إعداد : أحمد ضياء الشحات أبو المعاطي إبراهيم - ID : 803089485

NLP section : Task 2

Search for NLP applications.

- **Text Classification:**

Categorizing text into predefined classes (e.g., **sentiment analysis**, spam detection, topic modeling).

- **Named Entity Recognition (NER):**

Identifying and classifying named entities in text (e.g., people, organizations, locations).

- **Machine Translation:**

Converting text from one language to another.

- **Chatbots:**

Conversational agents that interact with users in a simulated manner.

- **Text Summarization:**

Condensing lengthy text into a shorter, informative representation.

- **Question Answering:**

Extracting answers to questions posed in natural language.

- **Paraphrasing:**

Rewriting text while preserving its meaning.

- **Speech Recognition:**

Transforms spoken words into text (e.g., virtual assistants) like Siri or Alexa.

- **Text Generation:**

Creates new text content (e.g., poems, code).

- **Information Retrieval:**

Locates relevant information from vast text data (e.g., search engines).

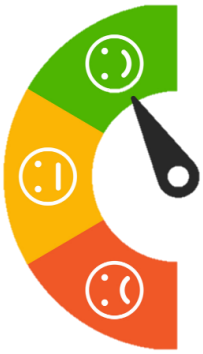
- **Automatic writing assistants**

analyze writing styles and suggest improvements like grammar checkers, plagiarism detection, and even content rephrasing for better clarity.

- **Text-to-speech(TTS):**

convert written text into spoken language. This is crucial for accessibility features and allows for the creation of audiobooks and other spoken content from written text.

Select one of the applications as your favorite one.



Sentiment Analysis

- **Core Functionality:** Determining the emotional tone of text (positive, negative, neutral).
- **Transferability:** Useful in social media monitoring, customer feedback analysis, and brand reputation management.
- **Example:** Analyze social media sentiment towards a new product launch.

Search for a suitable dataset according to the selected

Reddit Dataset

Reddit, a vast online platform teeming with user-generated content, offers a rich source of data for Natural Language Processing (NLP) applications. The sheer volume and variety of text present on Reddit make it a valuable resource for training and evaluating NLP models.

- **Reddit_Data.csv:** This specific file focuses on Reddit comments. It likely includes columns for:
- **clean_comment:** The actual content of the Reddit comment.
- **category:** The assigned sentiment label (positive, negative, or neutral).
    - 0 => Indicating it is a **Neutral** Comment
    - 1=> Indicating a **Positive** Sentiment
    - -1 =>Indicating a **Negative** Comment