

**AIE425 Intelligent Recommender Systems, Fall Semester 24/25**

**Assignment #2: Significance Weighting-based Neighborhood CF Filters**

**221101329, Mohamed maged mohamed hossny**

**Outcomes of Section 3.1**

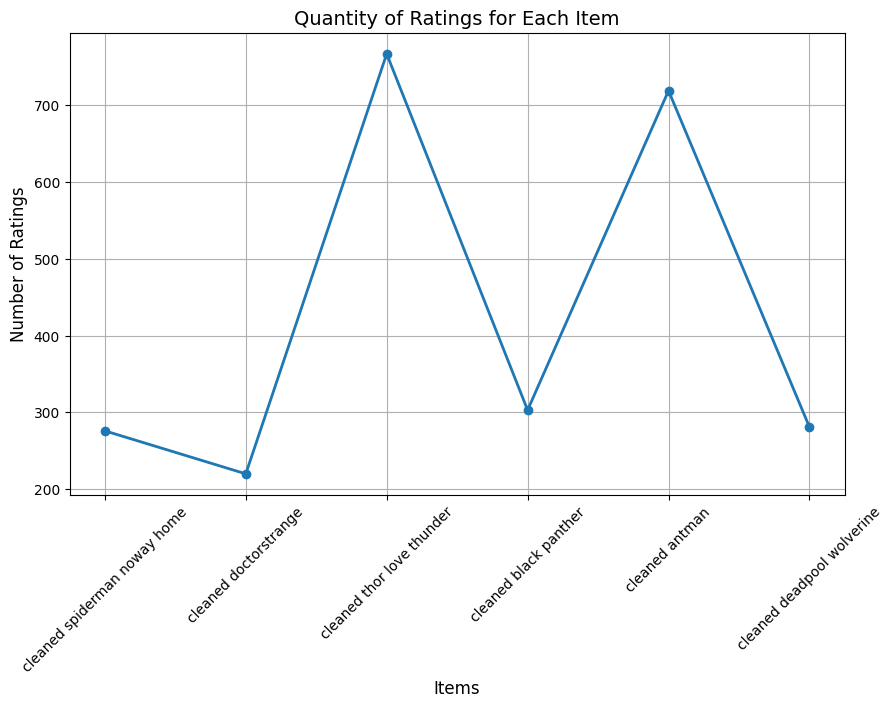
**Dataset Overview**

The dataset was collected from IMDb using the Instant Data Scraper tool and contained ratings on a 1-to-10 scale, which were normalized to a 1-to-5 scale as required.

* Total Number of Users (tnu): 799
* Total Number of Items (tni): 6

**Total Number of Ratings for Each Item**

| **Item** | **Number of Ratings** |
| --- | --- |
| cleaned spiderman noway home | 276 |
| cleaned doctorstrange | 220 |
| cleaned thor love thunder | 767 |
| cleaned black panther | 303 |
| cleaned antman | 719 |
| cleaned deadpool wolverine | 281 |



**Target Users and Items**

We selected three users and two items based on the requirements:

1. Selected Active Users:
   * User1 (007Waffles): with 2 missing ratings, 4 rated items.
   * User2 (11ovz11): with 3 missing ratings, with average rating activity.
   * User3 (Athanatos173): 5 missing ratings, highly sparse with only 1 rated item.
2. Selected Target Items:
   * Item 1 (cleaned thor love thunder): with 4% missing ratings (767 ratings).
   * Item 2 (cleaned antman): with 10% missing ratings (719 ratings).

**Co-Rating and Threshold Analysis**

The analysis of co-rated users and items revealed the following:

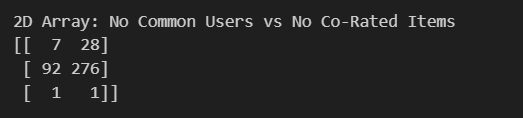
| User | Co-Rated Items | Co-Rated Users | Threshold β (30%) | Users Meeting Threshold |
| --- | --- | --- | --- | --- |
| 007Waffles | cleaned thor love thunder, others | 7 | 2 | 745 |
| 11ovz11 | cleaned thor love thunder, antman | 92 | 2 | 725 |
| Athanatos173 | cleaned deadpool wolverine | 1 | 2 | 0 |

