Sentiment analysis on Trip advisor hotel Reviews

Introduction

The project is mainly on sentiment analysis on hotel reviews for trip in the trip advisor is platform in which where we can get the perfect reviews and ratings based on that our project depends on feedbacks of customers and by that we can do our analysis on that to get insights from it

About Dataset:

The data set contains 20491 rows in it and there two columns which decides the data and sentiment analysis of it we have took this dataset from kaggle website

**Review:**

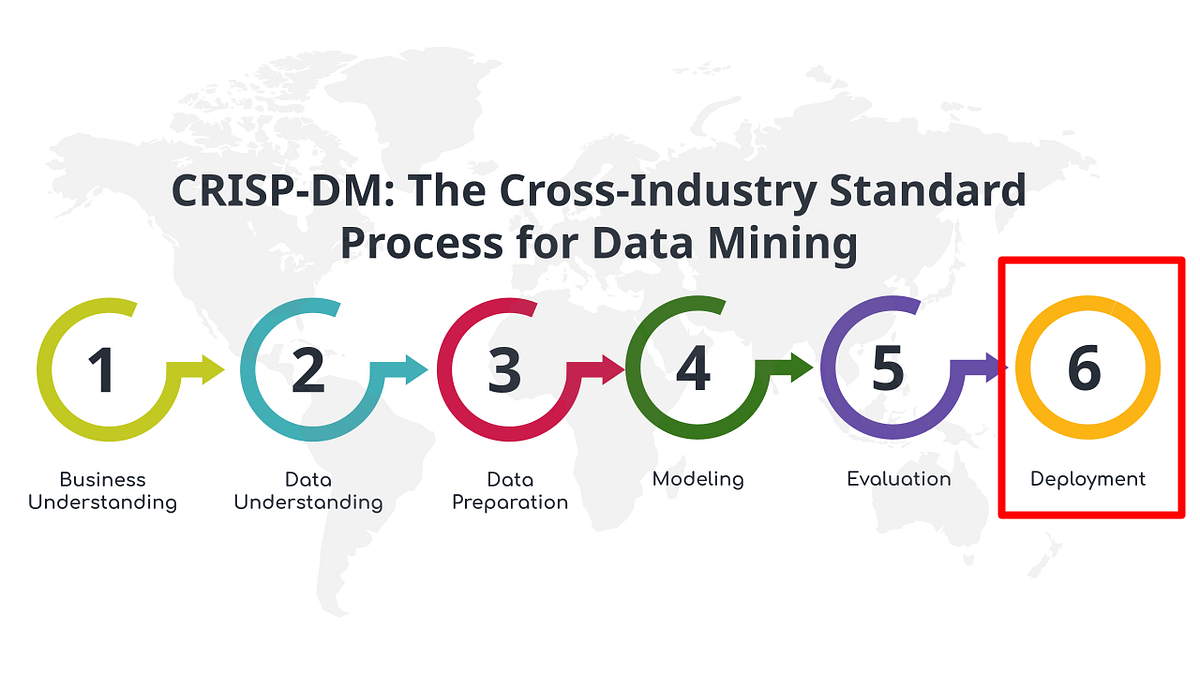
The review columns will help to know the comments and feedbacks from customers

