

2023

SENTIMENT ANALYSIS
OF PRODUCTS REVIEW
FOR E-COMMERCE
PLATFORM

Prepared by: Sonal Bisla



SUBMITTED TO:

DR. FARID SHIRAZI

DR. CENI BABAOGLU

# TABLE OF CONTENTS

1.	Intro	Introduction					
2.	Data	Data Description					
3.	Lite	Literature Review					
3	3.1	Introduction	5				
3	3.1.1	Amazon dataset	5				
3	3.1.2	Twitter Dataset	5				
3	3.1.3	General E-commerce platforms	6				
3	3.1.4	Review helpfulness and market value	6				
3	3.1.5	Data analytics and visualization	7				
3	3.1.6	1.6 Business analytics and innovation					
3	3.2	Conclusion	8				
4.	. EXPLANATORY DATA ANALYSIS (EDA)						
4	l.1	Introduction	9				
4	l.1.1	Data Sources	9				
4	l.1.2 D	Data Cleaning	9				
4	l.1.3 D	Data EXPLORATION	9				
	UNI	IVARIATE ANALYSIS	9				
	BIV	ARIATE ANALYSIS	9				
	TEX	XT ANALYSIS	10				
4	l.1.4 II	NSIGHTS AND VISUALIZATIONS	10				
4	l.1.5 C	CONCLUSION	10				
5.Methodology & Experiments							
5	5.1	Data Storage and Cleaning	11				
5	5.1.1	Experimentation with models	11				
5	5.1.2	Visualizations	11				
6.	Resu	ults	12				
7.	Disc	cussion	36				
7	7.1.1	LEARNINGS & FUTURE WORK	36				
7	7.1.2	CONCLUSION	36				
8	GIT	THUB REPOSITORY / CODE FILES LINK	37				
8	8.1	MY REPOSITORY LINK:	37				
8	3.2	Research Paper References	38				

## 1. INTRODUCTION

The ubiquity of e-commerce platforms has revolutionized the way businesses operate and customers shop. With this digital evolution, product reviews and ratings have emerged as powerful indicators of product performance and customer satisfaction. These reviews not only influence consumer purchase decisions but also provide businesses with invaluable insights into their product's market performance. However, the sheer volume of reviews generated on e-commerce platforms makes manual analysis a daunting task. Consequently, there is an increasing need for automated tools capable of analyzing and interpreting this rich source of information.

The proposed project, 'Sentiment Analysis of Product Reviews for E-commerce Platforms,' aims to fill this gap by implementing advanced Natural Language Processing (NLP) and deep learning techniques to automatically analyze and interpret the sentiments inherent in product reviews. The project focuses on datasets from two of the largest e-commerce platforms globally, namely Amazon and Flipkart.

The project's primary objective is to assess a product's market performance by leveraging sentiment analysis to evaluate customer reviews, ratings, and product prices. Such a methodology can provide crucial insights into customer behavior and preferences, enabling businesses to understand their product's strengths and weaknesses more effectively. In turn, this can guide strategic growth decisions, such as adjusting product prices based on consumer sentiment and competitive positioning.

Furthermore, the project addresses the research question: "How can sentiment analysis of product reviews and ratings on e-commerce platforms be used to assess products' performance?" Answering this question will not only enable manufacturers and sellers to gain a comprehensive understanding of customer perceptions but also provide a roadmap for product improvement, informed pricing strategies, and ultimately, business growth.

## 2. DATA DESCRIPTION

The datasets for this project are sourced from two of the most prominent e-commerce giants globally, Amazon and Flipkart. These platforms have been selected for their broad global reach, extensive product range, and the sheer volume of customer reviews they garner.

#### **Datasets**

#### 1. Flipkart Products Review Dataset

https://www.kaggle.com/datasets/mansithummar67/flipkart-product-review-dataset/discussion?select=flipkart\_product.csv

#### **Description**

✓ This Dataset contains information of Products Name, Price, Review, Rate, Summary for the Sentiment Analysis Purpose.

#### 2. Amazon Mobile Phone Reviews Dataset

https://data.world/promptcloud/amazon-mobile-phone-reviews

### **Description**

✓ This Dataset contains information of Product Title, Brand, Price, Rating, Review Text

## 3. LITERATURE REVIEW

#### 3.1 INTRODUCTION

In the E-commerce platform domain, understanding consumer sentiments through product reviews is essential. Sentiment analysis or opinion mining extracts subjective information like opinions, emotions, and attitudes from text. Major platforms such as Amazon, Flipkart, and eBay rely on customer reviews to gauge product quality, features, and user satisfaction. This literature review categorizes 30 papers into six categories: "Amazon Dataset", "Twitter Dataset", "General E-commerce Platforms", "Review Helpfulness and Market Value", "Data Analytics and Visualization", "Business Analytics and Innovation".

### 3.1.1 AMAZON DATASET

Five papers utilizing Amazon datasets for sentiment analysis were reviewed.

In the study "Deep Learning Sentiment Analysis of Amazon.com Reviews and Ratings", a deep learning model achieved 81.82% accuracy in detecting mismatched ratings by combining review and product embeddings. "Product Sentiment Analysis for Amazon Reviews" demonstrated that the BERT model excelled in predicting sentiment polarity of Amazon mobile phone reviews, with an accuracy of 94%. The research titled "Predicting online product sales via online reviews, sentiments, and promotion strategies" underscored the importance of variables such as review volume and sentiment in forecasting product sales. "HOW HELPFUL ARE COMPARATIVE REVIEWS FOR PREDICTING PRODUCT DEMAND?" found that text semantic similarity in comparative reviews held more weight than sentiment in influencing product sales. Lastly, "Model Comparison in Sentiment Analysis: A Case Study of Amazon Product Reviews" introduced the Hybrid Sequential Binary Classification (HSBC) method, outperforming other methods with an accuracy of 64.7%.

### 3.1.2 TWITTER DATASET

Two papers focused on Twitter sentiment analysis. Both the papers focus on sentiment analysis on Twitter data but employ different methodologies.

The paper "Big data approach for sentiment analysis of twitter data using Hadoop framework and deep learning" employed a Hadoop framework in conjunction with a deep recurrent neural network (RNN) to classify Twitter data into positive or negative sentiments with high accuracy, sensitivity, and specificity. On the other hand, "PROJECT REPORT SENTIMENT ANALYSIS ON TWITTER USING APACHE SPARK" utilized Apache Spark and the Twitter4j library to extract live tweets and perform sentiment analysis through a custom algorithm, effectively classifying tweets into positive, negative, and neutral categories.

Both methodologies present opportunities for adaptation to e-commerce product review sentiment analysis, with the first paper's deep learning approach potentially offering more accurate classification for complex

sentences, and the second paper's real-time data processing being applicable for live customer feedback analysis.

#### 3.1.3 GENERAL E-COMMERCE PLATFORMS

Four papers explored sentiment analysis across various e-commerce platforms.

The paper "Sentiment Analysis on Reviews of E-commerce Sites Using Machine Learning Algorithms" emphasized the importance of tailored algorithms for languages like Bangla and English, with Support Vector Machine achieving high accuracy in sentiment analysis. Meanwhile, "Sentiment analysis of online product reviews based on SenBERT CNN" introduced a novel model, SenBERT CNN, which combined BERT and Convolutional Neural Network for improved performance in analyzing e-commerce reviews, particularly in extracting deeper text features. "Sentiment Analysis of Air Conditioning Users' Comments on e-Commerce Platform Based on NLP Technology" employed Natural Language Processing to analyze user comments on air conditioning products, highlighting the importance of focusing on product performance and after-sales service. Lastly, "Sentiment Analysis in E-Commerce: A Review on The Techniques and Algorithms" provided a comprehensive overview of various sentiment analysis techniques and algorithms used in e-commerce, emphasizing their potential in customer relationship management despite challenges like sarcasm and word sense disambiguation.

### 3.1.4 REVIEW HELPFULNESS AND MARKET VALUE

This category contains four papers focusing on review helpfulness and market value.

