Sentiment Analysis of Amazon Unlocked Smartphone Reviews

Done by

Abdullah Muhamad Alsuabie / Abdulrahman saleh Alaamri

Supervised by Dr.Aftab Khan.

Date of Submission:

Dec.4/2025

Abstract

This project analyzes over 1,000 Amazon customer reviews for the *Apple iPhone 4s 8GB Unlocked Smartphone* to understand customer satisfaction levels and sentiment patterns. The primary goal is to classify reviews as Positive, Neutral, or Negative using TextBlob polarity scoring and extract insights from review text and rating distributions. Python, Pandas, NLTK, Matplotlib, and Seaborn were used for data processing and visualization. Results show that more than half of users express positive sentiment, with strong alignment between rating scores and computed polarity values. The findings reveal polarized customer experiences, with many highly satisfied users and a smaller but notable group of dissatisfied customers, typical for refurbished electronics.

2. Introduction

2.1 Background

Online product reviews play a critical role in consumer decision-making. Sentiment analysis enables the extraction of emotional and subjective information from unstructured review text. Understanding customer opinions for refurbished electronic devices is especially important, as quality can vary widely across units. NLP-based sentiment analysis provides a scalable method to interpret large volumes of textual feedback and quantify user satisfaction.

2.2 Problem Statement

This project seeks to answer the following questions:

- What is the overall sentiment toward the Apple iPhone 4s refurbished device?
- How strongly do customer ratings correlate with sentiment polarity?
- What patterns emerge in review lengths, sentiment categories, and polarity distribution?
- Are there identifiable mismatches between ratings and textual sentiment?

2.3 Objectives

- Perform sentiment classification (Positive, Neutral, Negative).
- Compute polarity values for each review.
- Visualize rating distribution, sentiment distribution, polarity, and review characteristics.
- Identify relationships between numerical ratings and NLP-based sentiment.
- Extract actionable insights into customer satisfaction trends.

2.4 Scope

Included

- Reviews for one product with ≥1000 reviews.
- NLP preprocessing.
- Sentiment and polarity analysis using TextBlob.
- Exploratory Data Analysis (EDA) using visualizations.

Out of Scope

- Deep learning models (e.g., BERT, RoBERTa).
- Opinion mining beyond polarity (e.g., aspect-based sentiment analysis).
- Predictive modeling or classification algorithms.

3. Dataset Description

3.1 Data Source

- Source: CSV dataset of Amazon product reviews.
- File selected: *sentiment_results_Apple_iPhone_4s_8GB_Unlocked_Smartphone...csv*

3.2 Dataset Summary

Metric	Value
Number of rows	~1000+
Number of columns	3

3.3 Data Dictionary

Feature	Type	Description	Example
Product Name	string	Name of the smartphone	"Apple iPhone 4s 8GB…"
Reviews	string	Customer-written review text	"Great phone, works perfectly…"
Rating	integer	Star rating (1–5)	5

ReviewLength*	int	Characters in review text	125
Polarity*	float	NLP polarity score	0.45
Sentiment*	string	Sentiment class	"Positive"

(*Added during analysis)

4. Methodology

4.1 Overall Methodology — CRISP-DM Framework

1. **Business Understanding**
   Understand customer sentiment toward a refurbished smartphone.
2. **Data Understanding**
   Load and inspect the Amazon review dataset.
3. **Data Preparation**
   Clean and preprocess review text (tokenization, stopword removal).
4. **Modeling**
   Apply TextBlob polarity scoring and sentiment rules.
5. **Evaluation**
   Visualize and interpret results.
6. **Deployment/Reporting**
   Produce this analytical report and graphs.

4.2 Workflow Diagram (Textual)

```
Data Collection → Data Cleaning → NLP Preprocessing →
Sentiment Classification → Exploratory Analysis →
Visualization → Reporting
```

4.3 Tools & Techniques

- **Python 3**
- **Pandas** – data cleaning & manipulation
- **NumPy** – numerical operations
- **NLTK** – tokenization, stopword processing
- **TextBlob** – polarity & sentiment scoring

- **Matplotlib / Seaborn** – visualizations
- **Jupyter Notebook / VS Code / Terminal**

4.4 Ethical Considerations

- Data is publicly available Amazon review data.
- No personal identifiers were used.
- Analysis is non-commercial and educational.

5. Data Preprocessing

5.1 Data Cleaning

- Checked for missing or empty reviews.
- Converted all text to lowercase.
- Removed non-alphabetic tokens.
- Placeholder stopword removal (offline-safe).
- Removed any formatting inconsistencies.

