

Industrial Internship Report on " Amazon Product Reviews and Sentiment Analysis using Python"

**Prepared by
SHWETHA S D**

Executive Summary

This report provides details of the Industrial Internship provided by upskill Campus and The IoT Academy in collaboration with Industrial Partner UniConverge Technologies Pvt Ltd (UCT).

This internship was focused on a Data Science problem statement related to customer review analysis. The project was completed within a duration of six weeks.

The project titled "Sentiment Analysis of Amazon Alexa Reviews Using Python and VADER" focuses on understanding customer opinions using Natural Language Processing (NLP). The project involved data loading, visualization of ratings, and sentiment classification into Positive, Negative, and Neutral categories.

This internship provided excellent exposure to real-world data analysis, Python programming, NLP, and visualization techniques. It was an overall valuable learning experience.

TABLE OF CONTENTS

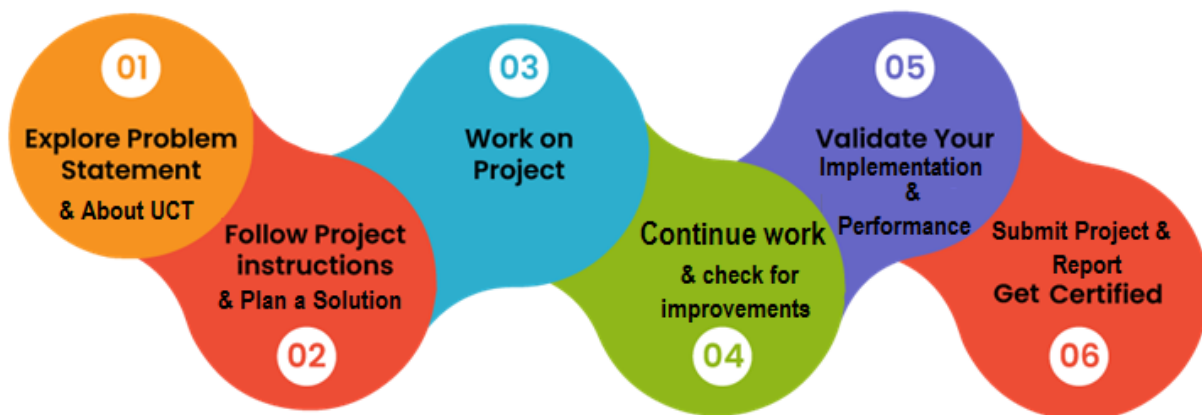
1	Preface	3
2	Introduction	4
2.1	About UniConverge Technologies Pvt Ltd	4
2.2	About upskill Campus.....	8
2.3	Objective	9
3	Problem Statement.....	10
4	Existing and Proposed solution	10
5	Proposed Design/ Model	11
5.1	High Level Diagram	11
6	Performance Test	12
6.1	Test Plan/ Test Cases	12
6.2	Test Procedure.....	12
6.3	Performance Outcome.....	12
7	My learnings.....	15
8	Future work scope	15

1 Preface

This internship program was conducted for a duration of six weeks. During this period, I worked on a Data Science-based project related to Sentiment Analysis. The internship helped me gain practical exposure to industry-level problem solving using Python and Machine Learning concepts.

The project involved working with real customer review data and applying sentiment analysis techniques to extract meaningful insights. This internship played a significant role in enhancing my technical skills, confidence, and understanding of real-world applications.

I would like to sincerely thank upskill Campus, The IoT Academy, UniConverge Technologies Pvt Ltd, my mentors, and all those who supported me throughout this internship.



2 Introduction

2.1 About UniConverge Technologies Pvt Ltd

A company established in 2013 and working in Digital Transformation domain and providing Industrial solutions with prime focus on sustainability and RoI.

For developing its products and solutions it is leveraging various **Cutting Edge Technologies** e.g. **Internet of Things (IoT), Cyber Security, Cloud computing (AWS, Azure), Machine Learning, Communication Technologies (4G/5G/LoRaWAN), Java Full Stack, Python, Front end** etc.



