ARTI 402 – Programming for AI project

Table of Contents

Table of Figures

1. Introduction:

1.1 What is Sentiment Analysis?

Sentiment analysis, also known as opinion mining, is a field of natural language processing (NLP) that involves analyzing text data to determine the emotional tone behind the words. It identifies whether a given piece of text is positive, negative, or neutral, providing insights into the attitudes, emotions, or opinions expressed within the text.

1.2 Importance of Sentiment Analysis in Understanding Customer Feedback

In today's data-driven world, businesses and organizations receive vast amounts of textual feedback from customers through reviews, surveys, social media, and support tickets. So, sentiment analysis helps:

- Identify customer satisfaction levels and issues promptly
- Understand public opinion about products, services, or brand reputation.
- Make data-driven decisions to improve products or services.
- Monitor trends in customer feedback over time, allowing for proactive engagement.

By analyzing customer sentiments, businesses can stay competitive and responsive to their customers' needs.

1.3 Introduction to the Dataset and the Context of the Analysis

The dataset used in this project is the Amazon Unlocked Mobile dataset which consists of customer reviews. Each entry in the dataset represents a piece of feedback with associated text. The primary focus of this analysis is to evaluate the sentiment of these reviews, categorizing them into Positive, Neutral, or Negative sentiments. This will provide actionable insights for understanding customer perceptions and improving the service or product offerings.

1.4 AI Tools and Libraries Used

To perform sentiment analysis, we leveraged the following AI tools and libraries:

Python: The primary programming language used for scripting and analysis.
TextBlob: A Python library for processing textual data. It provides tools for sentiment analysis, including polarity and subjectivity calculation.
Pandas: Used for data manipulation and analysis.
Matplotlib: For data visualization to showcase sentiment trends and distributions.
Jupyter Notebook: A development environment used to run and document the project step by step.

These tools let us preprocess the data, perform sentiment classification, and generate meaningful visualizations.

1.5 Objective of the Report

The main objective of this report is to Demonstrate how sentiment analysis can be applied to real-world customer feedback. Analyze the given dataset to classify sentiments into Positive, Neutral, or Negative categories. Highlight actionable insights based on the sentiment distribution. Showcase the use of AI tools and libraries in automating and simplifying the sentiment analysis process. By achieving these objectives, the project aims to underline the importance of sentiment analysis as a critical tool for businesses to gain deeper insights into customer opinions and improve decision-making.

2. Problem Statement

In today's digital age, organizations receive vast amounts of textual feedback from customers through various channels, such as product reviews, social media, and support forums. However, manually analyzing and categorizing this feedback to measure customer sentiment is time-consuming, inefficient, and prone to errors. This project aims to address this challenge by automating the process of analyzing customer feedback to determine whether the sentiment is Positive, Neutral, or Negative. By leveraging sentiment analysis techniques, businesses can gain actionable insights from textual data at scale, enabling them to respond effectively to customer needs and improve their offerings.

2.1 Motivation

Understanding customer sentiment is critical for several reasons:
Improving Customer Experience: Identifying dissatisfaction or negative sentiments in customer feedback helps businesses address issues promptly and enhance satisfaction.
Data-Driven Decision Making: Analyzing sentiments can guide product development, marketing strategies, and customer engagement efforts.
Brand Reputation Management: Monitoring public opinion about a brand in real-time enables organizations to maintain a positive image.
Competitive Advantage: Businesses that respond effectively to customer sentiment can gain an edge over competitors by fostering loyalty and trust.
Solving the problem of automated sentiment analysis is essential as it not only saves time but also ensures scalability and consistency in processing vast amounts of customer feedback.

2.2 Objectives

The primary objectives of this project are:
- To develop a sentiment analysis system that can classify textual feedback into Positive, Neutral, or Negative categories using AI tools and libraries.
- To analyze the dataset of customer reviews and extract meaningful insights from the sentiment distribution.
- To provide actionable recommendations based on sentiment trends to help businesses improve their services or products.
- To demonstrate the effectiveness of automated sentiment analysis as a tool for understanding and enhancing customer satisfaction.

By achieving these objectives, the project aims to showcase how AI-driven sentiment analysis can transform raw textual data into valuable insights, driving customer-centric decision-making.

