

Chapter 8: Natural Language Processing

Grade VIII • Notebook Notes

Topics: Introduction to NLP • Working of NLP • Components of NLP • NLP Tasks & Techniques • Applications of NLP

Learning Goal: Understand how computers process human language and how Natural Language Processing (NLP) is used in everyday applications.

1. Introduction to NLP

Natural Language Processing (NLP) is a branch of Artificial Intelligence that enables computers to process, understand, interpret and generate human language.

Human language is called **natural language** because it is the language people naturally use to communicate, such as English, Hindi and other languages.

In simple words: NLP helps computers work with human language.

Examples of NLP

- Voice assistants understanding spoken commands.
- Translation systems converting one language into another.
- Chatbots answering questions.
- Email systems detecting spam.
- Applications analysing opinions in reviews.

2. Why is NLP Needed?

Computers naturally work with data in machine-readable forms, while humans communicate using words, sentences and speech. NLP provides techniques that help computers process this language.

- Language can have different meanings depending on context.
- People may use abbreviations, slang or informal expressions.
- The same idea can be expressed in many different ways.
- Human language may contain ambiguity and grammatical variations.

3. Working of NLP

An NLP system generally processes language in several stages. The exact stages depend on the task, but a simple workflow can be understood as:

Stage	What happens?
1. Input	Text or speech is provided to the system.
2. Pre-processing	The input is cleaned and prepared for analysis.
3. Language Analysis	Words, sentences and their relationships are analysed.
4. Meaning / Pattern Identification	The system looks for useful information, intent or patterns.
5. Output	The system produces a response, classification, translation, summary or other result.

4. Example of NLP Working

Suppose a user types: **“Book a train ticket to Delhi tomorrow.”**

- The system receives the sentence as input.
- It identifies important words such as “book”, “train ticket”, “Delhi” and “tomorrow”.
- It determines that the user's intention is related to booking a train ticket.
- It extracts useful information such as the destination and date.
- The application can then provide an appropriate response or perform an action.

5. Components of NLP

NLP combines ideas from several areas of language and computer science. Important components include:

Component	Meaning
Tokenization	Breaking text into smaller units called tokens, such as words or punctuation marks.
Sentence Segmentation	Dividing a paragraph or document into individual sentences.
Part-of-Speech Tagging	Identifying words as nouns, verbs, adjectives, etc.
Named Entity Recognition	Identifying names of people, places, organisations and other entities.
Stemming	Reducing a word to a simpler root-like form by removing endings.
Lemmatization	Reducing a word to its meaningful dictionary form.

6. NLP Tasks and Techniques

An **NLP task** is a problem that an NLP system is designed to solve. Different tasks use different techniques.

Task	Purpose
Sentiment Analysis	Determines whether a piece of text expresses a positive, negative or neutral opinion.
Text Classification	Assigns text to categories, such as spam or not spam.
Machine Translation	Converts text from one language into another.
Text Summarization	Produces a shorter version containing important information.
Question Answering	Provides answers to questions expressed in natural language.
Speech Recognition	Converts spoken language into text.
Text-to-Speech	Converts written text into spoken language.
Natural Language Generation (NLG)	Generates human-readable language from data or information.

7. Sentiment Analysis

Sentiment Analysis is an NLP technique used to identify the opinion or emotional tone expressed in text.

Example: “The movie was excellent!” → Positive sentiment.

8. Lemmatization

Lemmatization reduces a word to its meaningful dictionary form, called its **lemma**. It considers the word and its linguistic role rather than simply removing characters.

Example: “**running**”, “**ran**” and “**runs**” may be related to the lemma “**run**”, depending on context.

9. Sentence Segmentation

Sentence segmentation is the process of dividing a piece of text into individual sentences.

Example: "I like Python. It is easy to learn." contains two sentences.

10. Applications of NLP

Application	Example
Virtual Assistants	Understanding voice commands and responding to users.
Chatbots	Answering questions and supporting users through conversation.
Machine Translation	Translating text between languages.
Email Filtering	Detecting spam and unwanted messages.
Search Engines	Understanding search queries and finding relevant results.
Grammar and Writing Tools	Checking grammar, spelling and style.
Sentiment Analysis	Studying opinions in customer reviews or social media text.
Accessibility	Speech recognition and text-to-speech can assist users.

11. NLP in Everyday Life

- When a phone converts speech into text.
- When an email service detects spam.
- When a translation application translates a sentence.
- When a chatbot answers a question.
- When a search engine interprets a user's query.
- When a writing application suggests corrections.

12. NLP, AI and Machine Learning

NLP is a branch of Artificial Intelligence. Many modern NLP systems use **Machine Learning** and especially **Deep Learning** to learn patterns from large collections of language data.

AI is the broader field → **NLP** focuses on human language → Machine Learning can provide methods for learning language patterns from data.

13. Advantages of NLP

- Allows people to communicate with computers using natural language.
- Can process large amounts of text quickly.
- Can automate repetitive language-related tasks.
- Can support translation and accessibility.
- Can help organisations understand customer feedback and opinions.

14. Challenges of NLP

- Words can have different meanings depending on context.
- Human language contains slang, abbreviations and spelling variations.
- Sarcasm and humour can be difficult for computers to interpret.

- Different languages have different grammar and structures.
- NLP systems can reflect errors or biases present in their training data.

15. Key Terms

Term	Meaning
NLP	Natural Language Processing; AI techniques for working with human language.
Token	A basic unit produced when text is split during tokenization.
Sentiment Analysis	Finding the opinion or emotional tone in text.
Lemmatization	Finding a word's meaningful dictionary form.
OCR	Optical Character Recognition; extracting text from images.
NLG	Natural Language Generation; producing human-readable language.

16. Quick Revision

- NLP stands for Natural Language Processing.
- NLP enables computers to process, understand and generate human language.
- NLP can work with both text and speech.
- Important NLP techniques include tokenization, sentence segmentation, stemming, lemmatization and named entity recognition.
- Important NLP tasks include sentiment analysis, translation, classification, summarization and question answering.
- NLP is used in chatbots, virtual assistants, search engines, translation systems and writing tools.
- Good NLP systems need suitable data and careful evaluation.

17. Practice Questions

1. What does NLP stand for? What does it do?
2. Why is NLP considered a branch of Artificial Intelligence?
3. Explain the basic working of an NLP system.
4. What is tokenization?
5. What is sentence segmentation?
6. What is the difference between stemming and lemmatization?
7. What is Sentiment Analysis? Give one example.
8. What does NLG stand for? What is its function?
9. Mention any four applications of NLP.
10. How can NLP be useful in everyday life?

Notebook Tip: Draw the basic NLP flow in your notebook: **Text/Speech** → **Pre-processing** → **Language Analysis** → **Meaning/Pattern** → **Output**.