Natural Language Processing Lab

Exercise 2:

Application: Sentiment Analysis

Dataset: Download the dataset from [Sentiment Analysis Dataset (kaggle.com)](kaggle.com)

Task:
- Explain the pipeline for developing sentiment analysis task.
- Perform cleaning and preprocessing of text.
- Generate representations using:
  - Bag of Words
  - TF-IDF
  - Continuous Bag of Words
  - Skip gram
  - Word2Vec.
- Classify the data using appropriate machine learning techniques to generate labels.
- Analyze the labels and explain the impact of embedding techniques in misclassification.
- Discuss the limitations of each embedding technique and explain the techniques that rectify it.

Input:
What is not to like about this product.
Not bad.
Not an issue.
Not buggy.
Not happy.
Not user-friendly.
Not good.
Is it any good?
I do not dislike horror movies.
Disliking horror movies is not uncommon.
Sometimes I really hate the show.
I love having to wait two months for the next series to come out!
The final episode was surprising with a terrible twist at the end.
The film was easy to watch but I would not recommend it to my friends.
I LOL'd at the end of the cake scene