Quiz 4

Online (In-class) Quiz –

1) [1 point] Select ALL of the statements that are TRUE about Inverse User Frequency.

A. Universally liked items are as useful in capturing similarity as less common items.

B. Inverse frequency fj = log(nj/n), nj is number of users who have rated item j, n is total number of users.

C. If everyone has rated item j, then fj is zero.

D. When transform ratings by multiplying the original rating by fj, less popular items will have greater effect on prediction

Ans.) C, D

2) [1 point] Which of the following Hybrid Recommendation Systems involves the concept of breaking ties using one of its components?

A. Feature Augmentation
B. Cascade Hybrid
C. Meta-Level Hybrid
D. Switching Hybrid

Answer. B

3) [3 Points] Consider the following table where rows are the Users and columns are the Items. The values corresponding to ratings.

	I1	**I2**	**I3**	**I4**
U1	2	1		3
U2	3	**?**	5	2
U3		4	2	3
U4	5	3	1	

Find the recommendation rating on I2 for U2(shown by **?**) using User-Based CF. You need to use Pearson Correlation for your computations.

W(1,2) = -1
Numerator = (2-2.5) (3-2.5) + (3-2.5)(2-2.5)
Denominator = ((2-2.5)^2+(3-2.5)^2)^0.5 * ((3-2.5)^2 + (2-2.5)^2)^0.5

W(2,3) = -1
Numerator = (5-3.5) (2-2.5) + (2-3.5)(3-2.5)
Denominator = ((5-3.5)^2+(2-3.5)^2)^0.5 * ((2-2.5)^2 + (3-2.5)^2)^0.5

W(2,4) = -1
Numerator = (3-4) (5-3) + (5-4)(1-3)

Denominator = ((4-3)^2+(2-3)^2)^0.5 * ((3-2)^2 + (1-2)^2)^0.5

P(2,2) = 3.33 + ((1-2.5)*-1 + (4-2.5)*-1 + (3-3)*-1) / (1 + 1 + 1)
= 3.33 + 0
= 3.33

4) [0.5 point] Jaccard Similarity works better than Pearson Correlation for User based Collaborative Filtering System.

False

5) (1.5 Points) Briefly describe the Cold Start Problem in recommender systems?

When a new user or item has just entered the system, it is hard to find similarities as there is not enough information to make good recommendations.
New item problem: can't be recommended until some users rate it. Also called "first-rater problem"
New users: not given good recommendations because of lack of rating or purchase history

6) (2 pts) What is the long tail effect and how do you address this problem in recommendation systems (1pt) ? Give an example (1pt).

Answer - Long tail effect: A long tail of some distributions is the portion of the distribution having a large number of occurrences far from the "head" or central part of the distribution. (The Long Tail is composed of a small number of popular items, the well-known hits, and the rest are located in the heavy tail, those that do not sell that well.)

Address: User-based/Item-based. Recommend according to user's interests, even if it's in long tail (not popular).
Example: Amazon generates 57% of its sales from long-tail searches.
Rubrics:
Explanation - 0.5 point
Address the problem - 0.5 point
Example - 1 point

7) (0.5 Point) Suppose we have three recommender systems - System A, System B and System C. System A has a RMSE value of 0.5, System B has a RMSE value of 0.8 and System C has a RMSE value of 0.2. Which recommender system is better at making recommendations?

a) System A
b) System B
c) System C
d) All are equally good

Ans (c)

8) (0.5 Point) Does the Pearson Correlation measure the extent to which two variables linearly relate? (True or False) -
True

Offline (Take-home) Quiz –

1) [1 points]What is a possible disadvantage of using Pearson Correlation for finding missing ratings? Which extension to Memory-based algorithm can be applied to overcome this?

> Pearson Correlation considers only co-rated items. This may sometimes neglect the global behavior reflected in a user's entire rating history.
>
> Using Default voting, we can limit the weight such pairs contribute while computing the final prediction.

3) (1 pts) Briefly explain the difference between feature augmentation and meta-level hybrid using examples.

Answer-

Feature Augmentation: Generate new feature by a contributing recommender as augment profile. Pass the augment profile to another recommender

Feature augmentation hybrid generates a new feature for each item by using the recommendation logic of the contributing domain

Ø E.g. use association rule mining over the collaborative data to derive new content features for content-based recommendation

! At each step, the contributing recommender intercepts the data headed for the actual recommender and augments it with its own contribution

Meta-Level: Use a model learned by one recommender as input for another. For example, a restaurant recommender used the naive Bayes technique to build models of user preferences.

