  - bc : Misspeak Correction
  - bd : Downplayer
  - bh : Rhetorical-Question Backchannel
  - bk : Acknowledgment
  - br : Signal Non understanding
  - bs : Reformulation
  - ….

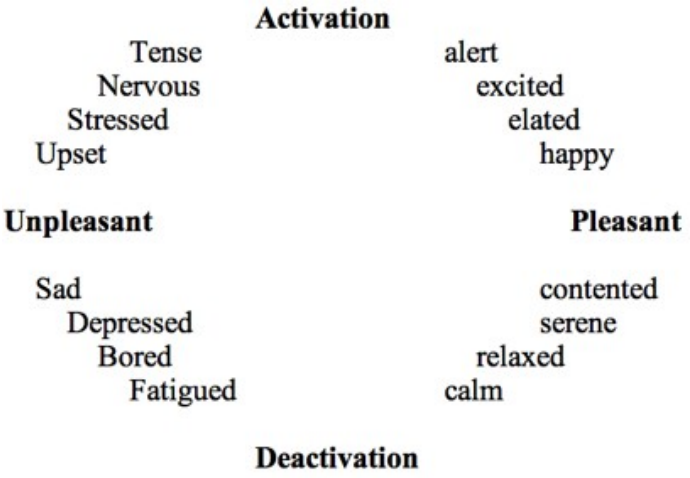
Topic 1: Dialog Act Classification

- Your best results should be submitted in a .txt file with the following format:
  - uttID dialog_act_label, e.g.
    - 2121_2 qw
    - 2121_11 ar
    - …
- Scoring:
  - Accuracy = #correct_predictions/#utt_in_test_set

Topic 2: Emotion Recognition from Speech

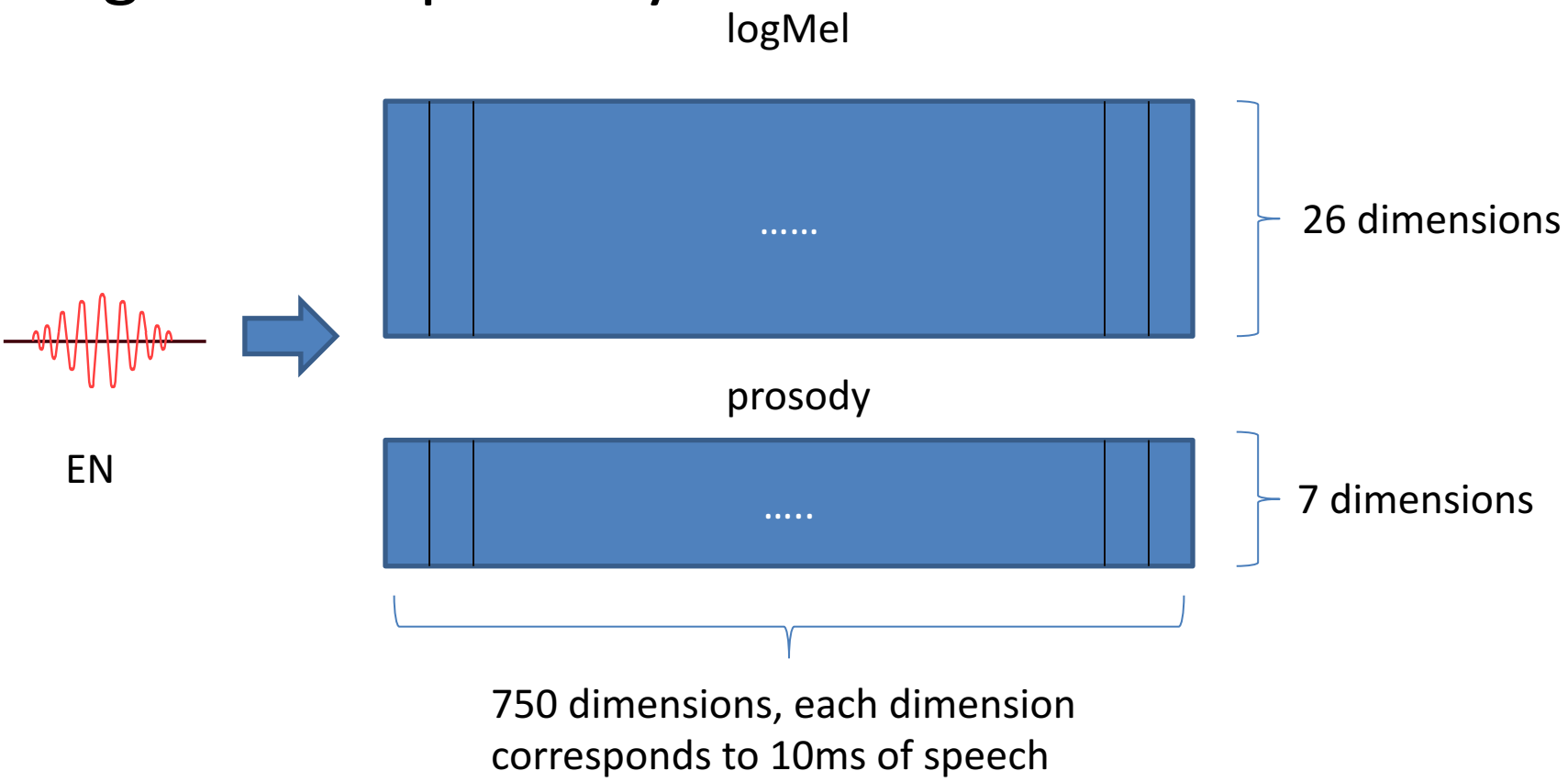
- Develop a classifier which can recognize emotion expressed in a given speech independent of the spoken language





