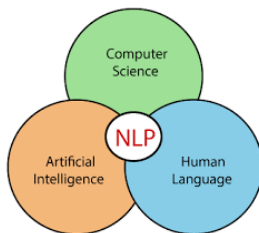# Natural Language Processing - An Introduction

July 17, 2023

Outline

- What is NLP
- Why to study NLP
- NLP Use Cases
- Fundamental NLP functions
- Python Libraries
- Language Modeling

The Goal and Prerequisites for this course

Objectives

- Students will learn the concepts and skills required to develop practical NLP applications.

- Students should have a solid foundation in NLP and be equipped with the necessary skills to tackle various NLP tasks, both in academic and industry settings.

Pre-Requisties

- The basic skills of Python
- Elementary probability and statistics

What is NLP

NLP - A definition

Natural Language Processing is concerned with the interactions between computers and human (natural) languages, in particular how to program computers to process and analyze large amounts of natural language data. [Wikipedia]

Goal of NLP

Ability to process and harness information from a large corpus of text with very little manual intervention

Why to study NLP

Text is the gigantic archive of human knowledge

Books, Research papers, Web pages, Scientific articles, News articles, Patents, Government documents, Medical documents, etc.

Rapid growth in the text from the variety of sources

Facebook, Twitter, Instagram, Quora, Stack Overflow, Product review, Service review, Movie review, news forums, and many more...

The Rapid Growth of data

Figure: The Data Growth Over the Internet - Its Believable[1]

---

[1]Source of image - https://www.go-globe.com/wp-content
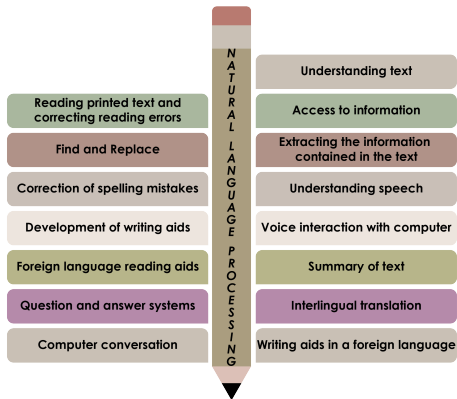
NLP Downstream Applications

Figure: The variety of tasks under the umbrella of NLP
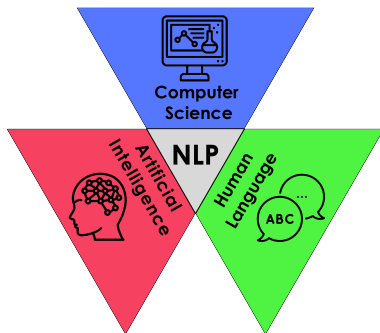
NLP Junction

Figure: What makes NLP

NLP - Use Cases

- Information Retrieval
- Text Summarization
- Language Modeling
- Spelling Correction

- Machine Translation
- Sentiment Analysis
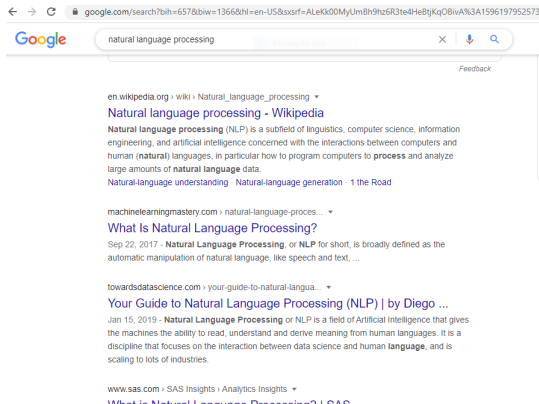- Speech Processing

Functions of NLP

Information Retrieval

Figure: Retrieving Relevant Information

NLP Use Cases - Text Summarization

- Machine-suggested reply of the email
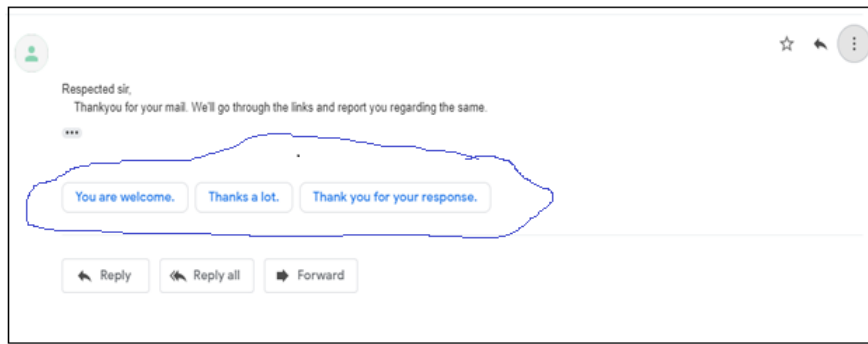- Abstractive summarization of the email text