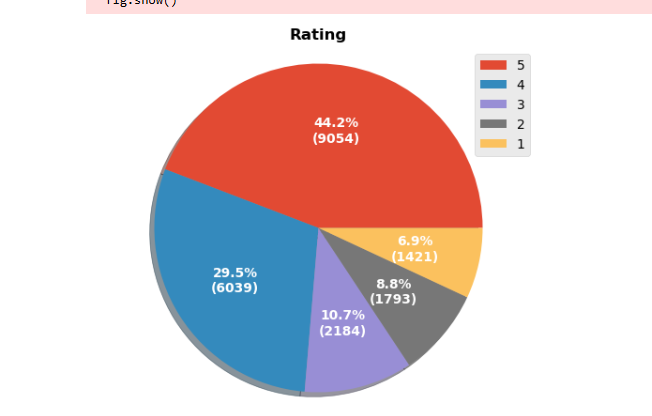
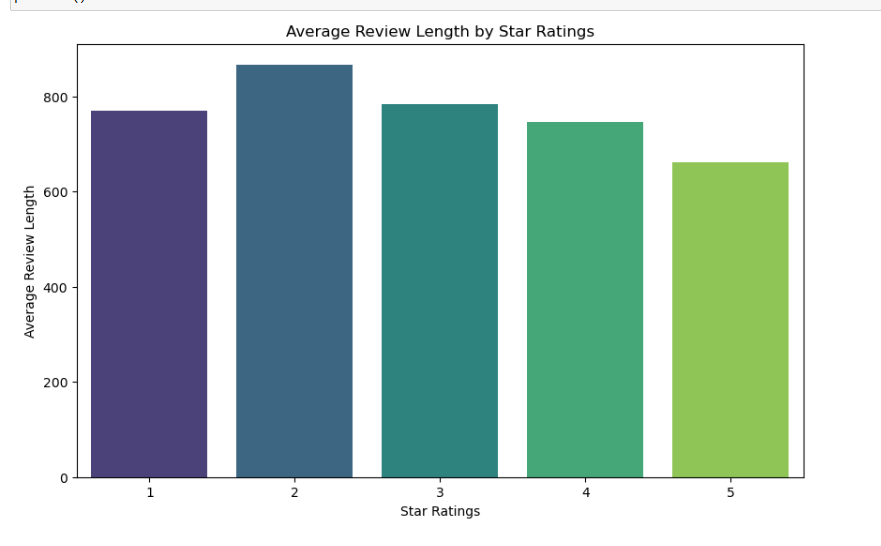
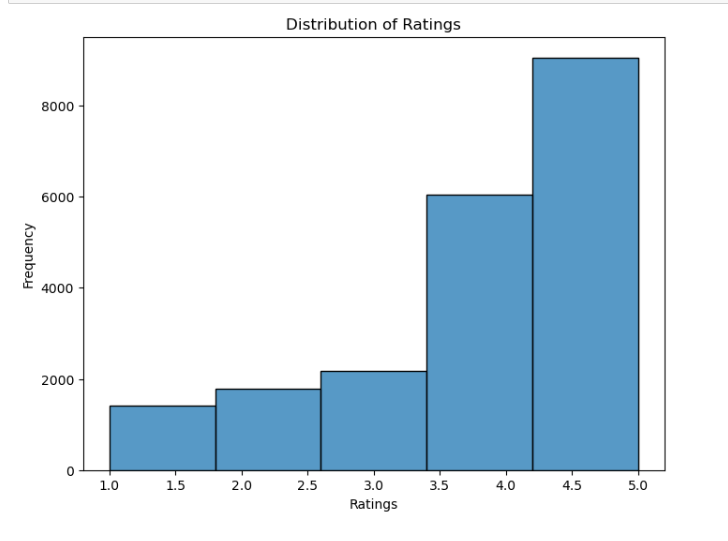
Ratings:

