Introduction to NLP
[Natural Language Processing]

Course Title: Artificial Intelligence

Course Code: cse-403

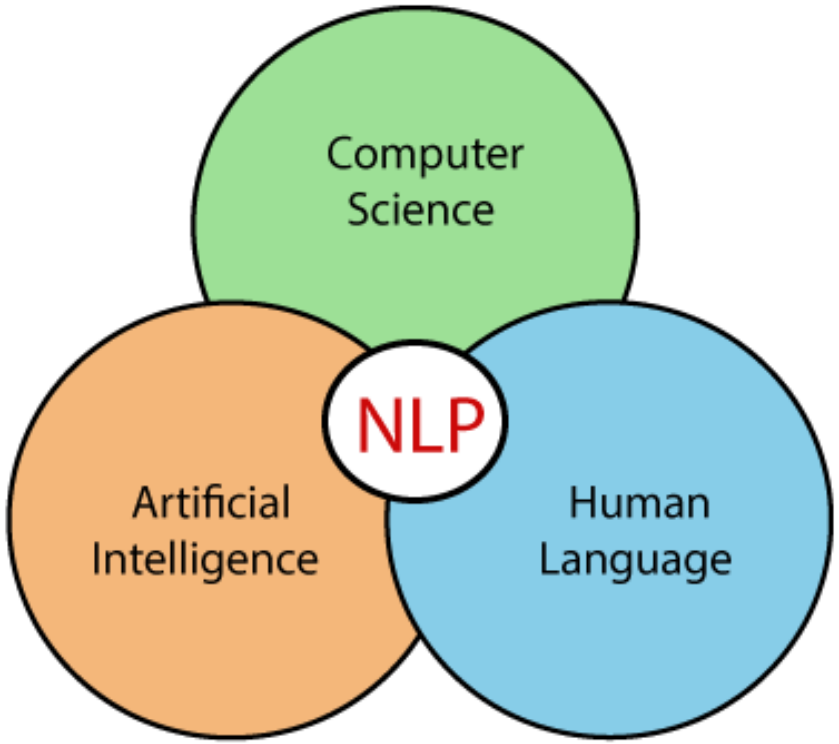
Khandaker Jannatul Ritu, Lecturer, Dept. of CSE(BAIUST)

Outlines

1. What is NLP?
2. Components of NLP
3. Applications of NLP
4. Phases of NLP
5. Why NLP is difficult?
6. Applications of NLP
7. How can you differentiate Artificial Intelligence, Machine Learning, and Natural Language Processing?

❑ What is NLP?

NLP stands for **Natural Language Processing**, which is a part of **Computer Science, Human language,** and **Artificial Intelligence**. It is the technology that is used by machines to understand, analyse, manipulate, and interpret human's languages. It helps developers to organize knowledge for performing tasks such as **translation, automatic summarization, Named Entity Recognition (NER), speech recognition, relationship extraction,** and **topic segmentation**.



What is NLP?

NLP stands for Natural Language Processing. It is the branch of Artificial Intelligence that gives the ability to machine understand and process human languages. Human languages can be in the form of text or audio format.

❑ Components of NLP

There are the following two components of NLP –

1. Natural Language Understanding (NLU)

Natural Language Understanding (NLU) helps the machine to understand and analyse human language by extracting the metadata from content such as concepts, entities, keywords, emotion, relations, and semantic roles.

NLU mainly used in Business applications to understand the customer's problem in both spoken and written language.

NLU involves the following tasks -

•It is used to map the given input into useful representation.

•It is used to analyze different aspects of the language.

2. Natural Language Generation (NLG)

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.

❑ Applications of NLP

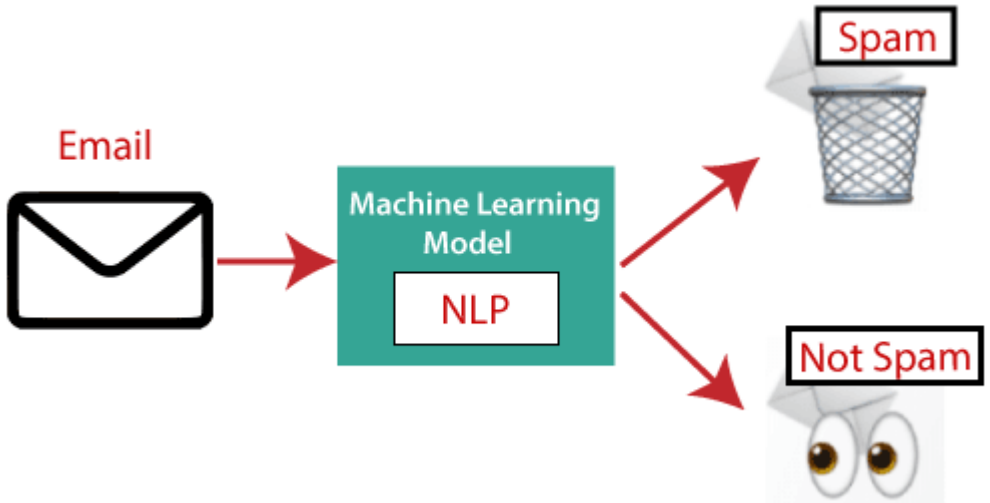
There are the following applications of NLP -

