ML TEAM - WEEK 3

Hey guys, Welcome to the TechOdessey ML event conducted by team DCS. Here you will be provided with a series of tasks you must complete within the given time. Ready to dip your toes in ml? Let's get started !!

This week's main aim is to get started with a Movie Recommendation System.

Here are your tasks:

1.Write short notes(max 3 lines in your own words) regarding the various types of recommender systems[This is to make sure you guys get familiar with the gerund that we will be using moving forward]

2. IMDB's weighted rating :

$$\text{Weighted Rating (WR)} = \left(\frac{v}{v+m} . R\right) + \left(\frac{m}{v+m} . C\right)$$

v is the number of votes for the movie
m is the minimum number of votes required to be listed in the chart
R is the average rating of the film
C is the mean vote across the whole report

We already have *v* and R present in the data
2.1 Calculate C
[Hint: C is the mean of the vote_average column]
2.2. Calculate *m* *(the minimum votes required to be listed in the chart)*

TECH ODYSSEY

[Hint: You can calculate this with the help of the quantile() function in python.

Essentially, you will be computing the 90th percentile cutoff,

That is: for a movie to feature in the charts, it must have more votes than at least 90% of the movies in the list.]

2.3 Find movies that have a higher vote-count than 'm'

[Hint: Use the loc function on the vote count column]

2.4 Create a python function to calculate the weighted rating using the formula and store that in a new column .

[Hint: You can make use of the apply function in pandas to effectively construct the function]

2.5 Sort the dataset using the newly computed Score column and display the top 20 movies with the highest score

[Hint: Use the sort_values() function to sort and use the head() function to get the top 20 movies]

3. HITS OF ALL TIME recommender system

3.1 Sort the dataset with the help of the popularity column and display the top 10 most popular movies with the help of a bar plot

[Hint: Use the sort_values function and use the plt.barh() function to print out the bar plot]

THANK YOU!