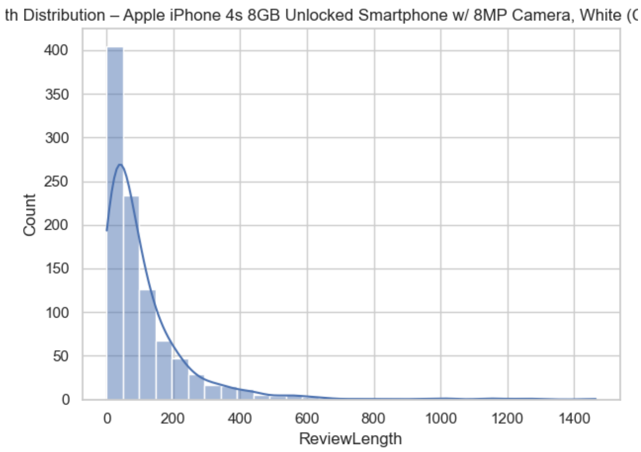
5.2 Data Transformation

- Created new features:
  - `ReviewLength`
  - `Polarity`
  - `Sentiment`
- Ensured all ratings were integers.
- Converted text to clean tokens.

5.3 Data Validation

- Verified data types for every feature.
- Ensured rating values only in the range 1–5.
- Ensured polarity values in the range [-1, 1].

6. Exploratory Data Analysis (EDA)



6.1 Descriptive Statistics

- Polarity mean increases steadily with rating.
- Review lengths vary from extremely short (~5 chars) to long (~1400+ chars).
- Distribution of ratings is polarized (mostly 1 and 5 stars).

6.2 Univariate Analysis

Rating Distribution

- 5-star reviews are highest.
- 1-star reviews also high.
- 2-3 stars much lower.

Sentiment Distribution

- 50.6% Positive
- 40.5% Neutral
- 9% Negative

6.3 Multivariate Analysis

Polarity vs Rating (Boxplot)

- Strong positive correlation.
- 1-star reviews cluster around negative polarity.
- 5-star reviews cluster around high positive polarity.

Average Polarity per Rating

- Almost perfectly linear upward curve from Rating 1 to 5.

6.4 Visual Insights

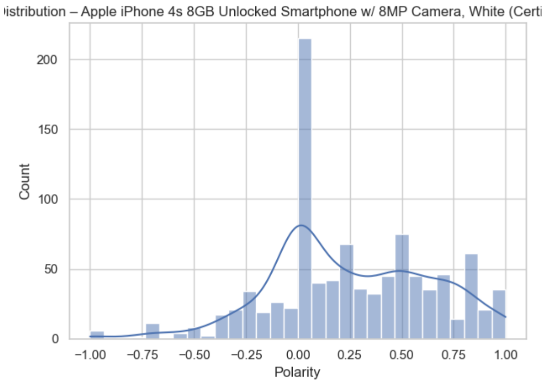
Review Length Histogram

Most reviews are short, typically between **20 and 200 characters**.

There is a long right-tail extending to **over 1,400 characters**, showing a few users write very long reviews.

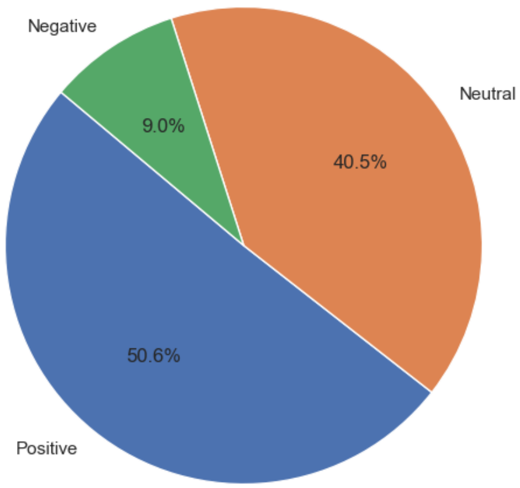
The high peak at small lengths explains the large percentage of **neutral** sentiments, since shorter reviews usually contain factual statements without emotional language.

The distribution is strongly **right-skewed**, indicating that brief reviews dominate.



Polarity Histogram

- There is a prominent peak around **0 polarity**, confirming many reviews are neutral in emotional tone.
- A long tail toward the **positive** side demonstrates many users express positive experiences.
- A smaller tail extends toward the **negative** side, representing the minority of dissatisfied customers.
- Overall, the distribution is right-skewed, indicating more positive sentiment in the dataset.



