NLP Tasks

Basic NLP (Natural Language Processing) tasks are foundational for understanding, processing, and generating human language. Here are some core tasks in NLP:

1. Tokenization

- Splitting text into smaller units such as words, sentences, or subwords.
- Example:
    - Input: "Natural Language Processing is fun!"
    - Output: ["Natural", "Language", "Processing", "is", "fun", "!"]

2. Part-of-Speech (POS) Tagging

- Assigning grammatical tags (e.g., noun, verb, adjective) to words in a sentence.
- Example:
    - Input: "I love NLP."
    - Output: [("I", "PRP"), ("love", "VBP"), ("NLP", "NN")]

3. Named Entity Recognition (NER)

- Identifying and classifying entities (e.g., people, organizations, locations) in text.
- Example:
    - Input: "Barack Obama was born in Hawaii."
    - Output: [("Barack Obama", "PERSON"), ("Hawaii", "LOCATION")]

4. Sentiment Analysis

- Determining the sentiment or emotion conveyed by text (e.g., positive, negative, neutral).
- Example:
    - Input: "I love this movie!"
    - Output: "Positive"

5. Lemmatization and Stemming

- Reducing words to their root forms.
- Example:
    - Input: "running", "ran", "runs"
    - Output (Lemmatization): "run"
    - Output (Stemming): "run"

6. Stop Word Removal

- Removing common words (e.g., "and", "is", "the") that do not add significant meaning.

- Example:
  - Input: "The cat is on the mat."
  - Output: ["cat", "mat"]

7. Text Classification

- Categorizing text into predefined classes or labels.
- Example:
  - Input: "This is a sports article."
  - Output: "Sports"

8. Language Modeling

- Predicting the next word in a sequence or assigning probabilities to sequences of words.
- Example:
  - Input: "The cat sat on the ___"
  - Output: ["mat"(0.8), "chair"(0.1), "floor"(0.1)]

9. Machine Translation

- Translating text from one language to another.
- Example:
  - Input: "Hello, how are you?"
  - Output: "Hola, ¿cómo estás?"

10. Text Summarization

- Generating a concise summary from a longer text.
- Example:
  - Input: "Natural language processing is a subfield of AI. It involves understanding and generating human language."
  - Output: "NLP is a subfield of AI for understanding and generating language."

11. Coreference Resolution

- Identifying when different words refer to the same entity.
- Example:
  - Input: "Alice said she would come."
  - Output: "Alice"-> "she"

12. Question Answering

- Answering questions based on a given context or knowledge.
- Example:

- Input: Context: "Paris is the capital of France."Question: "What is the capital of France?"
- Output: "Paris"

13. Word Embeddings

- Representing words as continuous vectors in a high-dimensional space.
- Example:
    - Input: "king", "queen", "man", "woman"
    - Output: Relationships such as "king - man + woman ≈ queen"

14. Spell Checking and Text Normalization

- Correcting misspellings and standardizing text (e.g., "u r"-> "you are").
- Example:
    - Input: "I luv NLP!"
    - Output: "I love NLP!"

15. Text Generation

- Generating coherent text from a model or input prompt.
- Example:
    - Input: "Once upon a time,"
    - Output: "Once upon a time, there lived a brave knight in a faraway kingdom."

These tasks serve as building blocks for more complex NLP systems and applications, such as chatbots, virtual assistants, and recommendation systems.