In "Extracting Reliability of the Products through User Reviews," researchers developed automated methods for generating reliable product rating scores through sentiment analysis and feature extraction of user reviews. "Effectiveness of Fine-tuned BERT Model in Classification of Helpful and Unhelpful Online Customer Reviews" showed that fine-tuning BERT can significantly improve the classification of helpfulness in online customer reviews. "Sentiment Analysis over Online Product Reviews: A Survey" provided an extensive overview of the techniques, challenges, and applications of sentiment analysis, emphasizing its importance in extracting insights from user-generated content. Lastly, "Impact of Positive vs. Negative e-Sentiment on Daily Market Value of High-Tech Products" revealed that negative electronic sentiment has a more significant impact on the market prices of high-tech products than positive sentiment, highlighting the importance for managers to pay close attention to negative feedback.

### 3.1.5 DATA ANALYTICS AND VISUALIZATION

This category contains thirteen papers focusing on different python packages, tableau tips etc for visualizing data.

The "Visualization System For Sentiment Analysis Using Textblob On Twitter" paper focuses on sentiment analysis in Bahasa Indonesia using Tweepy and TextBlob libraries, utilizing visual analytics to explore sentiment distribution geographically and achieving higher accuracies compared to SentiWordNet. Another paper, "Data Filtering and Visualization for Sentiment Analysis of Ecommerce Websites," performs sentiment analysis on customer reviews of Realme 6i, collecting and preprocessing data with UI path and Python, visualizing the results to aid in product improvement. "Visualization Resources: A Survey" provides a categorized collection of visualization resources for students, researchers, practitioners, and developers, offering recommendations as a starting point for exploration. "Sentiment Analysis, Tweet Analysis and Visualization on Big Data Using Apache Spark and Hadoop" conducts sentiment analysis on tweets using Apache Spark and Hadoop, comparing algorithms and visualizing user-related metrics with Power BI. "An Introduction to Data Visualization Tools and Techniques in Various Domains" highlights the importance of data visualization, providing an overview of techniques and practical examples for enhanced data understanding and communication. Lastly, "The State of the Art in Sentiment Visualization" surveys and categorizes sentiment visualization techniques, emphasizing evaluation and perception aspects, supporting researchers and practitioners in sentiment analysis and visualization. These papers contribute valuable insights, methodologies, and recommendations in the fields of sentiment analysis, data visualization, and related areas.

In the realm of Data Analytics & Visualization Methods, the following papers offer valuable contributions. "Exploring the Cultivation of Innovative Talents in the Era of Big Data and Cloud Computing" focuses on fostering innovation and enhancing interdisciplinary abilities through statistical and data analysis education, promoting cross integration between science and engineering. "Information Visualization: Perception for Design" delves into the science of perception and the human visual system, providing design principles to effectively convey information through well-designed visualizations. "E-commerce data analysis and visualization" investigates Pakistan's Largest E-commerce dataset to extract insights about popular categories, customer behavior analysis, and sales condition prediction to enhance business strategies. "The Use of Data Visualization in E-Commerce: A Review" critically examines data visualization in e-commerce, specifically emphasizing Clickstream data. It categorizes existing literature, summarizes visualization techniques, and offers recommendations for designers in this domain. "Survey on Sentimental Analysis and Visualization of Reviews" highlights the significance of sentiment analysis and visualization in the context of food ordering websites, exploring diverse algorithms and techniques while underscoring the importance of visualization tools. "DataSist: A Python-based library for easy data analysis, visualization, and modeling" introduces DataSist, a Python library that simplifies and automates the data analysis process, facilitating high-level functions for analyzing, mining, and visualizing large datasets. "A Study of Big Data Processing for Sentiments Analysis" delves into the tools, techniques, and technologies employed for processing big data, emphasizing sentiment analysis on social media platforms and its applications in streaming data analysis. These papers collectively contribute to the field of data analytics and visualization by providing insights into interdisciplinary education, effective information display, e-commerce analysis, sentiment analysis, and automated data analysis techniques.

### 3.1.6 BUSINESS ANALYTICS AND INNOVATION

This category contains two papers.

"NNBI: A Neural Network based Business Intelligence Dashboard for a Clothing Store" focuses on developing a Neural Network based Business Intelligence (NNBI) dashboard for a clothing store. It utilizes recurrent neural networks (RNN) to predict sales on a daily, monthly, and weekly basis, with Tableau as the chosen BI tool for interactive analysis and visualization. The NNBI dashboard integrates data from multiple sources, achieving high accuracy in sales forecasting and enabling effective decision-making. "Business and data analytics: New innovations in the management of e-commerce" explores innovative approaches to managing e-commerce through business and data analytics. The research covers various data types, including quantitative, qualitative, and multimedia data, and employs methods such as regression analysis, neural network analysis, and algorithm-based approaches. The studies encompass diverse areas such as customer behavior analysis, personalization systems, decision support tools, and ranking mechanisms, contributing to the understanding of data-driven strategies in e-commerce. These papers collectively contribute to the field of Business Analytics and Innovation by providing insights into predictive modeling, interactive dashboard design, and leveraging data for informed decision-making in the clothing retail industry and e-commerce management.

#### 3.2 CONCLUSION

Sentiment analysis in e-commerce is critical for extracting insights from consumer reviews. Advanced methodologies, including deep learning and NLP techniques, have shown effectiveness in sentiment classification. While positive sentiments are important, negative sentiments significantly influence consumer decisions and market values. Additionally, real-time sentiment analysis and consideration of various languages can be beneficial. The papers reviewed form a solid foundation for further research in sentiment analysis within the e-commerce domain.

## 4. EXPLANATORY DATA ANALYSIS (EDA)

#### 4.1 INTRODUCTION

This EDA aims to conduct a preliminary assessment on datasets of customer reviews and ratings of products from e-commerce platforms. Through sentiment analysis, the performance of products can be gauged by analyzing customer reviews, ratings, and prices. This allows businesses to understand their product's strengths and weaknesses and make strategic decisions for improvement and growth.

### 4.1.1 DATA SOURCES

- 1. Flipkart Products Review Dataset
  - Source: Kaggle
  - Features: Product Name, Price, Rating, Review, Summary Link to dataset
- 2. Amazon Mobile Phone Reviews Dataset
  - Source: data. World
  - Features: Product Title, Brand, Price, Rating, Review Text, Review Votes Link to dataset

### 4.1.2 DATA CLEANING

Data cleaning is an essential step to ensure the quality and reliability of the data. During this step:

- 1. Remove any duplicates and irrelevant data.
- 2. Handle missing values by imputation or removal.
- 3. Clean text data by removing punctuations, converting to lowercase, etc.

### 4.1.3 DATA EXPLORATION

#### UNIVARIATE ANALYSIS

- 1. **Distribution of Ratings**: Create pie chart to display the distribution of ratings.
- 2. **Price Distribution**: Utilize box plots to visualize the distribution of prices for the products.

#### **BIVARIATE ANALYSIS**

1. **Ratings vs. Price**: Explore the correlation between product prices and their respective average ratings using scatter plots.

2. **Sentiment vs. Rating**: Perform sentiment analysis on the reviews and explore its relationship with ratings. Use heatmap to represent the proportion of positive, negative, and neutral sentiments for different rating groups.

#### TEXT ANALYSIS

1. **Word Cloud**: Generate word clouds for positive and negative reviews to identify common words in each category.

## 4.1.4 INSIGHTS AND VISUALIZATIONS

- 1. Use Tableau or Python libraries (such as seaborn and matplotlib) for visualizing the patterns and trends in the data.
- 2. Present the distribution of ratings and how it correlates with price, along with a breakdown by brand. This helps in identifying if more expensive products or certain brands receive higher ratings.
- 3. Analyze and visualize the sentiments expressed in the review texts, how they relate to the product ratings, and their impact on review votes.
- 4. Present the results of keyword analysis and topic modeling to summarize the main concerns or praises in the reviews, and how these topics might affect product sentiment.
- 5. Visualize the aggregated sentiment scores by product and brand and analyze how they can impact market performance.

### 4.1.5 CONCLUSION

The EDA will provide insights into the customer reviews and ratings for products on e-commerce platforms. These insights will help businesses understand customer perceptions, which is valuable for product enhancement and pricing strategies. This preliminary analysis sets the foundation for further in-depth analysis and modeling.

## **5.METHODOLOGY & EXPERIMENTS**

#### 5.1 DATA STORAGE AND CLEANING

For managing and processing our large datasets, we utilize Apache Hadoop, an open-source software framework that allows for distributed processing of large data sets across clusters of computers. Given the scale of data from Amazon and Flipkart, Hadoop's ability to scale from a single server to thousands of machines, each offering local computation and storage, is crucial.

