Recommendation Systems Project

Data Description:

The dataset contains information about ratings from customers from about 130 restaurants. The rating information is spread across 3 columns - an overall rating, food rating and service rating.

Source (inside the RCdata folder)

Domain:

Restaurants and Foodservice industry

Context:

This dataset was used for a study where the task was to generate a top-n list of restaurants according to the consumer preferences.

Attribute Information:

- userID: unique identifier for a user
- placeID: unique place identifier
- rating: restaurant rating
- food_rating: food rating
- service_rating: service rating

Learning Outcomes:

- Popularity based model
- Collaborative filtering
- Singular vector decomposition

Objective:

Generate a top-n list of restaurants on consumer preference - Restaurant & Consumer data

Steps and tasks:

1. Import the necessary libraries

2. Read the data as a dataframe

3. Perform basic EDA which should answer the following (5 marks)
    a. What is the shape of the data?
    b. What are the data types of each attribute?
    c. Are there any missing values in the data?
    d. On what scale are the ratings recorded?
    e. What is the distribution of the 'rating' column
    f. Find the number of unique users and places from the data

4. Take a subset of the dataset to make it less sparse/ denser. For example, keep the users only who has given 12 or more number of ratings (5 Marks)

5. Build Popularity Recommender model. (10 marks)

6. Split the data randomly into a train and test dataset (2 marks)

7. Build Collaborative Filtering model (15 marks)

8. Evaluate the above model. (Once the model is trained on the training data, it can be used to compute the error (like RMSE) on predictions made

on the test data.) You can also use a different method to evaluate the models (10 marks)

9. Get top - K (K = 5) recommendations. Since our goal is to recommend new places to each user based on his/her taste, we will recommend 5 new places (10 marks)

10. Summarise your insights (3 marks)

Note:
If you are stuck at the model evaluation part of this project.

Please refer to below links -

1. https://surprise.readthedocs.io/en/stable/accuracy.html
2. http://surpriselib.com/ - Getting started, example