Overview of NLP

NLP (Natural Language Processing), a branch of AI (Artificial Intelligence), involves creating algorithms that allow computers to process human language. Processing in this case means taking linguistic input to give an output.

Natural language understanding deals with taking in input from a user while natural language generation deals with outputting a response.

Some examples of NLP applications include:

- Analyzing sentiment from customer reviews
- Automated assistant to address calls and requests
- Machines translations to translate from one written language to another
- Recommendation systems based on keywords from search or purchase history
- Automated email reply systems

There are three main approaches to NLP:

1. Rules-based Approach - 1960s: This approach has a programmer create rules using regular expressions and exceptions for the program to follow depending on the input. The problem with this approach is that these rules tend to not scale up to the complexity of human language. E.g: spell check, context-free grammar, Eliza chatbot.
2. Statistical and Probabilistic Approach - 1980s: This method involves counting words and finding their probabilities and sequences of words. This approach is where classical Machine learning falls into, and can outperform Deep Learning on smaller datasets. This approach has the problem of needing a moderate amount of data and processing power to be functional. E.g.: Machine Translation, Predictive text in search bars, more sophisticated chat boxes.
3. Deep Learning - 2010s: It evolved from neural networks when huge amounts of data became available, and processing power increased through GPUs and cloud computing. This approach has improved results in language translation, generation, understanding, and much more. However, it also needs more data and processing power than the statistical and probilitics approach (classical machine learning) and has way too much hype surrounding it.

My interest with NLP came along with my interest with AI and all of its related fields. Then I became more interested when, right before the beginning of this semester, I thought about ways of reforming the orthography of the English language. This increased my curiosity with NLP because I think it can be used for that purpose. What I think I would do with NLP technology is I would feed the system with words from a dictionary and their pronunciation according to the IPA. Based on some basic rules, it would decide if the spelling is okay or if it needs to be changed.