To assist in the cleaning and pre-processing of our data, we use PySpark, the Python library for Apache Spark. PySpark is a powerful tool for handling large-scale data in a distributed fashion and provides a variety of functionalities for data transformation, making it a suitable tool for preparing our datasets for further analysis.

#### 5.1.1 EXPERIMENTATION WITH MODELS

Python, known for its simplicity, readability, and the availability of powerful data manipulation and machine learning libraries, is our language of choice for model experimentation and prediction tasks.

To accomplish sentiment analysis, we utilize Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), both types of deep learning models. RNNs' strength lies in their ability to analyze sequential data, like sentences, making them ideal for our project's sentiment analysis. While CNNs are traditionally used for image processing, their application to text data allows us to capture local semantic features, thus enhancing the performance of our sentiment analysis.

In addition, we leverage Transformer models, like BERT, which have revolutionized natural language processing. Transformer models can capture more complex patterns in text data compared to traditional methods, leading to improved performance in tasks such as sentiment analysis.

### 5.1.2 VISUALIZATIONS

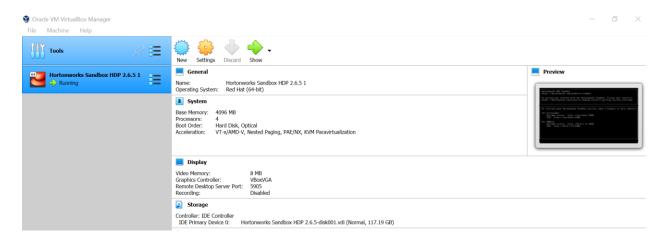
We employ Python libraries such as matplotlib, seaborn, and plotly for creating comprehensive data visualizations. These libraries enable us to create histograms, bar plots, scatter plots, etc., helping us visualize the distribution of sentiments and correlations between different variables effectively.

Further, we use Tableau, a robust data visualization tool, to create interactive dashboards and reports. Tableau allows for the visualization of sentiment scores for different products and the observation of trends over time, offering a more interactive and intuitive understanding of our findings.

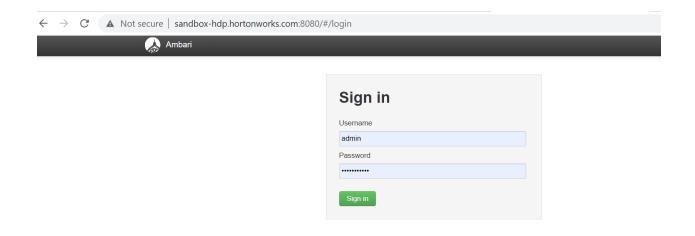
# 6. RESULTS

## Data cleaning task using Hadoop and PySpark.

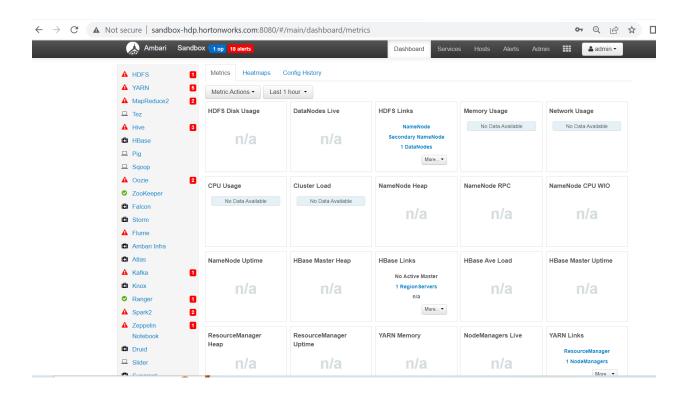




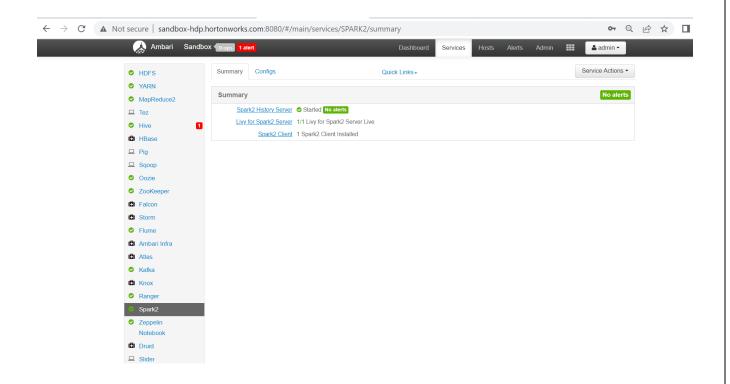
Run the virtual box first. VirtualBox allows you to create isolated, reproducible environments for Hadoop development and testing, and simulate multi-node clusters on a single machine.



Start the services on Ambari using the browser link <a href="http://sandbox-hdp.hortonworks.com:8080/#/login">http://sandbox-hdp.hortonworks.com:8080/#/login</a>



Services are in red alert, that means not started yet.



Services on Ambari started successfully.

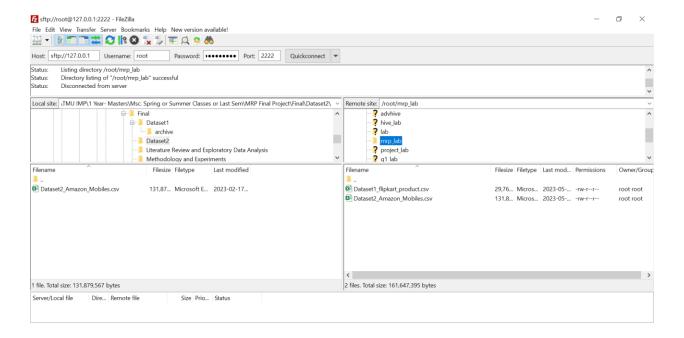
## ← → C ▲ Not secure | sandbox-hdp.hortonworks.com:4200

```
sandbox-hdp login: root
root@sandbox-hdp.hortonworks.com's password:
Last login: Wed May 10 19:45:50 2023 from 172.18.0.2
[root@sandbox-hdp ~]# hadoop fs -ls /
Found 11 items
drwxrwxrwx
                      hadoop
                                      0 2022-10-04 00:33 /app-logs
             - yarn
drwxr-xr-x
             - hdfs
                      hdfs
                                      0 2018-06-18 16:13 /apps
drwxr-xr-x
                      hadoop
                                      0 2018-06-18 14:52 /ats
            - yarn
drwxr-xr-x
             - hdfs
                      hdfs
                                      0 2018-06-18 14:52 /hdp
drwx----
            - livy
                      hdfs
                                     0 2018-06-18 15:11 /livy2-recovery
drwxr-xr-x
             - mapred hdfs
                                      0 2018-06-18 14:52 /mapred
            - mapred hadoop
                                    0 2018-06-18 14:52 /mr-history
drwxrwxrwx
drwxr-xr-x
             - hdfs
                      hdfs
                                     0 2018-06-18 15:59 /ranger
                                     0 2022-11-29 04:49 /spark2-history
drwxrwxrwx
             - spark
                     hadoop
drwxrwxrwx
             - hdfs
                      hdfs
                                      0 2022-11-06 12:02 /tmp
drwxr-xr-x
             - hdfs
                      hdfs
                                      0 2022-10-25 15:44 /user
[root@sandbox-hdp ~]# hadoop fs -ls /user/root
Found 15 items
drwx----
             - root root
                                  0 2022-11-29 00:00 /user/root/.Trash
drwxr-xr-x
                                  0 2022-11-27 05:40 /user/root/.hiveJars
             - root root
             - root root
                                  0 2022-11-04 03:09 /user/root/.sparkStaging
drwxr-xr-x
                              20177 2022-11-06 12:11 /user/root/2014-summary.json
-rw-r--r--
             1 root root
                              21368 2022-11-06 12:11 /user/root/2015-summary.json
-rw-r--r--
            1 root root
drwxr-xr-x
             - root root
                                  0 2022-11-28 04:10 /user/root/advhive
                                  0 2022-11-04 01:53 /user/root/file.txt
drwxr-xr-x
             - root root
                                  0 2022-11-27 08:21 /user/root/hive lab
drwxr-xr-x
             - root root
                                  0 2022-11-27 07:19 /user/root/hivelab
drwxr-xr-x
             - root root
drwxr-xr-x
                                  0 2022-11-06 12:14 /user/root/lab
             - root root
                                  0 2022-11-26 05:13 /user/root/project lab
drwxr-xr-x
             - root root
                                  0 2022-11-01 00:22 /user/root/q1 lab
drwxr-xr-x
             - root root
                                  0 2022-10-26 03:28 /user/root/q2 lab
drwxr-xr-x
             - root root
drwxr-xr-x
                                  0 2022-11-10 23:38 /user/root/sparklab
             - root root
                                  0 2022-11-27 05:39 /user/root/twitter
drwxr-xr-x
           - root root
[root@sandbox-hdp ~]# hadoop fs -mkdir /user/root/mrp lab
[root@sandbox-hdp ~]# mkdir /root/mrp_lab
```

