# AIE425: Intelligent Recommender systems, Fall Semester 24/25

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

### Name: Hamza Nashaat Abdelbaki

### ID: 2211003328

**Selected Company**: Spotify  
Spotify is a popular music streaming service that uses advanced recommender systems to enhance user experience through personalized playlists such as "Discover Weekly" and "Daily Mix." These systems rely on various feedback mechanisms, primarily implicit signals like track plays, skips, repeats, and listening time, rather than explicit ratings.

In addition to Spotify, several other companies across different domains rely on recommender systems to enhance user experience:

* Netflix: Recommender system for personalized movie and TV show suggestions based on viewing history and ratings.
* Amazon: Product recommendations based on user purchase history and browsing behavior.
* YouTube: Suggests videos using watch history, search queries, and user engagement (e.g., likes, shares).
* Spotify: Uses a music recommendation system to suggest songs and playlists such as "Discover Weekly" based on user listening habits.

In this report, Spotify has been chosen for its advanced music recommender system, which relies heavily on implicit feedback such as track plays, skips, and repeats.

**Customer Feedback Collection and Rating Type**

Spotify primarily collects implicit feedback rather than explicit ratings to understand user preferences. Implicit signals include:

* Track Plays: Indicates user interest in the track.
* Skips: Suggests disinterest if a track is skipped shortly after it starts.
* Repeats: Shows a strong positive preference when a user listens to the same track multiple times.
* Listening Time: Captures how much of the track was heard, from partial plays to full listens.

These implicit signals are crucial in Spotify’s recommendation system, as they help determine user preferences without the need for explicit ratings like stars or thumbs up/down. Additionally, Spotify may also use explicit feedback when available (e.g., users “liking” a song or adding it to their playlists).

### Data Collection and Preprocessing

### Spotify’s user interaction data is collected based on user behavior on the platform. The data includes listening times, skips, and repeats. The preprocessing steps involved:

### Handling Missing Data: Missing interactions were assigned a rating of 0, signifying no interaction between the user and a track.

### Scaling Listening Times: Non-integer listening times were converted into a 5-point scale to standardize the feedback across all users. For example:

### 5: Listened fully multiple times

### 4: Listened fully once or twice

### 3: Partially played

### 2: Skipped after a few seconds

### 1: Skipped almost immediately

### Normalization: The listening times were normalized to fit within the predefined scale, allowing easier integration with collaborative filtering methods.

### Each of these steps was crucial in converting raw interaction data into a structured format suitable for building a user-item matrix.

**1. Introduction**

The goal of this assignment is to explore the impact of significance weighting on the performance of neighborhood-based collaborative filtering (CF) methods. Two approaches are considered: user-based CF and item-based CF, evaluated using Cosine Similarity and Pearson Correlation. The assignment highlights the importance of similarity adjustments, comparing predictions before and after applying discount factors.

Collaborative filtering is widely used in recommendation systems to predict user preferences based on historical data. By introducing significance weighting, this assignment aims to refine these predictions and assess their practical impact.

**2. Data Preparation**

The dataset used in this assignment is derived from Assignment #1. The following preprocessing steps were applied to ensure compatibility with CF methods:

1. **Rating Scale Adjustment**: Ratings were normalized to a 1-to-5 scale to standardize user-item interactions.
2. **Handling Missing Values**: Placeholder values were assigned to represent items with no user ratings.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Username | Bad Guy | Billie Jean | Blinding Lights | Bohemian Rhapsody | Circles | Closer | Dancing Queen | HUMBLE. | Heat Waves | Hey Jude | Imagine |
| beatbuster | 1 | 1 | 3 | 1 | 5 | 5 | 1 | 4 | 3 | 4 | 4 |
| beatslover21 | 5 | 2 | 2 | 5 | 1 | 4 | 2 | 1 | 4 | 3 | 5 |
| groovemaster | 3 | 1 | 2 | 5 | 2 | 2 | 5 | 2 | 5 | 1 | 2 |
| melodymaker | 5 | 5 | 5 | 4 | 5 | 1 | 5 | 3 | 1 | 3 | 0 |
| musicfan92 | 2 | 2 | 0 | 0 | 1 | 2 | 1 | 1 | 1 | 4 | 0 |
| rhythmaddict | 0 | 5 | 2 | 2 | 4 | 5 | 2 | 4 | 3 | 4 | 5 |
| rhythmlove88 | 2 | 5 | 3 | 3 | 5 | 2 | 1 | 5 | 5 | 5 | 4 |
| soundexplorer | 5 | 5 | 2 | 3 | 0 | 2 | 3 | 0 | 1 | 3 | 3 |
| tunes\_lover1 | 1 | 4 | 3 | 3 | 5 | 1 | 0 | 1 | 2 | 4 | 5 |
| vibesonly22 | 5 | 2 | 2 | 2 | 3 | 2 | 4 | 3 | 5 | 5 | 3 |

