DSAIT4335 Recommender Systems

Final Project

Feedback session

- Feedback sessions for final project in week 8
- The duration is 20 minutes per group
- Enroll for timeslot in: https://queue.tudelft.nl/lab/9030
- Location: Building 28, 5th floor, room Kerkhoffs

Final Project

- **LO 5: Implement a recommendation algorithm to operate in a specific domain**
  - **Task 1:** Implement a weighted hybrid recommender by combining multiple recommendation models
- **LO 6: Evaluate the effectiveness of a recommendation model through an offline evaluation**
  - **Task 2:** Run offline experiments for both rating prediction and ranking tasks for each individual and a hybrid recommendation model
  - **Task 3:** Compare the performance of each with baselines
  - **Task 4:** Coefficient analysis: impact of each recommendation model on successful recommendation in a hybrid setting
- **LO7: Analyze the recommendation results with respect to known challenges and societal aspects**
  - **Task 5:** Performance analysis beyond accuracy: diversity, novelty, calibration, fairness and popularity bias

Task 1: Hybrid recommender

Implementing individual recommenders

- Implement several recommendation algorithms
  - Content-based recommender (CB)
  - User-based neighborhood method (UserKNN)
  - Item-based neighborhood method (ItemKNN)
  - Matrix Factorization (MF)
  - Bayesian Probabilistic Ranking (BPR)
- Use linear regression to find the optimal weights for each algorithm
- Prediction:
  - Compute prediction for each algorithm
  - Combine predictions using weights

Task 1: Hybrid recommender

Combining recommenders

- Recommendation components C_1 … C_k
  - Each component is a recommendation algorithm
- Overall prediction

$$score(u,i) = \sum_{j=1}^{k} \alpha_j score_{C_j}(u,i)$$

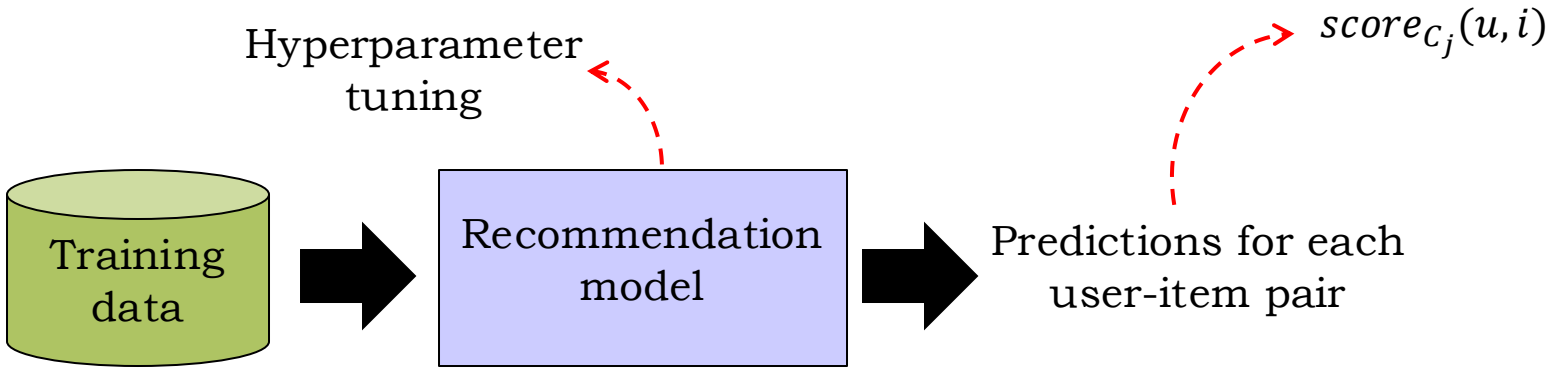
Prediction of component (algorithm) C_j

Coefficient weight, the degree to which component C_j contributes to final prediction

