Michael Gardner
Information Extraction
Assignment 3 Proposal

I'd like to use an LLM as the embedding (not sure which LLM yet, because I need to experiment which I run on my hardware in a timely manner, probably distilbert)
The LSTM we built was the best performer for me in assignment 1 and 2, so...
With that, I'd like to do an experiment of 3 different feature ideas:
1   Something simple with the just the 2 terms as features (likely to perform poorly)
2   Just the words between the 2 terms as features
3   The full sequence as features.

My hypothesis is that #2 will do the best, but I'd like to see.

And as an after-thought, if there's time, I can rerun with bigrams.