Introduction to Deep Learning

Introduction to
Deep Learning

Course Introduction

October 16, 2017

Sebastian Stober <sstober@uni-potsdam.de>
Jens Johannsmeier <johannsmeier@uni-potsdam.de>

Introductions

“Living Stats”

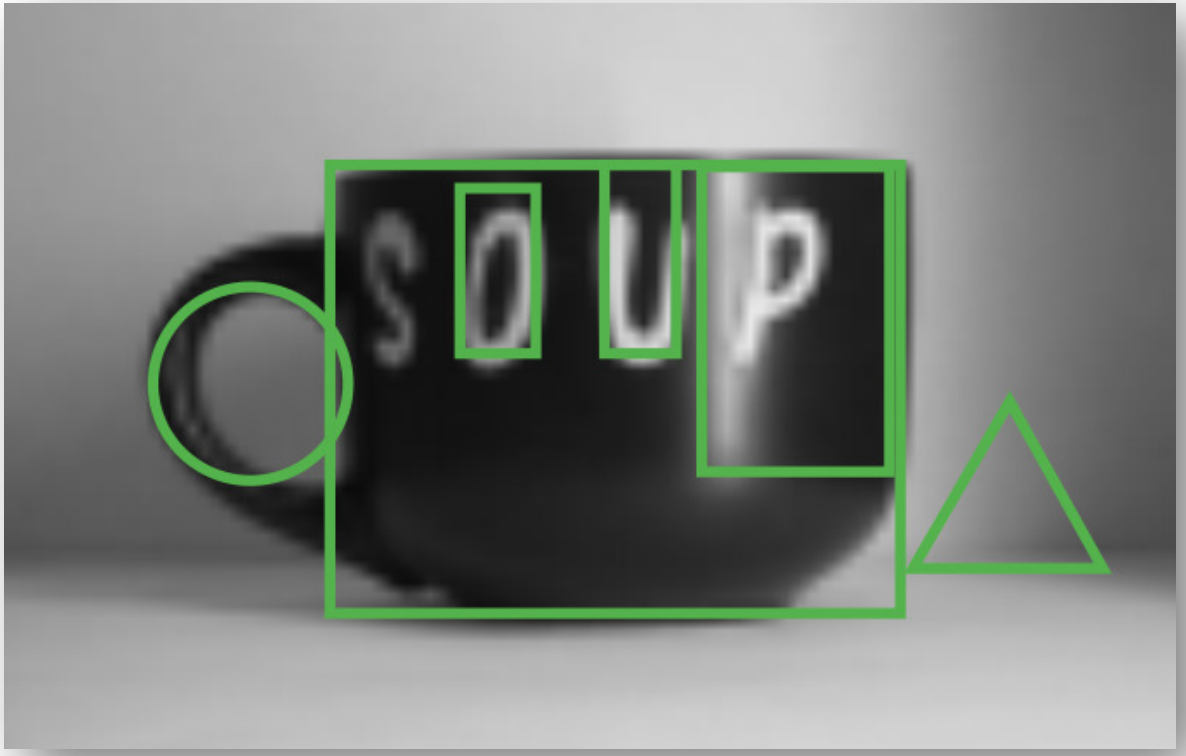
- How far away from home are you?
- How long have you studied already?
- How experienced are you in ML?
- How experienced are you in DL?

A Brief Intro to

Deep Learning

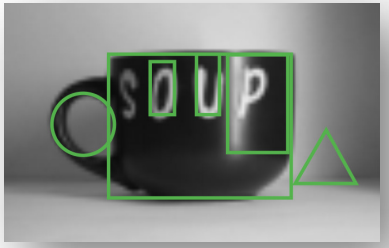
My 1st ML Project

- A coffee mug though the “eyes” of a road-sign detector in 2003:

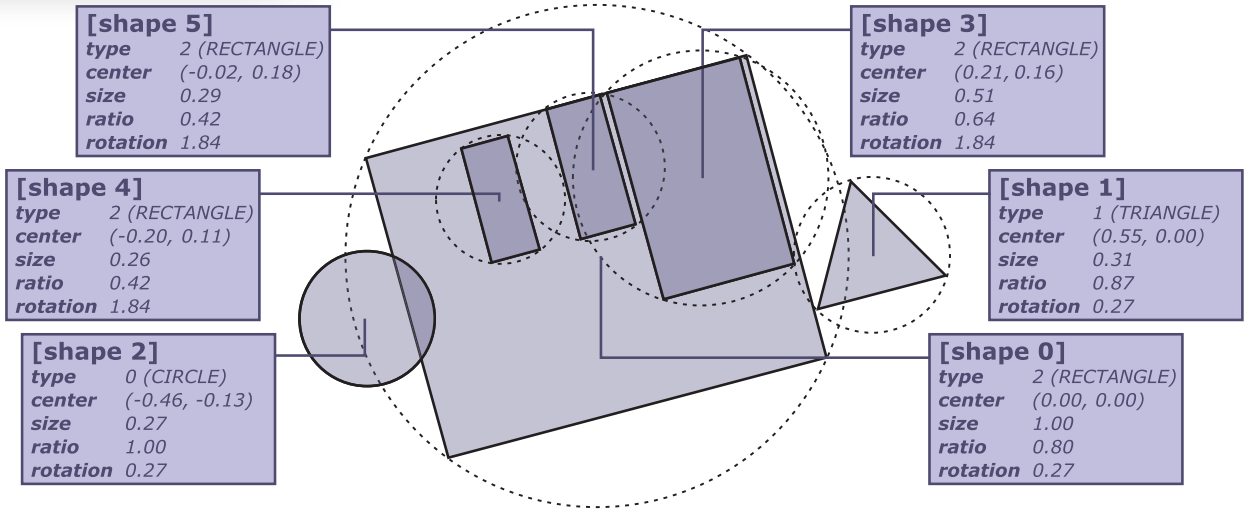


My 1st ML Project

- A coffee mug though the “eyes” of a road-sign detector in 2003:



representation
to be used by classifier



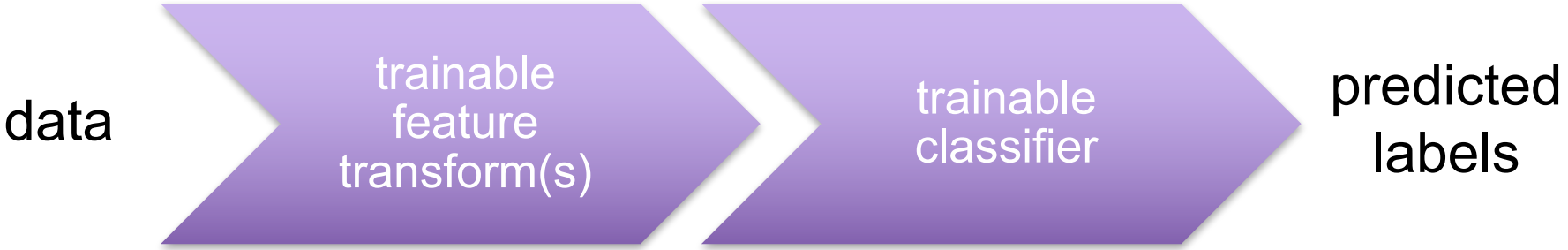
Typical Machine Learning Workflow (for Classification)

make use of domain knowledge from experts

train with labeled data (supervised)

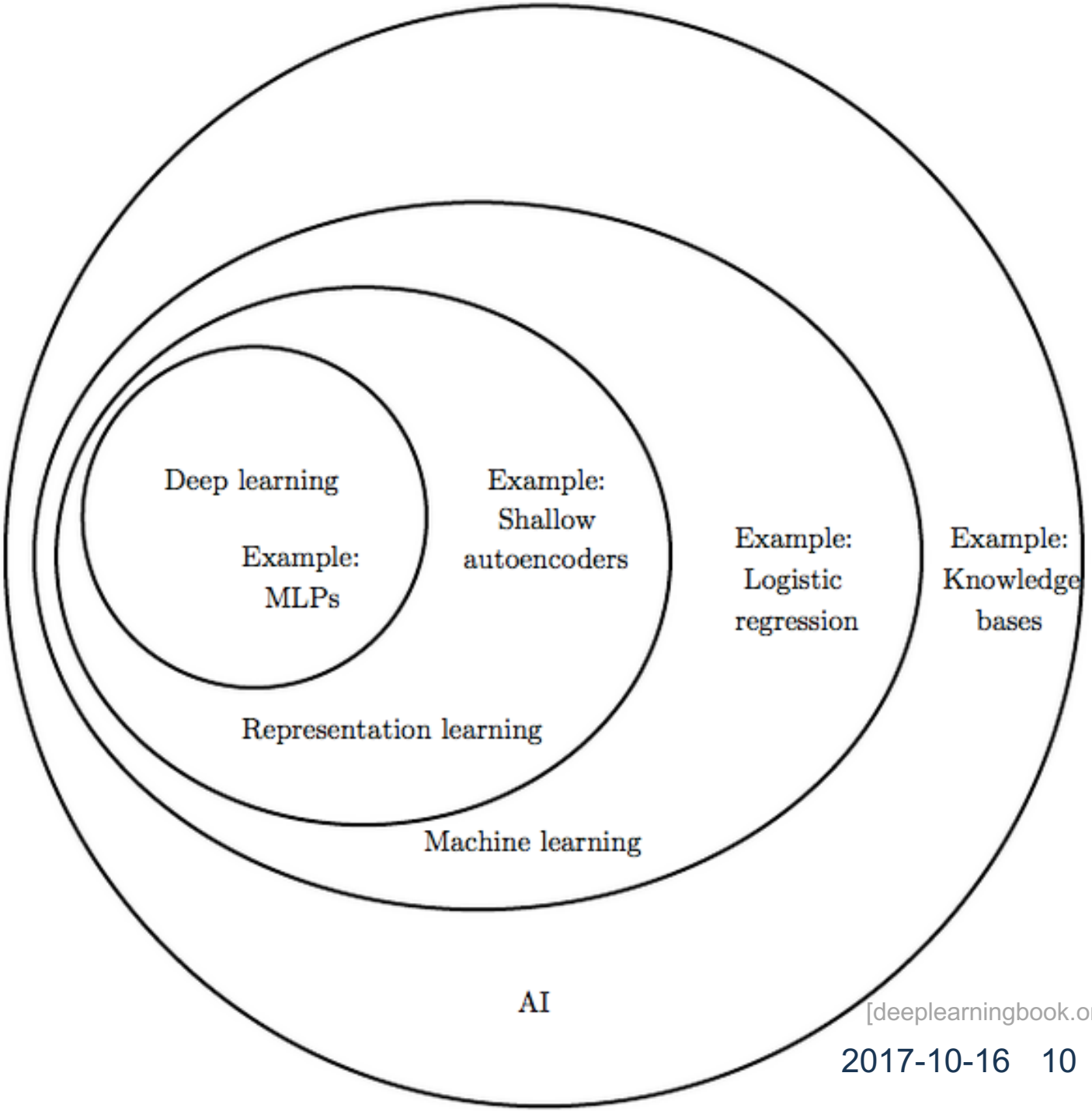
Typical Deep Learning Workflow (for Classification)