Login to HDP sandbox using the browser link <a href="http://sandbox-hdp.hortonworks.com:4200/">http://sandbox-hdp.hortonworks.com:4200/</a>. I checked the list of all the directories/data that is present at the location /user/root. Then, I made the directory using 'mkdir' on local system and 'Hadoop fs -mkdir' for making directory in hdfs.

```
[root@sandbox-hdp ~]# hadoop fs -ls /user/root
Found 16 items
drwx----
                                  0 2022-11-29 00:00 /user/root/.Trash
             - root root
                                  0 2022-11-27 05:40 /user/root/.hiveJars
drwxr-xr-x
             - root root
                                  0 2022-11-04 03:09 /user/root/.sparkStaging
drwxr-xr-x
             - root root
-rw-r--r--
             1 root root
                              20177 2022-11-06 12:11 /user/root/2014-summary.json
                              21368 2022-11-06 12:11 /user/root/2015-summary.json
             1 root root
-rw-r--r--
             - root root
                                  0 2022-11-28 04:10 /user/root/advhive
drwxr-xr-x
drwxr-xr-x
             - root root
                                  0 2022-11-04 01:53 /user/root/file.txt
                                  0 2022-11-27 08:21 /user/root/hive lab
drwxr-xr-x
             - root root
                                  0 2022-11-27 07:19 /user/root/hivelab
drwxr-xr-x
             - root root
drwxr-xr-x
             - root root
                                  0 2022-11-06 12:14 /user/root/lab
drwxr-xr-x
                                  0 2023-05-10 19:58 /user/root/mrp lab
             - root root
                                  0 2022-11-26 05:13 /user/root/project_lab
drwxr-xr-x
             - root root
                                  0 2022-11-01 00:22 /user/root/q1 lab
drwxr-xr-x
             - root root
                                  0 2022-10-26 03:28 /user/root/q2 lab
drwxr-xr-x
             - root root
                                  0 2022-11-10 23:38 /user/root/sparklab
drwxr-xr-x
             - root root
                                  0 2022-11-27 05:39 /user/root/twitter
drwxr-xr-x
             - root root
```

Directory created successfully with the name /mrp\_lab



Login to FileZilla and check if the directory created already. I now transferred/drag and dropped the desired files (Dataset1\_flipkart\_product.csv, Dataset2\_Amazon\_Mobiles.csv) into the remote site directory 'mrp\_lab'

Host: sftp://127.0.0.1

Username: root

Password: \*\*\*\*\*

Port: 2222

```
[root@sandbox-hdp ~]# cd /root/mrp_lab
[root@sandbox-hdp mrp_lab]# hadoop fs -put Dataset1_flipkart_product.csv /user/root/mrp_lab
[root@sandbox-hdp mrp_lab]# hadoop fs -put Dataset2_Amazon_Mobiles.csv /user/root/mrp_lab
[root@sandbox-hdp mrp_lab]# hadoop fs -ls /user/root/mrp_lab
Found 2 items
-rw-r--r-- 1 root root 29767828 2023-05-10 20:18 /user/root/mrp_lab/Dataset1_flipkart_product.csv
-rw-r--r-- 1 root root 131879567 2023-05-10 20:20 /user/root/mrp_lab/Dataset2_Amazon_Mobiles.csv
```

Now in this step, I changed the directory so that I can go to mrp\_lab folder. Once I enter to that directory/folder. I then copied the files from local system to hdfs using put command while the get command performs the visa-versa operation on the files.

Using Python version 2.7.5 (default, Apr 11 2018 07:36:10) SparkSession available as 'spark'.

Here in the above snapshot, we can see that by typing the pyspark, it landed me to the spark session where I can start typing my pyspark commands to get the desired outputs.

```
>>> from pyspark.sql import SparkSession
>>> from pyspark.sql.functions import col
>>> spark=SparkSession.builder \
       .appName("Sentiment Analysis") \
        .getOrCreate()
>>> flipkart df=spark.read \
        .format("com.databricks.spark.csv") \
         .option("header", "true") \
         .option("inferSchema", "true") \
         .load("hdfs:///user/root/mrp_lab/Dataset1_flipkart_product.csv")
>>> amazon df=spark.read \
       .format("com.databricks.spark.csv") \
        .option("header","true") \
       .option("inferSchema","true") \
       .load("hdfs:///user/root/mrp lab/Dataset2 Amazon Mobiles.csv")
>>> flipkart df.show()
  .-----
        ProductName | Price | Rate |
                                         Review
                                                            Summary
   -----+
|Candes 12 L Room/...|�??3,999| 5|
                                           Super! Great cooler.. ex...
|Candes 12 L Room/... | �??3,999 |
                            5|
                                         Awesome Best budget 2 fit...
|Candes 12 L Room/...|�??3,999|
                             3
                                             Fair The quality is go...
                            1  Useless product | Very bad product ...
|Candes 12 L Room/...|�??3,999|
|Candes 12 L Room/... |�??3,999|
                             3 l
                                             Fair
                                                      Ok ok product
|Candes 12 L Room/... | �??3,999 |
                             5
                                          Awesome The cooler is rea...
                             5 | Highly recommended | Very good product
|Candes 12 L Room/...|�??3,999|
|Candes 12 L Room/...|�??3,999|
                             3|
                                             Nice
                                                           Very nice
|Candes 12 L Room/...|�??3,999|
                                                    Very bad cooler
                             11
                                    Unsatisfactory
|Candes 12 L Room/...|�??3,999|
                             4
                                    Worth the money
                                                           Very good
|Candes 60 L Room/... |�??8,999|
                                     Great product | Beautiful product... |
                             5|
|Candes 60 L Room/...|�??8,999|
                             5 Mind-blowing purc...
                                                             Awesome
|Candes 60 L Room/...|�??8,999|
                                 Highly recommended
                                                               Good
                             5|
|Candes 60 L Room/... |�??8,999|
                             5|
                                         Brilliant|Wonderful product...|
|Candes 60 L Room/...|�??8,999|
                             5
                                    Classy product Nice air cooler, ...
|Candes 60 L Room/...|�??8,999|
                             5|
                                         Must buy!
                                                               Awsm
|Candes 60 L Room/...|�??8,999|
                             5|
                                         Fabulous! Nice product ???�|
|Candes 60 L Room/... | ♠??8,999|
                            5| Worth every penny| Great cooler..|
|Candes 60 L Room/...|�??8,999|
                             5
                                           Super!
                                                     Nice product
|Candes 60 L Room/...|�??8,999| 5|
                                    Great product
                                                       Good ???�|
+-----
only showing top 20 rows
```

```
>>> amazon_df.show()
+-----
     Product Name Brand Name Price Rating Reviews Review Votes
+----+
"""CLEAR CLEAN ES...
                        Samsung | 199.99 |
                                             5|I feel so LUCKY t...|
"""CLEAR CLEAN ES...
                                             4|nice phone, nice ...|
                        Samsung | 199.99 |
"""CLEAR CLEAN ES...
                                             5| Very pleased|
                         Samsung | 199.99 |
                                                                               0
"""CLEAR CLEAN ES...
                         Samsung | 199.99 |
                                             4 It works good but...
                                                                               01
"""CLEAR CLEAN ES...
                         Samsung | 199.99 |
                                             4 Great phone to re...
"""CLEAR CLEAN ES...
                         Samsung | 199.99 |
                                             1 I already had a p...
                                                                               1
"""CLEAR CLEAN ES...
                         Samsung | 199.99 |
                                             2 The charging port...
                                                                               0
"""CLEAR CLEAN ES...
                         Samsung | 199.99 |
                                             2|Phone looks good ...|
                                                                               01
"""CLEAR CLEAN ES...
                         Samsung | 199.99 |
                                             5|I originally was ...|
                                                                               01
"""CLEAR CLEAN ES...
                         Samsung | 199.99 |
                                             3|It's battery life...|
                                                                               0 l
"""CLEAR CLEAN ES...
                         Samsung | 199.99 |
                                             3 My fiance had thi...
                                                                               0
"""CLEAR CLEAN ES...
                         Samsung | 199.99 |
                                             5 This is a great p...
                                                                               0
"""CLEAR CLEAN ES...
                         Samsung | 199.99 |
                                             5 These guys are th...
"""CLEAR CLEAN ES...|
                         Samsung | 199.99 |
                                             1 I'm really disapp...
                                                                               1 |
"""CLEAR CLEAN ES...
                         Samsung | 199.99 |
                                             5|Ordered this phon...|
                                                                               1
"""CLEAR CLEAN ES...
                         Samsung | 199.99 |
                                             2|Had this phone be...|
                                                                               01
"""CLEAR CLEAN ES...
                                             5|I was able to get...|
                         Samsung | 199.99 |
                                                                               6
"""CLEAR CLEAN ES...
                         Samsung | 199.99 |
                                             5 I brought this ph...
                                                                               0
"""CLEAR CLEAN ES...
                        Samsung | 199.99 |
                                             4 I love the phone....
                                                                               11
"""CLEAR CLEAN ES...
                                            3 unfortunately Spr...
                        Samsung | 199.99 |
+-----
only showing top 20 rows
>>> flipkart_df=flipkart_df.dropna()
>>> amazon_df=amazon_df.dropna()
>>> from pyspark.sql.functions import regexp_replace
>>> flipkart_df=flipkart_df,withColumn('Review',regexp_replace(col('Review'),'[^a-zA-Z0-9\s]',''))
Traceback (most recent call last):
 File "<stdin>", line 1, in <module>
NameError: name 'withColumn' is not defined
>>> flipkart_df=flipkart_df.withColumn('Review',regexp_replace(col('Review'),'[^a-zA-Z0-9\s]',''))
>>> amazon_df=amazon_df.withColumn('Reviews',regexp_replace(col('Reviews'),'[^a-zA-Z0-9\s]',''))
>>> flipkart_df=flipkart_df.filter(~col('Review').contains('??'))
>>> amazon_df=amazon_df.filter(~col('Reviews').conatins('??'))
Traceback (most recent call last):
 File "<stdin>", line 1, in <module>
TypeError: 'Column' object is not callable
>>> amazon_df=amazon_df.filter(~col('Reviews').contains('??'))
>>> from pyspark.sql.functions import lower
>>> flipkart_df=flipkart_df.withColumn('Review',lower(col('Review')))
>>> amazon_df=amazon_df.withColumn('Reviews',lower(col('Reviews')))
>>> from pyspark.sql.functions import regexp_replace,col
>>> for column in flipkart df.columns:
      flipkart df=flipkart df.withColumn(column,regexp replace(col(column),'*','?''))
```

