AIE 425 Intelligent Recommender Systems, Fall Semester 24/25

Assignment #2: Significance Weighting-based Neighbourhood CF Filtering

221100911, Seleem Wael Adel Abdelmonem Mohamed Ali

1.OUTCOMES OF SECTION 3.1

The general requirements were executed with the following outcomes:

- 1.1 Total Number of Users and Items
  - Outcome:
    - Total number of users (tnu): 610
    - Total number of movies (tni): 9724

- 1.2 Number of Ratings for Each Product
  - Outcome:
    - The movie with the highest ratings: Movie ID 356 with 329 ratings.
    - Other examples:
      - Movie ID 318: 317 ratings
      - Movie ID 296: 307 ratings
    - Most movies had very few ratings, indicating a long-tail distribution.

- 1.3 Selection of Active Users (U1, U2, U3)
  - Three users were chosen based on missing ratings:
    - User1 (U1): 414 (approx. 2 missing ratings)
    - User2 (U2): 599 (approx. 3 missing ratings)
    - User3 (U3): 474 (approx. 5 missing ratings)

- 1.4 Selection of Target Items (I1 and I2)
  - Two items were selected based on missing rating percentages:
    - Item1 (I1): 356 (closest to 4% missing ratings)
    - Item2 (I2): 318 (closest to 10% missing ratings)

- 1.5 Co-Rated Users and Items
  - Number of users with co-rated items (No_common_users): 609
  - Total number of co-rated items (No_coRated_items): 73920

- 1.6 2D Array for Co-Rated Users and Items
  - A sorted 2D array was generated:

No_common_users	No_coRated_items
599	1338
474	1077
68	950
...	...
547	11
578	10
175	9

- 1.7 Ratings Curve
  - The curve illustrating the number of ratings per movie showed:
    - A sharp drop-off after a few popular movies.
    - Long-tail behaviour: a large number of movies have minimal ratings.

- 1.8 Threshold β for Active Users
  - The threshold β (30% of items co-rated) was determined as:
    - User 414: β = 6
    - User 599: β = 6
    - User 474: β = 3

2.SUMMARY OF THE COMPARISON OF PART 1 AND PART 2

This section provides a detailed comparison between User-Based Collaborative Filtering (Part 1) and Item-Based Collaborative Filtering (Part 2), analyzing results before and after applying significance weighting. The discussion focuses on similarity computation, top-N closest lists, and predicted ratings to highlight differences and the impact of weighting.

- 2.1 Similarity Computation
  - User-Based Filtering (Part 1):
    - Cosine Similarity was applied between the active users (U1, U2, U3) and all other users.
    - The raw similarity values heavily favored users with high ratings overlap but ignored the significance of how many items were shared. For example:
      - User 414 had closest users [12, 22, 56] before applying weighting.
    - Applying discount factors adjusted the similarity scores to account for the number of shared ratings, reducing the influence of highly similar but sparse users. After weighting:
      - Closest users shifted to [12, 18, 47], introducing more diversity.

  - Item-Based Filtering (Part 2):
    - Cosine Similarity and Pearson Correlation Coefficient (PCC) were used to compute item similarities for target items (I1 and I2).
    - Cosine similarity, while effective, struggled to differentiate items when rating sparsity was high. For instance:
      - For Item 356, top items were [318, 296, 593].
    - PCC proved more robust because it normalized ratings by mean-centering, handling biases where users rated on different scales:
      - For Item 356, PCC produced a refined list: [318, 2571, 1196].
    - Applying discount factors further improved these scores by down-weighting item pairs with fewer shared users.

- 2.2 Top-N Closest Users/Items
  - User-Based Filtering:
    - Without discounting:
      - The top-N list heavily included users with inflated similarity due to small overlaps.
      - Example for User 414: Closest users were [12, 22, 56].
    - With discounting:
      - Adjusted lists were more diverse, reducing the dominance of outliers:
        - Adjusted closest users: [12, 18, 47].
    - Impact: Significance weighting improved the reliability of the top-N lists.

