Yorick Chern
Homework 5 Report

Originally, I wanted to try out methods utilizing NLTK's part-of-speech tagging. However, that method failed miserably, and I decided to take a more straightforward route. For my diversity construction, I tackled the one problem the similarity method was facing–duplicates. Right before I performed the nearest neighbors selection algorithm, I preprocessed the training examples to ensure that there were no duplicates. Once all duplicates are removed, I added the embeddings to a list where cosine similarity is run. Then, the predictions continued as usual. This time, I was able to reach a 65% exact match with the diverse prompts, and this selection algorithm outperformed all other selection algorithms.

Skimming through the predictions, I noticed that the responses from the similarity promptings are very verbose, such as: The states that border m0 would be determined by using the "next_to_2" function to find the states… Therefore, the answer to the prompt would be the intersection of the set of states and the set of states adjacent to m0, which is represented by "intersection (state , next_to_2 (m0))" and "Answer: The biggest/largest city in m0 is not specified as there is no information provided about the cities in m0.". There are also a few responses that simply said it cannot infer further without additional information. I think this is a direct result of the similarity prompting having lots of duplicates.

Although the random sampled predictions seem more aligned with the representation we are targeting, the performance is actually worse than that of similarity prompting. In the end, the diversity prompts win because it combines the best of both worlds of random samples and similarity. It retained information from prompts similar to the utterance without the duplicants.