

README | MAYN Semantic Relations Software Project

Structure of this Repository:

Codebase

Codebase contains all our used codes for our project.

Preprocessing

In preprocessing there are two codes which we have used to restructure our raw data so that we can use it further.

sat_z_umformer.py
semeval_reader.py

Zero-Shot:

Zero-shot contains all the codebases that we used for our zero shot testing.

Performance

bert_predict_class.py
softmax_and_log_dicts.py
LOG_and_SOFTMAX_result.ipynb

Interpretability

token_importance_visualizer.ipynb

Fine-Tuning:

Performance

semeval_multiclassification.ipynb

Interpretability

Yi-Wan Code

Data

SemEval

In the SemEval folder, all the data we pulled from the original SemEval-2010 Task 8 [here](#) are included. We use training data to extract verbs for evaluating our zero-shot MLM and to fine-tune our Multi-Classfier Neural Network. We use testing data to evaluate performance.

Zero-Shot

Performance

Zero Shot Results Log csv:

Zero Shot Results Softmax csv:

Confusion Matrix rmd

Confusion Matrix html:

Interpretability

Nilu Data

Fine-Tuning:

Performance

Adri Data csv

Matthias confusion matrix rmd, html

Interpretability

Yi-Wan shapley data