Assignment 3: Performance Evaluation of Five Recommender Algorithms

Name: Shashank P C

Student ID: 4147336

E-mail: s4147336@student.rmit.edu.au

Contact: +61 413 138 751

Statement:-

I certify that this is all my own original work. If I took any parts from elsewhere, then they were non-essential parts of the assignment, and they are clearly attributed in our submission. I will show I agree to this honor code by typing "Yes": **YES**

TABLE OF CONTENTS

Introduction to Recommender Systems

- Recommender Systems suggest relevant items to users
- Commonly used in platforms like Netflix, Amazon, Spotify
- Predict user preferences based on past behavior or item similarity
- Aim to improve user engagement and satisfaction
- Core task: **predict user rating** $r_{a,I}$ for item i

Overview of Implemented Methods (M1–M5)

Method	Description	Prediction Logic	Fallback
Method 1	User Average	$\hat{r}_{a,i} = \bar{u}_a,$	-
Method 2	Item Average	$\hat{r}_{a,i} = \bar{t}_i,$	-
Method 3	User-KNN Collaborative Filtering	Weighted rating from **K** most similar users who rated item **i**	If none, use user average
Method 4	Item-KNN Collaborative Filtering	Weighted rating from **K** most similar items rated by user **a**	If none, use item average
Method 5	Hybrid method	$\hat{r}_{a,i} = \lambda \hat{r}^u_{a,i} + (1-\lambda)\hat{r}^t_{a,i},$	If none, use user average

Optimisations in Methods 3, 4, and 5

- **Method 3**:-
  - Manually computed cosine similarity between users
  - Avoided self-comparisons: skipped similarity(u, u)
  - Configured K=5 & selected Top-K similar users who rated the same item
  - Weighted average used for prediction
  - Fallback to user’s average if no neighbours rated the item.
  - Only predicted where test rating exists (test_ds > 0)

Optimised for: personalized predictions & sparse data handling

Optimisations in Methods 3, 4, and 5

- **Method 4**:-
  - Computed cosine similarity between items using all user ratings
  - Created a full item-to-item similarity matrix (1682×1682)
  - For each item, selected K (value set as 5) most similar items rated by user
  - Used weighted average of user’s ratings for prediction
  - Fallback to item’s average rating if no similar items found

Optimised for: stability with high item density & cold users

Optimisations in Methods 3, 4, and 5

- **Method 5**:
  - Blended predictions from Method 3 and Method 4
  - Used a weighted combination: $\hat{r}_{a,i} = \lambda \hat{r}^{u}_{a,i} + (1 - \lambda)\hat{r}^{t}_{a,i},$
  - Lambda set to 0.5 for equal weightage for user and item predictions
  - Combines strengths of both methods
  - More robust for cold-start or sparse scenarios

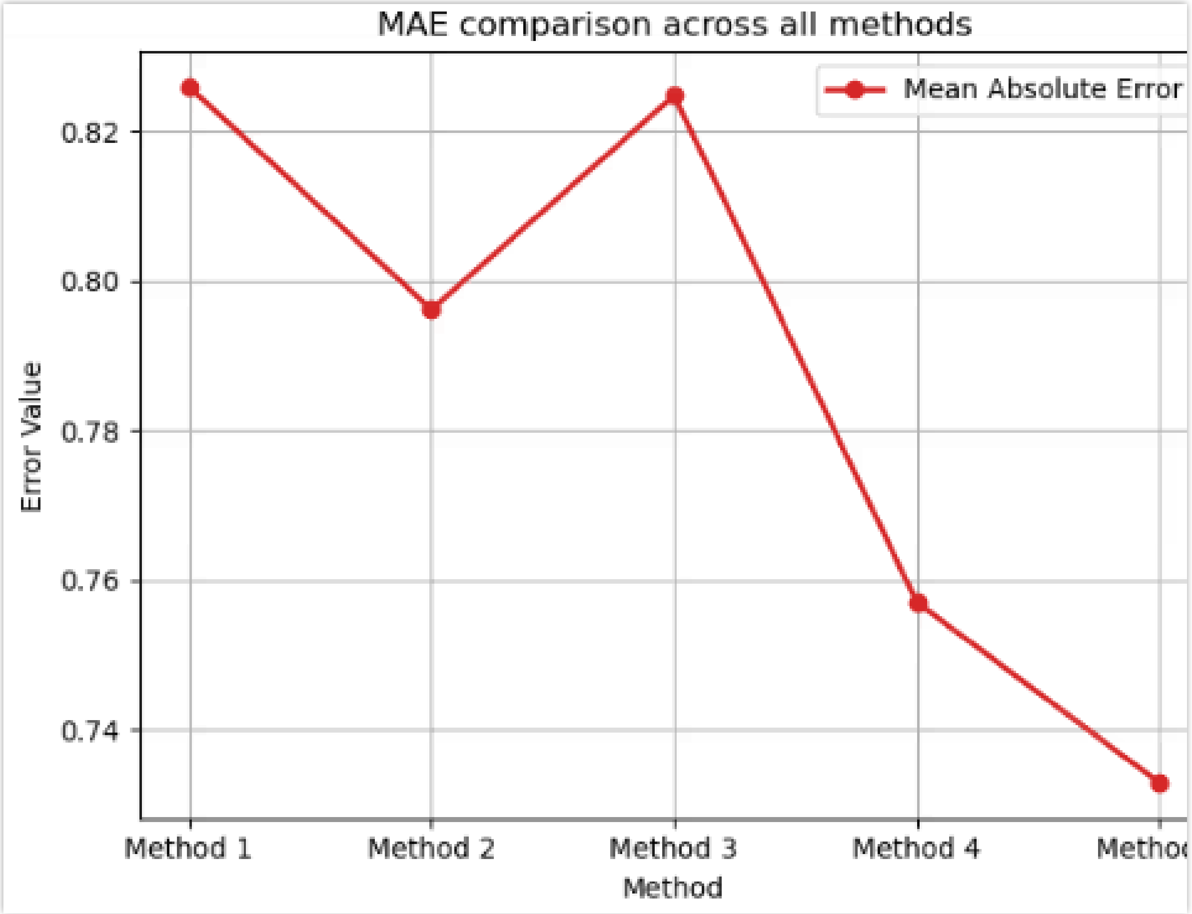
  Optimised for: balanced predictions and overall performance

