

Group Proposal

DATS 6203 – Natural Language Processing

Nov. 5th, 2021

Group members: Abdulaziz Gebril, Yuchen Ma, Yuan Meng

We decided to work on the Glue Benchmark problem, which contains Single-Sentence Tasks, Similarity and Paraphrase Tasks, and Inference Tasks. It is a widely used benchmark, so we can solve the same problem with different teams from all around the world to get more inspiration. The Glue Benchmark datasets include The Corpus of Linguistic Acceptability (CoLA), The Stanford Sentiment Treebank(SST-2), Microsoft Research Paraphrase Corpus for Similarity and Paraphrase Task (MRPC), Semantic Textual Similarity Benchmark (STS-B), Quora Question Pairs (QQB), MultiNLI Matched (MNLI), MultiNLI Mismatched, Question NLI (QNLI), Recognizing Textual Entailment (RTE) and Winograd NLI, Diagnostics Main (WNLI).

We will pick Pretrained NLP(Transformers Based Networks). It is a classical model, so we don't need to customize it. Also, we are planning to use packages of Pytorch, Keras, and Tensorflow due to their ability to build and customize neural network models.

The NLP tasks we are working on are listed below.

The Corpus of Linguistic Acceptability (CoLA), The Stanford Sentiment Treebank(SST-2), Microsoft Research Paraphrase Corpus for Similarity and Paraphrase Task (MRPC), Semantic Textual Similarity Benchmark (STS-B), Quora Question Pairs (QQB), MultiNLI Matched (MNLI), MultiNLI Mismatched, Question NLI (QNLI), Recognizing Textual Entailment (RTE) and Winograd NLI, Diagnostics Main (WNLI).

We will use different metrics for the related tasks performed. These metrics will include Accuracy, F1-score, Matthew's Corr and Pearson-Spearman Corr.

Here is our rough schedule for completing the project.

Task	Date
CoLA & SST-2	11/12/2021
MRPC, QQP & STS-B	11/19/2021
MNLI , QNLI, RTE & WNLI	11/30/2021
Report and Presentation	12/09/2021