

Natural Language Processing (NLP) and it's Applications in Healthcare

By: Rohit Pardasani

Staff Data Scientist

GE Healthcare

Topics:

- Fundamentals of NLP
 - Introduction
 - Basic Text Processing
- Sentiment Analysis with Logistic Regression
- Vector Space Models
- Basic Word Representations
- Applications of NLP in Healthcare
- Tiny Project – NLP on Healthcare Dataset



Introduction to NLP



- **Natural language processing (NLP)** is a subfield of linguistics, computer science, and artificial intelligence concerned with the **interactions between computers and human language**, in particular how to program computers to process and analyze large amounts of natural language data.

- Consider sentence:

I am happy because I am learning NLP.

Can computer ever understand the meaning of this sentence as you and I do ?



Some applications of NLP

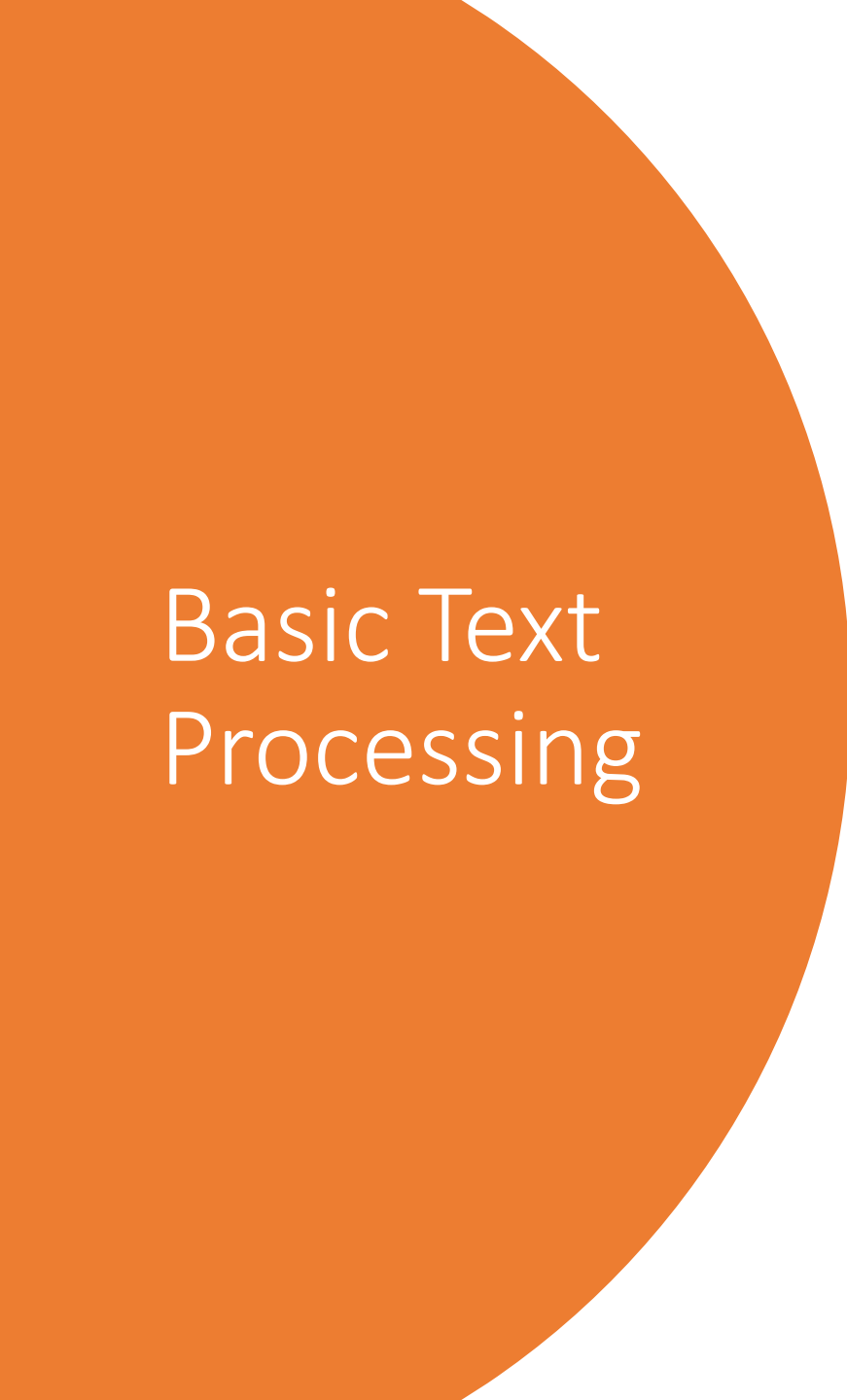
- Information extraction
- Spam detection
- Sentiment Analysis
- Machine Translation
- Question Answering
- Summarization
- Paraphrasing

Ambiguity
makes NLP
hard, but
also fun !!!

Teachers strike idle kids.

- Interpretation 1:
 - Teachers are on strike therefore kids are idle
- Interpretation 2:
 - Teachers hit kids who are sitting idle.



A large orange circle is positioned on the left side of the slide, partially cut off by the edge.

Basic Text Processing

- Regular Expressions
- Word Tokenization
- Lemmatization
- Stemming
- Removing stop words, punctuations



Regular Expressions

- A formal language for specifying text strings
- How can we search for any of these ?
 - woodchuck
 - woodchucks
 - Woodchuck
 - Woodchucks

Pattern: [wW]oodchuck

* Stands for 0 or more of previous characters

Oo*h !	Oh !	Ooh!	Oooh !	Ooooh !
--------	------	------	--------	---------

Other Text Processing

- Tokenization
 - They visited San Francisco. – More than one word can together form a single token.
- Lemmatization
 - Car, cars, car's, cars' - car
- Stemming
 - Automate, automates, automatic, automation - automat

Few Applications of NLP in Healthcare

- Clinical Documentation
- Speech Recognition
- Clinical Decision Support
- Accelerating Clinical Trial Matching
- Review Management and Sentiment Analysis