make use of abundant data and (GPU) compute power

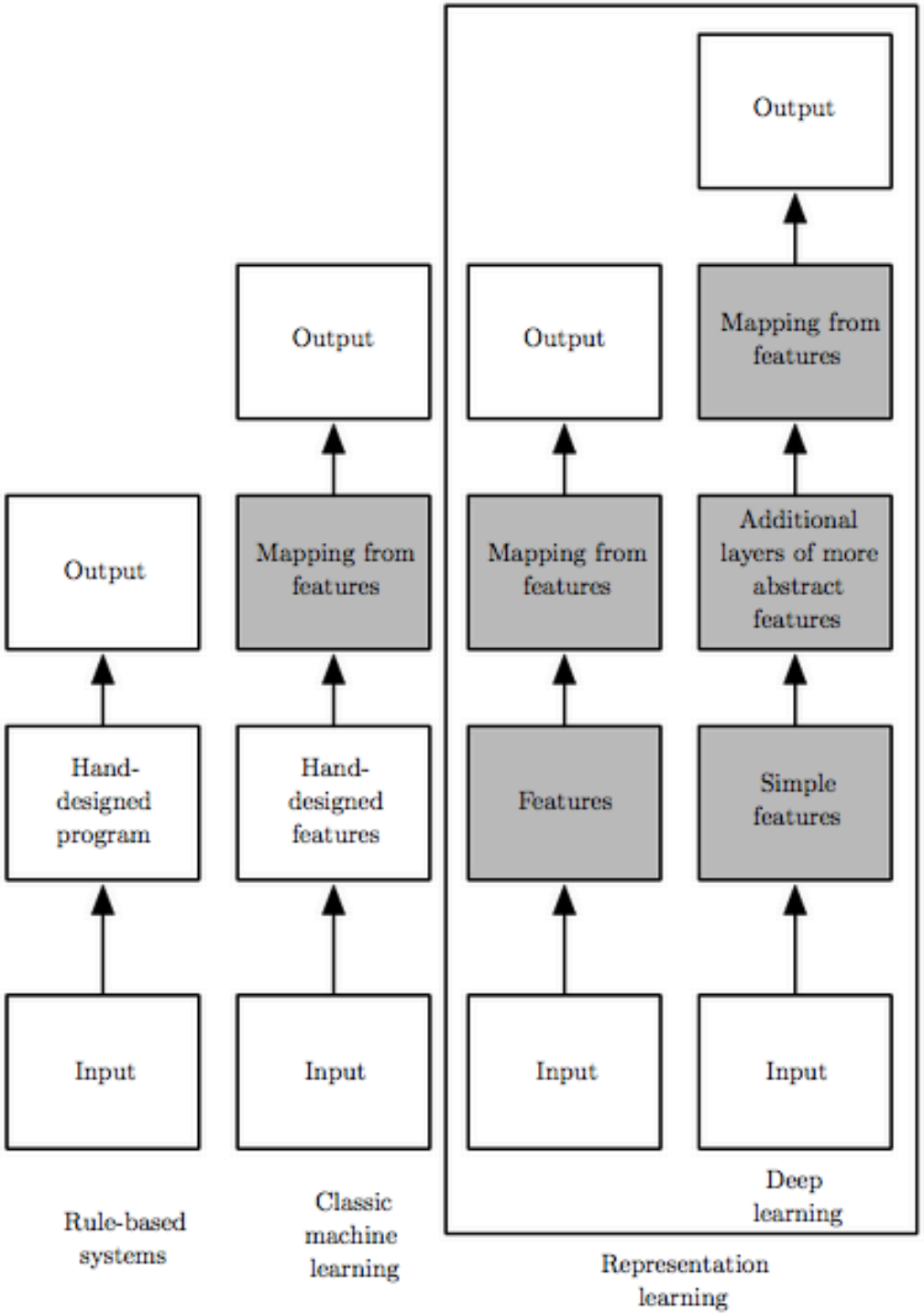


e.g., “pre-train” with unlabeled data (unsupervised)

train with labeled data (supervised)



[deeplearningbook.org]

