Week 3: Influencing Thought for Fun or Profit

Nate Bachmeier

TIM-8150: Artificial Intelligence

October 4, 2020

Northcentral University

# Influencing Thought for Fun or Profit

The NCU Political Action Committee (NPAC) seeks to promote its ideology through highly targeted censorship. Modern censorship does not restrict free speech; instead, it increases the noise and drowns the signal (Thomas, 2019). One critical challenge the organization faces is its ability to scale-out personalized communications with potential voters. Traditionally, businesses approach these problems by either hiring armies of people or resorting to mass marketing campaigns. However, NPAC lacks funding to employ a large staff, and modern spam filters reduce email blasts’ effectiveness. Instead, NPAC chooses to modernize its tactics and focus solely on Natural Language Processing (NLP) and social media graphs. “[…] NLP is an interdisciplinary field [that] studies and develops algorithms and systems, enabling computers to understand and perform tasks involving human language (Sintoris & Vergidis, 2017, p. 135).” NPAC plans to use these technologies to both parse free form text and also produce novel commentary. Maximizing the resonation of custom content with the audience requires a personalized voice. For instance, the manner that people speak in an academic forum differs from Facebook or Twitter. NLP language models can assist in these situations as well by adopting different vocabularies and alternative sentence structures.

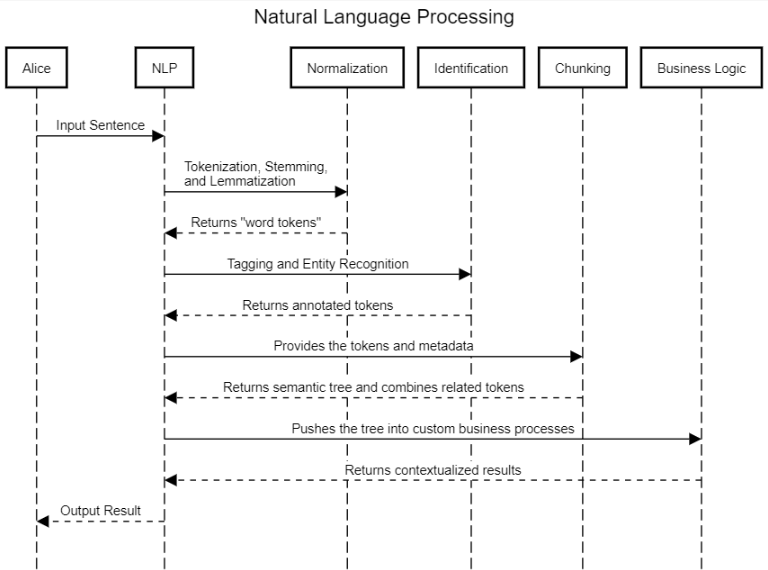
# How does NLP work

Natural Language Processing (NLP) sits at the intersection of artificial intelligence, human language, and computer science.

## Text Ingestion Process

NLP systems typically follow sentence normalization, token annotation and combining, and finally perform custom business logic (see Figure 1) (Edureka, 2018). Using strategies like Lemmatiziation and Stemming enables the parsers to reduce the variability between sentences, such as removing verb-tensing. Next, the words collect annotations by subsystems like Named Entity Recognition (NER) to discover the sentence’s critical components. After chunking related tokens together, the scenario-specific business logic can operate on a semantic representation of the text. Depending on the use-case, these steps will be massive subsystems or single lines of code.

Figure 1: NLP Process



## Business Logic

NLP appears across a wide range of use cases like language translation, speech-to-text, and sentiment analysis. In biology, animal brains accomplish these tasks through meshes of neurons that transmit signals across connected synaptic (transforming) and activation (filtering) links (Keller, Liu, & Fogel, 2016). Computer scientists mimic this behavior with Deep Learning on Neural Networks, which are essentially weighted graphs. Depending on *the* *architecture* of the graph, the system can perform tasks of varying complexity. For instance, sequence to sequence (seq2seq)

Table 1: Example Algorithms of NLP

|  |  |
| --- | --- |
| Algorithm | Description |
| seq2seq | Recurrent NN for a token sequence to sequence prediction |
| Embedding Space | Uses word distance (e.g., run versus walk versus sit) to derive context |
| LSTM | Uses long term and short term memory to increase context retention |
| Transformers | State of the art architecture using attention vectors, positional encoding, and feed-forward blocks |

## 