# SENTIMENT ANALYSIS OF AMAZON PRODUCT REVIEWS TO IMPROVE PRODUCT DESIGN AND FEATURE PRIORITIZATION

## 1. Background and Problem Statement:

A consumer electronics company released a new wireless earbud model on Amazon and received over 5,000 customer reviews within six months. While the overall rating was above average, the product team lacked clarity on specific design flaws and feature requests. The brand required a **text analytics and sentiment analysis approach** to uncover detailed customer insights, prioritize improvements, and guide future product iterations.

## 2. Objectives:

- To analyze the sentiment of Amazon product reviews using Natural Language Processing (NLP)
- To extract and categorize customer feedback into positive, negative, and neutral sentiments
- To identify common themes and feature-specific complaints or praise
- To generate design-level insights and rank product improvement priorities based on frequency and intensity of sentiment

## 3. Methodology:

**Data Source:**

- 5,218 reviews for the product collected via Amazon's product page scraper
- Data fields included: review title, full text, star rating, date, and helpfulness votes

**Tools Used:**

- **Python**: NLTK, spaCy, TextBlob for sentiment scoring; pandas for data wrangling
- **WordCloud** for keyword visualization
- **matplotlib/seaborn** for charting
- **Topic Modeling**: Latent Dirichlet Allocation (LDA) for extracting key themes

**Steps Followed:**

1. **Preprocessing:**
    - Tokenization, stop-word removal, lemmatization
    - Filtering reviews under 10 words to remove noise
2. **Sentiment Scoring:**
    - Assigned polarity scores using TextBlob:
        - Polarity > 0.1 → Positive
        - Polarity < -0.1 → Negative
        - Otherwise → Neutral
3. **Keyword Extraction and Frequency Analysis:**
    - Top phrases associated with negative reviews identified
4. **Topic Modeling:**
    - LDA applied on negative reviews to identify top complaint clusters

# 4. Results:

**Sentiment Distribution:**

- Positive: 62%
- Neutral: 23%
- Negative: 15%

**Top Positive Keywords:**

- "battery life," "sound quality," "comfort," "easy pairing"

**Top Negative Themes (From Topic Modeling):**

- **Charging Issues** (24% of negative reviews) → "won't charge," "battery not holding," "charging case faulty"
- **Bluetooth Connectivity** (18%) → "keeps disconnecting," "not connecting to iPhone"
- **Fit/Comfort Problems** (15%) → "hurts after long use," "falls off during workout"
- **Microphone Quality** (11%) → "voice not clear," "others can't hear me"

**Star Ratings vs Sentiment:**

- 1-star reviews were almost entirely negative (94%)

- Interestingly, 3-star reviews showed 40% negative sentiment, revealing hidden dissatisfaction even with average ratings

## 5. Interpretation and Insights:

- Despite an overall favorable star rating, **charging reliability** is a critical product weakness
- **Microphone issues** were underrepresented in star ratings but surfaced in textual complaints, signaling a blind spot in basic rating systems
- Positive feedback on “sound quality” and “battery life” can be leveraged in marketing
- Word frequency and polarity trends confirm that **post-sale experience** (charging, connectivity) shapes overall perception more than initial product unboxing

## 6. Recommendations:

- Redesign or quality check the charging case across units
- Update firmware to fix Bluetooth connectivity issues with newer smartphone models
- Add alternative ear tips to improve long-duration comfort
- Use microphone feedback to revise product positioning for calls or meetings
- Highlight positive themes like “battery life” and “easy pairing” in product listings and ads

## 7. Future Work:

- Perform monthly sentiment tracking post product redesign
- Extend analysis to competitor products for comparative positioning
- Use review trends to create automated product improvement feedback loops

## 8. Stakeholder Relevance:

**Academic:**

- Demonstrates applied use of NLP and topic modeling in consumer research
- Suitable for coursework in e-commerce analytics, business intelligence, and data science

**Corporate:**

- Practical framework for product teams to convert unstructured review data into concrete design actions
- Helps marketing and R\&D align efforts based on real customer voice