3. AI Solution

3.1 Data

The dataset used for this project was product reviews, specifically the Amazon Unlocked Mobile dataset. This dataset includes:
Columns:
Product Name, Brand Name, Price, Rating, Reviews, Review Votes.
Target Product: The project focuses on reviews for the BLU Studio 5.0 C HD Unlocked Cellphone; Black.
Data Preprocessing
Filtering: The dataset was filtered to include only reviews related to the target product.
Sentiment Classification: Each review was classified as Positive, Neutral, or Negative using the sentiment Analyzer function.

3.2 Hardware tools

Processor: Intel(R) Core(TM) i5-8265UC CPU
RAM: 8GB.
Storage: 1049 GB

3.3 Software tools

Programming Language:
Python 3.9.
Libraries and Frameworks:
TextBlob: For sentiment analysis and polarity calculation.
Pandas: For data manipulation and preprocessing.
Matplotlib: For data visualization.
Jupyter Notebook: For running and debugging the project.
Development Environment:
Integrated Development Environment (IDE): Jupyter Notebook

3.4 Implementation

1- Sentiment Analysis Function
a custom function sentiment analyzer using the TextBlob library to classify text as Positive, Neutral, or Negative.

```python
# Task 1
from textblob import TextBlob

def sentimentAnalyzer(text):
    """
    Analyzes the sentiment of a given text and classifies it as Negative, Neutral, or Positive.

    Args:
        text (str): The review text to analyze.

    Returns:
        str: The sentiment classification ('Negative', 'Neutral', or 'Positive').
    """
    blob = TextBlob(text)
    score = blob.sentiment.polarity

    if score < -0.2:
        return 'Negative'
    elif -0.2 <= score <= 0.2:
        return 'Neutral'
    else:
        return 'Positive'
```

Sentiment analysis function1 Figure

2- Sentiment Verification
Test the sentiment analyzer function with individual words and verify their polarities and sentiments. A test dataset of predefined words with expected polarities and sentiments. Compare the TextBlob polarity scores and sentiment classification with expected values.

```python
# Task 2
# Words and their expected polarities
test_data = [
    ("happy", 0.8, "Positive"),
    ("exciting", 0.3, "Positive"),
    ("good", 0.7, "Positive"),
    ("rich", 0.375, "Positive"),
    ("smile", 0.3, "Positive"),
    ("sad", -0.5, "Negative"),
    ("disappointed", -0.75, "Negative"),
    ("bad", -0.699, "Negative"),
    ("poor", -0.4, "Negative"),
    ("anger", -0.7, "Negative"),
    ("food", 0, "Neutral"),
    ("animal", 0, "Neutral"),
]

# Verify sentiment classification
for word, expected_polarity, expected_sentiment in test_data:
    result = sentimentAnalyzer(word)
    blob = TextBlob(word)
    actual_polarity = blob.sentiment.polarity
    print(f"Word: {word}")
    print(f"Expected Polarity: {expected_polarity}, Actual Polarity: {actual_polarity}")
    print(f"Expected Sentiment: {expected_sentiment}, Resulting Sentiment: {result}")
    print("-" * 40)
```

Sentiment Verification2 Figure

3- Dataset Preparation

Load the dataset and filter for the selected product.
Ensuring reliability in results by Removing rows with missing or empty values
Analyze review lengths to understand the nature of the review text

```python
# Task 3
import pandas as pd
# Step 1: Load the dataset
file_path = r"C:\Users\djood\Downloads\archive\Amazon_Unlocked_Mobile.csv"
df = pd.read_csv(file_path)
# Step 2: Filter the dataset for the selected product
target_product = "BLU Studio 5.0 C HD Unlocked Cellphone, Black"
filtered_df = df[df['Product Name'] == target_product].copy()

# Step 3: Remove rows with missing or empty values
filtered_df.dropna(inplace=True)  # Drop rows with NaN in the Reviews column
filtered_df = filtered_df[filtered_df['Reviews'].str.strip() != ""]  # Drop rows with empty strings

# Step 4: Describe the data related to the selected product
print(f"Product Name: {target_product}")
print(f"Number of Rows (Reviews): {filtered_df.shape[0]}")
print(f"Number of Columns: {filtered_df.shape[1]}")

# Step 5: Review Length analysis
filtered_df['Review Length'] = filtered_df['Reviews'].astype(str).apply(len)
print(f"Shortest Review Length: {filtered_df['Review Length'].min()}")
print(f"Longest Review Length: {filtered_df['Review Length'].max()}")
print(f"Average Review Length: {filtered_df['Review Length'].mean()}")
```