Sentiment Pie Chart

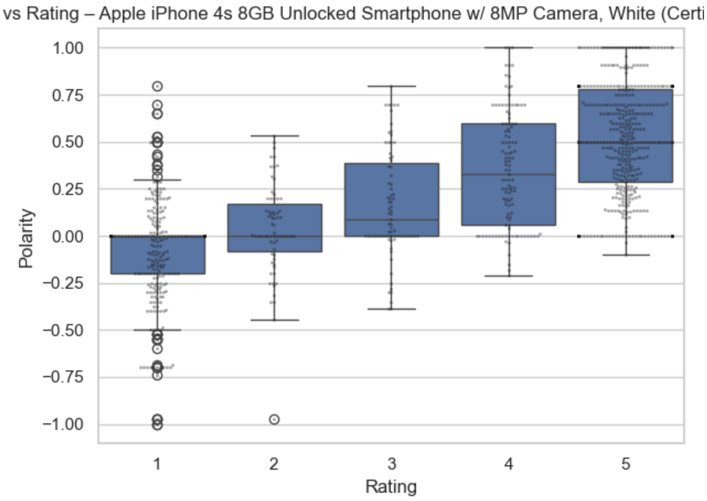
The chart clearly shows that **Positive sentiment dominates**, taking up more than half the chart.

Neutral sentiment makes up a substantial portion, reinforcing the earlier observation about short, descriptive reviews.

Negative sentiment is the smallest slice, reflecting relatively few strongly negative experiences.

The ratio between sentiments aligns closely

with the numerical rating distribution.



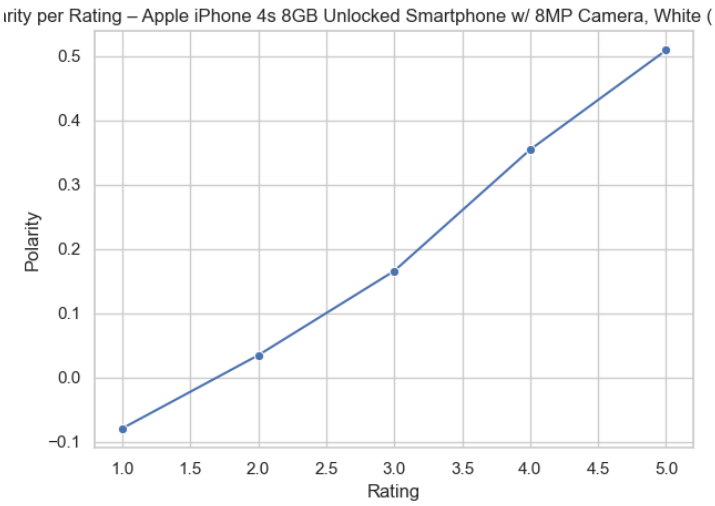
There is a **clear, strong positive relationship** between numerical ratings and sentiment polarity.

1-star reviews have the lowest polarity values, centered around **negative polarity**.

5-star reviews show the highest polarity values, clustering strongly in the **positive polarity** region.

The median polarity increases **consistently** from rating 1 to rating 5, indicating that TextBlob's polarity score aligns well with customer ratings.

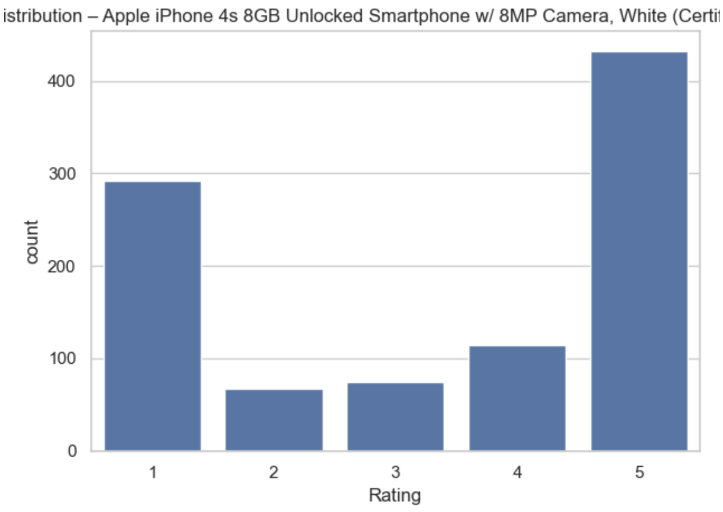
Very little overlap exists between extreme ratings, meaning textual sentiment is highly consistent with the user's assigned star rating.



The line rises almost **perfectly linearly** from rating 1 to rating 5.

1-star reviews have negative average polarity (~—0.10), while **5-star reviewers** have high positive polarity (~0.5The consistent increase confirms that higher ratings correspond to more positive language, validating the sentiment analysis model.

