**NLP Final Project Goal**

The project aims to develop a chatbot for summarizing text documents using two methods: a classical NLP-based approach (involving tokenization, POS tagging, Naive Bayes, and logistic regression for extractive summarization) and a neural network-based approach (using RNNs or LSTMs for abstractive summarization). The goal is to compare the effectiveness, efficiency, and scalability of these methods, with the chatbot enabling users to upload documents and choose between the two summarization techniques.

**Project Modules and Team Responsibilities**

The project will be divided into two main modules, with each team member focusing on a distinct aspect of the project:

1. Module 1: *Classical NLP-Based Summarization*  
   This module will focus on implementing the classical NLP techniques for extractive summarization. The main tasks will include:
   * Word tokenization and byte pair encoding (BPE) for handling rare and unknown words.
   * Word normalization, including stemming and lemmatization.
   * Part-of-speech tagging, named entity recognition (NER), and word representation using Word2Vec/GloVe embeddings.
   * Text classification with Naive Bayes or logistic regression to rank sentences for extractive summarization.
   * Evaluation using precision, recall, and other relevant metrics.

The team member responsible for this module will ensure that the classical NLP methods are properly integrated into the summarization pipeline.

1. Module 2: *Neural Network-Based Summarization*  
   This module will focus on training and deploying a neural network model for abstractive summarization. The main tasks will include:
   * Implementing an RNN or LSTM model to generate abstractive summaries.
   * Using dynamic word embeddings to adapt to different contexts.
   * Evaluating the model's performance using relevant metrics and backpropagation.

The second team member will work on building and fine-tuning the neural network model for summarization.

**Final Project Outcome**

The expected deliverables for this project are:

* A trained model for both classical NLP-based and neural network-based summarization techniques.
* A comparative analysis report that evaluates the strengths and weaknesses of the two approaches.