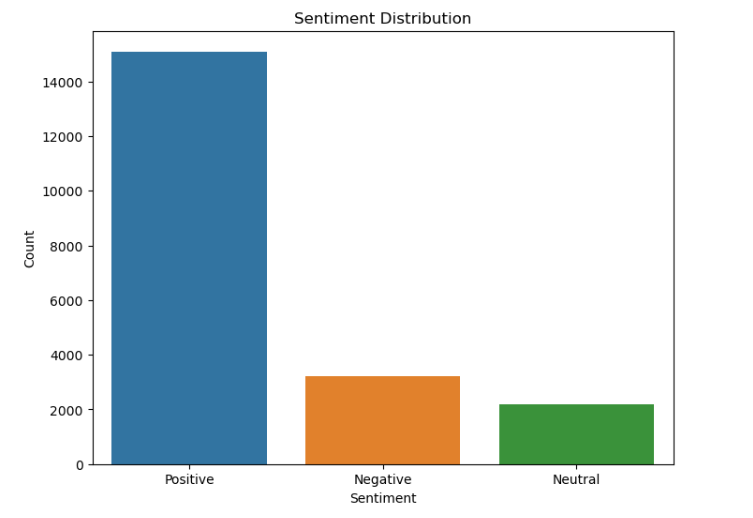
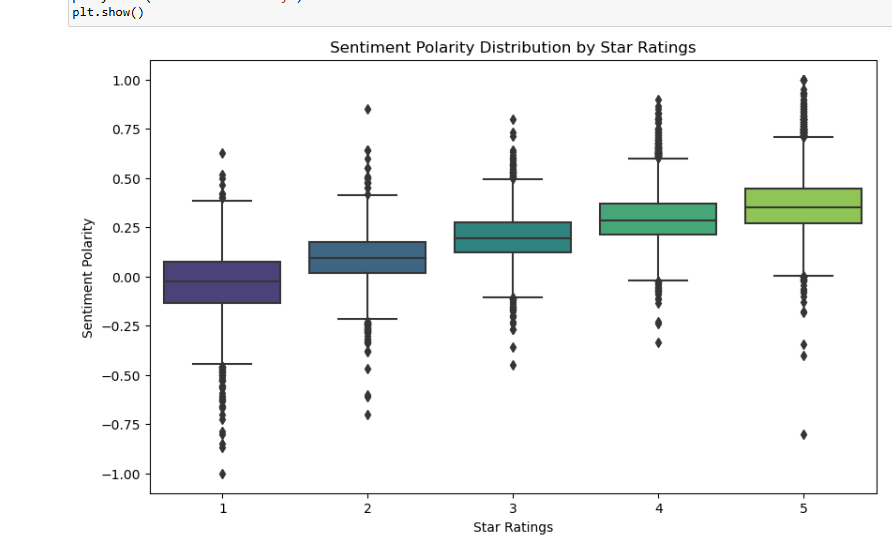
The ratings column will helps to know the ratings which are given by customers based on the hotel treating way and other reasons

The ratings are from 1 to 5 ratings are there and there are no null values in this dataset

**CRISP**







Pre processing

## Text Pre-Processing

The reviews texted is cleaned to remove the missing sentence and unwanted words in it and to get cleaned data

Tokenization

The tokenization is used to clean the data and break the words and make the meaning full

Stop words

* Commonly used words (stop words) that do not contribute to the sentiment are removed.