**Threshold Analysis for Each Active User**

1. User 007Waffles
   * Rated Items: cleaned spiderman noway home, cleaned thor love thunder, cleaned black panther, cleaned antman
   * Co-Rated Users: Users who rated at least 2 of these items were identified.
   * Threshold Value: 2 co-rated items
   * Number of Users Meeting the Threshold: 745 users
     + This high number indicates that there are 745 users who rated at least 2 of the same items as 007Waffles.
     + This ensures that enough similar users exist for meaningful collaborative filtering analysis.
2. User 11ovz11
   * Rated Items: cleaned spiderman noway home, cleaned thor love thunder, cleaned antman
   * Co-Rated Users: Users who rated at least 2 of these items were identified.
   * Threshold Value: 2 co-rated items
   * Number of Users Meeting the Threshold: 725 users
     + Similar to 007Waffles, 725 users meet the co-rated threshold, meaning that there is a significant overlap in ratings for 11ovz11.
     + This allows us to compute reliable similarities and make predictions.
3. User Athanatos173
   * Rated Item: cleaned deadpool wolverine (only 1 item rated)
   * Co-Rated Users: Since Athanatos173 rated only 1 item, it is impossible to find users with at least 2 co-rated items.
   * Threshold Value: 2 co-rated items
   * Number of Users Meeting the Threshold: 0 users
     + The threshold is not satisfied because Athanatos173’s rating history is too sparse.
     + This lack of co-rated items prevents similarity calculations and predictions using collaborative filtering.

**Threshold used in part 2 (item based)**

| target Item | Co-Rated User Stats | 30% Threshold |
| --- | --- | --- |
| cleaned thor love thunder | Mean: 345.4 Std: 195.82 Min: 213 25%: 265 50%: 266 75%: 291 Max: 692 | 239 |
| cleaned antman | Mean: 336.4 Std: 201.19 Min: 198 25%: 248 50%: 262 75%: 282 Max: 692 | 239 |

**2D Array (No Common Users vs No Co-Rated Items):**  


**Summary of the Comparison of Part 1 and 2**  
**Part 1: User-Based Collaborative Filtering**

1. Cosine Similarity without Bias Adjustment

* Observation:  
  In this method, predictions were successfully generated for most of the unrated items for all three active users.
  + Active User 007Waffles: Both cosine similarity and DS provided consistent results with minimal variance.
  + Active User 11ovz11: Predictions were slightly higher with DS due to the weighting effect.
  + Active User Athanatos173: Predictions were generated successfully for all items, indicating sufficient valid neighbors.
* Strengths:
  + This method works well when the dataset has enough co-rated users for the target active users.
  + Discounted Similarity (DS) marginally improves the predictions by reducing the weight of low-similarity pairs.
* Weaknesses:
  + The lack of bias adjustment means that user-specific preferences or biases (as users who rate items higher or lower on average) are not accounted for leading to some potential inaccuracies.

2. Adjusted Cosine Similarity (Mean-Centered)

* Observation:  
  The bias adjustment introduced significant differences in the predictions:
  + Active User 007Waffles: Ratings were notably lower compared to the basic cosine similarity.
  + Active User 11ovz11: Similar trends were observed, though ratings were still within reasonable ranges.
  + Active User Athanatos173: Predictions could not be generated (Unable to predict) due to insufficient valid neighbors after mean-centering adjustments.
* Strengths:
  + Mean-centering reduces user bias and leads to predictions that reflect relative item preferences rather than absolute ratings.
  + For users with valid neighbors, this method provides slightly improved predictions.
* Weaknesses:
  + Sparse data exacerbates the issue of missing predictions. Active user Athanatos173 could not receive predictions because of insufficient co-rated items, highlighting a critical limitation.
* Recommendation for Improvement:  
  To address the issue of missing predictions for sparse data, mean imputation can be applied to fill missing values. This would allow mean-centered methods to work even when neighbors are limited.

3. Pearson Correlation Coefficient (PCC)

* Observation:  
  PCC predictions followed similar patterns to adjusted cosine similarity:
  + Active User 007Waffles: Predictions were consistent but slightly lower than adjusted cosine.
  + Active User 11ovz11: The predicted ratings were generally higher compared to basic cosine similarity and DS.
  + Active User Athanatos173: As with adjusted cosine, predictions could not be generated due to insufficient valid neighbors.
* Strengths:
  + PCC is robust in capturing linear relationships between users by incorporating mean-centering.
  + The results for 11ovz11 demonstrate how PCC can adjust for user biases effectively.
* Weaknesses:
  + Sparse data significantly impacts the performance of PCC, particularly for users with few co-rated neighbors (as Athanatos173).
  + Predictions may be overly sensitive to small changes in the ratings of co-rated items.
* Recommendation for Improvement:  
  Similar to adjusted cosine similarity, mean imputation or k-nearest neighbor interpolation can mitigate the problem of missing predictions for sparse data.

