



Prof. R. Wattenhofer

Recurrent Neural Networks: Do you even generalize?

Like any traditional Machine Learning system, we usually train Neural Networks (NN) on training data and then we would like it to classify or predict on test data. While having a low training error is important, what we really care about is having a low test error. The difference between training- and test error tells us something about the *generalization* performance of our algorithm.

There are different ways to improve generalization: Getting more data, preventing the NN from overfitting the data, choose a better network architecture, etc. However, how to choose the best methods/hyperparameters to improve generalization is still not well understood, and is heavily based on best-practices, heuristics and trial-and-error.

In this thesis, we want to investigate different aspects of generalization in the context of *Recurrent Neural Networks (RNN)* which are mostly used for sequence classification/prediction. If you are interested in this topic, do not hesitate to contact us so that we can exchange our ideas!



Requirements: Interest in and willingness to study Machine Learning and Deep Learning. There will be weekly meetings to discuss progress and open questions.

Interested? Please contact us for more details!

Contacts

- Yuyi Wang: yuyi.wang@tik.ee.ethz.ch, ETZ 94
- Gino Brunner: gino.brunner@tik.ee.ethz.ch, ETZ G63