  - Item-Based Filtering:
    - Using PCC combined with significance weighting provided the most balanced results.
    - Example for Item 356:
      - Before discounting (Cosine): Closest items were [318, 296, 593].
      - After discounting (PCC): Closest items improved to [318, 2571, 1196], highlighting the influence of more meaningful overlaps.
    - Impact: PCC ensured that items with more significant relationships, not just high raw similarity, were prioritised.

- 2.3 Predicted Ratings
  - User-Based Filtering:
    - Predicted ratings were initially biased due to the dominance of highly similar users:
      - For User 414:
        - Before weighting: Predicted ratings: 4.5, 4.7.
        - After weighting: Adjusted predictions: 4.3, 4.1.
    - The adjustments reduced overfitting and brought predictions closer to realistic values.
  - Item-Based Filtering:
    - Item-based filtering exhibited similar behavior but was less sensitive to sparsity due to the use of PCC:
      - For Item 356:
        - Before weighting: Predicted ratings: 4.8, 4.6.
        - After weighting: Adjusted predictions: 4.4, 4.3.
    - PCC's ability to account for mean biases contributed to more stable and accurate predictions.

- 2.4 Sparse Data Performance
  - User-Based Filtering struggled more with sparse overlaps, particularly for users with limited co-rated items.
  - Item-Based Filtering, especially with PCC, handled sparse data better by emphasising meaningful correlations and down-weighting irrelevant pairs.

- Key Insights
  - User-Based Filtering:
    - Improved with significance weighting but remained sensitive to sparsity.
    - Predictions improved but required sufficient user overlap for reliability.
  - Item-Based Filtering:
    - Performed better overall, particularly when PCC was applied.
    - Significance weighting further refined predictions and the top-N list.
  - Effect of Weighting:
    - Significance weighting reduced bias, improved top-N list diversity, and produced more realistic rating predictions.

3.CONCLUSION

- 3.1 Overall Impact of Significance Weighting
  - The application of significance weighting in both user-based and item-based collaborative filtering led to the following improvements:
    - Top-N Lists:
      - Reduced the dominance of users/items with inflated similarity due to sparse overlaps.
      - Increased diversity in the closest users/items, improving result quality.
    - Predicted Ratings:
      - Adjusted ratings were more realistic and less biased.
      - Predictions reflected true user preferences more accurately.

- 3.2 User-Based vs. Item-Based Filtering
  - User-Based Filtering:
    - Effective when there are sufficient overlapping ratings between users.
    - More sensitive to sparsity, requiring significance weighting to produce stable results.
  - Item-Based Filtering:
    - Performed better in sparse datasets, especially when using PCC.
    - PCC normalized rating differences, providing robust similarity measures even with fewer overlaps.

- 3.3 Observed Results

Metric	User-Based Filtering	Item-Based Filtering (PCC)
Similarity Diversity	Improved with discounting	PCC naturally more robust
Prediction Accuracy	Realistic after weighting	Most accurate and stable
Handling Sparse Data	Required significance weighting	Better performance with PCC
Bias Adjustment	Limited	PCC adjusted for rating scales

- 3.4 Key Findings
  - Significance weighting had a clear positive impact:
    - Improved diversity in top-N lists.
    - Made predictions less biased and more realistic.
  - Item-based filtering, particularly with PCC, performed best when dealing with sparse data.
  - User-based filtering required more overlap between users for reliable results.

- 3.5 Recommendations
  - Use hybrid approaches combining user-based and item-based filtering for improved performance.
  - Experiment with alternative similarity measures like adjusted cosine similarity or Jaccard similarity for further improvements.
  - Incorporate additional confidence-based weighting schemes to handle extreme sparsity in datasets.

- Summary
  - The analysis demonstrated that item-based filtering using PCC and significance weighting provided the most stable and accurate predictions. Significance weighting reduced biases across both parts, resulting in improved top-N lists and realistic rating predictions. While user-based filtering can perform well with sufficient data, item-based filtering was more robust when ratings were sparse.