**Sentiment Analysis – Mini Project Report**

**1. Introduction**

This project aims to perform **basic sentiment analysis** on textual data using Natural Language Processing (NLP). The main goal is to identify whether a piece of text expresses a **positive**, **negative**, or **neutral** sentiment. This was a simple academic project intended to understand the workflow of sentiment classification using labeled data.

**2. Dataset Overview**

The dataset used in this project contains:

* **Text**: User comments or opinions in plain text.
* **Sentiment**: Labels such as *Positive*, *Negative*, or *Neutral*.
* **Source**: Data collected from open datasets available on platforms like **Kaggle**.

**3. Challenges Faced and How We Solved Them**

* **Missing Values**:  
  Some entries had missing text or scores. We handled this by:
  + Removing empty text entries
  + Filling missing numerical values with average values
* **Imbalanced Data**:  
  Some sentiments were more frequent than others. To fix this:
  + We applied **basic oversampling/undersampling** methods to balance the dataset

**4. Key Observations**

* Positive sentiments were more common in general reviews.
* Although it was a simple project, we saw how basic text preprocessing and classification can be powerful.

**5. Conclusion**

This project helped us learn how to perform **basic sentiment analysis** using Python and NLP libraries. It was not real-time, but it gave us a clear understanding of how text data can be cleaned, labeled, and analyzed for sentiment. Future scope can include real-time data handling, advanced deep learning models, and multilingual support—but for now, this project served as a solid foundation.