

Overview of NLP

[Download as PDF](#)

What is NLP?

NLP Stands for "Natural Language Processing." It is a branch of computer science and linguistics pertaining to the ability for computers to comprehend human language, whether it be text/speech/etc. The study of NLP goes as far back as the 1950s yet is still under thorough research today.

How does it relate to AI?

NLP is a subset of artificial intelligence (AI) which specifically focuses on the *understanding* of human language. AI is an umbrella term that refers to an entire set of systems that may emulate human intelligence.

NLU vs NLG

Natural language understanding (NLU) and natural language generation (NLG) are related in nature, yet have distinct classifications within NLP. NLU focuses on reading comprehension whereas NLG focuses on writing ability.

NLU utilizes the language's grammatical structure and semantic analysis to ascertain the meaning of a given sentence. Additionally, it captures the relationships between words/phrases in an ontology to better understand the intention behind a sentence.

NLG generates text (or speech) based on the given input data. Similar to NLU, NLG needs to take grammatical structure and syntax into account in order to properly formulate new sentences.

Modern Applications

Some common modern day use cases of NLP:

- Language translation
- Speech recognition

- Text autocomplete/autocorrect
- Text classification/sentiment analysis
- Virtual assistants/chat bots

Approaches to NLP

1. Rules Based

Rules based is one of the oldest approaches to NLP still used today. It works by creating a set of well defined rules pertaining to the given language and then matches the input to the set of rules to achieve the desired output. Rules based only works well in situations that can be fully encapsulated by a finite set of conditions. For example, a grammar or part of speech checker would work well in a rules based approach because they are typically defined by specific guidelines. Things like human language, on the other hand, are not good candidates for a rules based approach because human language is complex and constantly evolving.

Rules based approaches are "dumb" in nature wherein they don't truly develop an understanding of the given input. It simply searches for a condition and acts appropriately. With that being said, it is still used today because in situations where a set of rules is adequate enough to encompass the domain, rules based is faster and simpler compared to the alternative approaches.

2. Statistical/Probabilistic

Statistical approaches are the most widely used in NLP and work by attempting to apply mathematical principles to languages. It works by analyzing the data and generating statistics/probabilities for letters/words/phrases/etc given an input. This approach overtook rules based approaches because the computer learns based on input data rather than explicitly defined rules. It permitted solutions of higher scale which led to applications such as machine language translation or text prediction.

Machine learning is an example of a statistical approach that is widely used in all forms of AI. The machine learning algorithms learn by applying statistical and probabilistic methods to data and attempt to generalize the information to a broader scope. There are various different types of machine learning algorithms such as support vector machines (SVM) and neural networks.

3. Deep Learning

Deep learning based approaches directly stemmed from statistical approaches. As stated before, machine learning algorithms take an entirely mathematical (probabilistic) approach to artificial intelligence but are mainly associated with moderate to large amounts of available data. Deep learning, on the other hand, are neural networks that are built to handle huge amounts of data at

scale. Large amounts of hardware resources are required to process deep neural networks which makes it not as accessible as the other approaches, however, the results speak for themselves. Deep learning is the leading approach towards achieving a general intelligence AI, and is used in many NLP applications.

Today, NLP mainly utilizes deep learning techniques to achieve more realistic sounding speech synthesis. In the past, humans could, with relatively high certainty, tell whether a human or a robot was talking. With new approaches, speech synthesis has gotten good enough to fool people. In fact, in 2018 Google released an AI assistant that could make phone calls to local businesses wherein the business and the AI could hold an actual conversation with each other. (Though the assistant was still not truly intelligent but rather purpose built for that situation)

Personal Interest

I am interested in learning about and utilizing NLP because I see it as a major step toward bridging the gap between computers and humans. Computers only know/do what we tell them; they lack the ability think in the way that we do. However, NLP allows computers to begin to develop an understanding of human language and grants a means of natural communication that has never been possible before. While computers still aren't at the level of general intelligence, NLP is a great step forward to get to that point. NLP already has countless applications today and the use cases are bound to grow as more research is done.

References

- [1] I. Roldós, "NLP, AI, and Machine Learning: What's The Difference?," MonkeyLearn Blog, Jun. 09, 2020. <https://monkeylearn.com/blog/nlp-ai/>
- [2] IBM Cloud Education, "What is Natural Language Processing?," Ibm, Jul. 02, 2020. <https://www.ibm.com/cloud/learn/natural-language-processing>
- [3] E. Kavlakoglu, "IBM - United States," www.ibm.com, Nov. 12, 2020. <https://www.ibm.com/blogs/watson/2020/11/nlp-vs-nlu-vs-nlg-the-differences-between-three-natural-language-processing-concepts/>
- [4] A. Sharma, "Applications Of Natural Language Processing (NLP)," Analytics Vidhya, Jul. 08, 2020. <https://www.analyticsvidhya.com/blog/2020/07/top-10-applications-of-natural-language-processing-nlp/>

[5] "11 Natural Language Processing (NLP) Applications in Business," MonkeyLearn Blog, May 20, 2020. <https://monkeylearn.com/blog/natural-language-processing-applications/>

[6] K. Mazidi, Exploring NLP with Python | Building Understanding Through Code. 2019.