Plagiarism Scan Report

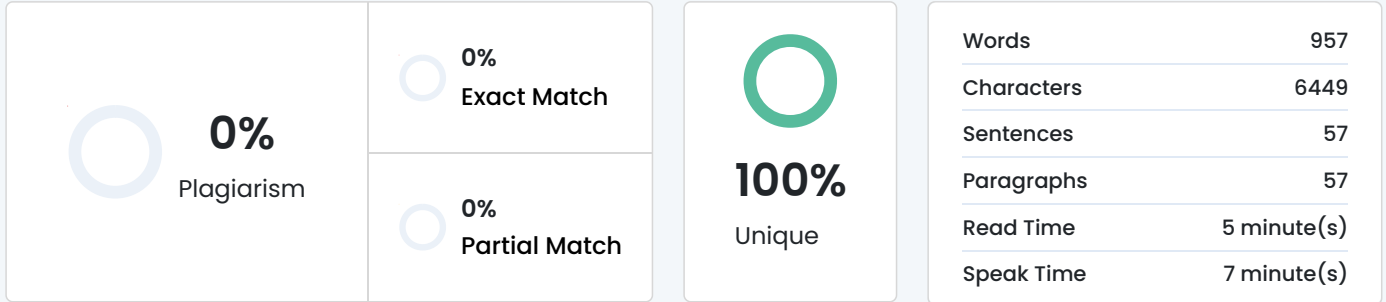|				
0% Plagiarism	**0%** Exact Match **0%** Partial Match	**100%** Unique	Words	957
			Characters	6449
			Sentences	57
			Paragraphs	57
			Read Time	5 minute(s)
			Speak Time	7 minute(s)

Content Checked For Plagiarism

AIE425 Intelligent Recommender Systems, Fall Semester 24/25

Assignment #3: Dimensionality Reduction methods

Student ID: A20000021, Full Name: Ahmed Ashraf Mohamed Ali

1. Summary

This section summarizes and compares the results obtained from the three methods used to predict missing ratings in a product-rating matrix. The three methods under evaluation are Part 1: PCA with Mean-Filling, Part 2: PCA with Maximum Likelihood Estimation (MLE), and Part 3: Singular Value Decomposition (SVD). The results will be compared based on the accuracy of predicting the missing ratings and the pros and cons of each method.

Part 1: PCA with Mean-Filling

In this method, missing ratings are replaced with the mean rating of the corresponding item before performing Principal Component Analysis (PCA). The dimensionality of the rating matrix is then reduced, and the missing ratings are predicted based on the reduced matrix.

Output Predictions:

- ETAMOON Wireless Gaming Controller (Top 5 peers): 3.81
- Xiaomi Redmi A3 (Top 5 peers): 3.81
- ETAMOON Wireless Gaming Controller (Top 10 peers): 4.06
- Xiaomi Redmi A3 (Top 10 peers): 4.06

Accuracy of Predicting Missing Ratings: The predicted ratings for both items are close to the mean value of the items, showing that the PCA with Mean-Filling method provides a moderate but relatively basic prediction. The results are not highly personalized, as they are strongly influenced by the mean rating of the items.

Pros:

- Simple and computationally efficient.
- Provides a good baseline for comparison with more advanced methods.

Cons:

- Lacks personalization: Ratings are influenced by the mean, which doesn't account for user-specific preferences.
- May introduce bias in the predictions, especially when item popularity varies.

Part 2: PCA with Maximum Likelihood Estimation (MLE)

The PCA with MLE method estimates the missing ratings using Maximum Likelihood Estimation based on the observed ratings of users interacting with the target items. This approach uses statistical modeling to provide a more personalized prediction compared to the Mean-Filling approach.

Output Predictions:
- ETAMOON Wireless Gaming Controller (Top 5 peers - MLE): 4.30
- Xiaomi Redmi A3 (Top 5 peers - MLE): 4.04
- ETAMOON Wireless Gaming Controller (Top 10 peers - MLE): 4.22
- Xiaomi Redmi A3 (Top 10 peers - MLE): 4.19

Accuracy of Predicting Missing Ratings: The predicted ratings from the MLE method are higher than those obtained from the Mean-Filling approach. The higher ratings for the items suggest that MLE better captures user-item interactions and provides more accurate predictions, especially when comparing the results to Part 1.

Pros:
- More personalized predictions compared to Mean-Filling.
- Uses observed ratings to generate better estimates for missing values.
- Reduces bias introduced by mean imputation.

Cons:
- Relies on sufficient data: The effectiveness of MLE decreases when the observed data is sparse.
- Requires additional computation to estimate likelihoods, making it computationally heavier than PCA with Mean-Filling.

Part 3: Singular Value Decomposition (SVD)

SVD is a matrix factorization technique that decomposes the ratings matrix into three matrices representing users, items, and latent factors. The missing ratings are predicted based on the latent factors learned from the observed ratings.

Output Predictions:
- ETAMOON Wireless Gaming Controller (SVD): 1.11
- Xiaomi Redmi A3 (SVD): 1.68

Accuracy of Predicting Missing Ratings: The results from the SVD method show notably lower predictions compared to both Part 1 and Part 2. This suggests that SVD may have overfitted the latent factors or the algorithm might have struggled to predict the ratings for these particular items.

Pros:
- Excels in capturing hidden relationships and patterns between users and items.
- Works well with sparse data by discovering latent features.
- Can provide better recommendations when tuned correctly.

Cons:
- Overfitting: The method may overfit to the training data if not properly regularized.
- Computationally intensive: Matrix factorization can be slow and requires more computational resources.
- Requires careful tuning.

2. Comparison

_____

3. Conclusion
- In conclusion, the results from all three methods highlight the varying levels of accuracy and computational demands associated with each approach.
- PCA with Mean-Filling is a simple and efficient method, but it lacks personalization, as it only uses the mean to fill missing ratings. While it offers a baseline prediction, it does not account for user-specific preferences, making it less effective in providing accurate ratings.
- PCA with Maximum Likelihood Estimation (MLE) offers more personalized predictions by leveraging observed ratings and modeling the likelihood of missing ratings based on these interactions. While it provides more accurate results than the Mean-Filling method, its reliance on sufficient data and higher computational cost could limit its effectiveness in some situations.
- Singular Value Decomposition (SVD), being a matrix factorization method, excels at uncovering latent factors that explain user-item interactions. However, it can suffer from overfitting and computational

complexity. Despite these challenges, when properly tuned, SVD can provide the most robust predictions, especially for datasets with rich interactions.

•   Impact of Matrix Factorization (SVD) Method: Matrix factorization methods like SVD are powerful tools in collaborative filtering because they are capable of uncovering latent factors that explain the underlying structure of user-item interactions. This method is particularly useful when the dataset is sparse, as it can still make accurate predictions based on the patterns discovered through factorization. However, the accuracy of the predictions depends heavily on tuning the number of latent factors and preventing overfitting, which requires both domain expertise and computational resources.

•   In summary, while SVD offers the most advanced and potentially accurate method, PCA with MLE provides a good balance of personalization and computational efficiency. PCA with Mean-Filling serves as a quick but less accurate approach. Choosing the right method depends on the specific requirements of the recommender system, such as the size of the dataset, the sparsity of the ratings matrix, and the computational resources available.

Matched Source

No plagiarism found