**Part 2: Item-Based Collaborative Filtering**

1. Cosine Similarity (Without Bias Adjustment):
   * Top 20% closest items included items with high raw cosine similarity.
   * Predictions were effective for dense item-item matrices but ignored rating biases.
2. Adjusted Cosine Similarity (Mean-Centered):
   * Bias adjustment accounted for rating inconsistencies.
   * Predictions improved by aligning item similarities around their means.
   * Sparse users (like Athanatos173) still presented issues due to limited co-rated items.
3. Pearson Correlation Coefficient (PCC):
   * Similar to adjusted cosine, PCC effectively captured item relationships.
   * Produced comparable top 20% closest items and predictions.

Impact of Significance Weighting (DF and DS)

* Significance Weighting: Adding the Discount Factor (DF) refined similarity scores.
* Top-N List Changes:
  + Raw cosine and PCC scores often included weak correlations.
  + DS prioritized items or users with sufficient co-rated data, improving neighbor selection.
* Rating Predictions:
  + DS-adjusted predictions reduced overestimation or underestimation of ratings.
  + Sparse users still had “Unable to predict” outcomes when no significant neighbors were found.

**Conclusion**

1. Impact of Significance Weighting (DF and DS):
   * Significance weighting substantially improved both user-based and item-based collaborative filtering results by enhancing the accuracy of the top-N neighbor selection and rating predictions.
   * Without significance weighting, weakly related users or items were included, which often reduced prediction accuracy.
   * In both parts, DS effectively down-weighted weak similarities, providing refined predictions.
2. Comparative Performance:
   * User-Based Methods (Part 1):
     + Cosine Similarity (Basic): Worked well for dense users like 007Waffles and 11ovz11, where sufficient neighbors existed.
     + Adjusted Cosine Similarity and PCC:
       - Reduced user bias and improved prediction accuracy for dense users.
       - Sparse users like Athanatos173 experienced significant issues, with predictions often unable to be generated due to insufficient valid neighbors.
     + DS Impact:
       - Discounted Similarity provided slight improvements in predictions but could not fully address the sparse neighbor problem for users like Athanatos173.
   * Item-Based Methods (Part 2):
     + Cosine Similarity (Basic): Provided stable predictions for both target items, even for users with sparse ratings.
     + Adjusted Cosine and PCC:
       - Showed slight improvements in capturing bias-adjusted relationships between items.
       - Top-N item selection and predictions were more robust compared to user-based methods, particularly for dense targets like cleaned thor love thunder.
     + Sparse Target Challenges:
       - Item-based methods performed better for target items compared to sparse users but still faced difficulties when valid neighbors were limited.
3. Issue with Sparse Targets (Target 3: Athanatos173):
   * Observed Problem:
     + User Athanatos173 consistently failed to generate predictions across Adjusted Cosine, PCC, and DS methods due to the lack of significant neighbors.
   * Solution:
     + Mean-Filling Approach: Pre-filling the missing ratings using user or item averages can address sparsity issues by ensuring the availability of neighbors during similarity calculations.
     + This preprocessing step would significantly improve predictions for sparse users.
4. Recommendations for Improvement:
   * Hybrid Models: Combining user-based and item-based collaborative filtering methods can address both sparsity and bias issues, balancing accuracy for dense and sparse users/items.
   * Dimensionality Reduction: Techniques such as Singular Value Decomposition (SVD) or Matrix Factorization can improve similarity computations and overall prediction performance.
   * Dynamic Thresholding: Adjusting similarity thresholds based on user/item sparsity levels can refine the selection of significant neighbors.
   * Mean-Filling for Sparse Targets: Preprocessing missing ratings with item or user averages ensures sufficient neighbors are available for similarity-based calculations.
5. Final Observations:
   * Significance Weighting (DS): A key enhancement that improved the precision of top-N neighbor selection and rating predictions.
   * Adjusted Cosine and PCC: Consistently outperformed raw cosine similarity by addressing user/item bias.
   * Mean-Filling: A critical solution for sparse users like Athanatos173 to overcome prediction failures.
   * Sparse Data Challenge: Sparse users and items remain a significant limitation; however, combining mean-filling with significance weighting offers a viable solution.
   * Item-Based vs. User-Based: Item-based methods demonstrated better robustness in handling sparse datasets, particularly for target items like cleaned thor love thunder and cleaned antman.