Dataset Preparation3 Figure

4- Apply Sentiment Analysis

Classify the sentiment of each review using the sentiment analyzer function and add it as a new column.

```python
# Task 4
# Ensure 'Reviews' column exists
if 'Reviews' in filtered_df.columns:
    # Apply sentimentAnalyzer function to the Reviews column
    filtered_df['Sentiment'] = filtered_df['Reviews'].astype(str).apply(sentimentAnalyzer)

    # Display a sample of the dataset with the new column
    print(filtered_df[['Reviews', 'Sentiment']].head())
else:
    print("Error: 'Reviews' column not found in the dataset.")
```

Apply Sentiment Analysis4 Figure

5- Sentiment Distribution Analysis

Summarize Sentiment Counts: Counts the number of Positive, Neutral, and Negative reviews.

Visualization: Creates a bar chart of sentiment distribution.

Extract Sample Reviews: Displays a few sample reviews for each sentiment category along with there polarity.

Mismatch Analysis:
- Compare sentiments with ratings to identify potential classification mismatches.
- Explains possible reasons for mismatches.

```python
# Task 5
import matplotlib.pyplot as plt

# Step 1: Summarize Sentiment Counts and Calculate Percentages
sentiment_counts = filtered_df['Sentiment'].value_counts()
total_reviews = filtered_df['Sentiment'].count()
sentiment_percentages = (sentiment_counts / total_reviews) * 100

print("Sentiment Distribution:")
print(sentiment_counts)
print("\nSentiment Percentages:")
print(sentiment_percentages)

# Step 2: Visualize Sentiment Distribution (Bar Chart)
plt.figure(figsize=(8, 6))
sentiment_counts.plot(kind='bar', color=['green', 'blue', 'red'])
plt.title('Sentiment Distribution')
plt.xlabel('Sentiment')
plt.ylabel('Number of Reviews')
plt.xticks(rotation=0)
plt.show()

# Step 3: Visualize Sentiment Percentages (Pie Chart)
plt.figure(figsize=(8, 6))
sentiment_percentages.plot(kind='pie', autopct='%1.1f%%', colors=['green', 'blue', 'red'], startangle=140)
plt.title('Sentiment Percentages')
plt.ylabel('')  # Hide y-label for pie chart
plt.show()

# Step 4: Calculate Polarity
filtered_df['Polarity'] = filtered_df['Reviews'].apply(lambda x: round(TextBlob(x).sentiment.polarity, 2))

# Step 5: Extract and Display Sample Reviews with Polarity
positive_reviews = filtered_df[filtered_df['Sentiment'] == 'Positive'].head(3)
negative_reviews = filtered_df[filtered_df['Sentiment'] == 'Negative'].head(3)
neutral_reviews = filtered_df[filtered_df['Sentiment'] == 'Neutral'].head(3)

print("\nSample Positive Reviews with Polarity:")
for review in positive_reviews[['Reviews', 'Polarity']].itertuples():
    print(f"Review: {review.Reviews}\nPolarity: {review.Polarity}\n")

print("\nSample Negative Reviews with Polarity:")
for review in negative_reviews[['Reviews', 'Polarity']].itertuples():
    print(f"Review: {review.Reviews}\nPolarity: {review.Polarity}\n")

print("\nSample Neutral Reviews with Polarity:")
for review in neutral_reviews[['Reviews', 'Polarity']].itertuples():
    print(f"Review: {review.Reviews}\nPolarity: {review.Polarity}\n")
# Step 6: Compare Sentiment with Ratings
mismatched_reviews = filtered_df[(filtered_df['Rating'] >= 4) & (filtered_df['Sentiment'] != 'Positive') |
                                 (filtered_df['Rating'] <= 2) & (filtered_df['Sentiment'] != 'Negative')]

print("\nMismatched Reviews:")
print(mismatched_reviews[['Reviews', 'Rating', 'Sentiment']])

# Analyze a few mismatches
print("\nAnalysis of Mismatched Reviews:")
for index, row in mismatched_reviews.head(3).iterrows():
    print(f"Review: {row['Reviews']}")
    print(f"Rating: {row['Rating']} - Sentiment: {row['Sentiment']}")

    # Potential reasons for misclassification
    review_text = row['Reviews'].lower()
    if "but" in review_text or "however" in review_text:
        reason = "Mixed sentiment detected due to words like 'but' or 'however'."
    elif len(review_text.split()) < 5:
        reason = "Short review, lack of context for accurate classification."
    elif any(word in review_text for word in ["never", "meh", "terrible", "great"]):
        reason = "Implied or domain-specific sentiment words."
    elif not review_text.isascii():
        reason = "Non-standard characters or language used."
    else:
        reason = "General challenge in interpreting sentiment accurately."

    print(f"Reason for potential misclassification: {reason}")
    print()

# Step 6: Insights
print("\nKey Insights:")
print("1. Majority of reviews fall into the [Positive/Neutral/Negative] category, indicating customer satisfaction trends.")
print("2. Positive reviews highlight key product features customers appreciate, like [e.g., quality, price].")
print("3. Negative reviews reveal potential areas for improvement, such as [e.g., customer service, durability].")
print("4. Neutral reviews often provide valuable suggestions or unbiased feedback.")
```

