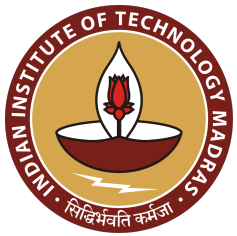
BSCS5002: Introduction to Natural Language Processing

Lecture 1 : Understanding Natural Language and Processing

Parameswari Krishnamurthy



Language Technologies Research Centre
IIIT-Hyderabad

param.krishna@iiit.ac.in

Natural Language

What is Natural Language?

- Natural language refers to the languages spoken by humans for everyday communication.
- It is different from:
  - Formal or artificial languages, such as programming languages like Python and Java
  - Constructed auxiliary languages, such as Esperanto and Interlingua
  - Non-human communication systems, such as the honey bee's dance and a cat's meow
- Examples of natural languages include English, Tamil, Hindi, Mandarin, Spanish, etc.

Characteristics and Challenges of Language

- Language consists of structured words and sounds, communicated through speaking, writing, and gestures.
- It is a human capacity for using complex systems to express thoughts and feelings.
- Language is used for exchanging knowledge and experiences.
- Language diversity varies greatly, often leading to communication barriers.
- Lingua francas or trade languages help overcome these barriers.
- About one-third of the world's languages are spoken in Asia and another third in Africa.

Natural Language Processing

What is NLP?

- Humans are masters of language! Can a machine do this?
- Natural Language Processing enables computer systems to understand, interpret, and generate human language.
- The outcome of NLP is the creation of language models.
  - deep knowledge and intelligent treatment of human languages
  - mimic human mind
- NLP combines computational linguistics with machine learning and deep learning models to process and analyze large amounts of natural language data.
- Applications of NLP include chatbots, sentiment analysis, machine translation, speech recognition, and more.