Figure: Gmail Smart Email Reply

NLP Use Cases - Smart Compose

- Supports predicting completion of sentence
- Gmail Smart compose
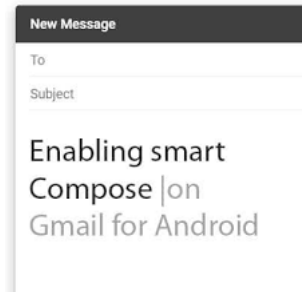- Real-time suggestions in Gmail that assists users in writing mails by reducing repetitive typing.

New Message

To

Subject

Enabling smart Compose |on
Gmail for Android

Figure: Gmail Smart Compose[2]

[2]Source of image- https://encrypted-tbn0.gstatic.com/
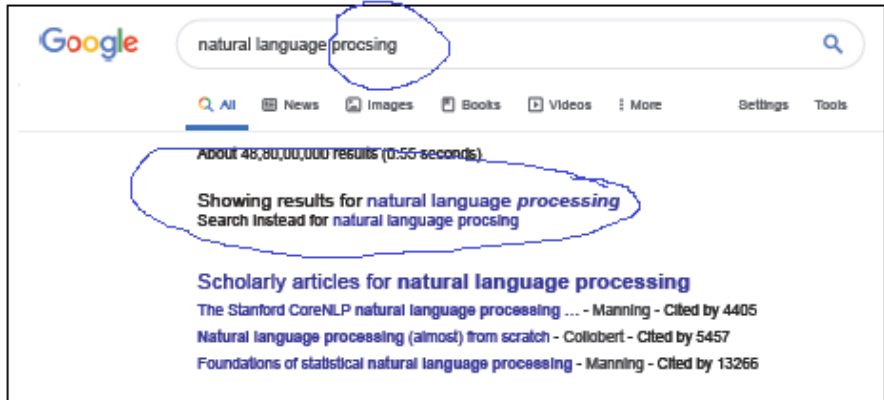
NLP Use Cases - Spelling Correction
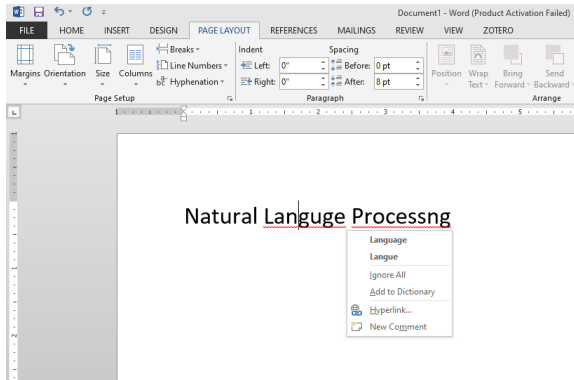

Figure: Google's Spelling Correction

NLP Use Cases - Spelling Correction

Figure: Word Processor - Spelling Correction
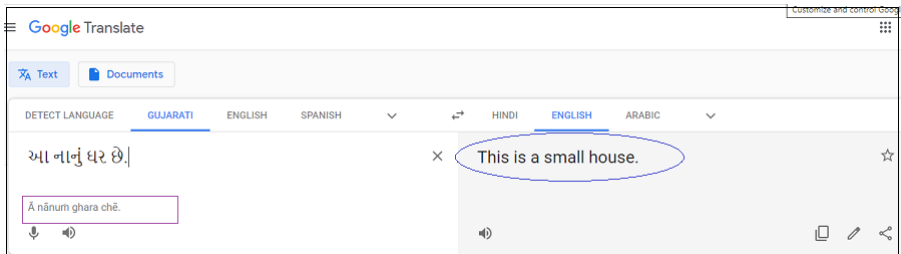
NLP Use Cases - Machine Translation

Figure: Google Machine Translation (Gujarati-English)