#### >>> flipkart\_df.show()

```
-----
    ProductName| Price|Rate| Review|
    -----
                             5| super|Great cooler.. ex...|
|Candes 12 L Room/...|????3,999|
|Candes 12 L Room/...|????3,999|
                           5
                                        awesome|Best budget 2 fit...|
|Candes 12 L Room/...|????3,999|
                             3 |
                                         fair|The quality is go...|
                               useless product|Very bad product ...|
|Candes 12 L Room/...|????3,999|
                             1
|Candes 12 L Room/...|????3,999|
                             3 |
                                         fair | Ok ok product
|Candes 12 L Room/...|????3,999|
                             5
                                         awesome The cooler is rea...
|Candes 12 L Room/...|????3,999|
                             5|
                               highly recommended | Very good product
|Candes 12 L Room/...|????3,999|
                             3|
                                            nice
                                                         Very nice
                                  unsatisfactory | Very bad cooler worth the money | Very good
|Candes 12 L Room/...|????3,999|
                             1
|Candes 12 L Room/...|????3,999|
                             4
                                                         Very good
|Candes 60 L Room/...|????8,999|
                                  great product|Beautiful product...
                             5
|Candes 60 L Room/...|????8,999|
                             5|mindblowing purchase|
                                                          Awesome
|Candes 60 L Room/...|????8,999|
                             5
                               highly recommended
                                        brilliant|Wonderful product...
|Candes 60 L Room/...|????8,999|
                             5
|Candes 60 L Room/...|????8,999|
                             5
                                   classy product Nice air cooler, ...
|Candes 60 L Room/...|????8,999|
                             5
                                        must buy
                                                             Awsm
                                        fabulous | Nice product ?????
|Candes 60 L Room/...|????8,999|
                             5|
                             5
|Candes 60 L Room/...|????8,999|
                               worth every penny| Great cooler..
                               super
                                                    Nice product
|Candes 60 L Room/...|????8,999|
                             5|
|Candes 60 L Room/...|????8,999|
                           5
                                                      Good ?????
                                    great product
+-----
```

only showing top 20 rows

```
>>> flipkart_df=flipkart_df.withColumn('Price',regexp_replace(col('Price'),'\?+',''))
>>> flipkart_df=flipkart_df.withColumn('Summary',regexp_replace(col('Summary'),'�$',''))
```

#### >>> flipkart\_df.show()

+	++	+	<u> </u>	++
ProductName	Price	Rate	Review	Summary
Candes 12 L Room/	3 <b>,</b> 999	5	super	Great cooler ex
Candes 12 L Room/	3,999	5	awesome	Best budget 2 fit
Candes 12 L Room/	:			The quality is go
Candes 12 L Room/	3,999	1	useless product	Very bad product
Candes 12 L Room/	3,999	3	fair	Ok ok product
Candes 12 L Room/	3,999	5	awesome	The cooler is rea
Candes 12 L Room/	3,999	5	highly recommended	Very good product
Candes 12 L Room/	3,999	3	nice	Very nice
Candes 12 L Room/	3,999	1	unsatisfactory	Very bad cooler
Candes 12 L Room/	3,999	4	worth the money	Very good
Candes 60 L Room/	8,999	5	great product	Beautiful product
Candes 60 L Room/	8,999	5	mindblowing purchase	Awesome
Candes 60 L Room/	8,999	5	highly recommended	Good
Candes 60 L Room/	8,999	5	brilliant	Wonderful product
Candes 60 L Room/	8,999	5	classy product	Nice air cooler,
Candes 60 L Room/	8,999	5	must buy	Awsm
Candes 60 L Room/	8,999	5	fabulous	Nice product ?????
Candes 60 L Room/	8,999	5	worth every penny	Great cooler
Candes 60 L Room/	8,999	5	super	Nice product
Candes 60 L Room/	8,999	5	great product	Good ?????
+	+	<b></b> +	<b></b>	++

only showing top 20 rows

```
+----+
     Product Name|Brand Name| Price|Rating| Reviews|Review Votes|
+----+
"""CLEAR CLEAN ES... | Samsung | 199.99 | 5 | i feel so lucky t... |
                                                                                     1
"""CLEAR CLEAN ES...| Samsung 199.99
                                              4|nice phone nice u...|
                                                                                     0 l
"""CLEAR CLEAN ES... | Samsung | 199.99 |
"""CLEAR CLEAN ES...| Samsung|199.99|
"""CLEAR CLEAN ES...| Samsung|199.99|
                                                5| very pleased|
                                                                                     01
                                                4|it works good but...|
                                                                                     0
"""CLEAR CLEAN ES...
                          Samsung | 199.99 |
                                                4 great phone to re...
                                                                                     0
"""CLEAR CLEAN ES...
                          Samsung | 199.99 |
                                                1|i already had a p...|
                                                                                     1
"""CLEAR CLEAN ES...
                          Samsung | 199.99 |
                                                2 the charging port...
                                                                                     0
"""CLEAR CLEAN ES...
                          Samsung | 199.99 |
                                                2|phone looks good ...|
                                                                                     0
"""CLEAR CLEAN ES...
                          Samsung | 199.99 |
                                                5|i originally was ...|
                                                                                     0
"""CLEAR CLEAN ES...
                          Samsung | 199.99 |
                                                3|its battery life ...|
                                                                                     0
"""CLEAR CLEAN ES...
                          Samsung | 199.99 |
                                                3 my fiance had thi...
                                                                                     0
"""CLEAR CLEAN ES...
                          Samsung | 199.99 |
                                                5 this is a great p...
                                                                                     0
                          Samsung | 199.99 |
"""CLEAR CLEAN ES...
                                                1|im really disappo...|
                                                                                     1
"""CLEAR CLEAN ES...
                                                5|ordered this phon...|
                          Samsung | 199.99 |
                                                                                     1
"""CLEAR CLEAN ES...Samsung | 199.99 |2 | had this phone be...|"""CLEAR CLEAN ES...Samsung | 199.99 |5 | i was able to get...|"""CLEAR CLEAN ES...Samsung | 199.99 |5 | i brought this ph...|"""CLEAR CLEAN ES...Samsung | 199.99 |4 | i love the phone ...|"""CLEAR CLEAN ES...Samsung | 199.99 |3 | unfortunately spr...
                                                                                     01
                                                                                     6
                                                                                     01
                                                                                     1
                                                                                     0 l
```