Topic 2: Emotion Recognition from Speech

- For each audio file, two feature sets are given: logMel and prosody



Topic 2: Emotion Recognition from Speech

- Given files:
  - data_logMel.train
  - data_prosody.train
  - data_logMel.valid
  - data_prosody.valid
  - data_logMel.test with anonymuous labels
  - data_prosody.test with anonymuous labels
- Each file contains a DataArray with filename (= dataID), features and labels
- A script is provided to read these files

Topic 2: Emotion Recognition from Speech

- Label set with four emotion classes
  - Angry
  - Happy
  - Sad
  - Neutral
- Some statistics:
  - Training: 5,531
  - Validation: 2576
  - Test: 3900

Topic 2: Emotion Recognition from Speech

- Your best results should be submitted in a .txt file with the following format:
  - dataID emotion_category, e.g.
    - data01 neutral
    - data02 happy
- Scoring:
  - Accuracy = #correct_predictions/#utt_in_test_set

Example Research Questions (I)

- Which optimizer works best?
- Comparison between different non-linear units
  - E.g. ReLU vs tanh
- Visualization
  - Attention weights using TensorBoard
  - Or other things
- Comparison of loss functions
  - Cross entropy vs. MLE
  - Cross entropy vs. hinge loss

Example Research Questions (II)

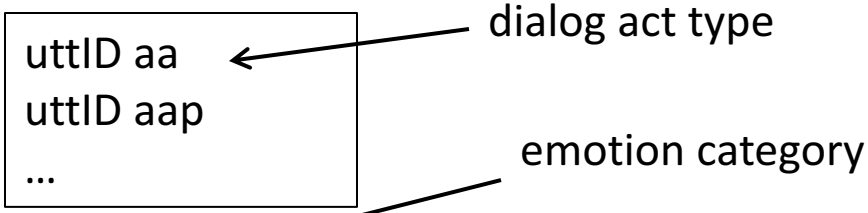
- Batch normalization:
  - Where to apply: before or after affine transformation
  - Relation to dropout
  - Relation to L2 regularization
- Relation between learning rate vs. batch size
- Adversarial Examples
  - How to poison a network

How To Submit Results

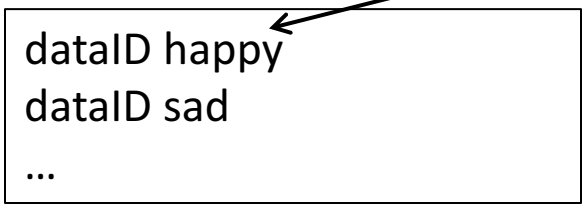
- Submit your results via **Ilias**
- Two dates:
  - February 01: Optional submission of results on test data
    - We will score the results and create a ranking
    - Idea: to get a feeling how your current model performs
  - February 08: Submission of results on test data and code
- Submission on February 01 is optional, submission on February 08 is mandatory to be able to write the exam!

What To Submit

- **Results on test data:**
  - Text file with one line per test item
  - Example for topic 1:



  - Example for topic 2:



- **Final running code:**
  - Python file (using Keras, PyTorch, Tensorflow… for neural network design and training - your choice)
  - Executable and understandable code
- **File names:**
  - matriculationnumber_lastname_topicX_result.txt
  - matriculationnumber_lastname_topicX_code.py

Thanks for listening!