ASSIGNMENT 11 – NATURAL LANGUAGE PROCESSING (NLP)

Name: Aman Nadeem
Roll No: 02
Course: Applied Data Science with AI
Topic: NLP Preprocessing

Objective

The goal of this task was to apply basic Natural Language Processing techniques such as tokenization, stopword removal, and TF-IDF vectorization on text data.

As the main project dataset is tabular, a small text dataset related to customer feedback was created to demonstrate the NLP pipeline.

Dataset Used

Sample customer feedback sentences such as:

- "The service is very bad and disappointing"
- "I am happy with the internet speed"
- "Billing charges are too high"

NLP Techniques Applied

1. Tokenization

Each sentence was split into individual words.

2. Stopword Removal

Common words like *is, the, and, with* were removed to clean the data.

3. TF-IDF Vectorization

Text was converted into numerical form using the TF-IDF method for further analysis or modeling.

	Customer_Feedback
0	The service is very bad and disappointing
1	I am happy with the internet speed
2	Customer care is not helpful
3	Very satisfied with the package
4	Billing charges are too high

Result

The output was a matrix representing the importance of each word in every sentence. This can be used later for sentiment analysis or text classification.

	Customer_Feedback	Tokens	Cleaned_Tokens
0	The service is very bad and disappointing	[The, service, is, very, bad, and, disappointing]	[service, bad, disappointing]
1	I am happy with the internet speed	[I, am, happy, with, the, internet, speed]	[happy, internet, speed]
2	Customer care is not helpful	[Customer, care, is, not, helpful]	[Customer, care, helpful]
3	Very satisfied with the package	[Very, satisfied, with, the, package]	[satisfied, package]
4	Billing charges are too high	[Billing, charges, are, too, high]	[Billing, charges, high]

Conclusion: This task successfully demonstrates an NLP preprocessing pipeline. Although the main project is numeric-based, this experiment shows readiness to handle text-based features in future datasets.

GitHub Link : https://github.com/amannadeem126/Customer-Churn-Prediction