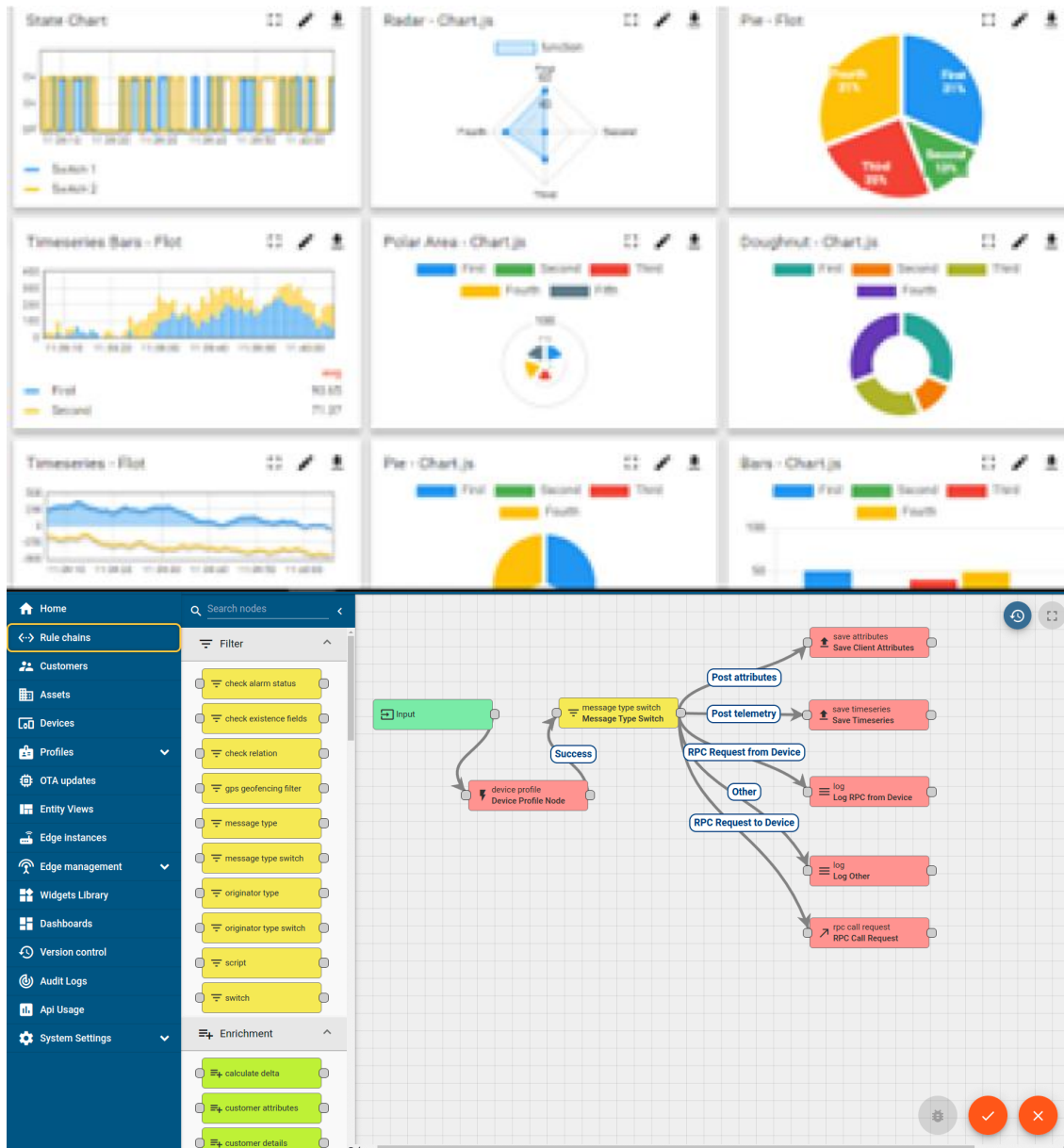
i. UCT IoT Platform()

UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable “insight” for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- It enables device connectivity via industry standard IoT protocols - MQTT, CoAP, HTTP, Modbus TCP, OPC UA
- It supports both cloud and on-premises deployments.

