



Investigating Transformers

Since their introduction in the context of machine translation, transformer models have been dominating the GLUE and SuperGLUE natural language processing benchmarks. A clear trend in transformer models is that bigger models tend to perform better, which was made abundantly clear by T5, the largest transformer model to date, when it outperformed all prior models by a large margin and almost achieved human performance on the new SuperGLUE task.

While there has been much work trying to investigate these models, we still understand relatively little about their inner workings and how they achieve this remarkable performance. In this thesis we continue prior work into investigating transformer models.

Requirements: Creativity and programming skills are an advantage. The student(s) should be able to work independently!

Interested? Please contact us for more details!

Contacts

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