only showing top 20 rows

```
>>> flipkart_df.write.csv('/user/root/mrp_lab/flipkart_data.csv', header=True)
>>> amazon_df.write.csv('/user/root/mrp_lab/amazon_data.csv', header=True)
```

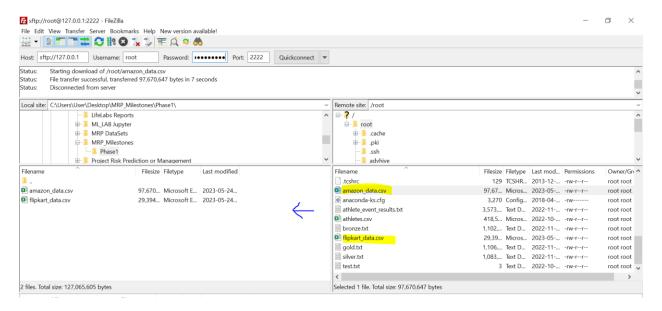
#### Wrote DataFrame to HDFS

```
[root@sandbox-hdp ~]# hadoop fs -getmerge /user/root/mrp_lab/flipkart_data.csv /C:/Users/User/Desktop/MRP_Milestones/flipkart_data.csv [root@sandbox-hdp ~]# hadoop fs -getmerge /user/root/mrp_lab/amazon_data.csv /C:/Users/User/Desktop/MRP_Milestones/amazon_data.csv [root@sandbox-hdp ~]#
```

Then, used the command hadoop fs -getmerge to merge and download the .csv file to my local machine

```
[root@sandbox-hdp ~]# hadoop fs -getmerge /user/root/mrp_lab/flipkart_data.csv /root/flipkart_data.csv [root@sandbox-hdp ~]# hadoop fs -getmerge /user/root/mrp_lab/amazon_data.csv /root/amazon_data.csv [root@sandbox-hdp ~]# hadoop fs -getmerge /user/root/mrp_lab/flipkart_data.csv /tmp/flipkart_data.csv [root@sandbox-hdp ~]# hadoop fs -getmerge /user/root/mrp_lab/amazon_data.csv /tmp/amazon_data.csv
```

Downloaded the files to the local file system on the server. The paths are now Unix-style and point to the server's local file system.



Now I have the files on my server's local file system (/tmp/flipkart\_data.csv and /tmp/amazon\_data.csv). The next step is to get these files to my local Windows machine. For this, I need to use a different tool, because Hadoop's get merge command can't transfer files from the server to my local Windows machine. Hence Used FileZilla tool to transfer the cleaned data files.

### **Experimentation with Models**

#### RNN\_Amazon

```
# Evaluate the model on the test set
| 8 | loss, accuracy = model.evaluate(X_test, y_test)
  print(f"Test Loss: {loss}")
  print(f"Test Accuracy: {accuracy}")
  Epoch 1/5
  Epoch 2/5
  Epoch 3/5
  230/230 [===========] - 21s 91ms/step - loss: 0.0926 - accuracy: 0.9735 - val_loss: 0.3165 - val_accuracy: 0.9078
  Epoch 4/5
  Test Loss: 0.3629966378211975
  Test Accuracy: 0.9026115536689758
```

### RNN\_Flipkart

```
✓ RAM
+ Code + Text
_{\rm n}^{\prime} [5] # Evaluate the model on the test set
 loss, accuracy = model.evaluate(X_test, y_test)
 print(f"Test Loss: {loss}")
 print(f"Test Accuracy: {accuracy}")
 Epoch 1/5
 Epoch 2/5
 Epoch 3/5
 Epoch 4/5
 Epoch 5/5
 Test Loss: 0.10982637107372284
 Test Accuracy: 0.9681565761566162
```

#### **CNN Amazon**

```
v
2m [4] print(f"Test Loss: {loss}")
    print(f"Test Accuracy: {accuracy}")
    [2 2 2 ... 2 2 2]
    Epoch 1/5
    230/230 [============ ] - 49s 164ms/step - loss: 0.5300 - accuracy: 0.7932 - val_loss: 0.3
    Epoch 2/5
    Epoch 3/5
    Epoch 4/5
    Epoch 5/5
    230/230 [============] - 5s 22ms/step - loss: 0.0495 - accuracy: 0.9874 - val_loss: 0.279
    115/115 [============] - 0s 3ms/step - loss: 0.2792 - accuracy: 0.9287
    Test Loss: 0.27921703457832336
    Test Accuracy: 0.928726851940155
```

### CNN\_Flipkart

```
✓ T4 RAM
Disk
+ Code + Text
# Evaluate the model on the test set
  loss, accuracy = model.evaluate(X_test, y_test)
  print(f"Test Loss: {loss}")
  print(f"Test Accuracy: {accuracy}")
  [2 2 2 ... 2 2 2]
  Epoch 1/5
  Epoch 2/5
  Epoch 3/5
  Epoch 4/5
  Epoch 5/5
  Test Loss: 0.10205527395009995
  Test Accuracy: 0.9722917079925537
```

### BERT\_Amazon

+ Code + Text

] print(classif	fication_repo	ort(y_test	, y_pred,	<pre>target_names=class_names))</pre>
	precision	recall	f1-score	support
positive	0.93	0.93	0.93	368
negative	0.96	0.95	0.96	249
neutral	0.98	0.98	0.98	1218
accuracy			0.97	1835
macro avg	0.96	0.96	0.96	1835
weighted avg	0.97	0.97	0.97	1835

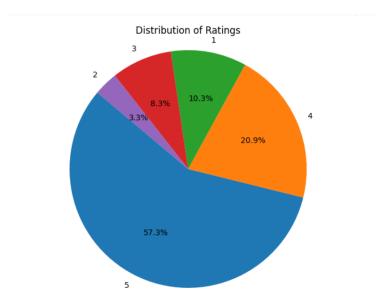
### BERT\_Flipkart

[49] print(classification\_report(y\_test, y\_pred, target\_names=class\_names)) precision recall f1-score support positive 0.94 0.95 0.94 280 negative 0.80 0.80 0.81 48 neutral 0.99 0.99 0.99 1672 0.98 2000 accuracy macro avg 0.91 0.92 0.91 2000 weighted avg 0.98 0.98 0.98 2000