It has features to

- Build Your own dashboard
- Analytics and Reporting
- Alert and Notification
- Integration with third party application(Power BI, SAP, ERP)
- Rule Engine



FACTORY WATCH

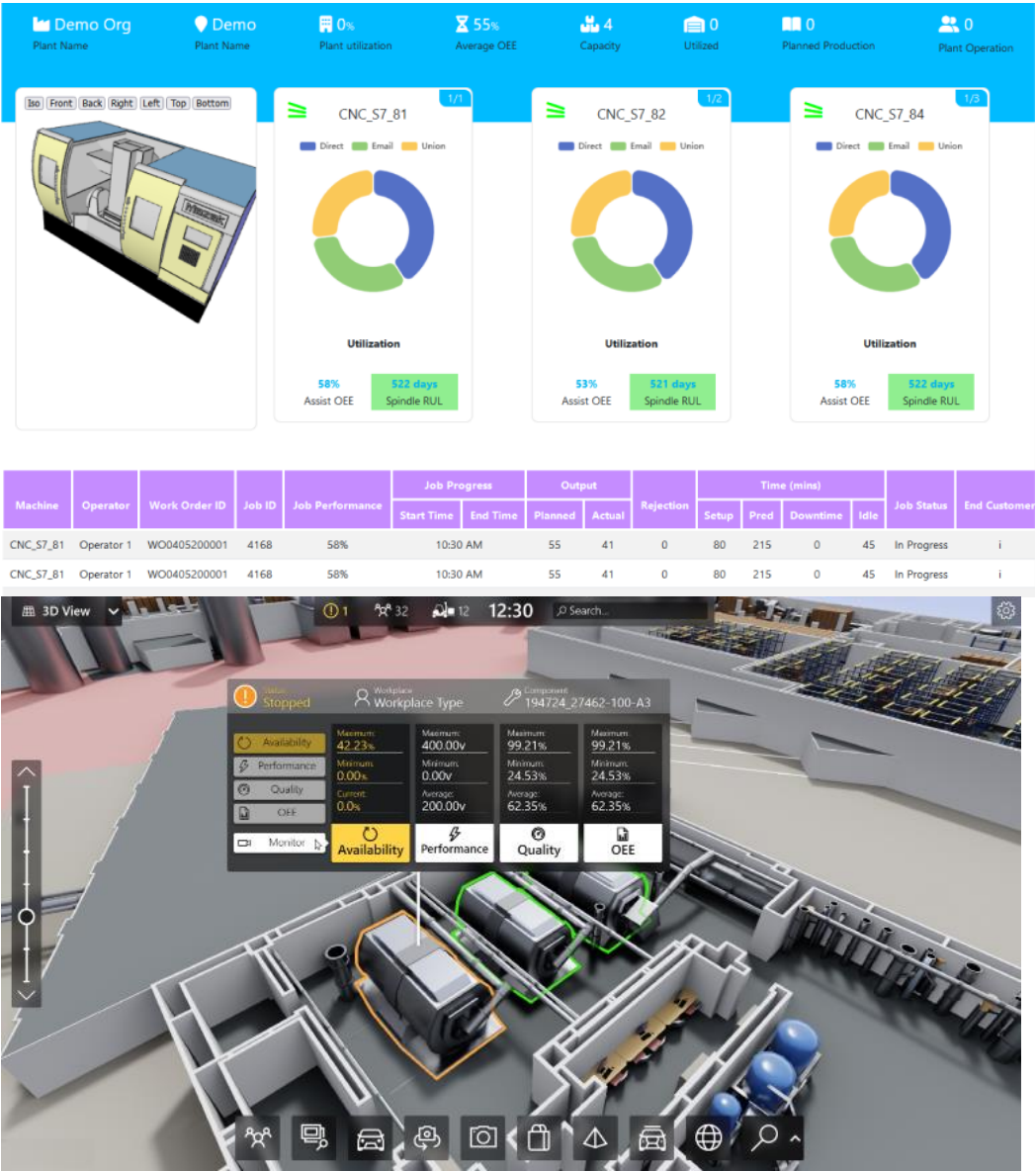
ii. Smart Factory Platform ()

Factory watch is a platform for smart factory needs.

It provides Users/ Factory

- with a scalable solution for their Production and asset monitoring
- OEE and predictive maintenance solution scaling up to digital twin for your assets.
- to unleash the true potential of the data that their machines are generating and helps to identify the KPIs and also improve them.
- A modular architecture that allows users to choose the service that they want to start and then can scale to more complex solutions as per their demands.

Its unique SaaS model helps users to save time, cost and money.



