

Language Modeling using an ensemble of Transformers

Supervisor : Haris Jabbar (jabbar@cis.lmu.de)

Examiner :

Level : Masters (Bachelors in exceptional circumstances)

Transformer models have revolutionized NLP and especially Language Modeling. Presently there are three different modalities of training these transformers :

- Masked Language Models (Bert etc.)
- Causal Language Models (GPT-2/3 etc.)
- Permutation Language Models (XLNet).

Arguably these three techniques encode different aspects of a language. To test this hypothesis, we want to use an ensemble of these models and evaluate it on downstream tasks. If the hypothesis is true, then we should see an improvement in evaluation metrics. Additionally, the project will make extensive use of interpretability techniques to analyze the effects of the three models.