List of tasks (**Unofficial**)

1. Train LSTM with custom and glove embs on a Toxicity detection dataset (binary) **~2 points**
2. Train a BERT-alike model on STSb dataset (semantic similarity) **~2 points**
3. Create custom RAG mechanism **~3 points**
4. Collect via prompt engineering a list of answers on questions for deep learning providing two answers one gentle and one toxic (Gordon Ramsay style) **~1 point**
5. Tune an LLM using DPO to always act like Gordon Ramsay! **~2 points**

1. Toxicity dataset

https://github.com/Sreyan88/Toxicity-Detection-in-Spoken-Utterances/blob/main/data

Contains already a train.csv, valid.csv and test.csv (ignore trigger_test.csv)

Use all examples (don't remove for balance)

Compute F1-score and Confusion Matrix for both LSTM with custom and glove embs

Use Tensorflow/Keras