

NLP ASSIGNMENT

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

Tokenization → Stop word Removal → Stemming → Lemmatization.

SOL:-

1) Custom Paragraph

```
my_para = """Technology is changing education rapidly.  
Online learning platforms, AI tutors, and interactive tools  
are making education more flexible and accessible.  
With NLP, teachers can analyze student feedback,  
and students can get instant answers to their questions."""  
  
print("Original Paragraph:\n", my_para)
```

Original Paragraph:
Technology is changing education rapidly.
Online learning platforms, AI tutors, and interactive tools
are making education more flexible and accessible.
With NLP, teachers can analyze student feedback,
and students can get instant answers to their questions.

2) Tokenization

```
import nltk  
nltk.download('punkt', quiet=True)  
from nltk.tokenize import word_tokenize  
  
words = word_tokenize(my_para)  
print("\nAfter Tokenization:\n", words)
```

After Tokenization:
['Technology', 'is', 'changing', 'education', 'rapidly', '.', 'Online', 'learning', 'platforms', ',', 'AI', 'tutors', ',',

3) Stop word Removal

```
▶ nltk.download('stopwords', quiet=True)
  from nltk.corpus import stopwords

  stop_words = set(stopwords.words('english'))
  no_stopwords = [w for w in words if w.lower() not in stop_words and w.isalpha()] |
  print("\nAfter Removing Stopwords:\n", no_stopwords)
```

After Removing Stopwords:
['Technology', 'changing', 'education', 'rapidly', 'Online', 'learning', 'platforms', 'AI', 'tutors', 'interactive', 'tools']

4) Stemming

```
▶ from nltk.stem import PorterStemmer
  ps = PorterStemmer()

  stemmed = [ps.stem(w) for w in no_stopwords]
  print("\nAfter Stemming:\n", stemmed)
```

After Stemming:
['technolog', 'chang', 'educ', 'rapidli', 'onlin', 'learn', 'platform', 'ai', 'tutor', 'interact', 'tool', 'make', 'educ', 'tools']

5) Lemmatization

```
nltk.download('wordnet', quiet=True)
nltk.download('omw-1.4', quiet=True)
from nltk.stem import WordNetLemmatizer
lm = WordNetLemmatizer()

lemmatized = [lm.lemmatize(w) for w in no_stopwords] # Lemmatize **original filtered words** for best results
print("\nAfter Lemmatization:\n", lemmatized)
```

After Lemmatization:
['Technology', 'changing', 'education', 'rapidly', 'Online', 'learning', 'platform', 'AI', 'tutor', 'interactive', 'tool', 'tools']

Q2: - Define NLP and its real time application in a specific domain base.

SOL:- Natural Language Processing (NLP) is a branch of Artificial Intelligence (AI) that focuses on enabling computers to read, understand, and respond to human language. It combines concepts from linguistics, computer science, and machine learning to process text and speech in a meaningful way. The main aim of NLP is to bridge the gap between human communication and computer understanding. NLP involves different techniques like tokenization, stop word removal, stemming, lemmatization, sentiment analysis, and text classification. Using these techniques, machines can handle tasks like translation, summarization, text generation, and more.

Domain: E-commerce

In the e-commerce field, NLP plays a very big role in improving customer experience and business operations. Some real-world applications include:

1. **Customer Support Chatbots** – E-commerce platforms use NLP-powered chatbots to answer customer queries instantly. For example, a customer may ask *“Where is my order?”* and the chatbot replies with the current order status. This saves time and reduces the workload on human support staff.
2. **Sentiment Analysis on Reviews** – Online shopping generates thousands of reviews daily. NLP algorithms analyze these reviews to figure out customer opinions about a product, whether they are positive, negative, or neutral. This helps companies to improve their products.
3. **Search Optimization** – When users type search queries like *“affordable black running shoes”*, NLP helps the system to understand the intent and show the most relevant products.
4. **Personalized Recommendations** – Based on browsing history and past purchases, NLP models analyze the type of products a customer likes and recommend similar items. This increases sales and improves customer satisfaction.

Example:

Companies like Amazon and Flipkart use NLP in almost every part of their platform. Their chatbots answer queries, their search system understands natural language, and their recommendation engine shows personalized products. This shows how NLP is transforming the e-commerce industry by making online shopping smoother and more interactive.

Q3: - What is NLU and NLG?

SOL: - Natural Language Understanding (NLU):

NLU is a subfield of NLP that mainly deals with understanding the meaning, intent, and context behind human language. Simply put, it helps machines “understand” what a person is trying to say. NLU focuses on breaking down sentences, identifying important entities, and removing confusion caused by synonyms or ambiguous words.

- **What NLU does:**

- Finds entities such as names, dates, products, and locations.
- Identifies the *intent* of the user (e.g., ordering food, booking a cab, checking weather).
- Handles different ways of saying the same thing.

- **Example:**

If a person says, “*Order me a pizza at 8 pm*”, the NLU system will extract:

- Intent: Order food
- Entities: Item = pizza, Time = 8 pm

So even if the user says, “*Can I get a pizza delivered by 8 tonight?*”, the system will understand the same meaning.

Natural Language Generation (NLG):

NLG is the other side of NLP, and it deals with producing meaningful sentences in natural language from structured data. Where NLU focuses on understanding, NLG focuses on creating human-like responses.

- **What NLG does:**

- Converts data or responses into natural sentences.
- Writes summaries, reports, or answers automatically.
- Generates text for chatbots, virtual assistants, and even news articles.

- **Example:**

If the input data is {Item: pizza, Time: 8 pm}, then NLG will generate the response:

→ “*Your pizza order has been placed for 8 pm.*”

This makes the system feel more natural and user-friendly, instead of replying with robotic or data-only responses.

In summary:

- **NLU** = Understanding human input (what the user means).
- **NLG** = Generating a human-like response (how the system replies).
Together, NLU and NLG make systems like chatbots, voice assistants (Alexa, Siri), and customer service bots work smoothly.