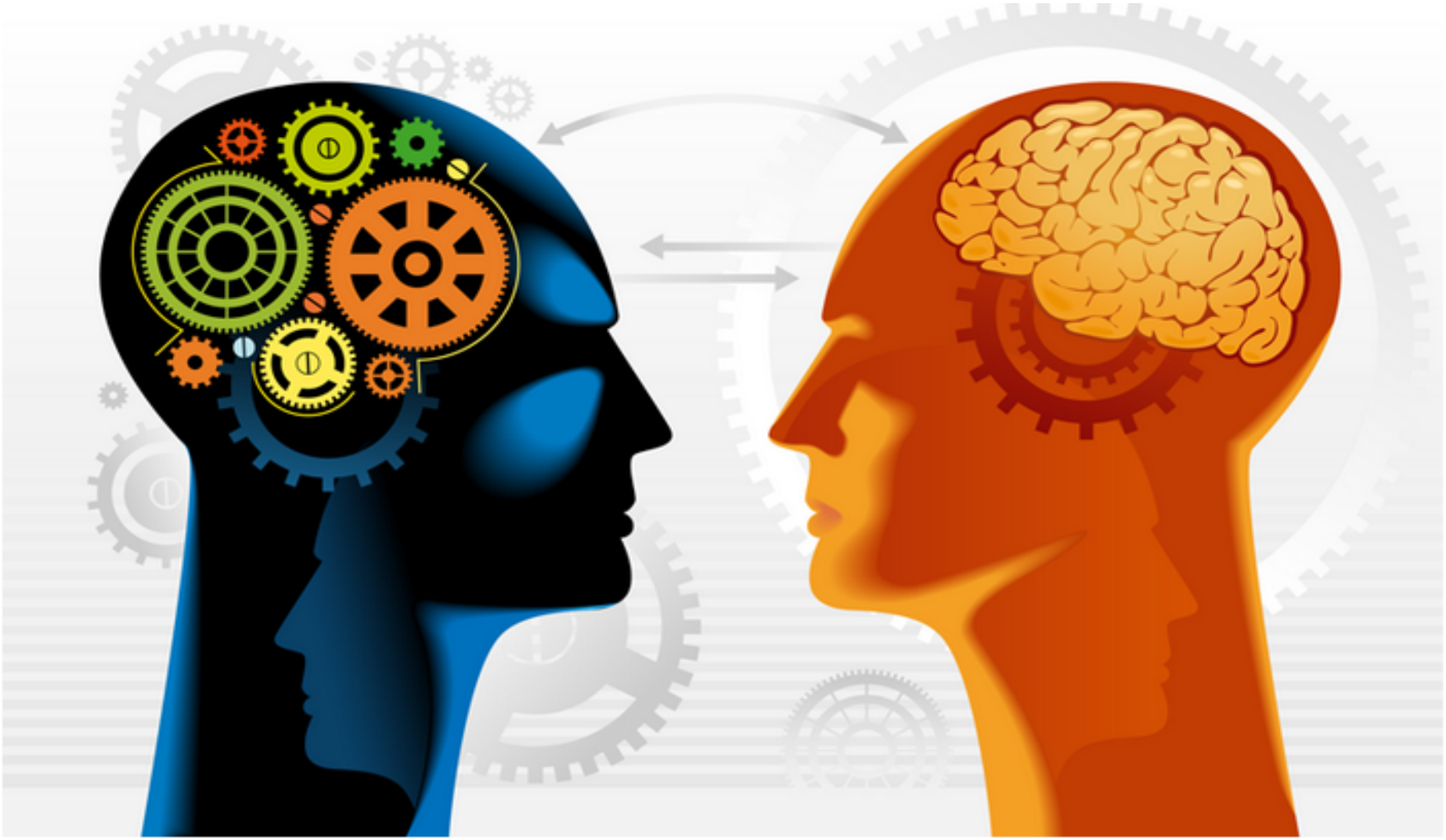
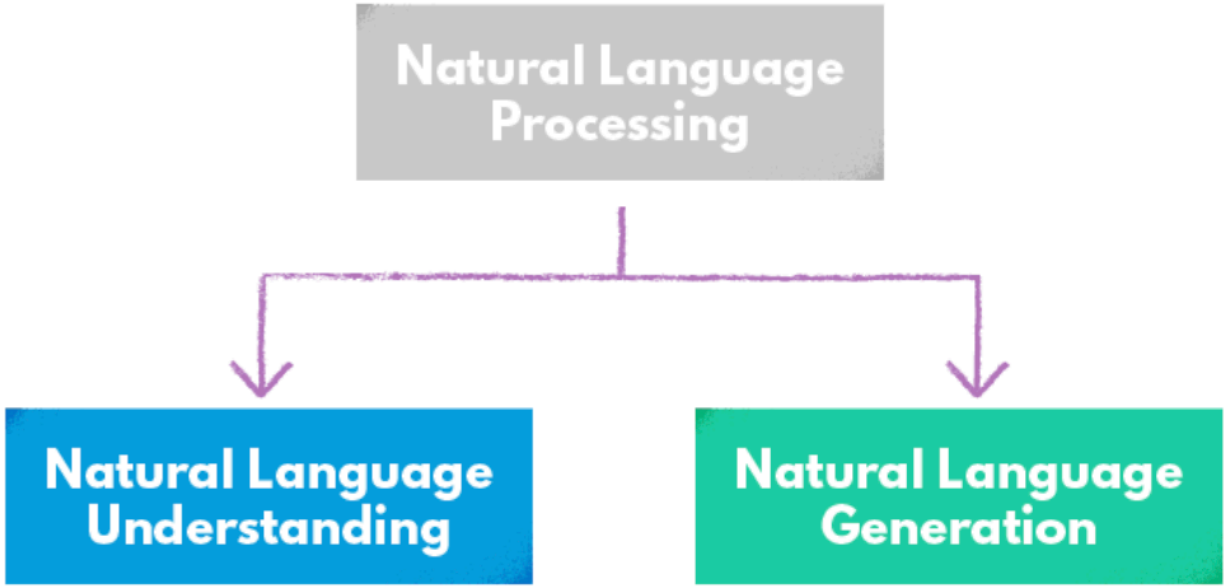


image source:https://towardsdatascience.com/introducing-natural-language-processing-nlp-series-1-aefbb69a20bc

NLP=NLU+NLG


Natural Language Processing
Natural Language Understanding
Natural Language Generation