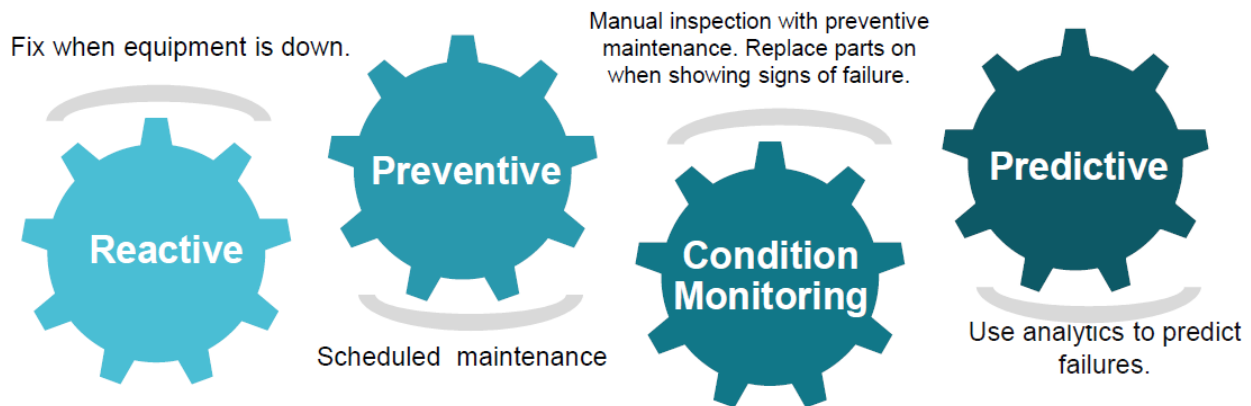


iii. LoRaWAN based Solution

UCT is one of the early adopters of LoRAWAN teschnology and providing solution in Agritech, Smart cities, Industrial Monitoring, Smart Street Light, Smart Water/ Gas/ Electricity metering solutions etc.

iv. Predictive Maintenance

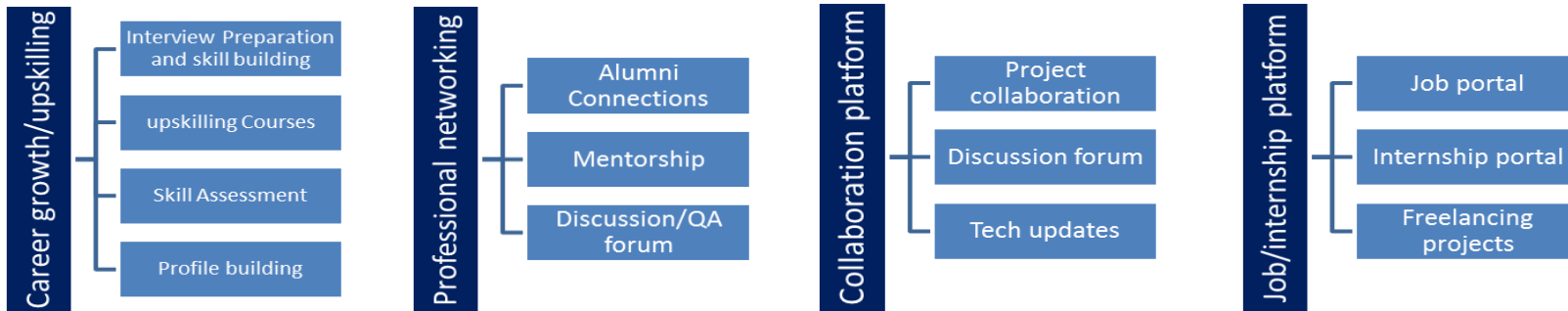
UCT isproviding Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



2.2 About upskill Campus (USC)

upskill Campus along with The IoT Academy and in association with Uniconverge technologies has facilitated the smooth execution of the complete internship process.

USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



2.3 The IoT Academy

The IoT academy is EdTech Division of UCT that is running long executive certification programs in collaboration with EICT Academy, IITK, IITR and IITG in multiple domains.

2.4 Objectives of this Internship program

The objective for this internship program was to

- ▣ get practical experience of working in the industry.
- ▣ to solve real world problems.
- ▣ to have improved job prospects.
- ▣ to have Improved understanding of our field and its applications.
- ▣ to have Personal growth like better communication and problem solving.

3 Problem Statement

In the assigned problem statement

The objective of this project is to analyze customer reviews of Amazon Alexa products using Python and Natural Language Processing techniques. The project focuses on extracting meaningful insights from real-world customer feedback by applying Sentiment Analysis to classify each review into Positive, Negative, or Neutral categories using the VADER sentiment analyzer. In addition to sentiment classification, the project also aims to visualize the distribution of customer ratings using graphical representations such as a pie (donut) chart. This helps in understanding overall customer satisfaction, identifying trends in user opinions, and supporting data-driven decision-making for product improvement and market analysis.

