An Overview of NLP

Add a document "Overview of NLP" in which you:

a. define NLP in your own words

NLP also known as Natural Language Processing is defined as a subfield of computer science and artificial intelligence. As the name suggests, Natural Language Processing processes languages so that computers and machines in addition to humans can interpret human languages. Natural Language Processing facilitates the communication between the bodies (people or things) generating or speaking a language and the bodies involved in listening and understanding the language that is spoken. There are three important use cases in Natural Language Processing (NLP): (1) rules-based approaches,(2) statistical and probabilistic approaches, and (3) deep learning.

b. describe the relationship between AI and NLP

Natural Language Processing, like Machine Learning, is a subfield of Artificial Intelligence.

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

Natural Language Understanding (NLU) refers to the party that processes the spoken language by the other party directly speaking to them. Natural Language Generation (NLG) however, is the generation of the spoken language. The party that attempts to understand the spoken language may or may not always understand the language that is spoken. Whereas, the party that attempts to formulate the spoken language has a duty to make sure that other parties (people,machines, etc.) understand the spoken language. For example, the telephone game consists of a network of people. A person who is source of this network (i.e. Person 1), is the one who will form the phrases in the language. All of the people within this network have to repeat exactly word for word what the source of the network said, until the phrase reaches its destination within the network. In other words, the remaining people who did not form the language repeat what the generator(Person 1) of the language said until all the members of the network have heard what the generator said. Person n, for example is the one who attempts to repeat back what Person 1 said in the beginning. Person n in this case is the person responsible for understanding what Person 1 said and repeating the language back. In the case of the telephone language, it is completely possible for the people besides Person 1 to misinterpret what Person 1. This leads to Person n completely misunderstanding the language. Hence, it is very important that the party forming or generating the language communicates it properly with the parties attempting to learn the language.

d. list some examples of modern NLP applications

Examples of modern NLP applications include the Smart Assistants (i.e. Google and Alexa), and Digital phone calls, and Grammar checkers. Smart Assistants such as Google and Alexa can store basic commands that the user might give such as "Alexa, set an alarm for 6 AM." Alexa understands the natural language generation of text formed by the user, and proceeds to display an alarm set for 6 AM. When making a Digital phone call to a company or organization, the phone may say, "this phone call may be recorded." The reason for this is because the organization to which you call may want to improve the experience for other customers. Grammar checkers are also very important as they are far

more powerful in determining the correct structure and format for papers and essays. Grammar checkers use NLP techniques to find ways in which people write formal prose. In essence, this enables grammar checkers to check for spelling mistakes, sentence flow, punctuation mistakes, etc.

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

The first approach in NLP, also known as the rules-based approach. Out of the three main approaches, rules-based approaches are the oldest methods used within NLP. Examples of the rules-based approach include using regular expressions to convert words in plural form to singular words, using context-free grammar and its set of rules employed to generate grammatically and syntactically correct sentences. The Eliza chatbot is a popular application that utilizes this rules-based approach. The Eliza chatbot acts as a therapist that charges nothing for therapy based services. The Eliza chatbot listens to the words spoken by the user and responds back with the appropriate response that a therapist uses. Between the 1960s and 1980s, the rules-based approach was considered a very popular approach towards processing languages in NLP. A major issue with the rules-based approach is that the rules that generate the context free grammar for the language don't account for complexities and ambiguities within human languages.

In the 1980s, scientists and mathematicians began to shift the focus of NLP towards providing mathematical approaches towards developing various texts. This set of mathematical techniques is known as Statistical and probabilistic approaches. Statistical and Probabilistic approaches are used to generate language models by finding the probabilities of words and sequences of words. In addition, machine learning algorithms such as Naive Bayes, Logistic Regression, SVMs, Decision Trees, and Neural Networks also learn from statistical and probabilistic approaches. The statistical and probabilistic approaches needs excellent processing capabilities and plenty of data. Without these two aspects, it is very difficult to perform NLP with statistical and probabilistic approaches.

A modern approach to NLP is deep learning. Deep Learning is the evolution of the neural network. Examples of algorithms utilizing Deep Learning include the recurrent neural networks, convolutional networks, and LSTMs. Utilizing a deep learning approach results in improved language translation, improved generation of formed language, and improved understanding of languages. Despite the many advancements deep learning has to offer, deep learning requires greater processing capabilities and data as opposed to that of statistical and probabilistic 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 am a Senior in college at the University of Texas at Dallas. I recently took the Introduction to Machine Learning course with Dr.Mazidi last year and I thoroughly enjoyed using the R and Python programming languages to code Classification and Regression Algorithms including but not limited to Linear Regression, Logistic Regression, Naive Bayes, SVM, Decision Trees,kNN, and k-means clustering. In this course, I would like to learn more about the relationship between Natural Language Processing and Machine Learning, as well as how Machine Learning programs are used to enhance Natural Language Processing. I am very excited to not only learn about a new topic (Natural Language Processing), but also to gain more experience in regards to coding in Python language. My goal by the end of this semester is to not only solidify Python programming, but also add to a repertoire of projects that I have worked on in the past.