

BIO-642

**State of the Art Topics in Neuroscience XIII**

Gerstner Wulfram

<b>Cursus</b>	<b>Sem.</b>	<b>Type</b>
Neuroscience		Opt.

Language of teaching	English
Credits	1
Session	
Exam	Term paper
Workload	30h
<b>Hours</b>	<b>25</b>
Courses	10
Project	15
<b>Number of positions</b>	<b>30</b>

**Frequency**

Only this year

**Summary**

The Loss Landscape of Neural Networks is in general non-convex and rough, but recent mathematical results lead provide insights of practical relevance. 9 online lectures, lecturers from NYU, Stanford, Shanghai, IST Austria, Google, Facebook, EPFL.

**Content**

Loss Landscapes of Neural Networks:

In practical applications, Deep Neural Networks are typically trained by walking down the loss surface using gradient descent augmented with a bag of tricks. One of the important practical insights has been that large, over parameterized, networks that have more parameters than necessary work better - one potential interpretation (but not the only one) is the 'lottery ticket hypothesis'. Obviously, the shape of the loss landscape is important when walking down. In recent years, research on the shape of the loss landscape has addressed questions such as: "Is there one big global minimum or many scattered small ones?", "Is the loss landscape rough or smooth?", "Should we worry about saddle points?", "Are there flat regions in the loss?", "How many saddle points are there?". While these look like questions for theoreticians, their answers might have practical consequences and lead to a better understanding of the role of over parameterization, pruning, and reasons of the bag of tricks. The aim of this workshop is to bring together researchers that have worked on these topics from different points of views and different backgrounds (Computer Science, Math, Physics, Computational Neuroscience), and build a community around these questions.

II. WORK LOAD of the students (1 credit = 25-30 hours of work)

A) 10 speakers over 2 days + discussion session (program attached). = 11 hours of work

B) Report preparation for course credit (14 hours of work)

1. Each participating student has to ask at least 2 questions for two different lectures
2. Each student has to submit a 5-page report

page 1: Select the best talk

- read the corresponding paper and watch the video again
- present the main result of the paper.
- how is this result shown (mathematical proof, main simulation figure etc)
- How could this result influence my own work/own thinking about artificial neural networks
- How could this result relate to biology/neuroscience

page 2: - For the same talk

- Why was it a good talk in term of presentation style/talking style?

- What are elements that I would like to copy for my next online talk?
- What could be further improved/what would I avoid.
- What is the question that I asked?
- What was the answer? Was the answer satisfying?

page 3/4: Select second best talk. Answer same points as page 1/2

page 5: Select a bad talk

- Why was it bad talk in term of presentation style/talking style?
- What could be further improved/what would I avoid?
- If this had been your PhD student, what is the advice that you would give?

### **Note**

By the end of this course you should be able to explain recent results on the shape and convergence properties of the loss landscape in neural networks.

### **Keywords**

Artificial Neural Networks, Deep Learning, Optimization, Gradient Descent, Loss landscape, Convergence, Convexity, Permutation Symmetry

### **Learning Prerequisites**

#### **Required courses**

A Master-level class on Artificial Neural Networks or Machine Learning or Optimization