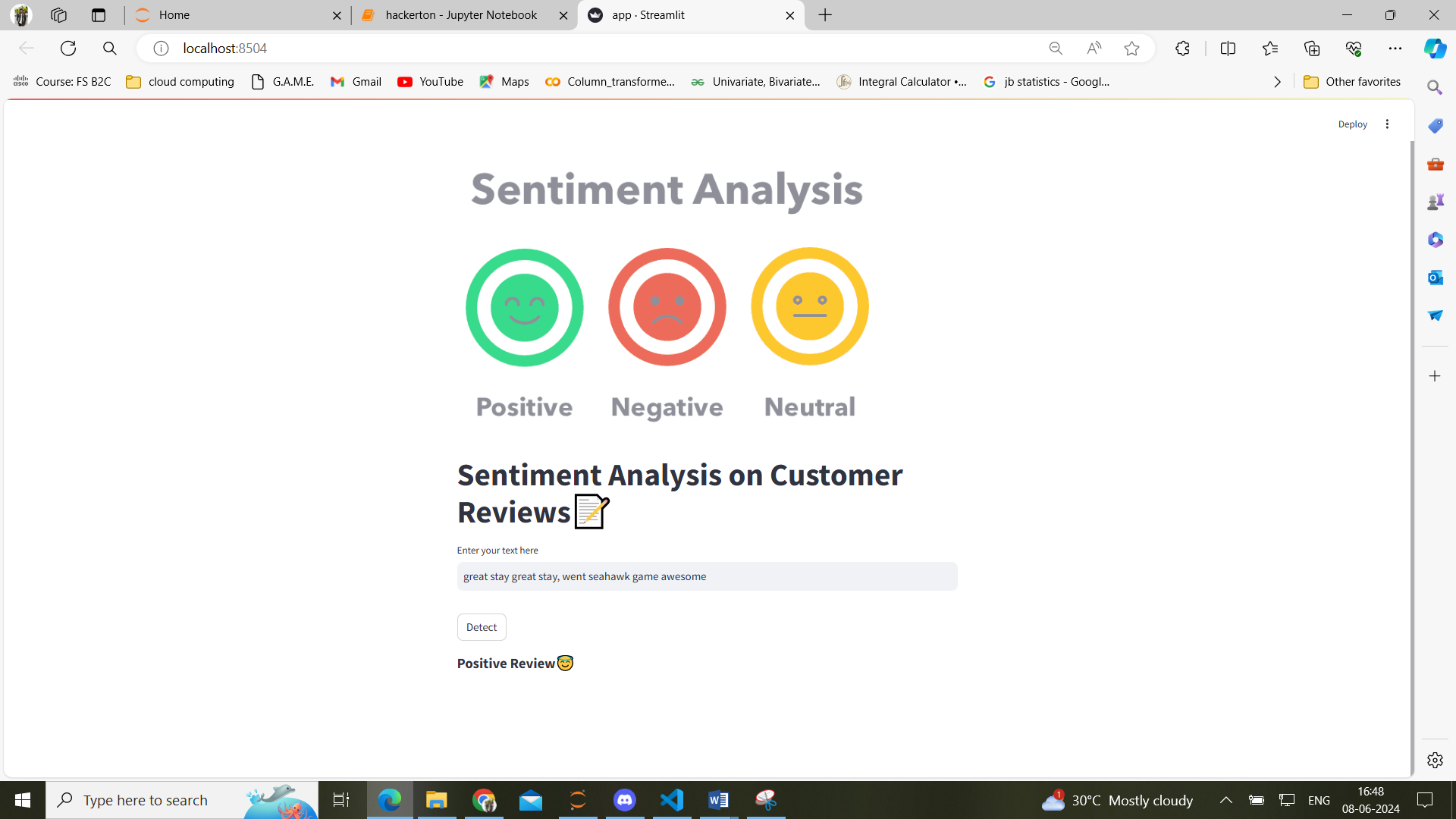
**Lemmatization:** Words are lemmatized to their base or root form.

**Vectorization:** The text data is transformed into numerical features using the Count Victories method, which converts the text into a matrix of token counts

Model training and evaluation

* The dataset is split into training and testing sets using a 75-25 split ratio to train the model on a portion of the data and evaluate its performance on unseen data.
* The models are trained and their hyper parameters are optimized using GridSearchCV and build models with Multinomial Naive Bayes, Logistic Regression, and XGBoost algorithms, which are commonly used techniques for text classification tasks.

DEPLOYMENT

The sentiment analysis model achieved an accuracy of 80% on the testing set, demonstrating its effectiveness in predicting sentiment from hotel reviews. Accuracy score metrics were computed for each sentiment class (positive, negative, neutral), providing insights into the model's performance across different sentiment categories

Applications

1. [Social media monitoring](https://monkeylearn.com/blog/sentiment-analysis-applications/#social-media)
2. [Customer support ticket analysis](https://monkeylearn.com/blog/sentiment-analysis-applications/#customer-support)
3. [Brand monitoring and reputation management](https://monkeylearn.com/blog/sentiment-analysis-applications/#brand-monitoring)
4. [Listen to voice of the customer (VoC)](https://monkeylearn.com/blog/sentiment-analysis-applications/#voc)
5. [Listen to voice of the employee](https://monkeylearn.com/blog/sentiment-analysis-applications/#voice-of-employee)
6. [Product analysis](https://monkeylearn.com/blog/sentiment-analysis-applications/#product-analysis)
7. [Market research and competitive research](https://monkeylearn.com/blog/sentiment-analysis-applications/#market)

**Conclusion:**

In conclusion, sentiment analysis of Trip Advisor hotel reviews offers valuable insights into customer sentiments and perceptions. By harnessing machine learning techniques, hotels can optimize their service offerings, enhance customer experiences, and ultimately drive business success.

