



Prof. R. Wattenhofer

Deep Learning: Investigation of Transformer Models

The field of Natural Language Processing has undergone an extraordinary progress since the introduction of Transformer models in 2017. Different variants of Transformers have been steadily pushing forward the state of the art in all language understanding tasks, to the point that currently they obtain super-human performance in most benchmarks. This is specially impressive given the complexity of human language. However, to this date it is not clear the reason why Transformers excel in language tasks: do they exploit statistical cues in the data or do they really understand language and represent it in deep and meaningful features?

In this project we build on our previous investigation of Transformers in order to acquire a better understanding of the inner workings of these models. We are interested in learning more about how non-linear these models are, how they cope with long-range dependencies and context and how they encode and propagate information. Achieving such an understanding is necessary to improve deep learning models in terms of performance, robustness or interpretability.

If this sounds interesting to you do not hesitate to contact us.

Requirements: Knowledge in Deep Learning, or solid background in Machine Learning. Implementation experience with TensorFlow or Pytorch is an advantage as well as knowledge in Natural Language Processing.

Interested? Please contact us for more details!

Contacts

- Damian Pascual: dpascual@ethz.ch, ETZ H93
- Gino Brunner: brunnegi@ethz.ch, ETZ H93