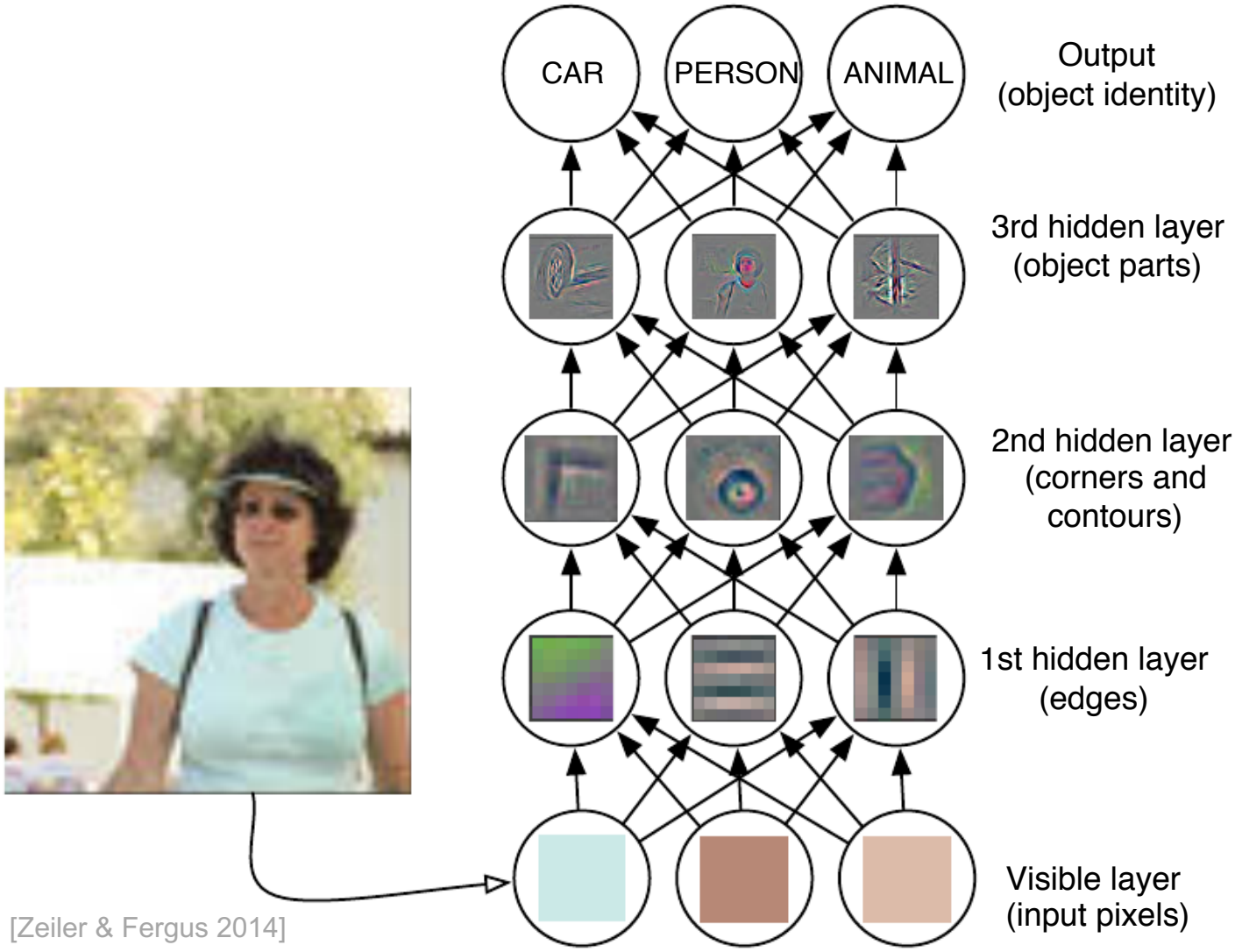


[deeplearningbook.org]

The Promise of Deep Learning

- learn suitable feature representations along with the actual learning task
- using a general-purpose learning procedure

An Example Deep Net



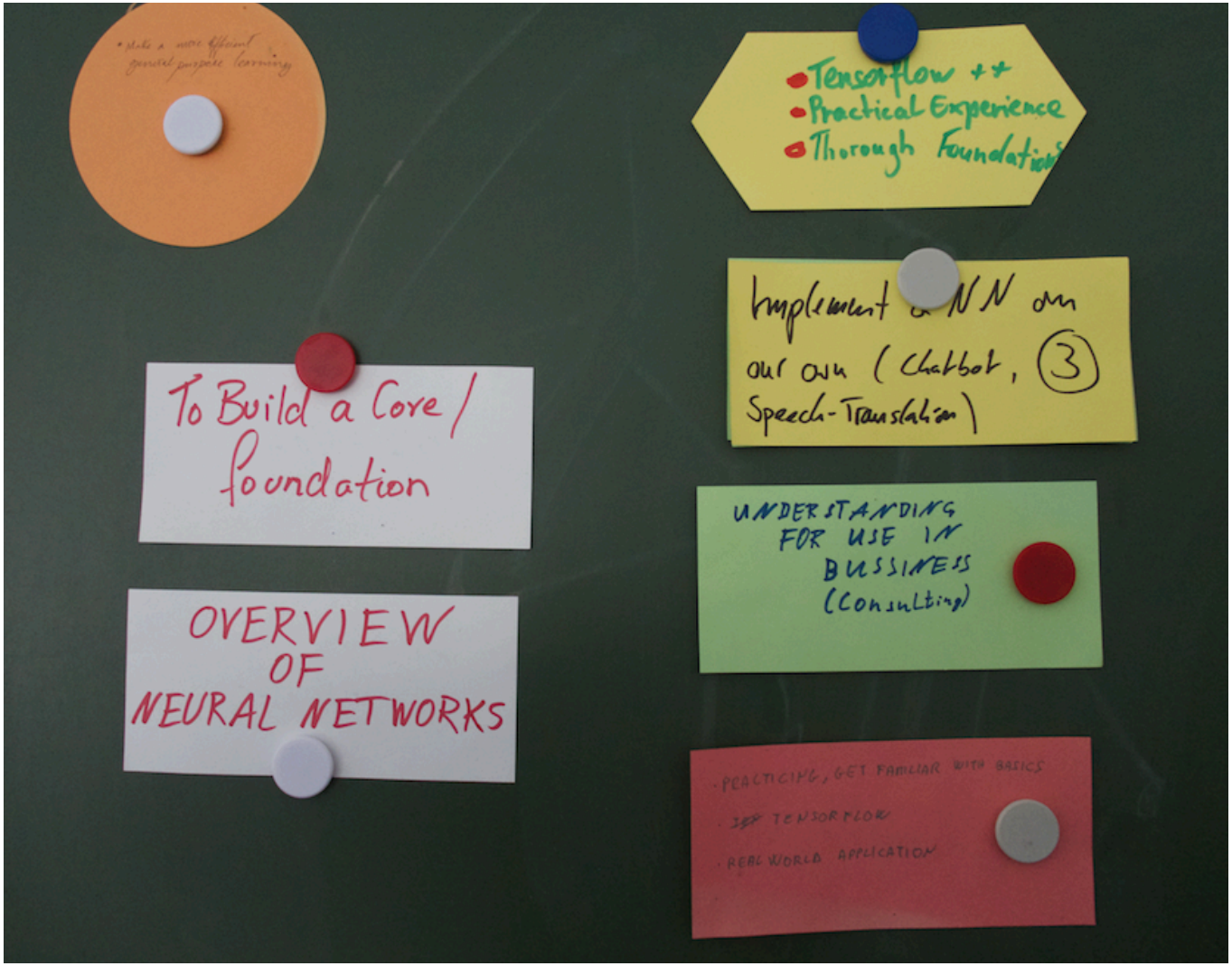
[Zeiler & Fergus 2014]