1. Question Answering

Question Answering focuses on building systems that automatically answer the questions asked by humans in a natural language.

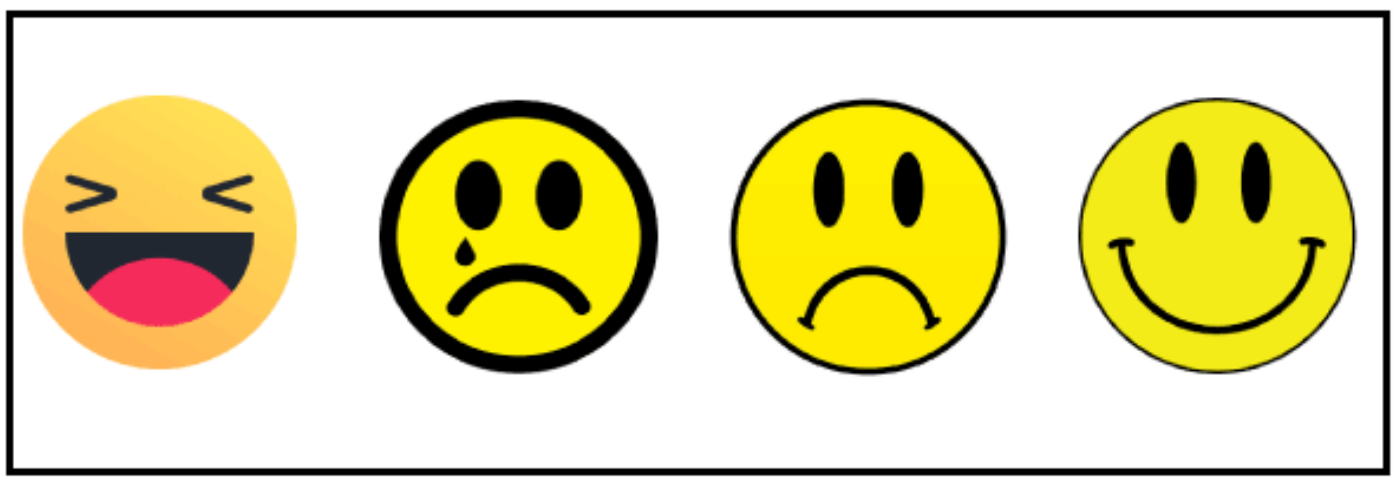
2. Spam Detection

Spam detection is used to detect unwanted e-mails getting to a user's inbox.



