**NLP ASSIGNMENT-1**

**Q1:-**  Take a custom paragraph, perform the entire pipeline and Print results at each step.

Tokenization → Stopword Removal → Stemming → Lemmatization.

**SOL:-**

1. Import Libraries

A screenshot of a computer screen

AI-generated content may be incorrect.

1. Download required resources

A screen shot of a computer

AI-generated content may be incorrect.

1. Custom Paragraph   
   A screen shot of a computer

   AI-generated content may be incorrect.
2. Tokenization

A screenshot of a computer

AI-generated content may be incorrect.

1. StopWord Removal

A white background with black text

AI-generated content may be incorrect.

1. Stemming

A white background with black text

AI-generated content may be incorrect.

1. Lemmatization

A white background with black text

AI-generated content may be incorrect.

**Q2: Define NLP and explain its real-time application in a specific domain.**  
**Answer:**  
Natural Language Processing (NLP) is a branch of Artificial Intelligence (AI) and linguistics that deals with the interaction between computers and human languages. Its goal is to enable machines to read, understand, and produce human language in a way that is both meaningful and useful.

NLP combines elements of machine learning, computational linguistics, and deep learning to process and analyze text and speech data.

**Domain Example: Healthcare**  
In the healthcare industry, NLP plays a significant role in making sense of large volumes of unstructured data to improve the quality of care and operational efficiency.

**Key Applications:**

1. **Processing Clinical Documents:** NLP helps in extracting important information such as patient symptoms, diagnoses, and prescribed treatments from medical records and physician notes.
2. **Virtual Health Assistants and Chatbots:** These systems use NLP to answer patient questions, book appointments, and offer initial health guidance.
3. **Predictive Analytics for Diseases:** By analyzing patient history, lab results, and other records, NLP tools can help in predicting potential health risks like heart disease, diabetes, or cancer.
4. **Speech-to-Text in Clinical Settings:** Physicians can dictate their notes verbally, and NLP software converts them into structured digital records.

**Real-World Example:**  
IBM Watson Health uses NLP technology to process clinical documents and medical research, helping doctors access relevant information quickly and make better treatment choices.

**Q3: What are NLU and NLG?**  
**Answer:**

**NLU (Natural Language Understanding):**  
NLU is a core component of NLP that focuses on enabling machines to comprehend and interpret the meaning and intent behind human language input.

**Key Capabilities:**

* Identifying named entities like people, places, and dates.
* Detecting user intent (e.g., making a request or asking a question).
* Resolving ambiguity and understanding context or synonyms.

**Example:**  
If a user says, *"Book me a flight to Delhi tomorrow morning,"* the NLU system might extract:

* **Intent:** Book flight
* **Entities:**
  + Destination: Delhi
  + Date: Tomorrow
  + Time: Morning

**NLG (Natural Language Generation):**  
NLG is the counterpart to NLU. It focuses on generating human-readable text or speech from structured data.

**Key Functions:**

* Translating data into coherent sentences.
* Producing summaries, responses, or reports automatically.

**Example:**  
Given the data {Destination: Delhi, Date: Tomorrow, Time: Morning}, an NLG system might generate the sentence:  
*“Your flight to Delhi is scheduled for tomorrow morning.”*