Six Wires
Karen Mazidi
CS 4395
September 3, 2022

Assignment 0

Overview of NLP

Natural language processing is the study of machine learning algorithms that interact and interpret human language patterns for various fields. Natural language processing is a subfield within artificial intelligence. We use AI techniques to process large sections of text to predict our chosen target.

Both natural language understanding and natural language generation deal with interactions with human language. Natural language understanding takes something like a document or a data set and uses it to process, train, and understand the given text. On the other hand, natural language generation takes some sort of input to produce a natural language blurb.

Some modern applications for NLP are translation apps, censor bots on social media, document scanners, tech support help bots, and much more!

The first and oldest approach to NLP is the "Rules Based" approach. In this approach, the algorithm has a set of rules it must follow to make predictions. It is less focused on learning over time and instead can only "respond" within a finite set of responses. Some examples of this are spell check and the Eliza chatbot. Within the chatbot, you can only get a set of prewritten responses, and the algorithm doesn't necessarily "understand" what you're feeding it in, rather it attempts to respond based on key words that are linked to rules. The overall issue with this implementation is that it can't understand the complexities of human language.

The second approach to NLP was the statistical and probabilistic approach. This implementation coincides more with basic machine learning. It uses statistical and probabilistic algorithms to build a model to make predictions. Now instead of rules, this approach is able to learn from the data fed to the algorithm. The only problem with this approach is that it requires a large enough data set to train the model well.

The third and most modern approach to NLP is deep learning. Instead of using traditional ML algorithms, we use deep neural networks to help improve upon statistical and probabilistic approach. This is used in things like real time translation, where the algorithm has been learning with large sets of data and performs well in action. The main issue with this is we need very large data sets to train the network on.

I hadn't had much of an interest into NLP until I took Machine Learning over the summer. I really enjoyed learning about how to create models and make predictions based on the

data fed in. I would like to use NLP to make some sort of text prediction algorithm similar to the text prediction on the iPhone.