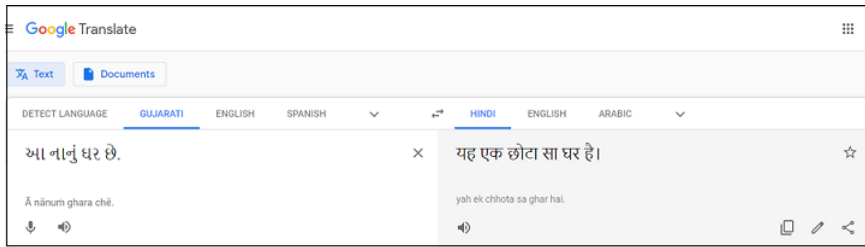Back to NLP Use cases

NLP Use Cases - Machine Translation

Figure: Google Machine Translation (Gujarati-Hindi)

Back to NLP Use cases

Sentiment Analysis

Defined in simple terms

Sentiment Analysis is the process of tagging a label of positive, negative or neutral to the short piece of text.
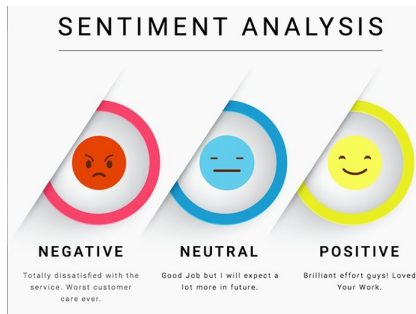

Figure: Sentiment Analysis[3]

[3]Source of image - https://www.kdnuggets.com/images/sentiment-fig-1-689.jpg

Sentiment Analysis

Basic Task

It just categorize the text to Positive or Negative

- I am happy as I got the first class. $\Rightarrow$ +VE
- I am extremely happy as I am top of the list. $\Rightarrow$ +VE
- He was sad due to his failure. $\Rightarrow$ -VE
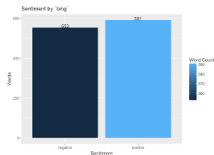- He was deeply sad due to his father death $\Rightarrow$ -VE.

Figure: Sentiment Analysis - The Basic Technique

Sentiment Analysis - An advanced Way

Figure: Sentiment Analysis - Advanced Version[4]

---

[4]source of image - https://encrypted-tbn0.gstatic.com

Sentiment Analysis - An advanced Way

Figure: Sentiment Analysis - Advanced Version[4]

Table: Emotions with Numerical Values

Emotion	Enumeration
Extreme Negative	-3
Very Negative	-2
Negative	-1
Neutral	0
Positive	1
Very Positive	2
Super Positive	3

Back to NLP Use cases

[4]source of image - https://encrypted-tbn0.gstatic.com

Chatbot

Figure: Chatbot - Automatic Customer Query Response[5]

Back to Goals of NLP

Well-known players

Figure: Google Assistant

Figure: Amazon Alexa

Figure: Microsoft Cortana

Fundamental Functions of NLP

Natural Language Processing is an interesting field of computer science.

Fundamental Functions of NLP

Natural Language Processing is an interesting field of computer science.

- Tokenization

Fundamental Functions of NLP

Natural Language Processing is an interesting field of computer science.

- Tokenization

- Stopwords
  Removal

Fundamental Functions of NLP

Natural Language Processing is an interesting field of computer science.

- Tokenization

- Stopwords
  Removal

- Punctuation
  Removal

Fundamental Functions of NLP

Natural Language Processing is an interesting field of computer science.

- Tokenization

- Stopwords Removal

- Punctuation Removal

- Stemming

Fundamental Functions of NLP

Natural Language Processing is an interesting field of computer science.

- Tokenization

- Stopwords Removal

- Punctuation Removal

- Stemming

- ['Natural','Language','Processing','is','an', 'interesting','field','of','computer','science','.']

Fundamental Functions of NLP

Natural Language Processing is an interesting field of computer science.

- Tokenization

- Stopwords
  Removal

- Punctuation
  Removal

- Stemming

- ['Natural','Language','Processing','is','an',
  'interesting','field','of','computer','science','.']

- ['Natural','Language','Processing',
  'interesting','field','computer','science','.']

Fundamental Functions of NLP

Natural Language Processing is an interesting field of computer science.

- Tokenization

- Stopwords Removal

- Punctuation Removal

- Stemming

- ['Natural','Language','Processing','is','an', 'interesting','field','of','computer','science','.']

- ['Natural','Language','Processing', 'interesting','field','computer','science','.']

- ['Natural','Language','Processing', 'interesting', 'field','computer','science']

Fundamental Functions of NLP

Natural Language Processing is an interesting field of computer science.

- Tokenization

  - ['Natural','Language','Processing','is','an', 'interesting','field','of','computer','science','.']