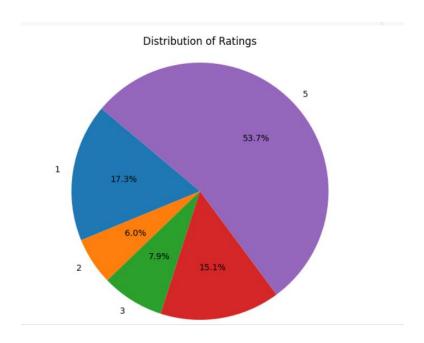
# **Visualizations:**

## 1) UNIVARIATE ANALYSIS:

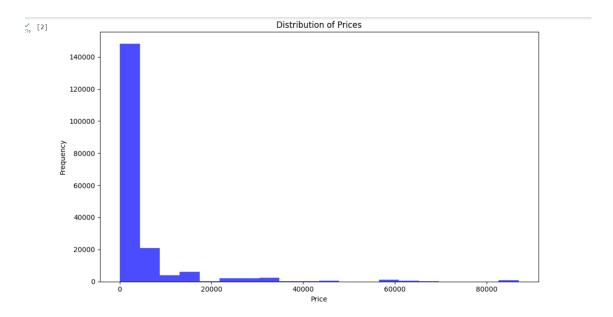
**Rating Distribution: Flipkart** 



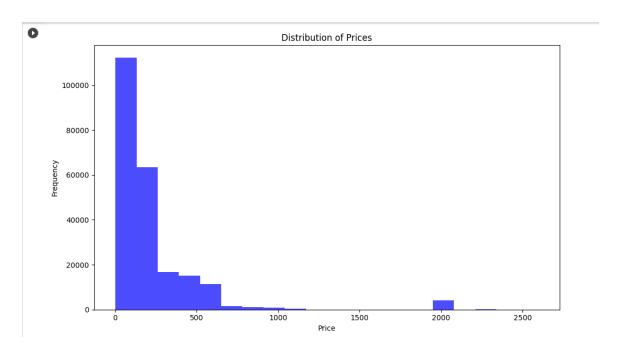
## **Rating Distribution: Amazon**



## **Price Distribution: Flipkart**

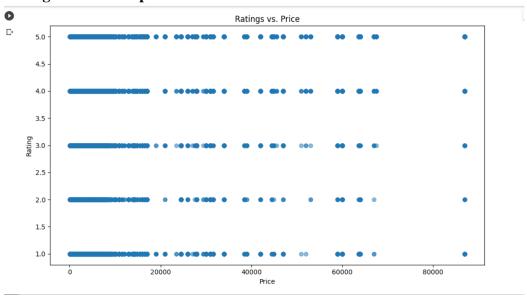


### **Price Distribution: Amazon**

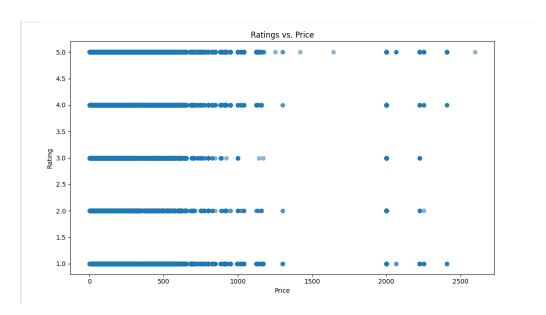


## 2) BIVARIATE ANALYSIS:

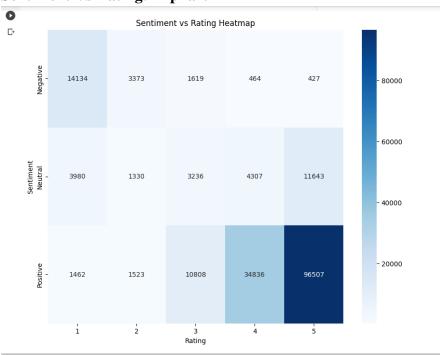
## **Rating Vs Price: Flipkart**



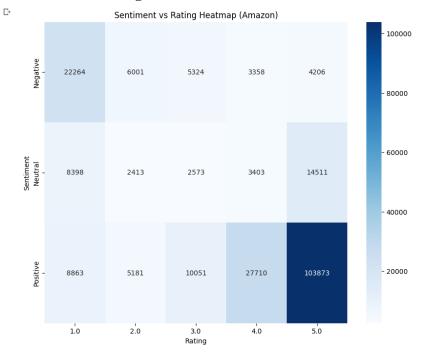
## **Rating Vs Price: Amazon**



## **Sentiment Vs Rating: Flipkart**



## Sentiment Vs Rating: Amazon

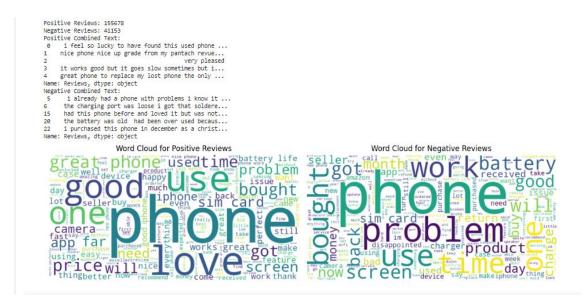


#### 3) TEXT ANALYSIS:

#### **Word Cloud: Flipkart**



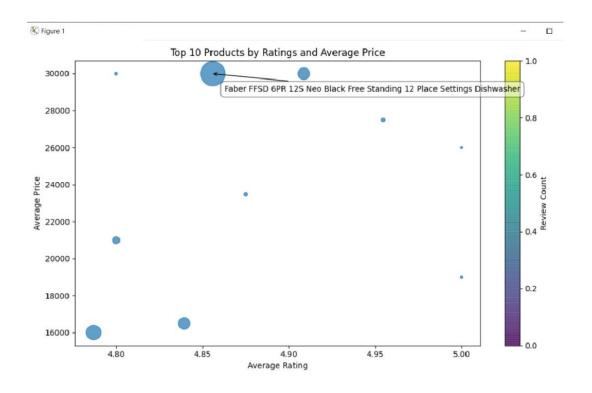
#### **Word Cloud: Amazon**

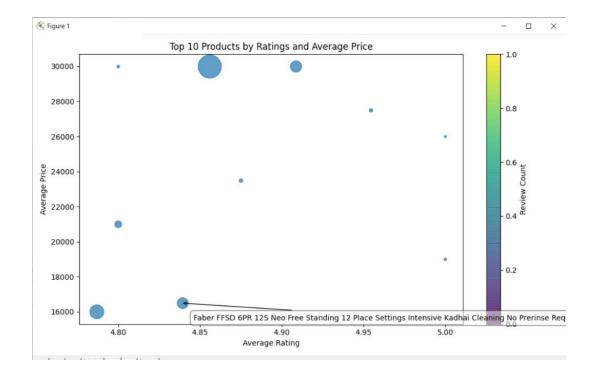


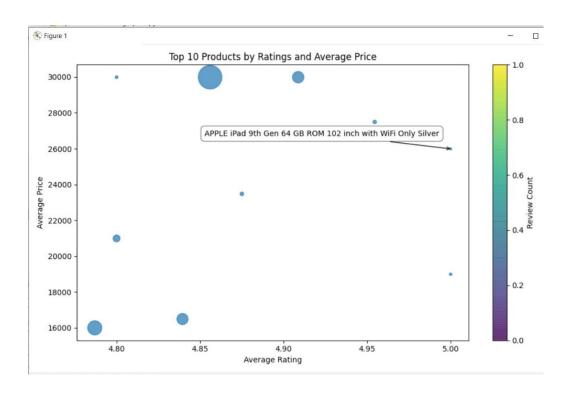
## 4) ADDITIONAL CHARTS

Top 10 Products based on the top ratings (4 and 5) and average pricing category: Flipkart.

