NLP Final Paper

By Ekta

In the module 2 of Week 2 we talked about Word Relatedness (or Word Association). We used this concept as base and incorporated it by using PMI (Pointwise Mutual Information) in our Project 1. Pointwise Mutual Information (PMI) is a statistical measure commonly used in Natural Language Processing (NLP) and information theory to quantify the association between two words in a corpus of text. It helps identify how often two words co-occur together in comparison to how often they would be expected to appear together by chance. We used Brown corpus and calculated PMI for each pair of words. We then used this information to check PMI for the pun sentence to get the location. However, one key issue which we faced in this approach was that brown corpus was not really a Pun related dataset so it did not have lot of pun words. The paper that we referenced for this approach augmented the dataset with Pun of the Day and some other datasets but due to time constraints we were not able to implement those. In the same module, we also learnt about word similarity which we also used as one of the metrics for pun location in project 1. We used cosine similarity between word embeddings to find the two most similar words to narrow down the potential list of pun words. In the Week 3 module we talked about Lesk Algorithm. Lesk relies on dictionary definitions of a word of interest and all dictionary definitions of the words surrounding it, known as a context window and picks the definition that has the highest overlap and obtains the sense of the target word. We used this algorithm for pun interpretation. We later moved to use Adapted Lesk algorithm which uses an online lexical database called WordNet where words are arranged semantically and grouped together into a relation called synsets because this allows us to capture a much wider range of word definitions and relationships than the original Lesk algorithm.

In the guest lecture by Geetanjali on the topic of Transformer Language Models, we discussed in detail about Transformer Architecture and also looked at Bert and Attention Layers. In Project 2, we used them for location and detection of heterographic puns. We reproduced the results from paper “The Boating Store Had Its Best Sail Ever”: Pronunciation-attentive Contextualized

Pun Recognition which used Bert Contextual encoding layers that captured the word and position information and phenome layer to capture phonetic information. It also used attention layers to capture the necessary information at both embedding and phenome layer. Our biggest motivation for picking this paper was to see the impact of attention layers on the NLP Tasks like pun detection and location.

For Project 3, we used the in-class activity by Damin on Alexa Skill Integration for our work on Alexa. However, instead of creating separate Intents, we created a global intent to handle our conversation flow. We found that using this approach gave us more fine grained control over the conversation and it also helped user have a dialog with Alexa.