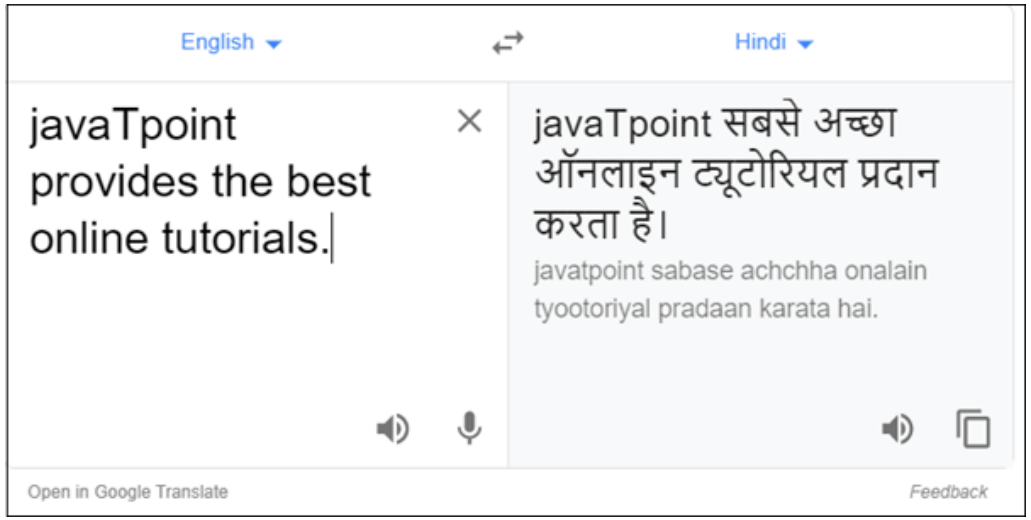
3. Sentiment Analysis

Sentiment Analysis is also known as **opinion mining**. It is used on the web to analyse the attitude, behaviour, and emotional state of the sender. This application is implemented through a combination of NLP (Natural Language Processing) and statistics by assigning the values to the text (positive, negative, or natural), identify the mood of the context (happy, sad, angry, etc.)



4. Machine Translation

Machine translation is used to translate text or speech from one natural language to another natural language.



Example: Google Translator

5. Spelling correction

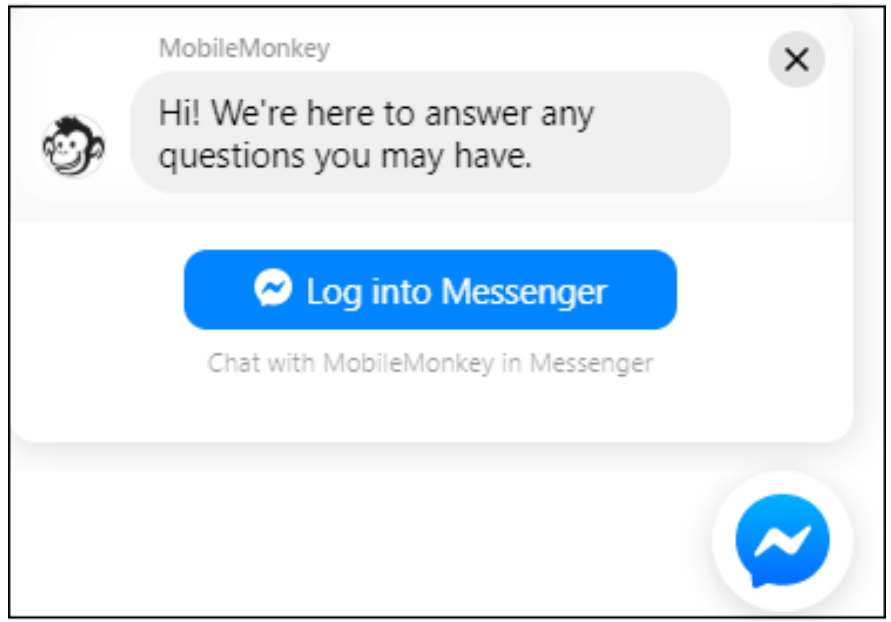
Microsoft Corporation provides word processor software like MS-word, PowerPoint for the spelling correction.

6. Speech Recognition

Speech recognition is used for converting spoken words into text. It is used in applications, such as mobile, home automation, video recovery, dictating to Microsoft Word, voice biometrics, voice user interface, and so on.

7. Chatbot

Implementing the Chatbot is one of the important applications of NLP. It is used by many companies to provide the customer's chat services.



8. Information extraction

Information extraction is one of the most important applications of NLP. It is used for extracting structured information from unstructured or semi-structured machine-readable documents.

9. Natural Language Understanding (NLU)

It converts a large set of text into more formal representations such as first-order logic structures that are easier for the computer programs to manipulate notations of the natural language processing.

❑ Phases of NLP

There are the following five phases of NLP:

1. Lexical Analysis and Morphological

The first phase of NLP is the Lexical Analysis. This phase scans the source code as a stream of characters and converts it into meaningful lexemes. It divides the whole text into paragraphs, sentences, and words.

2. Syntactic Analysis (Parsing)

Syntactic Analysis is used to check grammar, word arrangements, and shows the relationship among the words.

Example: Agra goes to the Poonam

In the real world, Agra goes to the Poonam, does not make any sense, so this sentence is rejected by the Syntactic analyzer.

3. Semantic Analysis

Semantic analysis is concerned with the meaning representation. It mainly focuses on the literal meaning of words, phrases, and sentences.