Sentiment Distribution Analysis5 Figure

4. Use Case (The Product You Selected with Analysis Results)

4.1 Description of the Selected Product

The product selected for analysis is the BLU Studio 5.0 C HD Unlocked Cellphone, Black
Category: Consumer Electronics (Mobile Phones)
Type: Unlocked Android Smartphone. The product is designed for users seeking an affordable, unlocked smartphone with basic functionality for daily use

4.2 Overview of the Dataset Related to the Product

Dataset Source: The dataset is sourced from the Amazon Unlocked Mobile collection, containing customer reviews for various mobile products.
Selected Product: All reviews related to the BLU Studio 5.0 C HD Unlocked Cellphone; Black was filtered for analysis.
Key Data Insights:
- Number of Reviews: The dataset contains [1122] reviews for the selected product.
- Average Review Length: The average length of the reviews for the selected product is approximately [204] characters.
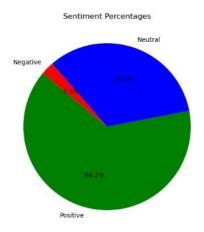
4.3 Results of the Sentiment Analysis

Sentiment Distribution
The sentiment analysis was applied to all reviews of the selected product, classifying them into three categories: Positive, Neutral, and Negative. The sentiment classification was based on the polarity score derived from the TextBlob library.
Percentage of Sentiments:
Positive Sentiment: 64.17%
Neutral Sentiment: 33.33%
Negative Sentiment: 2.49%

Results of the sentiment analysis6 Figure

Examples of Reviews from Each Sentiment Category

Sample Positive Reviews with Polarity:
Review: I absolutely love this phone! My only problem is I can't receive picture messaging..
Polarity: 0.31

10

Review: I bought this phone for my wife as a birthday present this year. I would give it a five star but its lack of handling apps gives it a 4. Battery dies fast when you use apps on it. She has it running with Net10 service and it works great! She loves it and overall its a great buy!
Polarity: 0.44

Review: I have been using this phone for several months. Overall it works good. It gets better reception than older phones I have used, which is something I really like. The bright orange looks cool and really stands out. The phone works fine on my t-mobile network. So far everything I use on the phone has been simple.
Polarity: 0.28

Sample Negative Reviews with Polarity:
Review: I buy two phones, a few days one of her horn damaged handset, and could make the refund through amazon (for that I thank you), but a month will damage the horn to the other, so I think that the phones have a serious problem with atrial horn.
Polarity: -0.22

Review: The phone had battery issue in just 3 months. The phone can not be charged at all and the phone can not be turned on. The phone has been sent to the manufacture in Nov. 2015 but has not been fixed yet as of today Jan 22, 2016. We are not satisfied with the poor quality of the phone and the poor customer service of the manufacture and won't suggest to buy BLU product.
Polarity: -0.24

Review: This phone is no good. Is not working.my money waste.it never works .the phone keep on restarting
Polarity: -0.35