4 Existing and Proposed solution

Existing Solution:

Traditionally, customer feedback is analyzed manually, which is time-consuming and inefficient for large datasets. It also lacks real-time insights.

Proposed Solution:

In this project, Python and Natural Language Processing (NLP) techniques are used to automate the sentiment analysis process. The VADER sentiment analyzer is used to classify customer reviews efficiently.

4.1 Code submission (Github link):

https://github.com/sdshwetha15-sketch/test_code.git

4.2 Report submission (Github link) :

[https://github.com/sdshwetha15-sketch /amazon-alexa-sentiment-analysis](https://github.com/sdshwetha15-sketch/amazon-alexa-sentiment-analysis)

5 Proposed Design/ Model

This project follows a structured workflow:

1. Data Loading from Amazon.txt file
2. Data Exploration using pandas
3. Rating Visualization using Matplotlib
4. Sentiment Analysis using VADER
5. Aggregation of sentiment results
6. Final sentiment classification

5.1 High Level Diagram

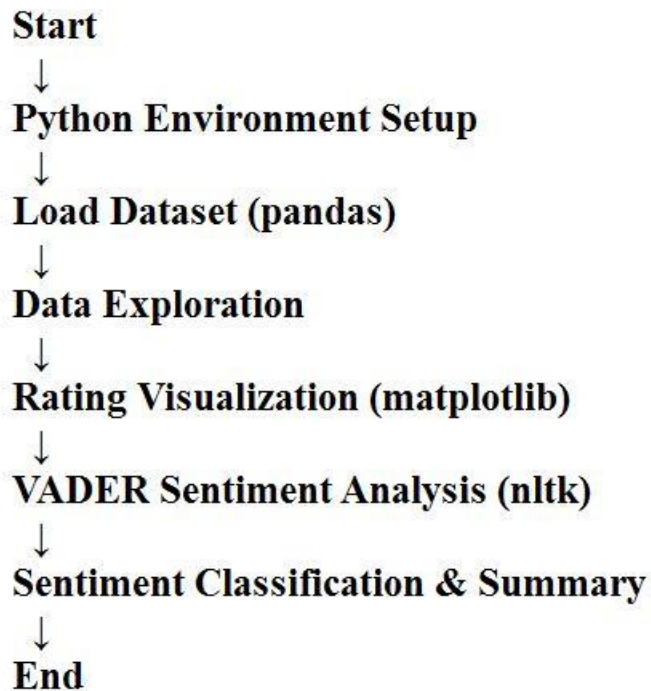


Figure 1: HIGH LEVEL DIAGRAM OF THE SYSTEM

6 Performance Test

Performance testing was conducted to evaluate the effectiveness of the sentiment analysis model.

Constraints Considered: • Text processing speed • Accuracy of sentiment scores • Dataset size handling

6.1 Test Plan/ Test Cases

Test Case 1: Verify correct file loading Test Case 2: Verify correct rating visualization Test Case 3: Verify correct sentiment classification

6.2 Test Procedure

The dataset was loaded using pandas, sentiment scores were calculated using VADER, and outputs were verified using visualization and printed results.

6.3 Performance Outcome

The system successfully classified sentiments and displayed the overall sentiment as Neutral based on total sentiment scores.

SCREENSHOTS

```
In [10]: 1 import pandas as pd
          2 import matplotlib.pyplot as plt
          3
          4 import nltk
          5 from nltk.sentiment.vader import SentimentIntensityAnalyzer
          6
          7
          8 sentiments = SentimentIntensityAnalyzer()
          9
```

```
In [9]: 1 data = pd.read_csv("Amazon.txt", delimiter="\t")
          2 print(data.head())
          3
```

	rating	date	variation \
0	5	31-Jul-25	Charcoal Fabric
1	5	31-Jul-25	Charcoal Fabric
2	4	31-Jul-25	Walnut Finish
3	5	31-Jul-25	Charcoal Fabric
4	5	31-Jul-25	Charcoal Fabric

