Title: Sentiment Analysis on Customer Reviews

1. Introduction Sentiment analysis, also known as opinion mining, is a key application of Natural Language Processing (NLP) that aims to determine the sentiment behind a given text. This project focuses on analyzing customer reviews to classify them as positive, negative, or neutral. Businesses can leverage sentiment analysis to understand customer opinions, improve products, and enhance customer satisfaction. The growing reliance on online reviews makes sentiment analysis a crucial tool for decision-making.

2. Problem Statement Customer feedback is essential for businesses to improve their services. However, manually analyzing thousands of customer reviews is time-consuming and inefficient. This project aims to automate sentiment analysis using NLP techniques to categorize reviews effectively. The primary challenge is handling sarcasm, negations, and context-dependent sentiments accurately.

3. Research Questions

- How accurately can NLP models classify customer sentiments in reviews?
- What preprocessing techniques improve sentiment classification accuracy?
- How does sentiment analysis help businesses in decision-making?
- What are the challenges in handling ambiguous or context-dependent sentiments?

4. Dataset

- **Dataset Name:** IMDB Movie Reviews Dataset (or Amazon/Yelp Reviews Dataset)
- **Source:** Kaggle, UCI Machine Learning Repository, or official company databases.
- **Properties:**
    - Number of reviews: 50,000 (IMDB dataset) or more depending on the source.
    - Attributes: Review text, sentiment label (positive/negative/neutral), metadata (date, user, rating).
    - Format: CSV/JSON
    - Preprocessed: Tokenized, cleaned (stopwords removal, stemming, lemmatization)

This project will apply NLP techniques such as tokenization, vectorization (TF-IDF, Word2Vec), and deep learning models (LSTM, BERT) to achieve optimal sentiment classification results.