End Term (Fall 2025)
DA 626: Recommendation System Design

Instructor: Dr Chiranjib Sur
Email: chiranjib@iitg.ac.in
Date: Nov, 2025

Total Marks: 80 marks

Instructions:

- The Answers must be brief and to the point. Bullets or tables or charts are highly encouraged as representations for answers.
- There should be a clear indication of the sub-parts of the answers. Like 2a) instead of just a). If you answer a question with just a), you will be awarded 0 credits.
- In case of any ambiguity, please mention (any) assumption explicitly and then answer the questions.

1. You are given the following true relevance values (0–3 scale):

Item	Relevence
A	3
B	2
C	2
D	1
E	0

Two ranking algorithms produce the following orderings:
- System X: [B, A, C, D, E]
- System Y: [A, C, B, D, E]

Use DCG@5 with the following equations:

$$DCG@k = \sum_{i=1}^{k} \frac{2^{rel_i} - 1}{\log_2(i+1)} \quad (1)$$

(a) Compute DCG@5 for both systems.

(b) Compute the ideal DCG (IDCG@5).

(c) Compute nDCG@5 for both systems.

(d) Explain why DCG incorrectly suggests System X is better, even though System Y ranks the most relevant item (A) earlier.

2. You are asked to design a content-based recommendation system for a news website (e.g., "DailyNews"). The website has thousands of articles with titles, body text, tags, authors, and publication dates.
   (a) What item features would you extract for the content-based model?
   (b) How would you represent the user profile?
   (c) Describe the end-to-end flow for generating a "Recommended for You" list for a logged-in user.
   (d) How would you handle fresh (very recent) articles that have no user interactions yet?

3. You are designing a recommender for a movie streaming platform. User–item interaction data (ratings, watch history) is sparse, and new users/movies arrive frequently.
   (a) Explain how you would handle the new user cold-start problem.
   (b) Explain how you would handle the new item cold-start problem.
   (c) Propose a hybrid architecture that combines collaborative filtering and content-based filtering to mitigate sparsity.
   (d) How would you gradually transition a new user from cold-start mode into a fully personalized CF-based mode?

4. You have built a prototype recommendation system for an e-commerce site and want to evaluate it properly before deployment.
   (a) Describe how you would set up an offline evaluation using historical interaction data. What metrics would you use and why?
   (b) Describe how you would set up an online A/B test once the model is deployed. What user-facing metrics would you track?
   (c) Why is it not enough to rely only on offline metrics like Precision@K and nDCG@K?

5. A video streaming platform (like Netflix/Hotstar) uses a simple popularity-based recommender on its homepage:
   - "Top Trending"
   - "Most Watched in Your Region"
   - "New Releases"

   However, user engagement has started declining. You are asked to recommend design improvements to the existing system.
   (a) Identify three major limitations of a popularity-based recommender.
   (b) Propose two algorithmic improvements to move toward personalized recommendations.
   (c) Suggest two UX/UI changes (front-end improvements) to surface more personalized content.
   (d) Explain how you would measure whether the improved recommender is successful.

40 13

12

6. A music app uses collaborative filtering to recommend songs and artists. However, new or niche artists rarely get recommended because they have few listens (cold-start).

   (a) Identify why collaborative filtering alone fails in this scenario.
   (b) Propose an improvement to better recommend new/niche artists.
   (c) Describe how you would incorporate content-based audio features into the system.
   (d) Suggest a strategy to balance exploration vs exploitation.

7. An e-commerce site uses a content-based recommender that relies only on product text descriptions. The quality of recommendations is mediocre.

   (a) Identify at least three weaknesses of using only text-based features for recommendations.
   (b) Suggest at least three new types of features that can be added to improve recommendations.
   (c) Propose an evaluation plan to ensure recommendations become more relevant.
   (d) Describe how you would incorporate contextual signals (time, season, location).

8. A social media app uses a simple algorithm that shows posts from followed users in reverse chronological order. Engagement is dropping, and users complain that they "miss important posts."

   (a) Identify limitations of reverse-chronological feeds.
   (b) Propose three improvements to make the feed more engaging.
   (c) How can the system incorporate user interaction signals (likes, comments, dwell time)?
   (d) Suggest a mechanism to avoid "echo chambers."

9. A recommendation system suggests top-10 items to a user. For a given user U, the system returns the following ranked list of items (from rank 1 to 10):

   Recommended list (R):
   [A, B, C, D, E, F, G, H, I, J]

   From ground truth (actual user behaviour), the items that are actually relevant to the user are:

   Relevant set (Rel):
   {A, D, F, K, L}

   Answer the following:

   (a) List which recommended items in the top-5 are relevant.
   (b) Compute Precision@5 and Recall@5.
   (c) Compute Precision@10 and Recall@10.
   (d) Briefly interpret what the values of Precision@5 and Recall@5 tell you about the system's performance for this user.

$\frac{TP}{TP+FN}$

10. A recommendation system returns the same ranked list of 8 items for all users:

System's ranked list (from rank 1 to 8):
[A, B, C, D, E, F, G, H]

Two users have the following sets of relevant items (ground truth):
- User 1 relevant items: {A, C, E, G}
- User 2 relevant items: {B, D, F}

Answer the following:

(a) For User 1, compute:
  o Precision@3 and Recall@3
  o Precision@5 and Recall@5

(b) For User 2, compute:
  o Precision@3 and Recall@3
  o Precision@5 and Recall@5

(c) Compute the average Precision@3 and average Recall@3 across the two users.

(d) Compute the average Precision@5 and average Recall@5 across the two users.

(e) Briefly comment on whether the system is more recall-oriented or precision-oriented at K=5 based on the averages.