194.035 Recommender Systems (VU 2,0) 2020S

**Final Exam**

1. Write a reasonable formula that computes the similarity between two songs based on the number of plays by users. For notation, you may use i and j to denote the two songs, and r\_{ui} to denote the number of plays of song i by user u. (LATEX FORMULA)

2. In I-I CF, recommendations can be obvious because recommended items are similar to the items the user has already rated. Can you explain why a non-similar item cannot be recommended? (OPEN QUESTION)

3. Cold-start problems only appear when collaborative filtering techniques are used. (TRUE/FALSE)

4. Collaborative filtering requires explicit feedback (e.g., ratings) from users on items. (TRUE/FALSE)

5. A larger neighborhood in U-U CF, implies a better accuracy of the recommendations. (TRUE/FALSE)

6. In the matrix factorization model that uses baseline estimates, what is the number of model parameters that need to be learned? Assume, n items, m users and latent feature dimensionality of k. (LATEX FORMULA)

7. Matrix factorization is another term for singular value decomposition. (TRUE/FALSE)

8. The goal of regularization is to ensure stochastic gradient descent converges to a local minimum. (TRUE/FALSE)

9. In content-based recommender systems, what are the prototype vectors in the Relevance Feedback approach (Rocchio’s method)? (OPEN)

10. Explain in one sentence the main difference between ranking accuracy and rating prediction accuracy metrics. (OPEN QUESTION)

11. A system recommends 12 items, among which only 4 are relevant. In total, there exist 8 relevant items. What is the F1 measure of this system? (NUMERIC ANSWER)

12. The Twitter RecSys Challenge 2020 is about achieving good rating prediction accuracy. (TRUE/FALSE)

13. In a precision-recall curve, the highest precision appears at the lowest recall level. (TRUE/FALSE)