Model Development Phase Template

Date	25 Jan 2025
Team ID	740678
Project Title	Amazon Kindle Store Reviews Analysis
Maximum Marks	5 Marks

Model Selection Report

The Model Selection Report for Amazon Kindle Store Reviews Analysis focuses on leveraging machine learning and natural language processing (NLP) techniques to extract insights from customer reviews of Kindle books. The objective is to understand customer sentiments, identify trending topics, detect fake reviews, and support authors and marketers with data-driven strategies.
.

Model Selection Report:

Model	Description
Model 1	**BERT (Bidirectional Encoder Representations from Transformers)** BERT is a transformer-based machine learning technique for NLP developed by Google. It has revolutionized many NLP tasks due to its bidirectional context understanding, allowing it to grasp the meaning of words based on all their surrounding words. **Key Components:** • **Contextual Word Embeddings:** Unlike traditional embeddings (e.g., Word2Vec), BERT captures context from both left and right of a word, providing deeper semantic understanding. • **Pre-trained Models:** BERT models pre-trained on large corpora like BooksCorpus and

Wikipedia are fine-tuned on the Kindle review dataset for tasks such as sentiment classification and fake review detection.

- **Fine-Tuning for Sentiment Analysis:**
  BERT is fine-tuned using labeled review data (positive, negative, neutral) to classify the sentiment of user reviews accurately.

- **Topic Modeling with BERT + BERTopic:**
  Using BERT embeddings with the BERTopic algorithm enables dynamic topic modeling to understand themes emerging in customer feedback over time.

- **Performance and Adaptability:**
  BERT models show high accuracy on sentiment classification tasks and can be adapted for aspect-based sentiment analysis to explore sentiments about plot, pricing, delivery, and writing style.