

Amol Perubhatla

AVP180003

CS4395.00

a). NLP is how we program or teach computers how to understand human languages.

Due to the complexity of human language and the structured nature of computing, teaching computers how to understand human language is very difficult. This is why many researchers have spent countless hours devising and creating new methods to program computers to perform this difficult task.

b). Artificial Intelligence is becoming an increasingly appealing tool to perform Natural Language Processing as traditional methods such as rules-based programming for processing human language are tedious and time consuming to program and even when properly done may not even be 80% accurate due to the nature of the human tongue. This is why training a computer with artificial intelligence to learn language much in the same way that humans learn language as children is seen as an apt tool for surmounting the gap that traditional Natural Language Processing tools lacked. With AI you build a model and feed it data so that it learns the grammar syntactical structure and vocabulary all on its own without requiring extensive oversight.

c). Natural Language Understanding is teaching a computer how to comprehend the meaning of any text, whereas Natural Language Generation is concerned with creating new text often this means summarizing an input document without losing the original meaning or generating new text with certain inputs.

d). Some good examples of modern Natural Language Processing applications are Google Translate. Google Translate is an excellent tool that can take language in one language and translate it into another language. This is a tool that has made communication internationally much easier. Another example of an NLP application is the auto-generated subtitling on YouTube videos. Instead of having a human perform the tedious task of watching a video and timing and transcribing the words spoken within. The computer can perform the task much faster and remarkably accurately.

e). The first and most antiquated method for performing NLP is something known as "rules-based" processing. This often requires extensive and tedious work by the programmer to build and test many rules to process language. A lot of work is put into matching patterns and identifying them based on the already programmed rules. If a pattern that wasn't programmed for is detected then you may receive a canned response or the program will fail to function properly. This is why rules-based NLP often fails when generalized. Examples of this approach includes regular-expressions and context-free-grammars.

The second method is machine learning which is where you design a model and test it by giving it input and outputs that it trains on to build a model that can respond to those inputs. Then you test it on data that it wasn't given previously to ensure that it isn't overfitted to the test data and can be generalized on new and novel text that it hasn't seen before. Applications for machine learning NLP include text inference, which is when apple suggests the next word in your text messages before you have even started typing it.

The third method is Neural Networks, this is the newest and most interesting method of Natural Language Processing. Neural Networks are designed to mimic the structure of the human brain. We already know that the human brain is really good at picking up new and disparate things. By building a model that mimics the neural network that sits between our ears, we can train a neural network for one task and apply it to another. An example of this is GPT-3 which is a neural network that has countless applications including NLP and image recognition.