
CHAPTER :1

INTRODUCTION

WHAT IS NLP?

- **Natural Language Processing**, usually shortened as **NLP**, is a branch of artificial intelligence that deals with the interaction between computers and humans using the **natural language**. The ultimate objective of NLP is to read, decipher, understand, and make sense of the human languages in a manner that is valuable.
- A few **examples** of **NLP** that people use every day are: Spell check. Autocomplete

TWO MAIN REASONS TO DEVELOP NLP

- 1. To develop automated tools for language processing.
- 2. To gain a better understanding of human language communication.

APPLICATIONS OF NAT. LANG. PROCESSING

- Machine Translation
- Database Access
- Information Retrieval
 - Selecting from a set of documents the ones that are relevant to a query
- Text Categorization
 - Sorting text into fixed topic categories
- Extracting data from text
 - Converting unstructured text into structure data
- Spoken language control systems
- Spelling and grammar checkers.

COMPONENTS OF NLP

❑ Challenges in NLP :

- ❑ Multiple intents in one question. When a customer asks for several things at the same time, such as different products, boost. ...
- ❑ Assuming it understands context and has memory. ...
- ❑ Misspellings in entity extraction. ...
- ❑ Same word - different meaning. ...
- ❑ Keeping the conversation going. ...
- ❑ Tackling false positives.

LANGUAGE AND GRAMMAR

- Grammer define the language,It consists of a set of rules that allows us to parse and generate sentences in a language.
- **Grammars and parsing**
- Syntactic categories (common denotations) in NLP
- np - noun phrase
- vp - verb phrase
- s - sentence
- det - determiner (article)
- n - noun
- tv - transitive verb (takes an object)
- iv - intransitive verb
- prep - preposition
- pp - prepositional phrase
- adj - adjective

EX:

POOJA PLAYS VEENA

VEENA IS PLAYED BY POOJA

- Transformational grammar has three components:
-

- 1. Phrase structure grammar
- 2. Transformational rules
- 3. Morphophonemic rules: these rules match each sentence representation to a string of phonemes.

1.PHRASE STRUCTURE GRAMMAR

2.TRANSFORMATIONAL RULES

- Which transform one phrase maker (underlying) into another phrase maker(derived).
- E.g:Chomsky's rule for relating active and passive sentences (as given in *Syntactic Structures*) is very similar, at first sight, to Harris's, discussed above. Chomsky's rule is:

□

3. MORPHOPHONEMIC RULES:

- These rules match each sentence representation to a string of phonemes.
- Ex:consider active sentence:
The police will catch the snatcher

APPLICATIONS OF NLP

1. Machine Translation

Machine translation (MT), process of translating one source language or text into another language, is one of the most important applications of NLP.

2. Automatic Summarization

In this digital era, the most valuable thing is data, or you can say information. However, do we really get useful as well as the required amount of information? The answer is 'NO' because the information is overloaded and our access to knowledge and information far exceeds our capacity to understand it.

3. Question-answering

Another main application of natural language processing (NLP) is question-answering. Search engines put the information of the world at our fingertips, but they are still lacking when it comes to answer the questions posted by human beings in their natural language.

4. Information Retrieval:

this is the concerned with identified document relevant to a user's query.

5. Automatic Summarization

In this digital era, the most valuable thing is data, or you can say information. However, do we really get useful as well as the required amount of information? The answer is 'NO' because the information is overloaded and our access to knowledge and information far exceeds our capacity to understand it.

INFORMATION RETRIVAL

- ❑ Information retrieval deals with unstructured data.
- ❑ The retrieval based on the content of the document rather than on its structure.
- ❑ Information retrieval (IR) may be defined as a software program that deals with the organization, storage, retrieval and evaluation of information from document repositories particularly textual information.
- ❑ The system assists users in finding the information they require but it does not explicitly return the answers of the questions. It informs the existence and location of documents that might consist of the required information. The documents that satisfy user's requirement are called relevant documents. A perfect IR system will retrieve only relevant documents.

With the help of the following diagram, we can understand the process of information retrieval (IR)

