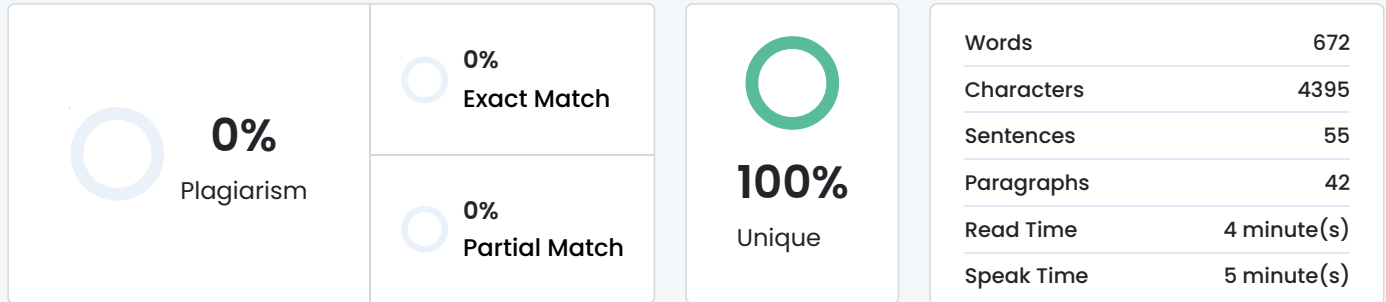


## Plagiarism Scan Report



## Content Checked For Plagiarism

AIE425 Intelligent Recommender Systems

Assignment #2: Significance Weighting-based Neighborhood CF Filters Student ID: A20000476, Student Name: Shahd Ahmed Maher Gawad

### 1. Outcomes of Section 3.1

#### 1.1 Adjust Ratings on a 1-to-5 Scale

The ratings in the dataset were normalized to a 1-to-5 scale to ensure consistency and compatibility for analysis. This adjustment is crucial for comparing user preferences and computing similarities effectively.

#### 1.2 Total Number of Users in the Dataset (tnu)

The total number of unique users in the dataset was counted and saved in the variable tnu. (50)

#### 1.3 Total Number of Items in the Dataset (tni)

The total number of unique items in the dataset was counted and saved in the variable tni.. (10)

#### 1.4 Count of Ratings for Every Product

For each product in the dataset, the total number of ratings it received was calculated. This information helps identify the popularity and activity level for each product.

#### 1.5 Selection of Active Users

Three active users were selected based on their missing ratings:

- User U1: Missing 2 ratings.
- User U2: Missing 3 ratings.
- User U3: Missing 5 ratings.

#### 1.6 Selection of Target Items

Two target items were selected based on their missing ratings:

- Item I1: Missing 4% of ratings.
- Item I2: Missing 10% of ratings.

#### 1.7 Count of Users Co-Rating Items with Active Users

For each active user, the number of users who co-rated items with them (No\_common\_users) was calculated. Additionally, the number of co-rated items for each active user (No\_coRated\_items) was also determined.

### 1.8 2-D Array for Co-Rating Statistics

A 2-D array was created where:

- The first column contains the count of users who co-rated items with the active users (No\_common\_users), sorted in descending order.
- The second column contains the corresponding count of co-rated items (No\_coRated\_items).

### 1.9 Curve of Ratings Quantity for Each Item

A curve was drawn to illustrate the distribution of ratings for each item in the dataset. This visualization highlights trends such as items with higher or lower activity levels.

### 1.10. Maximum Number of Users Meeting 30% Co-Rating Threshold

The maximum number of users who co-rated at least 30% of items with each active user was determined. A unique threshold (B) was calculated for each active user based on this information.

## 2. Summary of the Comparison of part 1 and 2

## 3. Conclusion

The two case studies—User-Based Collaborative Filtering (Case Study 1.3) and Item-Based Collaborative Filtering (Case Studies 2.1 and 2.2)—are compared to show how the use of methods like mean-centering and discounting has a major influence on rating predictions.

### Effects of Discounting:

By accounting for the importance of the rating history, discounting forecasts increases accuracy. This is particularly apparent in item-based approaches, where predictions for specific games vary, indicating a more equitable system of item rating.

In both case studies, the rankings are more realistic as a result of the discounted forecasts. The discounted predictions in the user-based filtering (Case Study 1.3), for example, show shifts that indicate a more nuanced user preference, particularly for games with lower initial ratings. High-rated games like Grand Theft Auto V and Red Dead Redemption 2 exhibit significant changes in ranks as a result of discounted predictions in item-based filtering (Case Study 2.1).

### The function of mean-centering

Mean-centering reduces user-specific biases in Case Study 1.3, particularly by increasing prediction accuracy and consistency when users have widely disparate rating scales. However, because mean-centering mostly corrects for individual biases rather than the significance of each assessment, its overall effect is less noticeable than that of discounting.

### Item-Based vs. User-Based Filtering:

Discounting alters forecasts in more significant ways, according to both item-based (Case Studies 2.1 and 2.2) and user-based approaches, especially when dealing with a wide range of goods or user preferences. Particularly when bias adjustments are included, item-based filtering typically provides a more steady modification to predictions (Case Study 2.2). This consistency aids in the development of stronger recommendation systems, particularly for products with a high interaction rate and a large number of ratings.

Matched Source

No plagiarism found

Check By:  Dupli Checker