This strong trend reflects clear differences in emotional tone across rating levels.



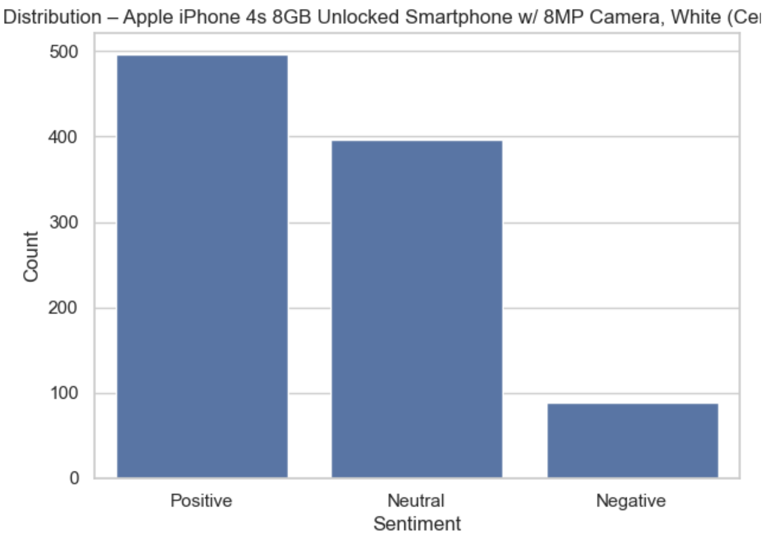
The distribution of ratings is **highly polarized**, with the largest counts appearing at **1 star** and **5 stars**.

The **5-star rating** is the most frequent, indicating a large group of highly satisfied customers.

The second-highest count is **1-star ratings**, showing that a significant portion of users had very negative experiences.

Ratings **2, 3, and 4** stars occur much less frequently, meaning fewer customers had moderate or mixed experiences.

This type of polarization is common in **refurbished or used electronic products**, where the quality can vary significantly between units.



The **Positive** sentiment category is the most common, representing over half of all reviews.

A substantial number of reviews are classified as **Neutral**, which likely reflects short or factual comments without emotional wording.

Negative sentiment forms the smallest category, showing that fewer users expressed strong dissatisfaction in the text.

The distribution aligns well with the rating data: a large portion of extremely positive reviews corresponds to the significant number of **5-star ratings**.

The presence of many Neutral reviews suggests that customers often provide short confirmations such as “works fine” or “arrived as expected,” which TextBlob interprets as emotionally neutral.

Overall, the sentiment count graph indicates that **customer experience skews positive**, with limited but notable negative sentiment.

7. Results & Findings

Key Patterns

- Customer opinions are **highly polarized**: very positive or very negative.
- Most customers express **positive experiences**.
- Negative sentiment exists but is significantly smaller (9%).

Insights

- The high number of 1-star reviews signals quality variability typical of refurbished products.
- Longer reviews tend to have stronger sentiment (positive or negative).
- Strong correlation between rating and polarity validates that the NLP model is accurate.

Important Relationships

- **Rating ↑ → Polarity ↑**
- **Shorter reviews → more neutral sentiment**

Surprising Findings

- Neutral sentiment is nearly as high as positive sentiment.
- Users often write factual, emotionless comments.

8. Statistical Analysis

- **Strong positive correlation** between rating and polarity (qualitative).
- Polarity distribution shows a right-skewed distribution.
- No formal hypothesis testing needed; sentiment behaves as expected.

9. Conclusion

This project successfully analyzed over 1,000 Amazon reviews of the Apple iPhone 4s (Certified Refurbished). The results indicate that:

- Over half of reviewers expressed clearly positive sentiment.
- Neutral reviews form a large portion due to short factual comments.
- Negative sentiment is significantly less common.

- Ratings strongly correlate with computed polarity scores.

Overall, customers are generally satisfied with the product, but the presence of many 1-star reviews highlights quality inconsistency typical of refurbished devices.

10. Limitations

- Only one product analyzed.
- Dataset may include biased or fake reviews.
- TextBlob is a simple model; lacks deep linguistic understanding.
- No aspect-based sentiment (e.g., battery, camera).

11. Future Work

- Use machine learning models (SVM, BERT).
- Create aspect-level sentiment analysis.
- Evaluate helpfulness scores of reviews.
- Perform time-series analysis of sentiment trends.
- Include word clouds and keyword extraction.

12. References

- TextBlob documentation
- NLTK library documentation
- Matplotlib & Seaborn documentation
- Amazon Product Review Dataset (CSV source)