Brief Project Proposal

Course: Deep and Reinforcement Learning Fundamentals

Team members: Saniya Ashraf (sa19bc) , Ameer Hamza (ah18r)

Problem Statement:

Our proposed project is a comparative study of two different NLP models: BERT [1] and XLnet [2]. BERT (Bidirectional Encoder Representations from Transformers), introduced in 2018, enables learning bidirectional contexts by maximizing the expected likelihood over all permutations of the factorization order. However, it neglects dependency between the masked positions and suffers from a pretrain-finetune discrepancy which led to the invention of XLNet (Generalized Auto-Regressive model), introduced in 2019. It overcomes that problem by autoregressive formulation [2].

Plan:

Our proposed plan is the analysis of BERT and XLnet; similarities, dissimilarities, pros and cons and situations where one can outperform the other. We want to explore data on which BERT can outperform XLnet and vice versa. We are also interested in understanding their inner workings. Furthermore, we would like to explore the other extended versions of the architectures aforementioned. We aim to outline possible improvements in the architectures, and provide a better understanding of the ideal architecture, given a problem.

References:

[1] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” arXiv:1810.04805 [cs], May 2019.
[2] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, and Q. V. Le, “XLNet: Generalized Autoregressive Pretraining for Language Understanding,” arXiv:1906.08237 [cs], Jan. 2020.