Course Rationale & Design

Learning Goals

- Think – Pair – Share
  1. Think about your personal learning goals for this course!
  2. Discuss with your neighbors & create a ranking!
  3. Name your most important one!

Learning Goals



Overall Learning Goals

- At the end of the course, you are able to …
  - **confidently** apply DL techniques to develop a solution for a given problem
  - follow recent DL publications and **critically** assess their contributions
  - formulate **hypotheses** and design & conduct DL experiments to **validate** them
  - **document** progress & design decisions for **reproducibility** and transparency

Please add your goals!

Disclaimer

this course may not be suitable for …

- mere credit collectors
- passive attendees
- remote students
- the lighthearted ;-)

end of withdrawal period:
November 20, 2017

No Free Lunch!

This is **not** how you will learn…

you will need to *participate*

The Nuremberg Funnel (1647)

Course Design

Preparation	In Class	Practical Work
(session summary)	"last episode on IDL" (3-min summary)	weekly exercise assignments (until Christmas)
reading • book chapters • papers • blogs	literature discussion / Q&A small-group activity (25-60 min)	channel discussions
channel discussions	project / assignments discussion / Q&A (25-60 min)	course project working in teams • 2-4 team members • scrum-style weekly sprints • team progress blog
weekly blog posts • new insights • hints, tricks & hacks • open questions		
3-6h per week →	**150h total**	← 3-6h per week

grading: *oral exam (20min)

Topics (Tentative)

- MLPs, Gradient Descent & Backpropagation
- Convolutional Neural Networks
- Recurrent/Recursive Neural Networks
- Auto-Encoders
- Regularization Techniques
- Advanced Regularization Techniques
- Introspection & Inception
- Optimization Techniques
- Advanced Training Strategies
- Deep Reinforcement Learning

Online Tools

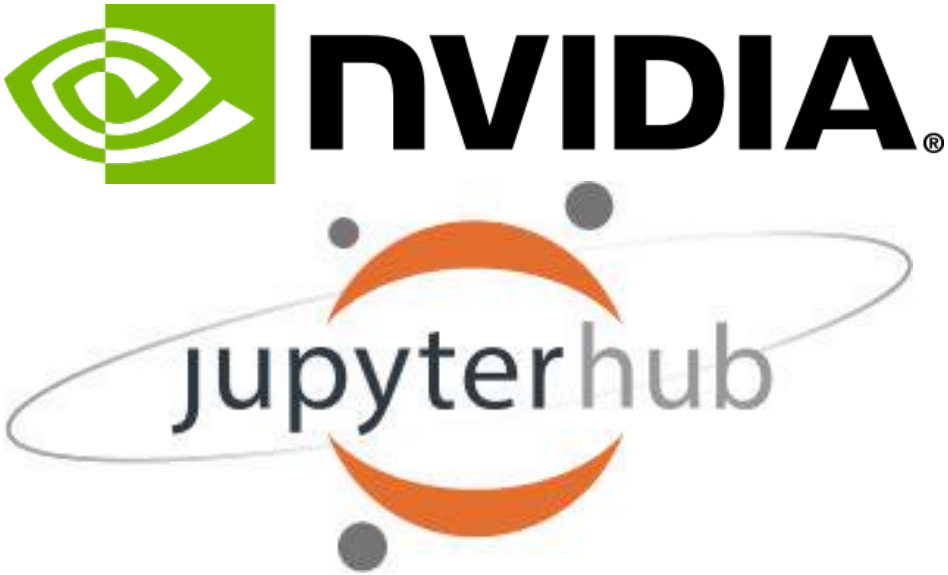
- Mattermost (~Slack) channel
  + Campus.UP workspace
  - channels / blogs (course / personal / team)
  - forum
  - wiki
  - messaging
- GPU compute environment (starting Jan.)
  - shell access & jupyterhub for notebooks

