AIE425 Intelligent Recommender Systems, Fall Semester 24/25

Assignment #2:

221101107, Abdelrahman Mohamed Mahmoud

Outcomes of Section 3.1

|  |  |
| --- | --- |
| Requirements | Result |
| Total users | 250 |
| Total items | 5 |
| Rating for movie1 | 247 |
| Rating for movie2 | 247 |
| Rating for movie3 | 247 |
| Rating for movie4 | 248 |

1. Users with Missing Ratings:

1 missing rating: user\_2

2 missing ratings: user\_18

3 missing ratings: user\_11, user\_67

No items had more than 4% or 10% missing ratings.

1. Co-Rating Statistics: The table below shows the co-rating results, sorted in descending order:

|  |  |
| --- | --- |
| Common Users | Co-rated items |
| 248 | 3 |
| 247 | 2 |
| 0 | 0 |
|  |  |
|  |  |

**Threshold Values (Significance Weighting)**: Threshold values (called β) for active users were calculated:

|  |  |
| --- | --- |
| Userid | Threshold |
| User\_11 | 247 |
| User\_67 | 247 |
| User\_2 | 246 |
| User\_18 | 0 |

**Cosine Similarity Matrix**: The pairwise cosine similarity between users was calculated. Below are the top similarities for user\_11:

|  |  |
| --- | --- |
| Similar user | similarity |
| User\_38 | 0.9587 |
| User\_237 | 0.9484 |
| User\_67 | 0.9469 |
| User\_2 | 0 |
| User\_18 | 0 |

**Summary of the Comparison of Part 1 and Part 2**

1. **Top-N Recommendations:**

|  |  |
| --- | --- |
| Parts | Description |
| No weighting | Used only similarity scores without checking how many ratings were shared. |
| With weighting | Used significance weighting to adjust scores based on co-rated items (threshold). |

**In Part 1**, some users had high similarity even with few shared ratings, which made the recommendations less reliable.

**In Part 2**, applying significance weighting gave more importance to strong relationships. This improved the accuracy of the recommendations.

3.Rating prediction

|  |  |
| --- | --- |
| Parts | Description |
| No weighting | Predictions were based on raw similarity scores. Sparse ratings caused errors. |
| With weighting | Significance weighting helped reduce errors by focusing on reliable ratings. |

**We conclude that :**

Predictions became more accurate in Part 2 because users with few shared ratings had less influence.

**Conclusion:**

To summarize, applying significance weighting had a big impact on the results:

1. **Top-N List**:

The recommendations were more accurate because the method gave more importance to users with many shared ratings.

1. **Rating Predictions**:

Predictions were better because users with very few shared ratings had less influence.

1. **Overall Improvement**:

Significance weighting made the method stronger by solving the problem of sparse ratings.

It improved the accuracy of both recommendations and rating predictions.

1. **My Suggestions for Improvement**:

Add time-based weighting to include when the ratings were given.

Collect more data (more users and items) to make the results even better.

In conclusion, significance weighting is a very useful method for improving collaborative filtering. It makes the recommendations and predictions more accurate, which is important for real-world applications