

# Natural Language Processing

# Introduction to NLP

- ▶ NLP deals with unstructured data
- ▶ ML and DL only uses structured data. We have to convert unstructured data to structured data
- ▶ NLP is a subfield of AI that combines the power of Linguistics and Computer Science to study the rules and structure of language and create intelligent systems capable of understanding analyzing and extracting meaning from text and speech
- ▶ In NLP, the goal is to make computers understand the unstructured text (human/natural language) and retrieve meaningful information from it
- ▶ The field is secondarily concerned with helping us come to a better understanding of human language



Artificial  
Intelligence

NLP

Human  
Language

Computer  
Science

# Importance of NLP

- ▶ The amount of data generated by us keeps on increasing day by day, raising the need for analyzing and documenting the data
- ▶ NLP enables computers to read this data, (texts or speech) extract meaning from it, and convey the same in human languages
- ▶ Not only the process is automated but also near accurate all the time

# Benefits of NLP

- ▶ Perform large-scale analysis: NLP helps machine automatically understand and analyze huge amounts of unstructured data (text data) like social media comments, online reviews, etc
- ▶ Automate processes in real-time: NLP tools can help machines learn to sort and route information with little to no human interaction – quickly, efficiently, accurately and around the clock
- ▶ Tailor NLP tools to your industry : NLP algorithms can be tailored to your needs and criteria, like complex, industry specific language- even sarcasm and misused words

# Challenges/Limitations of NLP

- ▶ Contextual words and phrases and homonyms (words same but meaning different)
- ▶ Synonyms (same meaning, different words)
- ▶ Irony and sarcasm and slangs
- ▶ Ambiguity in statements/speech
- ▶ Errors in text or speech
- ▶ Domain-specific language
- ▶ Understanding the context and extracting meaning
- ▶ Lack of research and development

Etc.

# Components of NLP

## Natural Language Understanding

- ▶ NLU helps machine understand and analyse human language by extracting the meta data from contents such as concepts, entities, keywords, emotion, relations and semantic rules
- ▶ NLU involves following tasks:
  - ▶ Mapping given input in natural language into useful representation
  - ▶ Different level of analysis required :
    - ▶ Morphological analysis
    - ▶ Syntactic analysis
    - ▶ Semantic analysis
    - ▶ Discourse analysis
    - ▶ Pragmatic analysis

# Components of NLP

## Natural Language Generation

- ▶ NLG acts as a translator that converts the computerized data into natural language representation
- ▶ It mainly involves :
  - ▶ Text planning, sentence planning and text realization
  - ▶ Producing output in natural language from some internal representation
  - ▶ Different level of synthesis required : deep planning (what to say), syntactic generation



# NLP vs NLU vs NLG

## **Natural Language Processing (NLP)**

- ▶ Seeks to convert unstructured language data into structured data format to enable machines understand speech and text and formulate relevant contextual responses
- ▶ It's subtopics are : NLU and NLG

## **Natural Language Understanding (NLU)**

- ▶ focuses on machine reading comprehension through grammar and context, enabling it to determine the intended meaning of a sentence

## **Natural Language Generation (NLG)**

- ▶ focuses on text generation or the construction of text in English or other languages, by a machine and based on a given dataset

# Applications of NLP

- ▶ Machine Translation
- ▶ Information retrieval
- ▶ Report generation
- ▶ Sentiment analysis
- ▶ Query answering
- ▶ Speech recognition
- ▶ Language detection
- ▶ Text generation
- ▶ Chatbot

Etc.