GPU Compute Servers

10 GPU Compute Servers

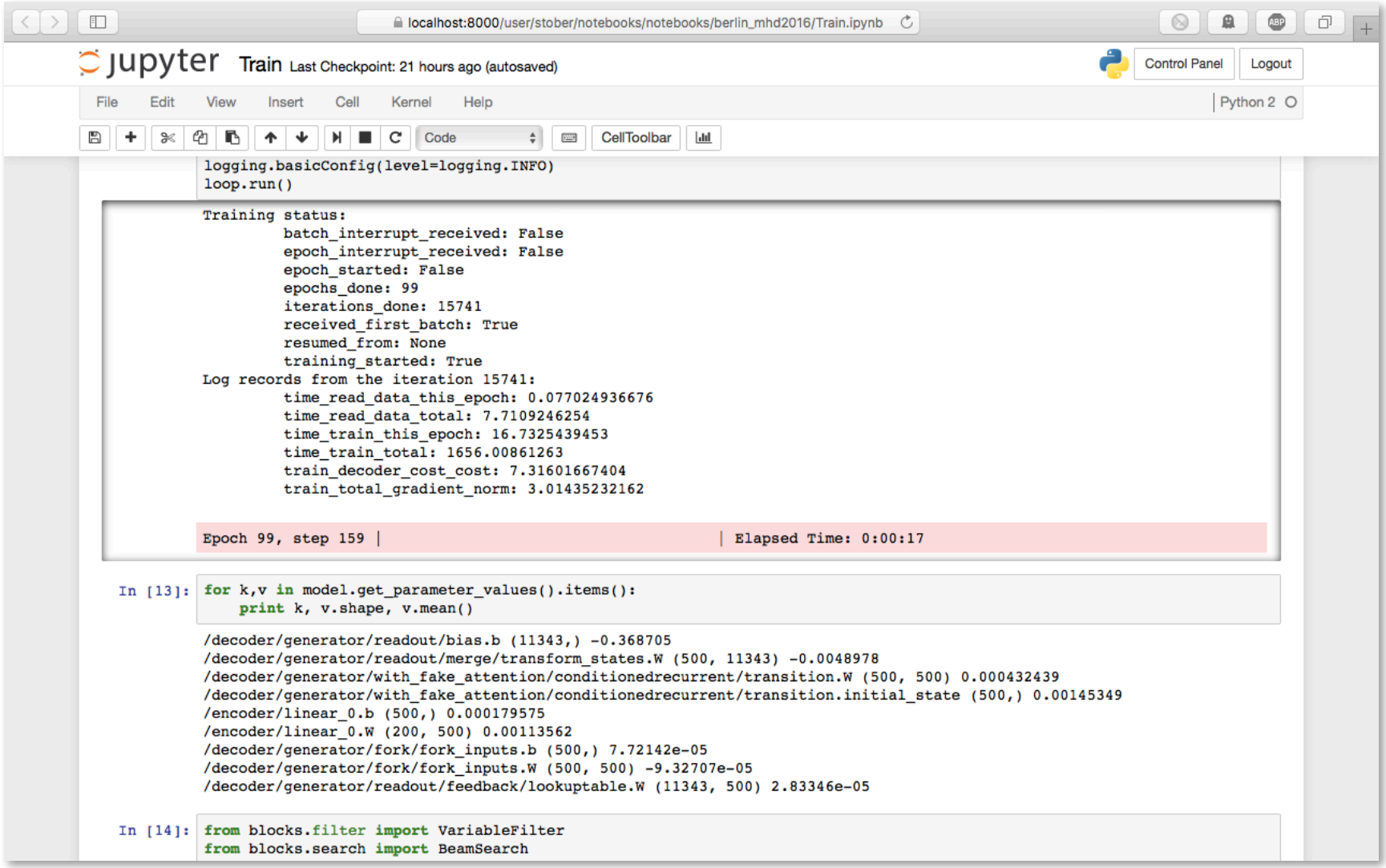
- 8 Pascal Geforce 1080 Ti GPUs
- 256 GB RAM
- 24 CPU cores
- jupyterhub server for notebooks
- fully dockerized