1. **User and Item Selection**:
   * Three active users were chosen, each with a varying number of missing ratings.
   * Two target items were selected based on their percentage of missing ratings.

Details of the dataset, user-item matrix, and selected users/items are provided in the **Outcomes** section.

**3. Methodology**

**3.1. User-Based Collaborative Filtering**

The user-based CF analysis comprises three case studies:

**Choose 3 active users:**

* One with 2 missing ratings. 2 missing users: beatbuster
* One with 3 missing ratings. 3 missing users: musicfan92
* One with 5 missing ratings. 5 missing users: rhythmaddict

**Threshold Setting**:

* beatslover21: 9 users meet the threshold of 9 co-rated items
* groovemaster: 9 users meet the threshold of 9 co-rated items
* musicfan92: 9 users meet the threshold of 9 co-rated items

**Case Study 1.1**: Cosine Similarity without mean-centering.

* Predicted Ratings (without DF): {'Industry Baby': 2.0, 'Midnight City': 5.0, 'Sunflower':1.0}
* Predicted Ratings (without DF): {'Levitating': 2.0, 'drivers license': 5.0}
* Predicted Ratings (without DF): {'Blinding Lights': 2.0, 'Bohemian Rhapsody': 5.0, 'Imagine': 5.0}

**Case Study 1.2**: Cosine Similarity with mean-centering.

* Predicted Ratings (with Mean-Centering): {'Industry Baby': 2.0, 'Midnight City': 5.0, 'Sunflower': 1.0}
* Predicted Ratings (with Mean-Centering): {'Levitating': 2.0, 'drivers license': 5.0}
* Predicted Ratings (with Mean-Centering): {'Blinding Lights': 2.0, 'Bohemian Rhapsody': 5.0, 'Imagine': 5.0}

**Case Study 1.3**: Pearson Correlation Coefficient.

* Predicted Ratings (using PCC): {'Industry Baby': 2.0, 'Midnight City': 2.0, 'Sunflower': 3.0}
* Predicted Ratings (using PCC): {'Levitating': 2.0, 'drivers license': 5.0}
* Predicted Ratings (using PCC): {'Blinding Lights': 2.0, 'Bohemian Rhapsody': 5.0, 'Imagine': 5.0}

**3.2. Item-Based Collaborative Filtering**

The item-based CF analysis mirrors the structure of user-based CF with the following case studies:

**Choose 2 target items:**

* + Imagine : 20% missing
  + Old Town Road: 10% missing
* **Case Study 2.1**: Cosine Similarity without mean-centering.
* **Case Study 2.2**: Cosine Similarity with mean-centering.
* **Case Study 2.3**: Pearson Correlation Coefficient.

**3.3. Significance Weighting**

For both user-based and item-based CF methods, thresholds were determined based on the number of co-rated items. The discount factor was calculated as:

This factor was applied to adjust similarity scores, emphasizing the reliability of users/items with higher overlap.

**4. Results**

A graph with blue lines and white text

Description automatically generated

**5. Comparison and Analysis**

The analysis of Part 1 (user-based CF) and Part 2 (item-based CF) highlights the nuanced differences in the application and effectiveness of collaborative filtering methods. The key findings are summarized below:

5.1. Similarities Across Parts 1 and 2

1. Impact of Significance Weighting:
   * In both user-based and item-based CF, applying significance weighting refined similarity scores by reducing the influence of unreliable neighbors.
   * This adjustment enhanced the accuracy of top-N recommendations and rating predictions, particularly for sparse datasets.
2. Similarity Measures:
   * Cosine Similarity was effective in identifying aligned patterns but was sensitive to rating scale variations.
   * Pearson Correlation mitigated biases introduced by differing rating scales and provided robust predictions across both approaches.
3. Mean-Centering:
   * The inclusion of mean-centering in both methods improved prediction reliability by accounting for individual user or item biases.