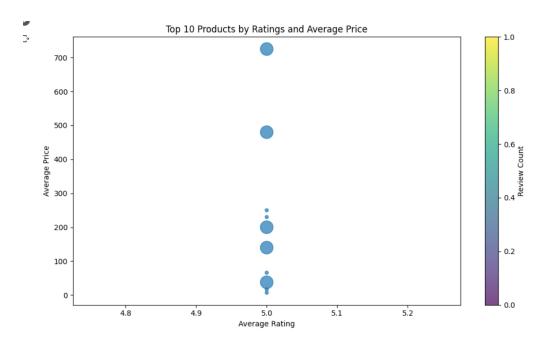


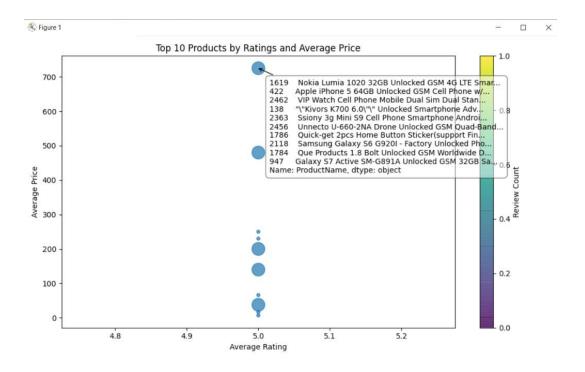


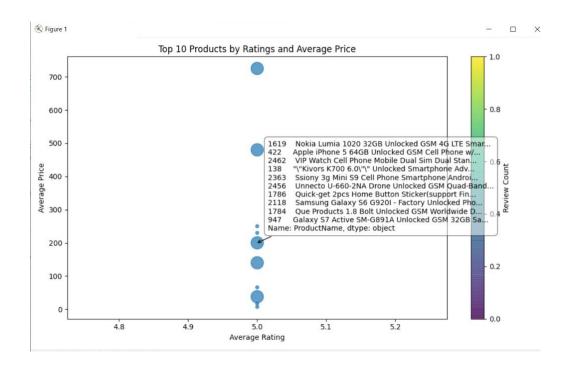


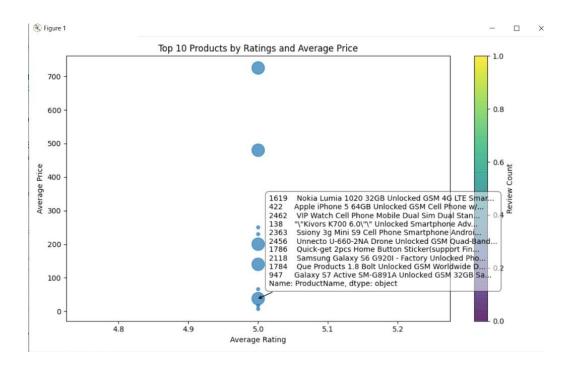


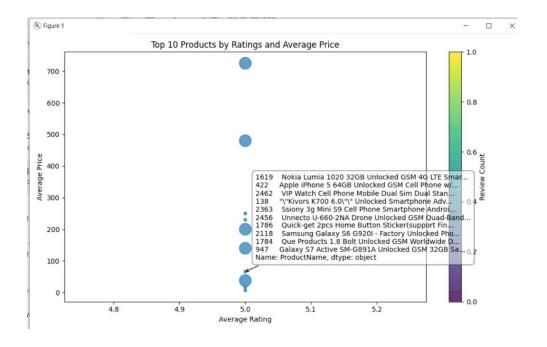
Top 10 Products based on the top ratings (4 and 5) and average pricing category: Amazon.











## 7. DISCUSSION

### 7.1.1 LEARNINGS & FUTURE WORK

In this project, BERT emerged as the most accurate model, underscoring its capability in handling the intricacies of text-based sentiment analysis. The insights gleaned from the sentiment analysis proved invaluable for informing pricing strategies and pinpointing areas for product improvement. By utilizing data from multiple e-commerce platforms, the analysis achieved a well-rounded perspective on customer sentiment. This research also marked a shift from traditional intuition-based decision-making to a more empirical, data-driven approach. For future endeavors, the project sets the stage for exploring real-time sentiment tracking and aspect-based sentiment analysis to gain even deeper consumer insights.

#### 7.1.2 CONCLUSION

This project confirmed that sentiment analysis is a vital tool for evaluating product performance on e-commerce platforms. High accuracy achieved via the BERT model emphasizes its utility in capturing consumer sentiment. The insights gathered are pivotal for both current assessment and future strategic decisions. The method proves beneficial in areas from pricing strategy to product enhancement. The approach paves the way for a more data-driven, consumer-centric business model.

# 8 GITHUB REPOSITORY / CODE FILES LINK

The code for this project is available on GitHub.

## 8.1 MY REPOSITORY LINK:

https://github.com/BislaSonal/Sentiment-Analysis-of-Products-Review

The repository contains a README file with detailed instructions on how to run the code and reproduce the results.

#### 8.2 RESEARCH PAPER REFERENCES

#### **Amazon Dataset Category**

https://arxiv.org/abs/1904.04096

https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3886135

https://core.ac.uk/works/8302737

https://www.researchgate.net/publication/344478746\_HOW\_HELPFUL\_ARE\_COMPARATIVE\_REVIEWS\_FOR\_PREDICTING\_PRODUCT\_DEMAND

https://www.researchgate.net/publication/365491589 Model Comparison in Sentiment Analysis A Case Study of Amazon Product Review s

#### **Twitter Dataset Category**

https://www.researchgate.net/publication/340976678 Big\_data\_approach\_for\_sentiment\_analysis\_of\_twitter\_data\_using\_Hadoop\_framework\_and\_deep\_learning

https://www.researchgate.net/publication/320625064\_PROJECT\_REPORT\_SENTIMENT\_ANALYSIS\_ON\_TWITTER\_USING\_APACHE\_SPAR\_K

### **General E-commerce Platforms Category**

https://www.researchgate.net/publication/360819202 Sentiment Analysis on Reviews of E-commerce Sites Using Machine Learning Algorithms

https://www.researchgate.net/publication/345681485\_Sentiment\_analysis\_of\_online\_product\_reviews\_based\_on\_SenBERT-CNN

https://www.researchgate.net/publication/366161045 Sentiment Analysis of Air Conditioning Users%27 Comments on e-Commerce Platform Based on NLP Technology

https://www.researchgate.net/publication/339513566\_Sentiment\_Analysis\_in\_E-Commerce\_A\_Review\_on\_The\_Techniques\_and\_Algorithms

#### **Review Helpfulness and Market Value Category**

https://www.researchgate.net/publication/277247969 Extracting reliability of the products through user reviews

https://europepmc.org/article/PMC/PMC9051848

 $\frac{\text{https://core.ac.uk/works/127220297?source=1\&algorithmId=15\&similarToDoc=17826331\&similarToDocKey=CORE\&recSetID=1daf429b-fb31-4b93-863a-d33aca5f77bc\&position=4\&recommendation_type=same_repo\&otherRecs=8302737,17809015.8490770,127220297,19014509}{\frac{1}{2}}$ 

https://www.researchgate.net/publication/267846365 Impact of Positive vs Negative e-Sentiment on Daily Market Value of High-Tech Products

#### **Data Analytics & Visualization Category**

https://www.semanticscholar.org/paper/Visualization-System-For-Sentiment-Analysis-Using-Gormantara/83a971cbea2e9699e5470733386fe6dfa84c2cab

 $\underline{https://www.semanticscholar.org/paper/Data-Filtering-and-Visualization-for-Sentiment-of-Keerthana-titlering-of-Keerth$ 

Prasannakumar/9fa1e9b3ae3481fd2bc79f6ba95a1e51d68c47dd

https://journals.sagepub.com/doi/pdf/10.1177/14738716221126992

https://iopscience.iop.org/article/10.1088/1757-899X/1099/1/012002

https://www.researchgate.net/publication/370593444\_An\_Introduction\_to\_Data\_Visualization\_Tools\_and\_Techniques\_in\_Various\_Domains

https://www.semanticscholar.org/paper/The-State-of-the-Art-in-Sentiment-Visualization-Kucher-

Paradis/2af1587a1f2b0b03e1a17d6a17cc4ef7b1d07c95

https://www.clausiuspress.com/article/5673.html

https://www.semanticscholar.org/paper/Information-Visualization%3A-Perception-for-Design-Ware/b989b192be547648594e5d7f8474f6da99ca4840

https://www.researchgate.net/profile/K-M-Zubair/publication/366020486\_E-

commerce Data Analysis and Visualization using Random Forest Regression and Prophet model/links/638e4f7e095a6a777406cbb7/E-commerce-Data-Analysis-and-Visualization-using-Random-Forest-Regression-and-Prophet-model.pdf

https://www.cs.ubc.ca/~tmm/courses/547-22/projects/yaman/report.pdf

https://www.researchgate.net/profile/Gajanan-

Bhole/publication/362517195 Survey on Sentimental Analysis and Visualization of Reviews/links/62ededc70b37cc344774b211/Survey-on-Sentimental-Analysis-and-Visualization-of-Reviews.pdf

https://arxiv.org/abs/1911.03655

https://www.researchgate.net/profile/Hari-

Rawat/publication/348118082\_A\_Study\_of\_Big\_Data\_Processing\_for\_Sentiments\_Analysis/links/614980233c6cb310697fe0db/A-Study-of-Big\_Data-Processing-for-Sentiments-Analysis.pdf

#### **Business Analytics and Innovation Category**

 $\underline{https://jisrc.szabist.edu.pk/ojs/index.php/jisrc/article/view/10/8}$ 

https://ink.library.smu.edu.sg/sis\_research/1752/