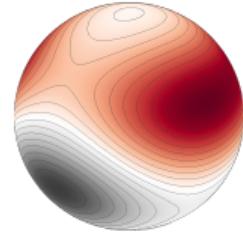




UNIVERSITÄT  
HEIDELBERG  
ZUKUNFT  
SEIT 1386



# Organizational Meeting

## Seminar: Mathematical Machine Learning

Winter Semester 2023/24

# Introduction



Roland Herzog  
[roland.herzog@iwr.uni-heidelberg.de](mailto:roland.herzog@iwr.uni-heidelberg.de)

[scoop-teaching@uni-heidelberg.de](mailto:scoop-teaching@uni-heidelberg.de)



Leonie Kreis  
[leonie.kreis@iwr.uni-heidelberg.de](mailto:leonie.kreis@iwr.uni-heidelberg.de)

## General

- Language: English
- Intended for Master students of Mathematics or Scientific Computing

# Concept of the Seminar

- Read and analyze scientific papers.
- Many-to-Many Role-Playing Format from  
<https://colinraffel.com/blog/role-playing-seminar.html>:
- One paper per meeting.
- Take a specific role.
- Discuss the paper together.

## Roles: Author

Assume that you are the author of the paper. Present your work.  
**Motivate** your research and **present your results**. Keep in mind that you only have 5 minutes, so focus on the main points. Split up the paper with your partner so that you do not repeat each other.

## Roles: Reviewer

Assume that the paper has not been published yet and is currently submitted to a top conference where you've been assigned as a peer reviewer. Complete a **full review** of the paper answering all prompts of the official review form of the top venue in this research area, e.g. <https://www.jmlr.org/reviewer-guide.html>. This includes recommending whether to **accept or reject** the paper.

## Roles: Archaeologist

Some time in the future: This paper was found buried under ground in the desert. You're an archaeologist who must determine where this paper sits in the **context** of previous and subsequent work. Find and report on one **older** paper cited within the current paper that substantially influenced the current paper and one **newer** paper that cites this current paper.

## Roles: Hacker

You're a hacker who needs a demo of this paper ASAP. **Implement** a small part or simplified version of the main algorithm in the paper on a small dataset or toy problem. Prepare to share the core code of the algorithm to the class and demo your implementation. Do **not** simply download and run an existing implementation – though you are welcome to use (and give credit to) an existing implementation for “backbone” code.

## Roles: Private Investigator

You are a detective who needs to run a background check on the **paper's authors**. Where have they worked? What did they study? What previous projects might have led them to working on this one? What motivated them to work on this project? Feel free to contact the authors, but remember to be courteous, polite, and on-topic.

# Grading

- Attending **all** meetings.
- Participation in class
  - In the beginning every person gives a 5 minute report from their role perspective.
  - Afterwards we have a scientific discussion of the paper.

# Mathematical Machine Learning - Challenges

## Ideal Setting:

Find global minimizer of

$$\min_{f \in \mathcal{F}} \mathbb{E}_{(X,Y) \sim \mathcal{D}} [\tilde{\ell}(f(X), Y)]$$

$\mathcal{F}$ : function class

## Realistic Setting:

Find (local) minimizer of

$$\min_{\theta \in \Theta \subset \mathbb{R}^m} \frac{1}{n} \sum_{i=1}^n \ell(f_\theta(x_i), y_i)$$

often with first-order methods

Estimation error

Approximation error

Optimization error

# Papers

1. Xavier Glorot and Yoshua Bengio. "Understanding the difficulty of training deep feedforward neural networks". In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. Vol. 9. Proceedings of Machine Learning Research. PMLR, May 2010, pp. 249–256. URL: <http://proceedings.mlr.press/v9/glorot10a.html>
2. Jimmy Ba and Rich Caruana. "Do deep nets really need to be deep?" In: *Advances in Neural Information Processing Systems*. Ed. by Z. Ghahramani et al. Vol. 27. NIPS'14. Curran Associates, Inc., 2014. arXiv: 1312.6184. URL: [https://proceedings.neurips.cc/paper\\_files/paper/2014/file/ea8fcd92d59581717e06eb187f10666d-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2014/file/ea8fcd92d59581717e06eb187f10666d-Paper.pdf)
3. Hao Li et al. "Visualizing the loss landscape of neural nets". In: *Advances in Neural Information Processing Systems*. Ed. by S. Bengio et al. Vol. 31. Curran Associates, Inc., 2018. arXiv: 1712.09913. URL: [https://proceedings.neurips.cc/paper\\_files/paper/2018/file/a41b3bb3e6b050b6c9067c67f663b915-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2018/file/a41b3bb3e6b050b6c9067c67f663b915-Paper.pdf)
4. Chiyuan Zhang et al. "Understanding deep learning (still) requires rethinking generalization". In: *Communications of the ACM* 64.3 (Feb. 2021), pp. 107–115. DOI: [10.1145/3446776](https://doi.org/10.1145/3446776)
5. Benjamin Grimmer. *Provably faster gradient descent via long steps*. 2023. arXiv: 2307.06324

## Timeline

- Today: Distribution of roles
  - Accepting a role is the **binding** registration for the seminar.
- Starting in November: 5 meetings à 90 minutes
- weekday and time: **decided today**
  - Monday 16:30-18:00
  - Tuesday 16:30-18:00
  - Monday 14:00-15:30
  - Tuesday 14:00-15:30
  - Thursday 14:00-15:30

# Assigning Roles

	Paper 1	Paper 2	Paper 3	Paper 4	Paper 5
Author	1	9	7	5	3
Author	2	4	6	8	10
Reviewer	3	1	9	7	5
Reviewer	4	6	8	10	2
Archaeologist	5	3	1	9	7
Archaeologist	6	8	10	2	4
Hacker	7	5	3	1	9
Hacker	8	10	2	4	6
Private Investigator	9	7	5	3	1
Private Investigator	10	2	4	6	8

# What's next?

- Resources are online:  
<https://scoop.iwr.uni-heidelberg.de/teaching/2023ws/seminar-mathematical-machine-learning/>
- These slides will be shared on the website.

# Questions?