

Amazon Reviews Sentiment Analysis Report

Prepared By: Gangishetty Saketh Date: 07/02/2026

1. Executive Summary This project performs sentiment analysis on Amazon product reviews using Natural Language Processing (NLP) and Machine Learning techniques. The dataset was cleaned and preprocessed, then analyzed to understand review trends, customer sentiment distribution, and frequent keywords.

After preprocessing, the reviews were categorized into three sentiment classes: Positive (Score ≥ 4) Neutral (Score = 3) Negative (Score ≤ 2) A total of 20,000 reviews were processed to ensure fast execution while maintaining meaningful results. Three classification models were trained and evaluated: Logistic Regression Naive Bayes Support Vector Machine (SVM) The results show that SVM generally performs best for text classification using TF-IDF features.

2. Text Analysis & Sentiment Insights 2.1 Data Cleaning and Preprocessing The following steps were applied to clean the reviews: Combined Summary + Text into a new column called full_review Converted text to lowercase Removed HTML tags Removed punctuation and special characters Removed stopwords using NLTK Lemmatization using WordNet Lemmatizer A new column clean_text was created after preprocessing.

2.2 Sentiment Distribution The dataset was labeled based on review scores: Positive reviews were the majority Neutral reviews were comparatively fewer Negative reviews were present but less than positive reviews This shows that most customers gave good ratings.

2.3 WordCloud Analysis WordClouds were generated for each sentiment category: Positive WordCloud: shows words like good, great, love, best, excellent Neutral WordCloud: shows mixed words such as product, okay, average, fine Negative WordCloud: shows words like bad, poor, waste, disappointed, broken This provides a clear idea of what customers like and dislike.

2.4 Bigram Frequency Analysis Bigrams (two-word combinations) were extracted using CountVectorizer. This helps identify common phrases such as: good quality highly recommend poor quality not worth Bigram analysis provides stronger insights than single word frequency.

3. Classifier Performance & Findings 3.1 Feature Extraction TF-IDF Vectorizer was used with: max_features = 5000 This converts cleaned text into numerical vectors for machine learning models.

3.2 Models Used Three models were trained:

1. Logistic Regression Works well for text classification Fast and efficient
2. Naive Bayes Very fast Works well for word-frequency based classification
3. Support Vector Machine (SVM) Strong performance for high-dimensional text data Best suited for TF-IDF features

3.3 Model Evaluation Each model was evaluated using: Accuracy Score Classification Report (Precision, Recall, F1-score) Confusion Matrix (for SVM)

3.4 Key Finding The SVM model generally performs best for sentiment classification because: It handles sparse text vectors effectively It separates classes more accurately than other models A confusion matrix heatmap was plotted to visualize model performance across all classes.

4. Topic Modeling & Keyword Extraction (LDA) Topic Modeling was performed using Latent Dirichlet Allocation (LDA): num_topics = 5 passes = 10 This extracted 5 main topics from customer reviews. Example topic patterns include: Product quality and durability Packaging and delivery experience Value for money Customer satisfaction Complaints and defects This helps identify what customers talk about most.
5. Trend Analysis (Sentiment Over Time) A time-based trend analysis was created by generating a synthetic date range. The sentiment counts were grouped month-wise and plotted. Trend Insights Positive sentiment remains consistent across time Negative spikes can indicate customer issues Useful for tracking product performance and customer satisfaction trends
6. Named Entity Recognition (NER) Named Entity Recognition was performed using spaCy: Model used: en_core_web_sm Entities extracted include: ORG (Organizations) PRODUCT PERSON GPE (Locations) DATE This can be useful for extracting brands, product names, and locations mentioned in reviews.
7. Actionable Insight Generation Top 10 words from positive and negative reviews were extracted using word frequency counts. Positive Words (Examples) good great love best excellent Negative Words (Examples) bad poor waste broken disappointed Business Suggestions Improve issues related to the top negative

keywords Promote features related to top positive keywords If negative sentiment rises in a month, prepare corrective actions and customer support response

8. Reporting & Visualization The following visualizations were generated:

8.1 Star Rating Distribution A count plot was created to show: How many reviews exist for each star rating (1–5)

8.2 Interactive Plot Plotly histogram was generated: Ratings vs Sentiment Color-coded by sentiment category This helps explore rating and sentiment relationships interactively.