

Group 4 Project Proposal

Mariko McDougall, Arathi Nair, Stefani Guevara

What problem did you select and why did you select it?

We chose to use the GLUE benchmark dataset because we think it could be challenging and we're interested in competing against others in the class.

What database/dataset will you use?

We'll be using the GLUE Datasets (9 total tasks).

What NLP methods will you pick from the concept list? Will it be a classical model or will you have to customize it?

We'll be using XLNet, Electra, and DeBERTa, using the transformers via Hugging Face. We'll use the classical version of the models.

What packages are you planning to use? Why?

We're planning to use the transformers package, datasets package, in addition to numpy, os, torch, etc. These packages, particularly the transformer package will allow us to quickly download the models and pretrained networks and jump-start of the project.

What NLP tasks will you work on?

We'll work on text classification on all 9 GLUE tasks.

How will you judge the performance of the model? What metrics will you use?

The performance of the models will be judged using the accuracy metrics respective for each task, as shown by the below code:

```
metric_name = "pearson" if task == "stsb" else "matthews_correlation" if task == "cola" else "accuracy"
```

Provide a rough schedule for completing the project.

Nov 12: Read Background information; initial setup

Nov 1: Write Section 1 and section 2 of the report

Nov 19: Try Basic LSTM and Transformers

Nov 26: Lock in models

Dec 3: Hyperparameter tuning done

Dec 3: Most writing in progress

Dec 5: Finish report

Dec 9: Presentation & Report Due