# An overview of Natural Language Processing from a relatively inexperienced student.

This document/blogpost/markdown sheet/assignment should not be a document about what natural language processing is. The writing here should not be used as factual information, but as an expression of my opinions.

To obtain a pdf of this page, click here. Please note that this is not a 1 to 1 representation of this page and is rendered using `pandoc` and `texlive`.

### Natural Language Processing, An overview

Natural Language Processing is a field of computing that is interested in the computational processing of human language. It is a subset of artificial intelligence. Where general artificial intelligence is concerned with general concepts such as image recognition, pathfinding, and other organic concepts, Natural Language Processing is specific with it's focus in the spoken and written language. It is not Natural Language Generation, where language is generated for use in some other application (eg: ChatGPT), but they are similar in the fact that they both require understanding of the underlying language being written/said (ex: "ies thue conroy desion or belefus?" is not a valid, understandable sentence). Some modern examples of Natural Language Processing are the input and contextual aspect of ChatGPT, paper summarization AIs like QuillBot, and Dall-E/Stable Diffusion's text input.

### The Approaches To Natural Language Processing

The first approach to Natural Language Processing was a "rules-based" approach, where a computer would use some predefined rules in order to understand what was being said. This proved useful in primitive examples, such as spell checking and context-free grammars, but did not scale up to actual human speech. The next advancement was using statistical models that relied on automatically computed weights and word frequency which did prove more useful than the rules-based in getting contextual language understanding. This needed even more processing power than the "rules-based" method. The third, and currently most modern method, is deep learning, which has drastically improved the translation and understanding of language using "a large pile of linear algebra." This requires immense processing power in order to actually compute an answer.

### My Personal Interest in Natural Language Processing

This has long been one of my own personal desires as I have wished I would be able to understand some form of human input beyond a simple "(Y/N)" input in my programs. I want them to be actually smart, and not some glorified decision tree. I also want to find a way to turn sentences into media, as Meta, Google, and OpenAI have, as I suck at doing my own generation.