- Stopwords Removal

  - ['Natural','Language','Processing', 'interesting','field','computer','science','.']

- Punctuation Removal

  - ['Natural','Language','Processing', 'interesting', 'field','computer','science']

- Stemming

  - ['Natural','Language','Process', 'interest', 'field','computer','science']

Why NLP is a Challenging task

- Lexical Ambiguity - Same word in the same sentence having different meaning homonyms
- Word and its meaning depends on the context
- Multiple interpretations of the sentence - Multiple Interpretation
- Complex in nature (Huge vocabulary sometimes...)
- Idioms, Sarcasm makes NLP more difficult idioms
- Neologism - Ever expanding vocabulary
  - Google, Retweet, Photoshop, Skype, Unfriend etc.
- Non-standard English e.g. C U → See You

 Go to Python Libraries

Homonyms - Multiple meanings of the same word

- "Deer is dear, but dear".

Homonyms - Multiple meanings of the same word

- "Deer is dear, but dear".
- Deer (the animal) is dear (lovely) but dear (expensive).

Homonyms - Multiple meanings of the same word

- "Deer is dear, but dear".
- Deer (the animal) is dear (lovely) but dear (expensive).

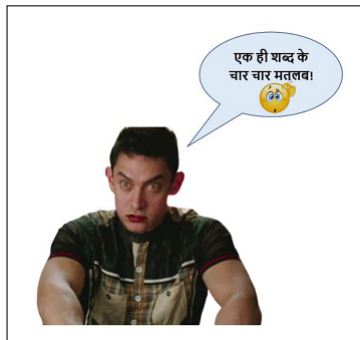Figure: Homonyms - The same words have multiple meaning[6]

[6]Source of Image-https://www.filmibeat.com/img/ pk-aamir-anushka.jpg
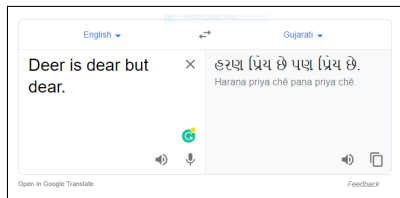
Google Translation Homonyms

Figure: Google translation - English to Gujarati
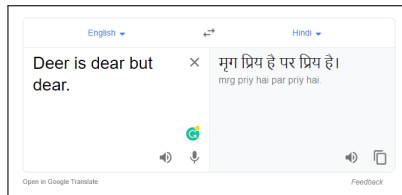
Figure: Google translation - English to Hindi

Multiple interpretation

I saw a man with binoculars.

Multiple interpretation

I saw a man with binoculars.

Figure: I saw the man with the help of
a binocular.[7]

---

[7]Source of image - https://2.bp.blogspot.com/ sherlock lores.jpg

Multiple interpretation

I saw a man with binoculars.
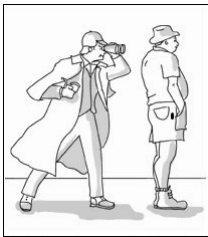
Figure: I saw the man with the help of a binocular.[7]

Figure: The man was using the binocular and I saw him.

Back to Why NLP is a challenging

[7]Source of image - https://2.bp.blogspot.com/ sherlock_lores.jpg

Idioms

- Now, the ball is in your court.

Idioms

- Now, the ball is in your court. The meaning is.... You have to handle it now.

Idioms

- Now, the ball is in your court. The meaning is.... You have to handle it now.

- It's not a rocket science.

Idioms

- Now, the ball is in your court. The meaning is.... You have to handle it now.

- It's not a rocket science. The meaning is... It's not complicated.

Idioms

- Now, the ball is in your court. The meaning is.... You have to handle it now.
- It's not a rocket science. The meaning is... It's not complicated.
- We are sailing in the same boat.

Idioms

- Now, the ball is in your court. The meaning is.... You have to handle it now.

- It's not a rocket science. The meaning is... It's not complicated.

- We are sailing in the same boat. The meaning is... We are having the same problem.

Back to Why NLP is a challenging

Sarcasm

- When something bad happens:

Sarcasm

- When something bad happens:
- And one says.... That's just what I needed today!

Sarcasm

- When something bad happens:
- And one says.... That's just what I needed today!
- When someone does something wrong:

Sarcasm

- When something bad happens:
- And one says.... That's just what I needed today!
- When someone does something wrong:
- Very good; well done! (With a different facial expression)

Back to Why NLP is a challenging

Python Libraries

- Natural Language ToolKit (NLTK)
- CoreNLP
- spaCy
- Gensim

Thank You All.......