

## Overview of NLP

### **a. define NLP in your own words**

NLP or Natural Language Processing is the machine processing of human language. An example of NLP, is when a machine is given English characters and derives meaning and context from the words, just as we humans do when communicating. NLP still fails to grasp the full complexity of human language, however, and can run into some difficulties when it comes to idiomatic language, sarcasm, and hidden motives.

### **b. describe the relationship between AI and NLP**

NLP is a subfield of AI and NLP programs often use machine learning, specifically in statistical and deep learning approaches. This allows the program to train on a moderate to large amount of human language data and then find patterns. Using these patterns it is able to process and generate human language.

### **c. write a sentence or two comparing and contrasting natural language understanding and natural language generation**

When two humans speak with each other they use natural language understanding to understand what the other person is saying and natural language generation to formulate spoken responses. They go hand in hand in order to create effective communication.

### **d. list some examples of modern NLP applications**

Chatbots, Tiktok Algorithm, Google/Alexa Voice Assistant, Email spam filters, Language translation apps, Sentiment analysis, Predictive text, and Search results are all examples of modern day NLP applications .

### **e. write 3 paragraphs describing each of the 3 main approaches to NLP, and list examples of each approach**

The oldest NLP technique is the Rules-based approach. The Rules-based approach simply uses a set of predefined rules, created by a programmer, to process and generate human language. Some examples include a shallow chatbot called Eliza that used regular expressions and hard coded responses, grammar checkers using predefined rules, and a spam email identifier using predefined rules. A disadvantage to the Rules-based approach is that it doesn't scale up to the complexity of the human language.

In the 1980s a Statistical approach to NLP was developed. This approach calculates the frequency and probability of words from a corpus to form a language model which it uses to process and generate human language. Some examples include Google translate, predictive text, and entity recognition. The disadvantage of statistical approach is that it requires a moderate amount of data and good processing power.

In the 2010s the Deep Learning approach became prominent and was regarded as the most powerful approach. The Deep Learning approach involves training large neural networks on a huge corpus to process and generate human language 'intelligently.' Some examples of the Deep Learning approach include chatGPT, a chatbot which can consider the context of a conversation and remember what a user has previously said, and a sentiment analysis system, which uses neural networks to learn relationships between words and emotions. The Deep Learning approach significantly improved results in language processing and generating, however it comes at the cost of needing a large amount of data and powerful processing. In specific situations the rules based and statistical approaches can outperform deep learning and many modern day NLP projects will use a combination of all three approaches.

**f. write a paragraph describing your personal interest in NLP and whether/how you would like to learn more about NLP for personal projects and/or professional application**

I first became interested in NLP when I messed around with chatGPT. I realized that my job might be taken by this thing in the future so it would be in my best interest to learn the inner workings or at least the foundations of how something like chatGPT is created. I think it would be cool to learn NLP to create personal projects like my own spam email filter or chatbot. I'm not sure if I would want a career in NLP yet, but I am open to the idea and excited to learn more about AI in the future.