Overview of Natural Language Processing (NLP)

1. **Define NLP**
   - In general, NLP means creating algorithms that allow computers to process human language. These algorithms have to process human language usually in a form of speech or text. By doing so, the NLP software would be able to mimic 2 actives in a human-to-human dialog which is natural language understanding and natural language generation. Some of examples of NLP are: to listen and recognize the key word in the conversation, to provide a suggestion on the search box when user typing the keyword, etc.

2. **Relationship between AI and NLP**
   - NLP is a branch of AI, as is ML. In a complex NLP project, some component may be purely NLP, others purely ML, and others in the general category of AI.

3. **Natural Language Understanding vs Natural Language Generation**
   - Both natural language understanding and natural language generation happen simultaneously in a human-to-human dialog. "*Natural language understanding*" means that each party understood what other person said, "*Natural language generation*" mean create the formation of spoken responses.

4. **Examples of modern NLP applications**
   - To identifying spam words in emails
   - To create a word cloud based on word counts
   - To classifying product previews as positive or negative
   - To create an automated agents which are able to understand human speech with a wide variety of accents and regional distinction.

5. **Three main approaches to NLP**
   - *Rules-based approaches*: is the oldest techniques in NLP. As the name suggested, the techniques use rules to convert plural forms of words to singular ones which involve a few regular expressions and a list of exceptions. It can also involves context-free grammar, which lists production rules for sentences. However, the problem of this approach is it does not scale up to the complexity of human language. Some examples: spell check, context-free-grammar, Eliza chatbot.
   - *Statistical and probabilistic approaches*: use mathematical techniques to processing text. Simply couting words and finding the probalitites of words and sequences of word lead to the useful language model. However, this approach requires a moderate amount of data and good processing power. Some examples: predictive text as when you type a query into a search bar and received suggestions for the most likely phrase you are typing, can be used as a part of machine translation system.
   - Deep learning: use neural networks to process data when large amounts is available. It also requires a powerful processing though GPUs and cloud computing. Some of

algorithms can be used in deep learning is recurrent neural networks, convolutional neural networks, LSTMs. However, this technique will require a huge amount of data and powerful processing which not everyone has access to. Some of examples: improve language translation, create a human like automatic assistants, improve language generation and language understanding.

6. **Personal interest in NLP**
    -   I was always very curious about how digital assistants such as Siri and Alexa work. So I watched some videos on speech recognition on youtube. Most of them did a good on how to manipulate the speech and text by using API to communicate with the AI from website. However, they never go into detail on how the AI was built. The NLP give me an idea on how the actual programs were based on. I hope as the class advances during the semester, I could have a better understanding of the algorithms and techniques which are currently used to develop these digital assistants.