5.2. Differences Between Parts 1 and 2

1. Nature of Predictions:
   * User-based CF predictions were more dynamic, depending on user interaction patterns.
   * Item-based CF provided more stable recommendations, particularly for items with consistent user engagement.
2. Sensitivity to Sparse Data:
   * User-based CF faced challenges with sparse data, as fewer co-rated items between users reduced the reliability of similarity scores.
   * Item-based CF was less sensitive to sparsity, as item profiles were often shared among multiple users, creating denser item-item matrices.
3. Computational Complexity:
   * User-based CF required more intensive calculations for large datasets due to the variability in user interactions.
   * Item-based CF computations were comparatively efficient, as item similarities were pre-computed and reused.

5.3. Impact of Significance Weighting

The addition of significance weighting in both parts showed a measurable improvement in:

1. Top-N Recommendations: Recommendations were more precise, with prioritized neighbors contributing to higher accuracy.
2. Rating Predictions: Predicted ratings closely matched actual user preferences, reducing errors caused by outliers.

In conclusion, while both user-based and item-based CF benefit from significance weighting, the method’s impact is more pronounced in user-based CF, where it addresses the inherent sparsity and variability in user interactions. [Placeholder summarizing comparisons made in Part 1 and Part 2, with emphasis on the impact of significance weighting on the top-N list and rating predictions.

**6. Critical Analysis and Discussion**

The results indicate that applying significance weighting enhances the reliability of similarity measurements in collaborative filtering, particularly in datasets with sparse interactions. Key findings include:

1. **Cosine Similarity vs. Pearson Correlation**:
   * Cosine Similarity effectively identifies users or items with aligned preferences, but it is sensitive to variations in rating scales.
   * Pearson Correlation mitigates the bias caused by rating scale differences, providing robust predictions when user or item interactions are diverse.
2. **Impact of Mean-Centering**:
   * Mean-centering improves prediction accuracy by accounting for user-specific biases, especially in datasets with consistent rating patterns.
   * The effect is less pronounced when user-item interactions are highly varied.
3. **Effectiveness of Significance Weighting**:
   * Significance weighting reduces the influence of unreliable neighbors, prioritizing those with higher co-rated items.
   * This adjustment is particularly beneficial in sparse datasets, where similarity scores may otherwise be misleading.

Overall, the integration of significance weighting improves prediction accuracy and enhances the interpretability of similarity scores.

**7. Pros and Cons of Each Method**

**Cosine Similarity**

**Pros:**

* Simplicity in implementation.
* Effective for datasets with consistent rating magnitudes.
* Suitable for identifying general preference alignment.

**Cons:**

* Sensitive to variations in rating scales.
* May overstate similarities for users/items with differing rating patterns.

**Pearson Correlation**

**Pros:**

* Normalizes rating scales, reducing bias.
* Ideal for datasets with varied user or item rating behaviors.

**Cons:**

* Requires sufficient data for accurate normalization.
* Computationally intensive compared to Cosine Similarity.

**Significance Weighting**

**Pros:**

* Enhances reliability by prioritizing neighbors with more co-rated items.
* Improves prediction accuracy in sparse datasets.

**Cons:**

* Threshold setting can be dataset-specific and requires fine-tuning.
* Adds computational complexity to similarity calculations.

**8. Conclusion**

The study demonstrated the critical role of significance weighting in improving neighborhood-based collaborative filtering models. Adjusting similarity scores using discount factors enhanced prediction accuracy, particularly in sparse datasets. While Cosine Similarity performed well for general alignment, Pearson Correlation provided robustness against rating scale variations.

Future work could explore hybrid models combining user- and item-based CF, along with advanced weighting schemes to further refine recommendations.