Task 1: Hybrid recommender

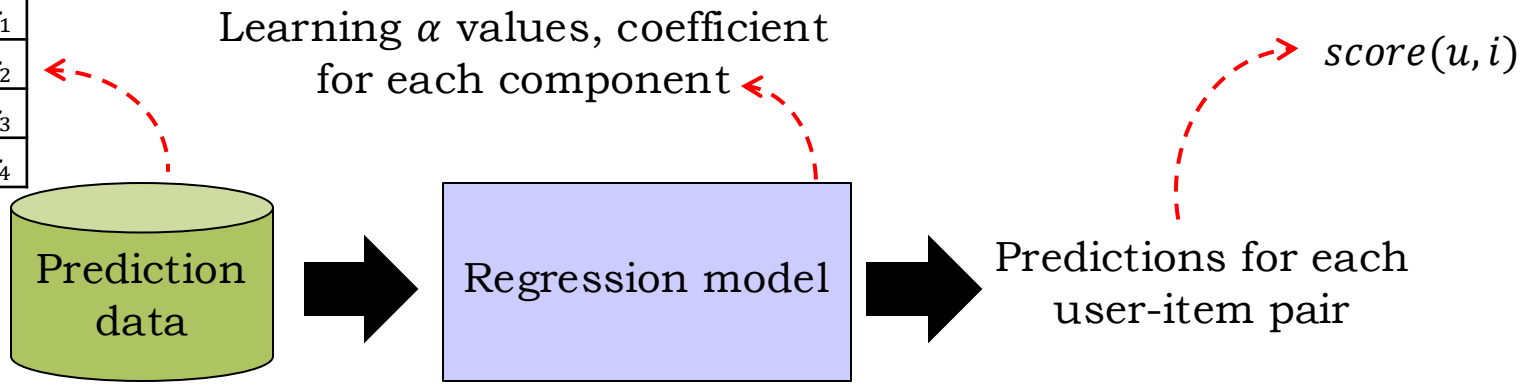
Parameter learning

Step 1



Step 2

$u1$	$i1$	C_1^{u1i1}	…	C_k^{u1i1}	r_1
$u1$	$i2$	C_1^{u1i2}	…	C_k^{u1i2}	r_2
$u2$	$i1$	C_1^{u2i1}	…	C_k^{u2i1}	r_3
$u2$	$i2$	C_1^{u2i2}	…	C_k^{u2i2}	r_4



Task 2: Evaluation of Effectiveness

Experiments: rating and ranking prediction

- Rating prediction task
  - Build/Fit *content-based*, *user-based*, *item-based*, and *matrix factorization* model
  - Evaluate each model in terms of RMSE
  - Build and evaluate a hybrid model with all four models
- Ranking task
  - Build/Fit *content-based*, *user-based*, *item-based*, *matrix factorization*, and *BPR* model
  - Evaluate each model in terms of Precision, Recall, and NDCG
  - Build and evaluate a hybrid model with all five models

Tune each model to find the best-performing setting

Task 3: Evaluation of Effectiveness

Experiments: Comparisons with baselines

- Implement the following baselines
  - Rating prediction task
    - Average ratings of target item
    - Mean hybrid: average prediction of all components
  - Ranking task
    - Random recommender
    - Most popular recommender
    - Mean hybrid: generating top-k recommendation by computing average relevance score obtained from all components
- Compare and discuss the single model, hybrid models, and baselines in terms of accuracy and non-accuracy metrics

Task 4: Evaluation of Effectiveness

Experiments: Coefficient analysis

- Analyze the coefficients of regression model (hybrid model) for both rating prediction and ranking tasks
  - Which models contribute the most to prediction?
- Where is each recommendation model successful in delivering accurate recommendation?
  - For which user groups each recommendation model results in the highest accuracy?

Task 5: Evaluation beyond accuracy

- Evaluate both single models and hybrid model using the following metrics:
  - Diversity (intra-list diversity)
  - Novelty (surprisal)
  - Calibration
  - Fairness metrics
- Discuss your observations comparing the models in terms of both accuracy and non-accuracy metrics

Deliverables

- Report
- Presentation
- Peer evaluation

Report including the codes

- A jupyter-notebook template is shared
  - Follow the instructions in jupyter-notebook file
- It is up to you how to structure your answer for each task
  - This improves the readability of your codes and answers
  - Try to be creative in organizing your report (jupyter-notebook), e.g., adding comments to your codes, concise and clear description of observations, etc
- Use proper means of visualization for explaining your observations and answers

Report including the codes

- The submission is due on **October 27, 2025, 23:59**

Presentation

- Presentations will take place in 3-7 of November
  - Time slots of 45 minutes for the whole week will be proposed
  - Each group must enroll in a time slot
  - Be first to enroll: first come, first served!
- 9 minutes presentation by all team members
  - Each student must present
  - 3 minutes per student
  - Individual presentation will inform the individual grade

Presentation
Q&A

- The presentation will be followed by Q&A
  - Individual examination
- Each Student will be asked a number of questions
  - About the project and general recsys topics
- Interaction with each student will inform the individual grade

Peer evaluation

- You will evaluate your teammates from different aspects including contribution, responsibility, communication, etc
- Your final grade will be affected if both teammates confirm that you did not perform well in the group

Consistent results

Important note

- Due to ties, the outputs might be different in multiple runs
- Make sure to set the **seed=10** for consistency in experiments

Next session …

- No class October 16, 2025
- Feedback sessions in Week 8
  - Make sure to enroll for timeslots: https://queue.tudelft.nl/lab/9030