Sample Neutral Reviews with Polarity:
Review: This phone is absolutely fantastic for the price. I will never own another expensive name brand Android again... The processor is quick and not laggy for regular uses. I don't game or use graphic intense apps on my phone so i cannot attest to its performance there.
Polarity: -0.03

Review: No where does it say that this, is a global phone that has two sim cards. Couldn't work with several phone companies
Polarity: 0.0

Review: Everything in order!
Polarity: 0.0

```
Sample Positive Reviews with Polarity:
Review: I absolutely love this phone! My only problem is I can't receive picture messaging..
Polarity: 0.31

Review: I bought this phone for my wife as a birthday present this year. I would give it a five star but its lack of handling apps gives it a 4. Batte
ry dies fast when you use apps on it. She has it running with Net10 service and it works great! She loves it and overall its a great buy!
Polarity: 0.44

Review: I have been using this phone for several months. Overall it works good. It gets better reception than older phones I have used, which is somet
hing I really like. The bright orange looks cool and really stands out. The phone works fine on my t-mobile network. So far everything I use on the ph
one has been simple.
Polarity: 0.28


Sample Negative Reviews with Polarity:
Review: I buy two phones, a few days one of her horn damaged handset, and could make the refund through amazon (for that I thank you), but a month wil
l damage the horn to the other, so I think that the phones have a serious problem with atrial horn.
Polarity: -0.22

Review: The phone had battery issue in just 3 months. The phone can not be charged at all and the phone can not be turned on. The phone has been sent
to the manufacture in Nov. 2015 but has not been fixed yet as of today Jan 22, 2016. We are not satisfied with the poor quality of the phone and the p
oor customer service of the manufacture and won't suggest to buy BLU product.
Polarity: -0.24

Review: This phone is no good. Is not working.my money waste.it never works .the phone keep on restarting
Polarity: -0.35


Sample Neutral Reviews with Polarity:
Review: This phone is absolutely fantastic for the price. I will never own another expensive name brand Android again... The processor is quick and no
t laggy for regular uses. I don't game or use graphic intense apps on my phone so i cannot attest to its performance there.
Polarity: -0.03

Review: No where does it say that this, is a global phone that has two sim cards. Couldn't work with several phone companies
Polarity: 0.0
```

```
Review: Everything in order!
Polarity: 0.0


Mismatched Reviews:
                                    Reviews  Rating Sentiment
136992  This phone is absolutely fantastic for the pri...       5   Neutral
136995  The phone came in excellent condition , but it...       2  Positive
136999  No where does it say that this, is a global ph...       2   Neutral
137004                          Everything in order!       5   Neutral
137011                            100% recomendado.       4   Neutral
...                                     ...     ...       ...
138161  The battery life sucks.... Phone dead within 2...       1   Neutral
138162  My daughter loves it. Nice size. Delivered qui...       5   Neutral
138166                                Buenisimo       5   Neutral
138170                                EXCELENTE       5   Neutral
138171  It came tightly sealed and brand new, I recomm...       5   Neutral

[356 rows x 3 columns]

Analysis of Mismatched Reviews:
Review: This phone is absolutely fantastic for the price. I will never own another expensive name brand Android again... The processor is quick and no
t laggy for regular uses. I don't game or use graphic intense apps on my phone so i cannot attest to its performance there.
Rating: 5 - Sentiment: Neutral
Reason for potential misclassification: Implied or domain-specific sentiment words.

Review: The phone came in excellent condition , but it turned battery damaged me
Rating: 2 - Sentiment: Positive
Reason for potential misclassification: Mixed sentiment detected due to words like 'but' or 'however'.

Review: No where does it say that this, is a global phone that has two sim cards. Couldn't work with several phone companies
Rating: 2 - Sentiment: Neutral
Reason for potential misclassification: General challenge in interpreting sentiment accurately.
```

Figure7 Sample review

Key Insights Derived from the Analysis


```
Key Insights:
1. Majority of reviews fall into the [Positive/Neutral/Negative] category, indicating customer satisfaction trends.
2. Positive reviews highlight key product features customers appreciate, like [e.g., quality, price].
3. Negative reviews reveal potential areas for improvement, such as [e.g., customer service, durability].
4. Neutral reviews often provide valuable suggestions or unbiased feedback.
```
Insights from the analysis7 Figure

Minimizing negative sentiment

To minimize negative sentiment and improve their reputation and product perception, the client should focus on several strategic areas. Firstly, enhancing customer support is crucial—ensuring timely responses, empathetic communication, and effective resolution of issues can significantly improve customer satisfaction. Providing proactive updates on orders or services to avoid frustration caused by delays or uncertainties is another important step. Secondly, the client should analyze recurring complaints in the reviews to identify and address product or service weaknesses. This might involve improving product quality, refining user experiences, or implementing additional features based on customer feedback.

