Project Initialization and Planning Phase

Date	11 November 2024
Team ID	SWTID1727420425
Project Title	analysis of amazon review using nlp technique
Maximum Marks	3 Marks

Project Proposal (Proposed Solution) template

The goal of this project is to analyze a large dataset of Amazon product reviews using advanced Natural Language Processing (NLP) techniques to gain insights into customer sentiment, trends, and product feedback. This analysis will help businesses and customers understand product performance and improve decision-making. The project will focus on sentiment analysis, keyword extraction, and review classification.

Project Overview	
Objective	The objective is to leverage NLP techniques to analyze Amazon reviews for extracting sentiment, trends, and actionable insights.
Scope	The scope includes applying NLP techniques to analyze Amazon reviews for sentiment analysis, topic modeling, trend identification across various product categories.
Problem Statement	
Description	Manually analyzing Amazon's vast reviews is inefficient, requiring automated solutions to extract insights, identify trends, and filter irrelevant feedback.
Impact	This project enables businesses to automate insights extraction from Amazon reviews, driving better customer satisfaction, product development, and marketing strategies.
Proposed Solution	
Approach	The approach involves applying NLP techniques like sentiment analysis, topic modeling, and text classification to automate the extraction of insights from Amazon reviews.

Key Features	topic modeling, keyword extraction, and review classification to derive actionable insights from Amazon reviews.

Resource Requirements

Resource Type	Description	Specification/Allocation
Hardware		
Computing Resources	AWS, Google Cloud, or Microsoft Azure for data storage, processing, and model training,	e.g., 16-32 GB
Memory	RAM , GPU/CPU	e.g., 32 GB
Storage	high-speed SSD storage	e.g., 500 GB
Software		
Frameworks	Python frameworks	e.g., Flask
Libraries	NLP libraries	e.g., SpaCy
Development Environment	software tools, libraries, hardware resources	e.g., Zira, Git,Vs
Data		
Data	• **Data Sources:** Public datasets (Amazon, Yelp, IMDb), web scraping, APIs (Twitter, News), proprietary/internal data, and crowdsourced data. • **Data Size:** Small-scale (thousands of entries), medium-scale (tens of thousands), large-scale (millions to billions). • **Data Format:** Text files, CSV, JSON, XML, Parquet, TFRecord — depending on how data is stored or retrieved	e.g., user details, review rating