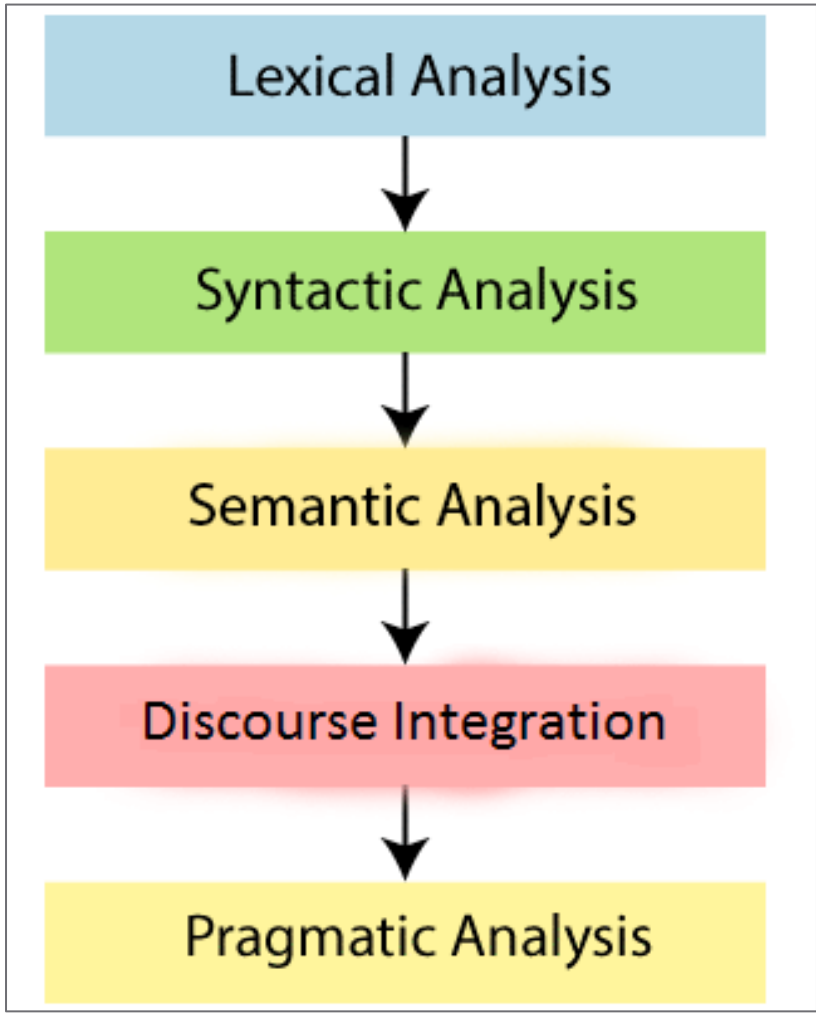
4. Discourse Integration

Discourse Integration depends upon the sentences that proceeds it and also invokes the meaning of the sentences that follow it.

5. Pragmatic Analysis

Pragmatic is the fifth and last phase of NLP. It helps you to discover the intended effect by applying a set of rules that characterize cooperative dialogues.

For Example: "Open the door" is interpreted as a request instead of an order.



❑ Why NLP is difficult?

NLP is difficult because Ambiguity and Uncertainty exist in the language.

Ambiguity

There are the following **three** ambiguity –

1. Lexical Ambiguity

Lexical Ambiguity exists in the presence of two or more possible meanings of the sentence within a single word.

Example:

Manya is looking for a **match**.

In the above example, the word match refers to that either Manya is looking for a partner or Manya is looking for a match. (Cricket or other match)

2. Syntactic Ambiguity

Syntactic Ambiguity exists in the presence of two or more possible meanings within the sentence.

Example:

I saw the girl with the binocular.

In the above example, did I have the binoculars? Or did the girl have the binoculars?

3. Referential Ambiguity

Referential Ambiguity exists when you are referring to something using the pronoun.

Example: Kiran went to Sunita. She said, "I am hungry."

❑ Applications of NLP

•Text and speech processing like-Voice assistants – Alexa, Siri, etc.
•Text classification like Grammarly, Microsoft Word, and Google Docs
•Information extraction like-Search engines like DuckDuckGo, Google
•Chatbot and Question Answering like:- website bots
•Language Translation like:- Google Translate
•Text summarization

❑ How can you differentiate Artificial Intelligence, Machine Learning, and Natural Language Processing?

Artificial Intelligence	Machine Learning	Natural Language Processing
Artificial Intelligence is a technique used to create smarter machines and computers.	Machine Learning is a term used for systems that learn from experience.	Natural Language Processing or NLP is the set of systems that can understand the languages used by humans and process these languages to make them understood by computers.
Artificial Intelligence requires human intervention. Without human intervention, it is not possible to create intelligent machines.	Machine Learning doesn't require human intervention. It purely involves the working of computers and machines.	Natural Language Processing uses both computer and human languages to work properly.