

NLP, also known as Natural Language Processing, is the branch of computer science that involves the creation of algorithms that let computers process human languages in a way that allows them to understand text and spoken words like how humans can.

The relationship between AI and NLP is that NLP is a branch of AI that allows for machines to understand human languages. They both can focus on using natural language as the means to interact between machines and humans. As in, AI uses NLP to understand and translate the human language.

Natural Language Generation (NLG) and Natural Language Understanding (NLU) are both subsets of Natural Language Processing, where the applications of both NLG and NLU need to take into consideration language rules, which are based upon semantics, syntax, lexicons, and/or morphology, in order to make decisions on how to have the appropriately phrased response. However, natural language generation is the process for computers to create written responses, based on a dataset, in a human language, while natural language understanding is focused on computer reading comprehension.

Some examples of modern NLP applications include machine translation, automated email reply suggestions, spam detection technologies, chatbots, virtual agents, and text summarization.

Of the 3 main approaches to NLP, one historical approach, which was popular in the 1960's up until the 1980's, is the Rules Based approach. This approach uses rules that are applied to text or techniques such as regular expressions and context free grammar to tackle NLP related tasks by focusing on pattern matching or parsing. Since human languages are complex, using this approach in a scaled-up way was difficult and often had low precision but had high performance in specific situations. However, it can be used for text processing purposes. Some examples of Rules Based approach include regular expressions, context free grammar, and the Eliza Therapist Simulator.

The 2<sup>nd</sup> approach to NLP is one that became popular in the 1980's, known the Statistical and Probabilistic approach. This approach involves finding the probabilities and sequences of words as well as counting them, where these statistical and probability methods led to language models that can be used by machine learning algorithms. This approach is useful when used with moderate to large amounts of data that is available for training. This training happens using a model that has defined parameters, which is followed up with it being fit on test data. Examples of this approach include language models used for translation programs, SVM, Decision Trees, Logistic Regression, Naive Bayes.

For the 3<sup>rd</sup> approach to NLP, it includes the modern approach known as deep learning. This approach can be similar to the mentioned 2<sup>nd</sup> approach, but it uses a large training data set, where into the neural network raw parameters are fed, and as these networks will learn important features, the manipulation of raw data to produce the wanted variables for a model is usually skipped. In addition, this approach was evolved from when neural networks had

access to large amounts of data and the needed increased processing power to handle it. Some examples of this approach include recurrent neural networks, LSTMs, and convolutional neural networks.

I have a budding personal interest in NLP the more I learn about it. For now, the potential growth of this personal interest has yet to be determined. I am fascinated by the usefulness of NLP which lays the foundation for my personal interest in NLP. As for whether I would like to learn more about NLP for personal projects and/or professional application, the answer to that would be most likely to be a yes for now. As for how I would like to learn about NLP for personal projects and/or professional application, that would involve researching about it in my free time, as a hobby, in a nonacademic setting.