



2 Questions

Total Marks: 110.0

1 SQL Question

1. Retail stores

+ 10.0

1 Machine Learning Question

2. Identify key aspects of a Review

+ 100.0

Question 2

Max. Marks 100.00 ?

Identify key aspects of a Review

E-commerce has revolutionized retail experience through multiple insights that it provides customers and retailers. Customer' reviews are considered as feedback by retailers about their product and services. These feedback allow them to understand the industry better and enhance the efficiency and quality of their product.

A thorough analysis of a review can provide insights to retailers and E-commerce platforms about improvements that are required for better customer satisfaction. Analysis of such million reviews are handled by the E-commerce team manually and it is a very time-consuming task. This impacts the business of retailers as they get delayed feedback.

Your task is to create a **simplified** Machine Learning and/or Deep Learning approach to analyze the feedback & identify the key problems that are highlighted. You are also required to curate creative visualization reports that can be shared with retailers to aid the understanding of potential improvement areas for sellers and the E-commerce team.

Task: Your tasks are as follows:

1. Predict the 'topic' label for reviews under the **test.csv** file
2. Share your documented analysis of data, along with the source code, in **Jupyter** notebook file (.ipynb)

Data description

Columns	Description
Review Text	Full description of reviews
Review Title	Short description of reviews
Topic	Problem type

Data volume:

- **Train.csv:** 5960 x 3
- **Test.csv:** 2553 x 2
- **Submission.csv:** 2553 x 3 (Refer 'Sample Submission.csv' [here](#))

Submission guidelines: You are required to upload:

1. The **submission.csv** file that contains your prediction labels such as **Upload Prediction File**
2. A **.zip** file that contains your **.ipynb** or **Jupyter** notebook file under the **Upload Source Files** labels
 - The notebook file must contain a well-documented analysis of data, logic, and reasons

Evaluation metric:

Machine Learning models: The evaluations will be based on the **top-n** accuracy metrics.

The top-n accuracy in this case is calculated by checking if the expected 'topic' is present in the predicted 'topics'. For each unique 'review text', 'review title', *n* is the expected number of topics per unique review.

To know more about this evaluation metric, click [here](#).

$$score = (Top - n \text{ accuracy per review}) / n$$

Data analysis:

- Algorithms, approaches, and logic used during the selection of models
- Creative in-depth researches (visualizations and reports) that are targeted at the following attributes:
 - Key aspects of product and services that customers require
 - Users' underlying intentions and reactions concerning those aspects

[Download dataset](#)[New Submission](#)[All Submissions](#)**Upload File**

3-amazon-reviews-ovr-imb.csv

[Change File](#)

Submission ID: 29341786

7 seconds ago


Result

✔ Accepted

Score











54.44820

Upload Source Code

 Upload File



Your Answer

B *I* |     |     |  

Write your answer here

Preview

Submit