Difference: in meta-level hybrid, contributing recommender completely replaces the original knowledge source with a learned model that the actual recommender uses

5) (2 pts) Briefly explain main idea and Plan of Action for content-based related video recommendations.

Main idea:

Recommend videos to customer x that are similar to previous videos rated highly by x. Requires characterizing the content of videos in some way

Plan of Action:

- Construct item profiles: Explicit features in a database, discovering features in documents, Tags. Create vectors representing items
- Construct user profiles: Create vectors with same components that describe user's preferences
- Recommend items to users based on content: Calculate cosine distance between item and user vectors

6) [4 points] Consider the following table with the rows as Users(A,B,C) and columns as Movies(1,2,3). Values represent movie ratings:

	1	2	3
A	2	3	4
B	1	5	3
C	3	2	5

1. Calculate the Cosine Similarity Score between user A, user B and user C for the Features of movie rating [1 points].
2. Calculate the normalized ratings [1 point].
3. What is the new centralized Cosine Similarity Score between A and B in scaled ratings [1 point]?
4. What is the relationship between Pearson correlation and cosine similarity? In other words,under what circumstances are these two measures equivalent? Briefly explain. [1 point]

a) Cosine Similarity

$$\text{Cos(A,B)} = \frac{2*1 + 3*5 + 4*3}{\sqrt{2^2+3^2+4^2}\sqrt{1^2+5^2+3^2}} = \frac{29}{\sqrt{29}\sqrt{35}}$$

$$\text{Cos(A,C)} = \frac{2*3 + 3*2 + 4*5}{\sqrt{2^2+3^2+4^2}\sqrt{3^2+2^2+5^2}} = \frac{32}{\sqrt{29}\sqrt{38}}$$

$$\text{Cos(B,C)} = \frac{1*3 + 5*2 + 3*5}{\sqrt{1^2+5^2+3^2}\sqrt{3^2+2^2+5^2}} = \frac{28}{\sqrt{35}\sqrt{38}}$$

b) Normalized Ratings

User/Item	**Average Rating**	**1**	**2**	**3**
A	*3*	*-1*	*0*	*1*
B	*3*	*-2*	*2*	*0*
C	*10/3*	*-1/3*	*-4/3*	*5/3*

c) Centralized Cosine Similarity:

Cos(A,B) = $\frac{-1*-2 + 0*2 + 1*0}{\sqrt{-1^2+0^2+1^2}\sqrt{-2^2+2^2+0^2}} = \frac{2}{\sqrt{2}\sqrt{8}}$

d) Pearson Correlation is given as :

$$w_{u,v} = \frac{\sum_{i\in I}(r_{u,i} - \bar{r}_u)(r_{v,i} - \bar{r}_v)}{\sqrt{\sum_{i\in I}(r_{u,i} - \bar{r}_u)^2}\sqrt{\sum_{i\in I}(r_{v,i} - \bar{r}_v)^2}}$$

Cosine Similarity is given as :

$$\cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n}(A_i)^2} \times \sqrt{\sum_{i=1}^{n}(B_i)^2}}$$

In cosine similarity, when the cosine vectors are normalized by subtracting the vector means, then they are equivalent to pearson correlation

Rubrics:
Part a) - 1 Correct (1 Mark), 2 Correct (1.5 Marks), 3 Correct (2 Marks)
Part b) - Atleast 50% of table is correct - Give 0.5 Marks
Part c) - Correct answer (1 marks)
Part d) (1 mark) for correct explanation.

7) [1 point] Circle ALL of the statements that are TRUE about Content-based Approach.

A. It is able to recommend to users with unique tastes
B. It is not able to recommend unpopular items
C. It never recommends items outside user's content profile
D. It is widely used because finding the appropriate features is easy

8) (1 Points) Describe how default voting can be used as an extension to memory based algorithms?

• By reducing the weight of users that have fewer than 50 items in common
• Use average of the clique (small group of corated items) as a default voting to extend a user's rating history
• Use a neutral or somewhat negative preference for the unobserved ratings and then computes similarity between users on the resulting ratings data.