

# IT492: Recommendation Systems

## Lab Assignment - 04

[Deadline: 24th April 2022, Sunday 11:59 PM]

This assignment involves generating explanations for recommendations and a conversational recommendation system.

### Marking scheme and requirements

Full marks will be given for (1) working, readable, reasonably efficient, documented code that achieves the assignment goals and (2) for providing appropriate answers to the questions in your Google colab file (name format: LA04\_rollnumber\_name) submitted via Google Classroom on the **assigned dataset only**. Please find the dataset allotted to you below.

### Please refer to the Dataset Allocation

Sr No	Student Id	Student Name	Student email	Registration Type	For LA-04
1	202118004	ABHISHEK SINGH	202118004@daiict.ac.in	AUDIT	Movielens 20M (Grouplens 2016)
2	202018004	P SARAN PANDIAN	202018004@daiict.ac.in	AUDIT	
3	202018026	AAKANKSHA SHAH	202018026@daiict.ac.in	AUDIT	
4	202111002	SHARMA HARSH DHARMENDRAKUMAR	202111002@daiict.ac.in	AUDIT	
5	202111029	GORASIYA RAGHAV NARESH	202111029@daiict.ac.in	AUDIT	
6	202018042	ABHIJEET KUMAR	202018042@daiict.ac.in	REGULARADD	Food Reviews (Kaggle 2019)
7	202111010	KEVIN JITENDRABHAI JADIYA	202111010@daiict.ac.in	REGULARADD	
8	202111035	VANSH RAHUL BHANJIBHAI	202111035@daiict.ac.in	REGULARADD	
9	202111048	MANSURI PINJARA MOHAMMED JUNED HANIFBHAI	202111048@daiict.ac.in	REGULARADD	
10	202112030	ARPITHA SREENIVASAN	202112030@daiict.ac.in	REGULARADD	
11	201801466	PARMAR SIDDHRAJ YOGESHBHAI	201801466@daiict.ac.in	REGULARADD	
12	202121004	SANDHYA KUMARI	202121004@daiict.ac.in	REGULARADD	last.FM (Hetrec 2011)
13	202116003	AMBUJ MISHRA	202116003@daiict.ac.in	REGULAR	
14	202116004	ARPITA NEMA	202116004@daiict.ac.in	REGULAR	
15	202116008	RAHUL KUMAR	202116008@daiict.ac.in	REGULAR	
16	202116009	RAHUL THAKUR	202116009@daiict.ac.in	REGULAR	
17	202116011	ROHAN BAGHEL	202116011@daiict.ac.in	REGULAR	
18	202116001	ABHISHEK YADAV	202116001@daiict.ac.in	REGULAR	
19	202116002	AKSHAY KAUSHIK	202116002@daiict.ac.in	REGULAR	

### Links to download the Datasets

- *last.FM (Hetrec 2011)*
- *Movielens 20M (Grouplens 2016)*
- *Food Reviews (Kaggle 2019)*

# IT492: Recommendation Systems

## Please adhere to the lab policy on the course website

- Cite resources and give credit where it's due. If you happen to discuss the questions with your peers, please mention your collaborators in your report/assignments.
- Acts of plagiarism will not be tolerated and will result in a straight ZERO for that assignment.
- Students who don't submit their assignment by 26th April 2022, Tuesday 11:59 PM will simply get ZERO.

## Main Assignment (15 Marks in Total)

### Dataset Analysis (1 Mark)

1. Explore the dataset and present a summary of the same (e.g. number of users and items, minimum/maximum/average number of ratings for an item, information apart from ratings present in the dataset). Plot a histogram of item ratings to visualize the distribution.

### Explaining Recommendations (3+5 = 8 Marks)

2. In a user-user collaborative filtering setting, generate recommendations using the k-Nearest Neighbor model. Generate user-based explanations for five users by plotting the histograms of the active user's neighbors' ratings for the items in their profile. Also, find out that can we explain such recommendations using items? If yes, then how? If no, then why not?
3. Use the Naive Bayes (or Logistic Regression) model for generating content-based recommendation. For any five users, by using tag-preference and tag-relevance formulation, generate top-3 tags (i.e. features) that can be shown to the user as an explanation (Useful resources: [[Tagsplanation by Vig et al.](#)]). Also, can you define your own formulations for ranking tags (based on their affinity with the item to which they are assigned and the user who have assigned them)?

### Conversational Recommendation (6 Marks)

4. For a user,
  - a. Choose an item from her test-set, we call it *target*.
  - b. Find the top-5 neighbours of the *target*.
  - c. Consider the 5th neighbor (the farthest one) as a *seed* (i.e. a start point), recommend top-3 items to the user using k-NN based recommendation algorithm.

## IT492: Recommendation Systems

- d.* Now, out of the 3 recommendations, select the one which is closest to the 1st neighbor (the nearest one), and treat it as a *seed* for the next round of recommendations.
- e.* Repeat the steps *c* and *d* for 5 rounds of recommendations or until you find the *target* as one of the top-3 recommendations in a round, whichever comes earlier.

Repeat steps *a-to-e* for 100 users and find out the overall *hit-rate*. Also, show the *hit-rate* histograms over the 5 rounds of recommendations.