	verified_reviews	feedback
0	Love my Echo!	1
1	Loved it!	1
2	Sometimes while playing a game, you can answer...	1
3	I have had a lot of fun with this thing. My 4 ...	1
4	Music	1

```
In [4]: 1 ratings = data["rating"].value_counts()
        2 number = ratings.index
        3 quantity = ratings.values
        4
```

```
In [5]: 1 custom_colors = ["skyblue", "yellowgreen", "tomato", "blue", "red"]
        2
        3 plt.figure(figsize=(5, 5))
        4 plt.pie(quantity, labels=number, colors=custom_colors)
        5
        6 central_circle = plt.Circle((0, 0), 0.5, color='white')
        7 fig = plt.gcf()
        8 fig.gca().add_artist(central_circle)
        9
        10 plt.rc('font', size=12)
        11 plt.title("Amazon Alexa Reviews", fontsize=20)
        12 plt.show()
        13
```



```
In [6]: 1 data["Positive"] = [sentiments.polarity_scores(str(i))["pos"] for i in data["verified_reviews"]]
2 data["Negative"] = [sentiments.polarity_scores(str(i))["neg"] for i in data["verified_reviews"]]
3 data["Neutral"] = [sentiments.polarity_scores(str(i))["neu"] for i in data["verified_reviews"]]
4

In [7]: 1 x = sum(data["Positive"])
2 y = sum(data["Negative"])
3 z = sum(data["Neutral"])
4

In [8]: 1 def sentiment_score(a, b, c):
2     if (a > b) and (a > c):
3         print("Positive")
4     elif (b > a) and (b > c):
5         print("Negative")
6     else:
7         print("Neutral")
8
9 sentiment_score(x, y, z)
10

Neutral
```

7 MY LEARNINGS

- LEARNED HOW TO USE PANDAS FOR DATA ANALYSIS
- UNDERSTOOD HOW SENTIMENT ANALYSIS WORKS USING NLP
- GAINED EXPERIENCE WITH MATPLOTLIB FOR VISUALIZATION
- LEARNED HOW TO WORK WITH REAL-WORLD DATASETS
- IMPROVED PYTHON PROGRAMMING AND DEBUGGING SKILLS

8. FUTURE SCOPES

- **APPLY MACHINE LEARNING-BASED SENTIMENT CLASSIFIERS:**

IN FUTURE, TRADITIONAL RULE-BASED SENTIMENT ANALYSIS USING VADER CAN BE REPLACED OR ENHANCED WITH MACHINE LEARNING MODELS SUCH AS NAÏVE BAYES, SUPPORT VECTOR MACHINES (SVM), OR DEEP LEARNING MODELS LIKE LSTM AND BERT. THESE MODELS CAN IMPROVE SENTIMENT PREDICTION ACCURACY BY LEARNING FROM LARGE VOLUMES OF LABELED REVIEW DATA AND UNDERSTANDING COMPLEX LANGUAGE PATTERNS, SARCASM, AND CONTEXT MORE EFFECTIVELY THAN LEXICON-BASED METHODS.

- **PERFORM SENTIMENT ANALYSIS PER RATING CATEGORY:**

THE PROJECT CAN BE EXTENDED TO ANALYZE SENTIMENTS SEPARATELY FOR EACH STAR RATING (1 TO 5). THIS WILL HELP IN UNDERSTANDING HOW CUSTOMER EMOTIONS DIFFER ACROSS VARIOUS RATING LEVELS AND ALLOW MORE DETAILED INSIGHTS INTO CUSTOMER BEHAVIOR AND SATISFACTION TRENDS FOR DIFFERENT PRODUCT CATEGORIES.

- **DEPLOY THE PROJECT AS A WEB-BASED DASHBOARD:**

THE SYSTEM CAN BE DEPLOYED AS AN INTERACTIVE WEB APPLICATION USING FLASK, DJANGO, OR STREAMLIT. THIS WOULD ALLOW USERS TO UPLOAD REVIEW DATA, VISUALIZE RESULTS IN REAL TIME, AND INTERACT WITH CHARTS AND SENTIMENT SUMMARIES THROUGH A USER-FRIENDLY INTERFACE.

- **INTEGRATE REAL-TIME REVIEW ANALYSIS:**

THE MODEL CAN BE CONNECTED TO LIVE DATA SOURCES SUCH AS E-COMMERCE PLATFORMS OR APIS TO PERFORM REAL-TIME SENTIMENT ANALYSIS ON INCOMING CUSTOMER REVIEWS, ENABLING INSTANT BUSINESS INSIGHTS.

S