

# GLUE

Individual Final Report

**Jongchan Kim**

## Introduction

The General Language Understanding Evaluation (GLUE) benchmark is a collection of resources for training, evaluating, and analyzing natural language understanding systems ([source](#)). The GLUE benchmark consists of nine natural language processing (NLP) tasks, all of which contain sentences or sentence-pairs used to assess model performance across a wide range of linguistic phenomena found in natural language. A leaderboard exists to compare model performance against fellow contestants across the globe. As a group, we tested various NLP transformer models with varying hyperparameters to test their capabilities and learn to work with the hugging face *transformers* module.

## Description of your individual work.

Fine tuning of RoBERTa and DeBERTa was performed on STSB and QNLI dataset.

Also prepare and present on BERT, RoBERTa, and DeBERTa architectures

Add some conclusions

## Describe the portion of the work that you did on the project in detail

Presentation file 5 – 11 pages (BERT, RoBERTa, DeBERTa)

Fine tuning of RoBERTa and DeBERTa was performed on STSB and QNLI dataset.

Different batch sizes and learning rate

## Results

Fine tuning of RoBERTa and DeBERTa was performed on STSB and QNLI dataset.

DeBERTa	QNLI	Acc: 0.92
RoBERTa	QNLI	Acc: 0.92

DeBERTa	STS-B	Pearson: 0.90, Spearman: 0.90
RoBERTa	STS-B	Pearson: 0.91, Spearman: 0.90

### Summary and conclusions

RoBERTa and DeBERTa showed pretty good but similar performance.

Batch size and epoch was not enough to maximize the performance of both models, especially DeBERTa.

**Calculate the percentage of the code that you found or copied from the internet.**

25 %

### References.

- [1] He, Pengcheng, et al. "Deberta: Decoding-enhanced bert with disentangled attention." arXiv preprint arXiv:2006.03654 (2020).
- [2] Liu, Yinhan, et al. "Roberta: A robustly optimized bert pretraining approach." arXiv preprint arXiv:1907.11692 (2019).
- [3] Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." arXiv preprint arXiv:1810.04805 (2018).