Team: DL Rookies
Project Title: Adapter on RoBERTa
Project Summary:
Pre-trained language models can be fine-tuned to achieve good performance for most natural language tasks. However, due to the potentially extremely large size of parameters, it is time-consuming and challenging to train multiple-task models, as well as store and share trained models. It is proved that domain and task adaptive pretrainings of RoBERTa models before task fine-tuning are able to produce better results [1]. Comparatively speaking, adapters require fewer parameters, allow sequential training, and they are modular and easily shareable. This project aims to leverage adapter transformers [2] to create domain and task specific adapters, and explore whether they are able to produce similar performance as the domain and task adaptive pretrainings prior to the standard fine-tuning [1].

Approach:
Based on our preliminary research, adapter-based architectures have the potential in adapting the pre-trained model to downstream tasks of specific domains. We intend to build our own adapter using AdapterHub [2][5] and perform domain adaptation described in [1] on the same computer science domain corpora. We expect to obtain at least competitive performance compared to the fine-tuned RoBERTa as in the paper [1].
As a stretch goal, we aim to modify our adapter and training strategy to achieve competitive performance compared to domain-adaptive pre-trained and task-adaptive pre-trained RoBERTa, and compare datasets on other different domains.
Another stretch goal is to conduct an ablation study on the sizes and configurations of adapters, and explore how they affect task performance.

Resources/Related Work:
[1] "Don't Stop Pretraining: Adapt Language Models to Domains and Tasks", Gururangan et al.
[2] "AdapterHub: A Framework for Adapting Transformers", Pfeiffer et al.
[3] "AdapterFusion: Non-Destructive Task Composition for Transfer Learning", Pfeiffer et al.
[4] "Attention is all you need", Vaswani et al.
[5] Pytorch RoBERTa https://pytorch.org/hub/pytorch_fairseq_roberta/
[6] "Learning multiple visual domains with residual adapters", Rebuffi et al.
[7] "Parameter-Efficient Transfer Learning for NLP", Houlsby et al.

Datasets:
ACL-ARC Measuring the Evolution of a Scientific Field through Citation Frames

SciERC [Multi-Task Identification of Entities, Relations, and Coreferencefor Scientific Knowledge Graph Construction](#)
IMDB [Large Movie Review Dataset](#)
HELPFULNESS [Amazon Review Data](#)

Team Members:
Yifan Liu
Anni Zhou