# Assignment 1: Text Preprocessing and Feature Engineering

In this assignment, you will explore various text preprocessing techniques and feature engineering methods commonly used in natural language processing (NLP).

Tasks- :

Stopwords Removal: Implement stopwords removal using N- LTK.

Tokenization: Tokenize the given text into words or sente- nces.

Stemming and Lemmatization: Apply stemming and lemmatization on the t- okens.

POS Tagging: Perform Part-of-Speech tagging on the - tokens.

TF-IDF Vectorization: Convert the preprocessed text into TF-IDF - vectors.

One-Hot Encoding: Encode the tokens using one-hot - encoding.

Bag of Words: Create a bag of words representation of- the text.

Unigram, Bigram, n-gram: Generate unigram, bigram, and n-gram representations of the text.

Dataset:

Use any text dataset of your choice, such as movie reviews, news articles, or any other ## text corpus.

Deliverables:

Write a Python script that performs the above tasks on the given text dataset. Include comments to explain each step clearly. Test your script on a sample text dataset and ensure it produces the desired outputs for each task.