Human NLU & NLG

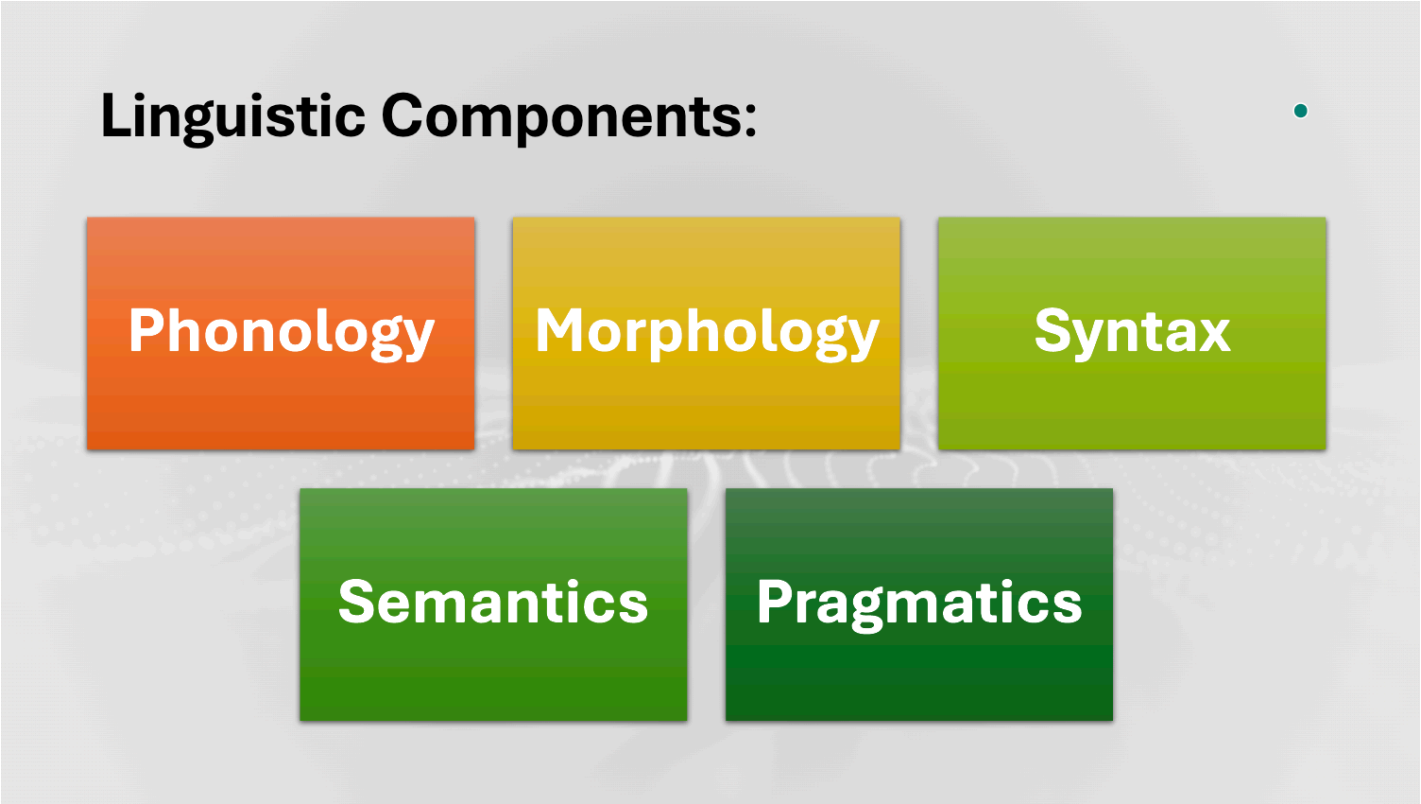
Human NLU & NLG

CONCEPTUALIZATION

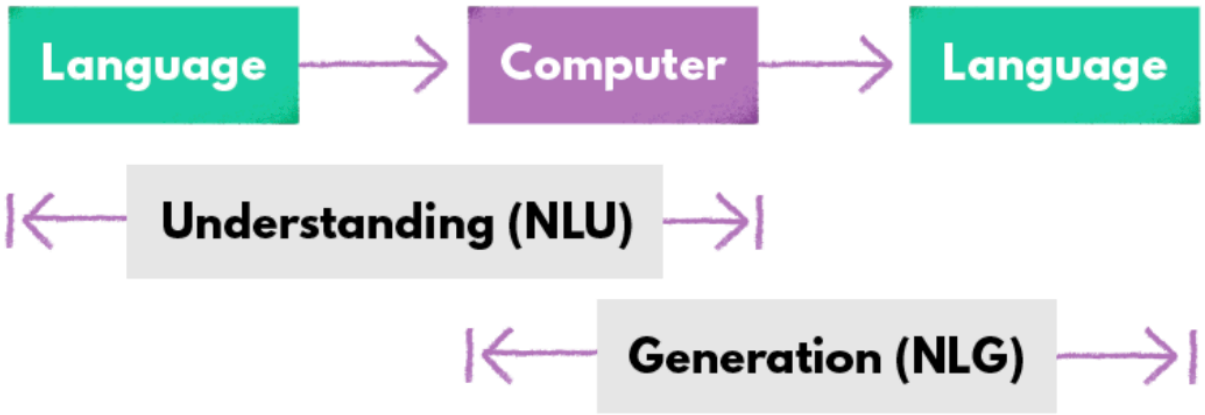
FORMULATION

ARTICULATION

Linguistic Components



Machine NLU & NLG



Major Challenges

UNDERSTANDING CONTEXT

LIMITED REASONING

BIAS IN TRAINING DATA

Key differences between Human vs Machine in NLP

Human:

- Deeper Understanding
- Common Sense Reasoning
- Context Adaptation
- Emotional Intelligence
- Creativity
- Learning from Minimal Data

Machine:

- Data Processing
- Complex Statistical Patterns
- Efficiency in Specific Tasks
- Consistency
- Scalability
- Multilingual Capabilities

NLP Applications

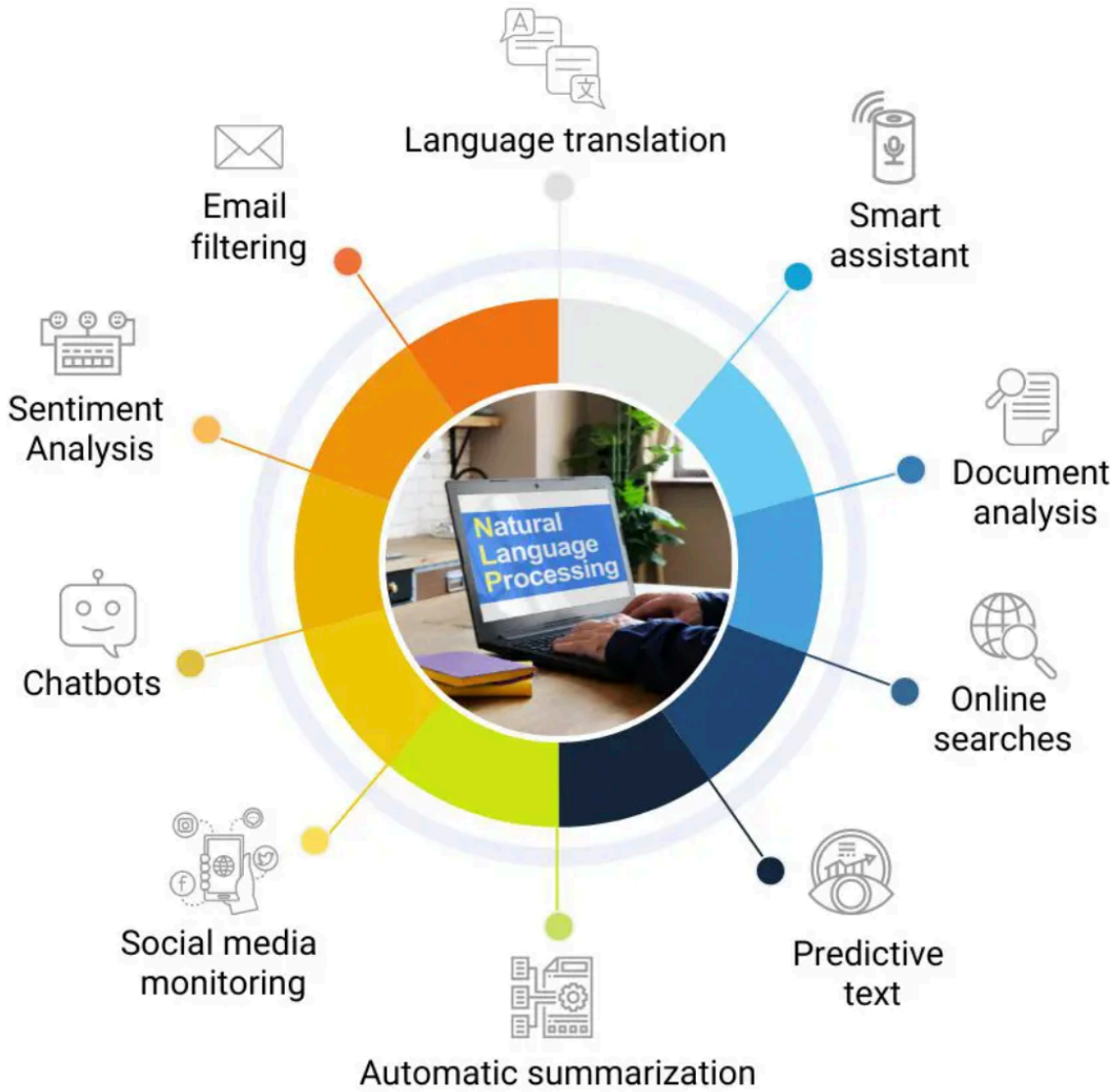


image source: https://datasciencedojo.com/blog/natural-language-processing-applications/