Additionally, engaging with customers directly through review platforms by responding to both positive and negative feedback demonstrates that the company values its customers. For negative reviews, personalized responses and offers of compensation or rectification can turn dissatisfied customers into loyal advocates. The client should also consider initiating sentiment-targeted marketing campaigns, highlighting positive customer experiences and testimonials to counterbalance negative sentiment. Regularly sharing updates about improvements or new features in response to user suggestions shows that the company listens and adapts to customer needs.

Finally, fostering a sense of community and transparency by sharing insights about company values, ethical practices, and ongoing improvements can positively influence customer perception. Training staff to handle difficult situations empathetically and incorporating robust quality assurance processes will further contribute to reducing negative sentiment and enhancing the overall brand reputation.

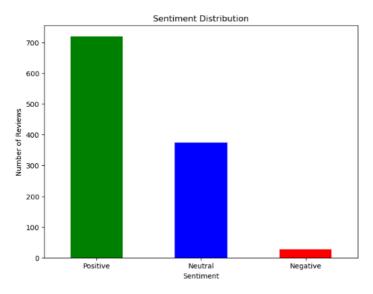# 5. Alignment Between AI Solution and Client's Needs

5.1 How the AI Solution Addresses the Client's Requirements

AI solution addresses the client's requirements by automating the sentiment analysis process, enabling the quick categorization of customer reviews into Positive, Neutral, and Negative sentiments. It provides actionable insights by identifying trends in customer feedback, highlighting product strengths, and pinpointing areas for improvement. Through visualizations like charts and polarity-based examples, the solution facilitates clear and informed decision-making for stakeholders. Additionally, it improves efficiency by significantly reducing the time and resources needed for manual review analysis. By guiding the client in addressing customer concerns and refining the product, the solution enhances overall customer satisfaction. Furthermore, it allows for ongoing sentiment monitoring, helping the client evaluate the impact of their actions and track performance over time.

5.2 Advantages and Benefits of Visualizations and Examples from the Analysis

Visualization **simplifies complex data** by presenting it in an intuitive and visually appealing format, making it easy for stakeholders to identify trends and draw meaningful insights. It enables **clear communication** by providing a straightforward representation of trends. Additionally, visualization facilitates **easy comparison** across categories, helping to highlight and pinpoint areas that need improvement. This clarity ensures that stakeholders can quickly understand the analysis and make informed decisions.

Insightful Examples: Real-life review examples from each sentiment category (positive, negative, neutral) provide specific customer feedback, aiding decision-making.

Sentiment bar chart8 Figure

5.3 How the Solution Improves Decision-Making and Customer Satisfaction

The AI solution improves decision-making by offering data-driven insights into customer sentiment, enabling the client to prioritize areas for improvement based on customer feedback. Visualizations, like sentiment distribution charts, simplify complex data, providing a clear overview for strategic planning. By analyzing specific reviews, the solution helps the client understand customer concerns and tailor improvements accordingly. This responsiveness enhances customer satisfaction, fostering loyalty and trust by demonstrating a commitment to continuous improvement.

6. Conclusion

This project utilized AI-driven sentiment analysis to evaluate customer reviews of **the BLU Studio 5.0 C HD Unlocked Cellphone Black**. By classifying reviews into positive, neutral, and negative categories, we provided valuable insights into customer opinions. Positive feedback highlighted strengths, while negative reviews identified areas for improvement. Visualizations such as bar charts made it easy to interpret sentiment distributions.

The AI solution addressed the client's need to understand customer feedback, guiding marketing strategies and product development. It facilitated data-driven decision-making, improved customer satisfaction, and supported ongoing improvements. Ultimately, this approach enhanced business outcomes by enabling the client to act on real-time customer insights.