KU AICL434 Assignments

## Assignment 1.1: NLP Preprocessing Basics

**Objective**: Apply core NLP preprocessing techniques and demonstrate them via an interactive web interface.  
**Tasks**:

* Implement tokenization, lemmatization, stemming, POS tagging, and Named Entity Recognition (NER) using NLTK or spaCy.
* Compare lemmatization and stemming with at least 10 examples and explain the differences.
* Create a REST API that exposes these preprocessing functions.
* Build a simple demo web app where users can input text and view the processed output interactively.

## Assignment 1.2: Word Embeddings & Visualization

**Objective:** Explore and apply word embedding techniques with an interactive component.  
**Tasks:**

* Use TF-IDF or GloVe embeddings on a small custom corpus.
* Visualize embeddings using dimensionality reduction techniques like t-SNE or PCA.
* Develop a REST API to compute and return embeddings for input words.
* Create a simple web app that lets users enter words and see their embeddings and nearest neighbors.

## Assignment 1.3: Seq2Seq Summarization with LSTM

**Objective**: Implement and evaluate a basic sequence-to-sequence model for text summarization.  
**Tasks**:

* Build an encoder-decoder model using LSTM layers for abstractive text summarization.
* Use a small dataset (e.g., 100–200 news articles or custom data).
* Train and evaluate the model.