# Customer Feedback Analysis for E-commerce

**1. Team Composition**

The project will be conducted by Parth Maniar and Shubham Patel.

**2. Statement of the Problem and Solution Approach**

* **Problem Statement:** E-commerce platforms accumulate vast amounts of customer feedback in the form of reviews and ratings. However, extracting meaningful insights from this unstructured textual data is challenging. The goal is to analyze customer feedback to identify common sentiments, key concerns, and areas for improvement to enhance customer satisfaction and inform business strategies.
* **Solution Approach:** I plan to develop a natural language processing (NLP) pipeline that processes customer reviews to perform sentiment analysis and topic modeling. This involves cleaning the data, analyzing sentiments, extracting prevalent themes, and presenting the findings through visualizations.

**3. Data Source and Acquisition**

* **Data to Use:** I will utilize publicly available customer reviews from e-commerce platforms like Amazon.
* **How to Get It:** The data will be obtained through:
* **Existing Datasets:** Accessing datasets like the Amazon Customer Reviews Dataset available on Kaggle or AWS Open Data Registry.
* **Web Scraping:** If necessary, scraping additional data using tools like Scrapy, adhering strictly to ethical guidelines and the website's terms of service.

**4. Algorithms and Techniques**

* Data Preprocessing: Tokenization, Stop-word removal, Stemming or lemmatization
* Sentiment Analysis:
  + Implementing machine learning classifiers such as Naive Bayes or Support Vector Machines.
  + Exploring deep learning models like Recurrent Neural Networks (RNNs) or Transformers for improved accuracy.
* Topic Modeling: Uing Latent Dirichlet Allocation (LDA) to identify common themes within the reviews.
* Aspect-Based Sentiment Analysis: Analyzing sentiments related to specific product features or aspects.
* Visualization Tools: Utilizing libraries like Matplotlib or Seaborn to create insightful charts and graphs.

**5. Measuring Success**

* Quantitative Metrics: Evaluating model performance using accuracy, precision, recall, and F1-score for sentiment classification.
* Qualitative Assessment:
  + Validating the relevance and usefulness of the extracted topics and insights.
  + Gathering feedback from potential stakeholders on the applicability of the findings.
* Impact Evaluation: Assessing whether the insights can lead to actionable business strategies that improve customer satisfaction.

**6. Expected Deliverables at the End of the Term**

* Comprehensive Report: Detailed documentation of the methodologies used, analyses performed, results obtained, and conclusions drawn.
* Functional NLP Pipeline: A working prototype that can process and analyze customer feedback data.
* Visual Presentations: Graphs and charts depicting sentiment distributions, key topics, and other relevant insights.
* Final Presentation: An oral presentation summarizing the project, showcasing the findings, and discussing potential business implications.

------------------------

This project aims to bridge the gap between unstructured customer feedback and actionable business insights by leveraging advanced NLP techniques. The end goal is to provide a tool that can help e-commerce businesses better understand their customers and improve their products and services accordingly.