

Sentiment Analysis of YouTube Comments on Samsung S24 Ultra

NAME: ARYAN KUMAR SINGH

ROLL: M22BBAU0399

SECTION: B

1. Introduction

- **Overview of the Dataset and Business Problem:**
 - The dataset comprises YouTube comments on videos related to the Samsung S24 Ultra. This data reflects public opinion and immediate reactions to the product.
 - The business problem is to understand customer sentiment towards the S24 Ultra, identifying key areas of satisfaction and dissatisfaction. This analysis can inform product improvement, marketing strategies, and customer service.
 - **Predictive Task:** To predict the sentiment (positive, negative, neutral) of YouTube comments regarding the Samsung S24 Ultra. This allows for automated monitoring of public opinion and a proactive response to emerging trends.

2. Methodology

- **Data Collection:**
 - YouTube comments were scraped from relevant videos.
- **Data Preprocessing:**
 - **Removing Duplicates:** Ensured each comment is unique, preventing bias.
 - **Handling Missing Values:** Rows with missing data were removed to ensure reliability.
 - **Tokenization:** Comments were split into individual words (tokens) for analysis.
 - **Lemmatization/Stemming:** Words were reduced to their base form (e.g., "running" to "run") for standardization.
 - **Removing Stop Words and Special Characters:** Common words (e.g., "the", "and") and special characters were removed to focus on meaningful terms.
- **Sentiment Analysis:**
 - **NLTK and TextBlob:** Employed for text preprocessing and sentiment scoring (polarity and subjectivity).
 - **Sentiment Categorization:**
 - Positive: Polarity score > 0
 - Negative: Polarity score < 0

- Neutral: Polarity score = 0

- **Model Building:**

- **Logistic Regression:** A statistical model used to predict the probability of a comment belonging to a particular sentiment category.
- **Decision Tree:** A tree-like model that splits data based on features to classify sentiment.

- **Model Validation:**

- **Accuracy, Precision, Recall, F1-score:** Standard metrics were used to evaluate the performance of both models.

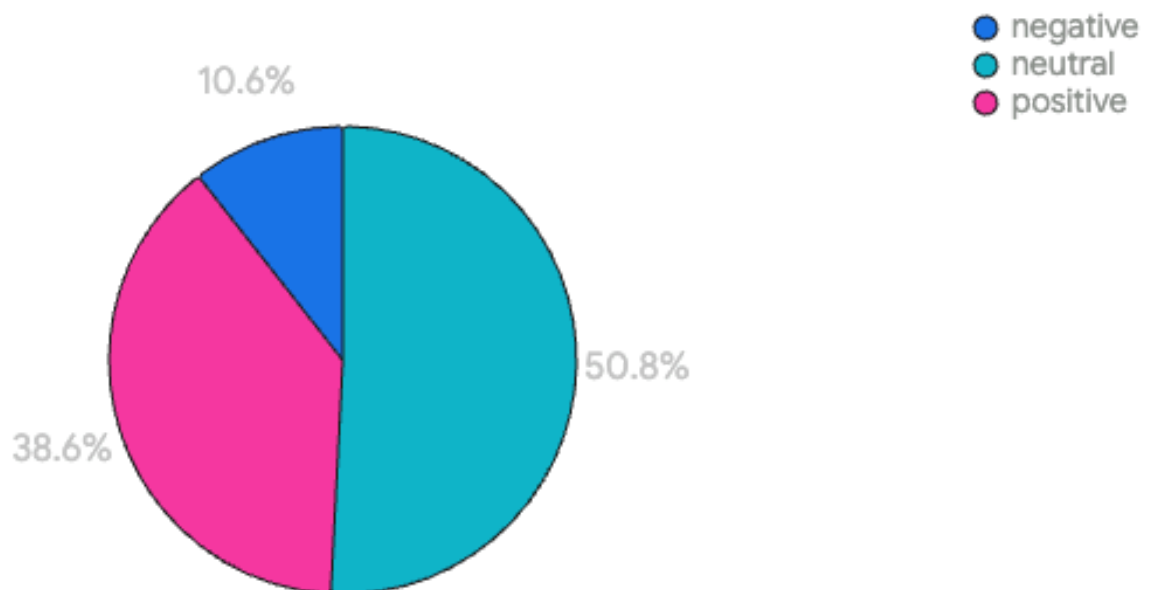
3. Findings

- **Model Performance:**

- **Logistic Regression:** Achieved an accuracy of 87.73%. Showed strong performance across all sentiment categories, especially for neutral and positive sentiments.
- **Decision Tree:** Achieved an accuracy of 84.34%. Performed well but slightly lower than logistic regression, particularly in identifying negative sentiment.

- **Sentiment Distribution:**

Sentiment Distribution of YouTube Comments



- Sentiment Distribution: Analysis revealed that 50.8% of the comments were neutral, 38.6% were positive, and 10.6% were negative.

- **Key Insights:**

- **Positive Comments:**

- Users frequently praised the camera quality, highlighting its improvements over previous models and its ability to capture stunning photos and videos.

- The new AI features, particularly the live translate function, were well-received, with many comments expressing excitement about their potential applications.
- The overall design and performance of the phone were also commended.
- Example: "This camera is a game-changer! The low-light photos are incredible."
- **Negative Comments:**
 - Concerns were raised about the battery life, with some users reporting that it drains quickly with heavy usage.
 - The price point was a barrier for some, with comments comparing it unfavorably to competitors like Apple.
 - A few comments mentioned software glitches or issues with the user interface.
 - Example: "Battery life is a real letdown. I have to charge it twice a day."
- **Neutral Comments:**
 - Many neutral comments were questions about specific features or comparisons with other phone models.
 - Some neutral comments expressed a wait-and-see approach, indicating that users were interested in the phone but wanted to see more reviews before making a decision.
 - Example: "How does the S24 Ultra's camera compare to the iPhone 15 Pro?"

4. Recommendations

- **Product Development:**
 - Address the concerns about battery life by investing in research and development to enhance battery performance.
 - Continue to innovate and improve the camera, building on the positive feedback received.
 - Ensure software updates are timely and reliable to address any glitches or user interface issues.
- **Marketing and Communication:**
 - Emphasize the strengths of the S24 Ultra, such as the camera and AI features, in marketing campaigns.
 - Acknowledge the concerns about battery life and price and proactively address them through clear communication and customer support.
- **Customer Service:**
 - Use sentiment analysis to identify and prioritize customer service requests.
 - Develop targeted responses to address common issues and concerns.
 - Invest in improving the responsiveness and effectiveness of customer service teams.
- **Continuous Monitoring:**
 - Implement a system for ongoing sentiment analysis of YouTube comments and other social media platforms to track changes in public opinion and adapt strategies accordingly.