Data Collection and Preprocessing Phase

Date	3 october 2024
Team ID	LTVIP2024TMID24974
Project Title	Analysis Of Amazon Cell Phone Reviews
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Report:

The data for the project will be collected from publicly available customer reviews on Amazon. This includes user ratings, comments, and metadata for cell phone products. The data will be gathered via web scraping or from pre-existing datasets on platforms like Kaggle or public APIs.

The identified raw data source is an Amazon cell phone review dataset in CSV format, which is publicly accessible. It contains all necessary features such as review content, ratings, and product metadata for analysis.

Data Collection Plan:

Section	Description
Project Overview	This project aims to analyze customer reviews of cell phones sold on Amazon, applying machine learning techniques to extract insights regarding user satisfaction, common issues, and product performance trends.
Data Collection Plan	The data will be collected from Amazon's public product review dataset, focusing specifically on cell phone reviews. Methods include web scraping or accessing public datasets via APIs.
Raw Data Sources Identified	Amazon review dataset containing cell phone review texts, ratings, and metadata such as review date and product information.

Raw Data Sources Report:

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	This dataset contains reviews, ratings, and metadata of various cell phones sold on Amazon, including customer feedback, review dates, star ratings, and product information.	https://www.kaggle.com/grikomsn/amazon-cell-phones-reviews	CSV	5 MB	Public