

Manuel Salado Alvarado
CS 4395.001
Professor Karen Mazidi

Natural Language Processing: A Brief Overview

I agree that Natural Language Processing (NLP) should be renamed to Human Language Technologies because being natural doesn't imply being a human when human language is the primary application of NLP. As far as I know, we are not using NLP to decipher animal noises to add meaning to it. Anyways, I believe NLP can be defined as algorithms developed to better allow the computer to decipher the nuances of the human language in order to extract meaning or generate a response [1].

Both Machine Learning and NLP are branches of Artificial Intelligence. Sometimes the solution to the problem domain is strictly NLP, or sometimes strictly ML, but it is likely to include some AI techniques in the mix regardless [1].

[According to IBM](#), Natural Language Understanding uses “[grammatical structure of the sentence] and semantic analysis of text and speech to determine the meaning of a sentence.” Which is juxtaposed with Natural Language Generation because NLG is the system emulating a human response by summarizing a body of text, or responding as a chatbot [2].

Some examples of NLP applied are Siri, Alexa, ok Google, any chatbot, the power summary of long emails on outlook, the automated support chats on websites, etc [2].

NLP first started as a rules-based approach in which the system is designated a set of rules to follow in order to parse input text. For example, plural nouns are capitalized, or adjectives are stuck next to nouns. This would be more useful if our human language is scalable, but it is not. There are many nuances regarding human speech and a set of rules cannot fully visualize the regional dialects, sarcasm, inside jokes and the likes of them [1].

Statistical and probabilistic approaches were initially around in the 1980's in which a mathematical approach was taken to gather the probabilities of certain words occurring in a sentence. This was used as part of a larger system or as a dedicated counting of words in a single body of text. This vein of NLP is similar to ML in that it uses statistical analysis to predict an outcome based on the input variables. An example of this would be that “I like my eggs poached” which a rules-based approach might infer that I enjoy thieving eggs from a farm in the dead of night whereas a statistical approach might notice a pattern in which poached eggs are a manner of cooking eggs [1]

Deep learning is more widely used when there are very large data sets available because they evolved from neural networks which require a lot of input data to set the weights on the nodes. Recurrent neural networks, convolutional NN, LSTM's and more are all denoted by Mazidi in [1]. Often times, all three approaches are used in unison.

My personal interest in NLP can be sourced in my sister's occupation. She is currently in school to get a Masters in information technology so she can become a librarian. It would be invaluable

to her if I could help her sort through bodies of text so that they could be better represented at her workplace. Imagine if research papers that don't have abstracts could get an abstract generated and then lightly tinkered. Other than that, I like the idea of using NLP for video games in which the cookie cutter responses that NPC's give you are a little more varied because they were produced from a response the user gives. I also really liked this one game when I was growing up where you were supposed to type in what you wanted to do such as "move forward" "open chest" "Use magic" and I can imagine the potential for a game like that in which everything is created on the fly from a database of a thousand fantasy books. On a more realistic note, it would be interesting to learn more so that I can better understand how the human language works in order to use it to learn a new language and contrast the NLP differences.

[1] Mazidi, Karen. *Exploring NLP with Python: Building Understanding through Code*

[2] <https://www.ibm.com/blogs/watson/2020/11/nlp-vs-nlu-vs-nlg-the-differences-between-three-natural-language-processing-concepts/>