… more infos later

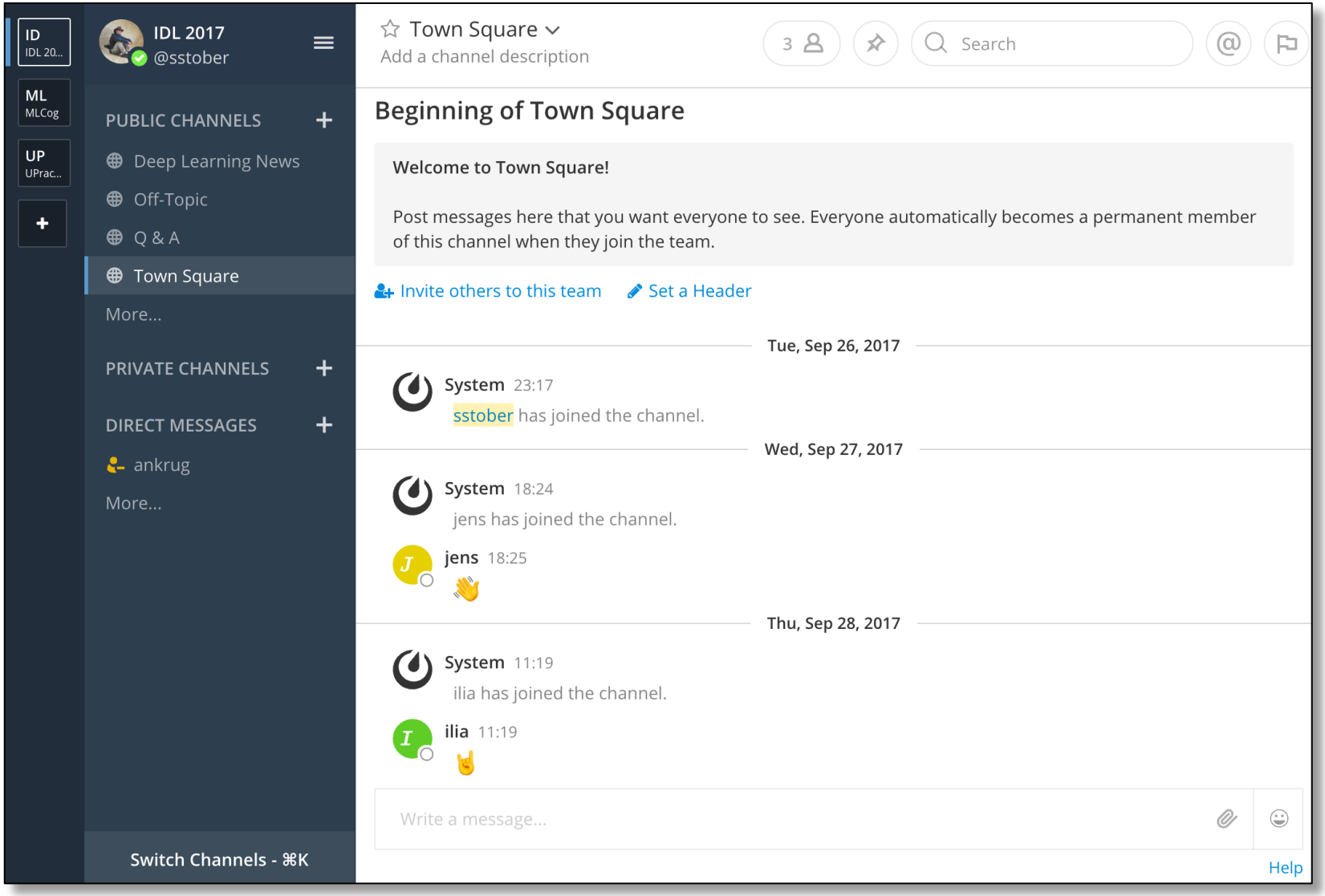


https://jupyter.org/

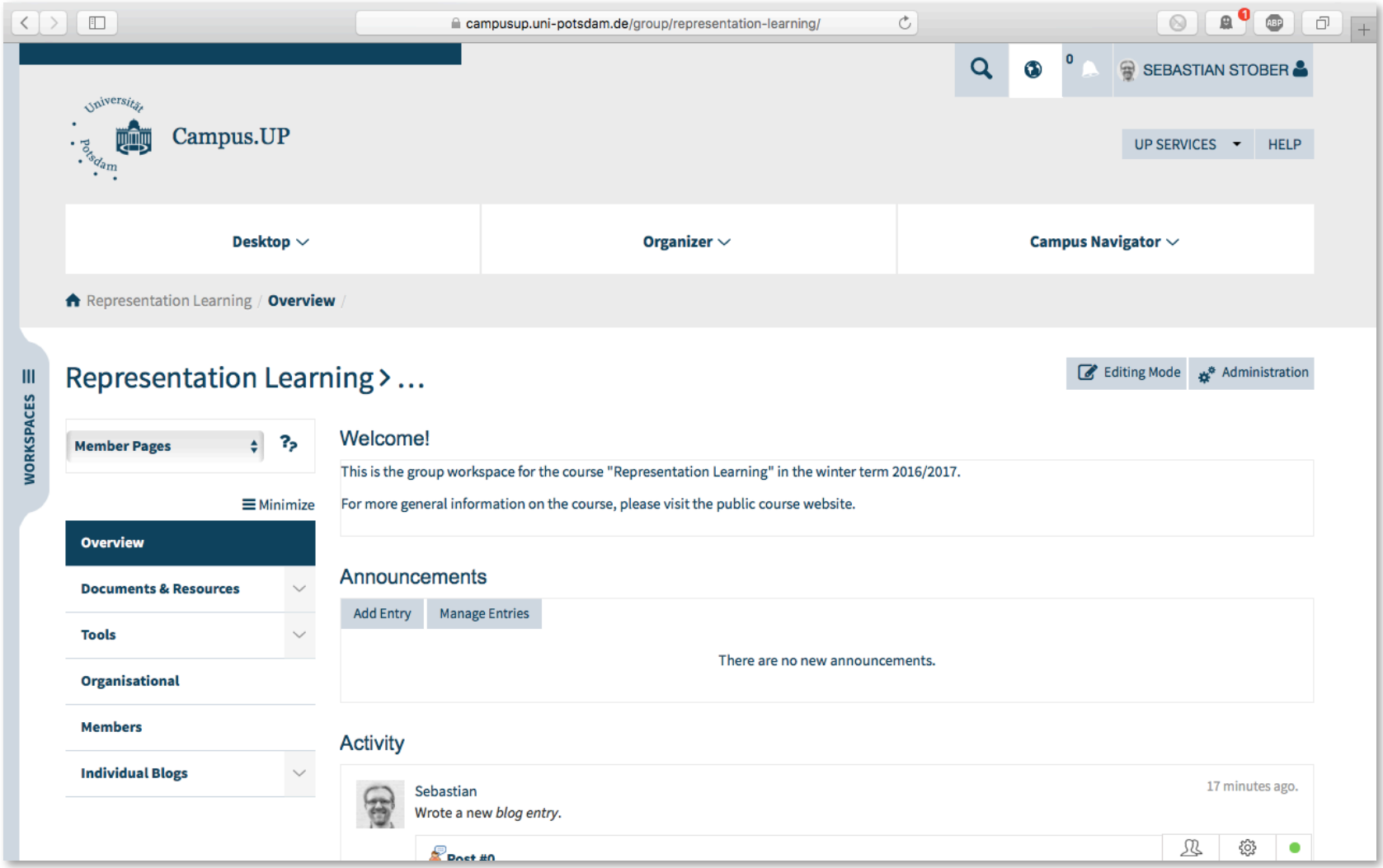
Jupyterhub


localhost:8000/user/stober/notebooks/notebooks/berlin_mhd2016/Train.ipynb
jupyter Train Last Checkpoint: 21 hours ago (autosaved)
Control Panel
Logout
File Edit View Insert Cell Kernel Help
Python 2
Code
CellToolbar
logging.basicConfig(level=logging.INFO)
loop.run()
Training status:
batch_interrupt_received: False
epoch_interrupt_received: False
epoch_started: False
epochs_done: 99
iterations_done: 15741
received_first_batch: True
resumed_from: None
training_started: True
Log records from the iteration 15741:
time_read_data_this_epoch: 0.077024936676
time_read_data_total: 7.7109246254
time_train_this_epoch: 16.7325439453
time_train_total: 1656.00861263
train_decoder_cost_cost: 7.31601667404
train_total_gradient_norm: 3.01435232162
Epoch 99, step 159 | | Elapsed Time: 0:00:17
In [13]:
for k,v in model.get_parameter_values().items():
print k, v.shape, v.mean()
/decoder/generator/readout/bias.b (11343,) -0.368705
/decoder/generator/readout/merge/transform_states.W (500, 11343) -0.0048978
/decoder/generator/with_fake_attention/conditionedrecurrent/transition.W (500, 500) 0.000432439
/decoder/generator/with_fake_attention/conditionedrecurrent/transition.initial_state (500,) 0.00145349
/encoder/linear_0.b (500,) 0.000179575
/encoder/linear_0.W (200, 500) 0.00113562
/decoder/generator/fork/fork_inputs.b (500,) 7.72142e-05
/decoder/generator/fork/fork_inputs.W (500, 500) -9.32707e-05
/decoder/generator/readout/feedback/lookuptable.W (11343, 500) 2.83346e-05
In [14]:
from blocks.filter import VariableFilter
from blocks.search import BeamSearch

Mattermost Channel



Campus.UP Workspace



Session Summaries

last episode on
“Introduction to Deep Learning”
…

- rotating job!
  (2 persons per session,
  assignment by poll)

- short summary blog post (in course blog)
  + 3-min intro recap at next session
  - key topics
  - results of the discussion
  - optional photos

Your Personal Channel

- document your learning / project progress
