



Prof. R. Wattenhofer

Analysing the Memory of Recurrent Neural Networks

Deep Artificial Neural Networks have achieved state-of-the-art performance in many different domains. When it comes to working with sequential data (text, speech, music, time-series), Recurrent Neural Networks (RNN) and its more capable variants are the name of the game.

Whenever we want our network to have some kind of “memory”, i.e., the predictions it is making now depend on past data, Long short-term memory (LSTM) networks have been applied very successfully, as they solve the Vanishing Gradient problem that prevents traditional RNNs from learning long-term dependencies. However, it is still not very well understood how large an LSTM needs to be in order to capture the necessary long-term dependencies of a given dataset. In other words: How do we figure out how much memory we need for a specific task, and how do we then need to choose our hyperparameters? How well can LSTMs “memorize” rare events? How fast do they “forget” them again?



Finding answers to these and other questions is the topic of this thesis. If this sounds interesting to you, do not hesitate to contact us so we can have a chat.

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

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