  - one post per week
  - share your experiences!
  - visible only to course participants

- examples:
  - https://deeprandommumbling.wordpress.com/
  - http://bartvanmerrienboer.nl/#blog
- guidelines:
  - https://www2.uwstout.edu/content/profdev/rubrics/blogrubric.html

Q & A Channel

- guide for what is covered in class
  deadline: Monday morning 7am

- do not hesitate to post questions!
  (If you got one, you are probably not the only one!)

- post a comment if you know the answer

Contribute!

- ask – in your channel and Q&A
- comment / like / rate
- answer
- document
  - hints, tricks & hacks
- recommend
  - additional readings (papers, blogs, etc.)
- give (constructive) feedback

Course Project

Course Project

- vision: speech-base interaction

real systems:

- Siri (Apple)
- Alexa (Amazon)
- Cortana (Microsoft)
- Google Home
- Skype Translate
- …

fictional characters:

- J.A.R.V.I.S. (Iron Man)
- Samantha (Her)
- Jane (Ender's Game)
- ...

Course Project

- Automatic Speech Recognition (ASR)
  - Beat the baseline system!

J. Kunze; L. Kirsch; I. Kurenkov; A. Krug; J. Johannsmeier & S. Stober. **Transfer Learning for Speech Recognition on a Budget**. In: *2nd Workshop on Representation Learning for NLP at the Annual Meeting of the Association for Computational Linguistics (ACL'17)*, 2017.

https://arxiv.org/abs/1706.00290

https://github.com/transfer-learning-asr/transfer-learning-asr

Course Project

- collaborative effort
- “coopetition” (cooperative + competition)
- multiple teams of 2-4 students:
  - discuss ideas and form team in December
  - self-organized (heterogeneous if possible)
  - scrum-like approach
  - focus on different goals / aspects / strategies

Team Channels (January/February)

- weekly progress reports for course project
  - similar to scrum
  - compare original goals with outcomes
    - What has worked well?
    - What did not work / had to be changed?
  - outline plan for next week
    - What would you like to try / investigate next?
- can be written up by one designated team member or in turns