Effect on K on prediction performance

Method	K	MAE	RMSE
Method 3 (User KNN)	3	0.8593	1.0953
	5	0.8247	1.0431
	7	0.8074	1.0207
	10	0.795	1.004
Method 4 (Item KNN)	3	0.7783	1.0249
	5	0.757	0.9845
	7	0.7499	0.9691
	10	0.7443	0.9612
Method 5 (Hybrid)	3	0.7478	0.9574
	5	0.7329	0.9362
	7	0.7272	0.9281
	10	0.7234	0.9237

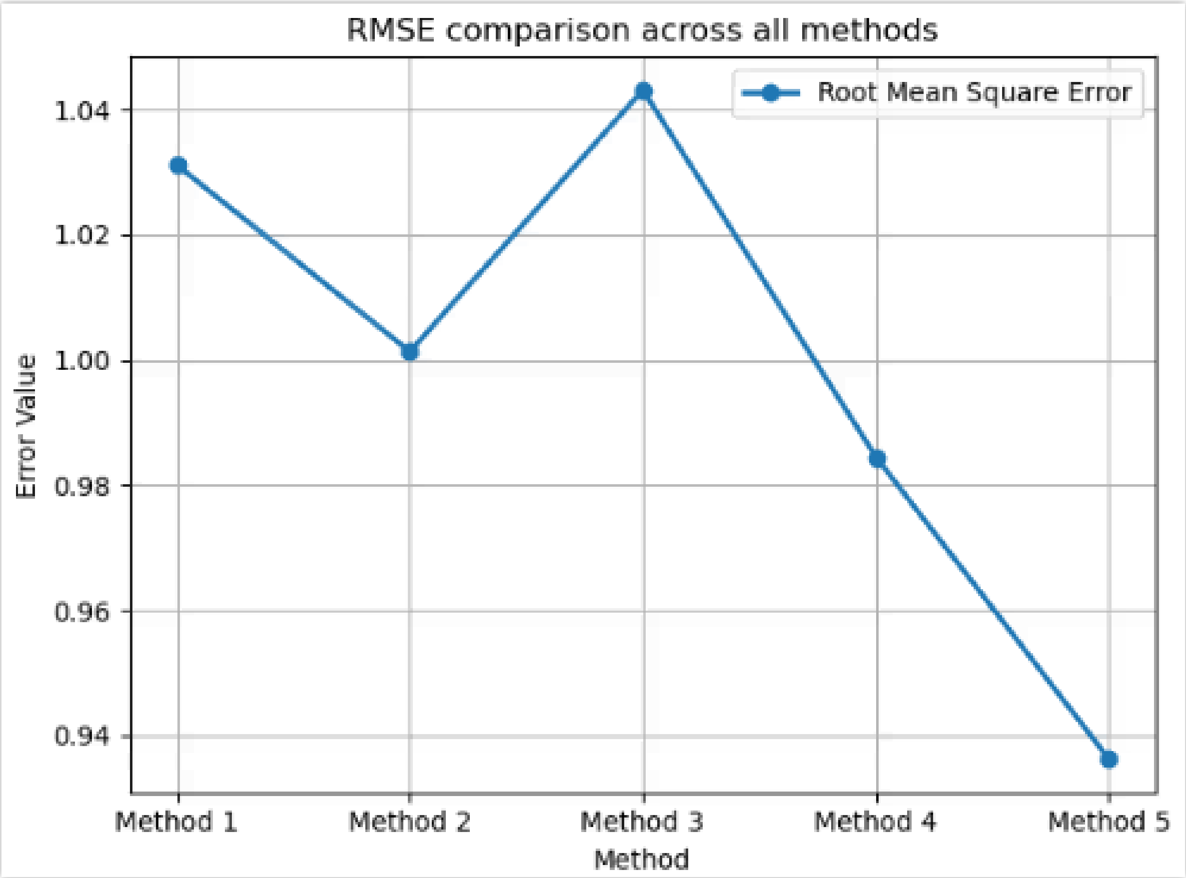
Performance Comparison (MAE)

- Lower MAE indicates more accurate predictions
- Method 5 (Hybrid) outperforms all other approaches
- Item-based (Method 4) performs better than user-based (Method 3)
- Methods 1 and 2 (global averages) are least accurate
- Hybrid model leverages strengths of both user and item similarities



Performance Comparison (RMSE)

- RMSE decreases overall from Method 1 to Method 5
- Method 3 (User-KNN) shows a small performance drop — possibly due to sparse user overlaps
- Method 4 (Item-KNN) improves over both baseline and Method 3
- Method 5 (Hybrid) achieves the lowest RMSE, showing best overall performance
- Combining user and item information in Method 5 helps reduce prediction error and improve robustness



Why Some Methods Perform Better

- Method 1 and Method 2 use global averages → limited personalization.
- Method 3 (User-KNN) underperforms slightly due to sparse user overlap in test set.
- Method 4 (Item-KNN) performs better — items are more densely rated, making similarity more reliable
- Method 5 (Hybrid) consistently achieves the lowest error in both MAE and RMSE.
- Hybrid model balances user behavior and item similarity → more robust across sparse or cold-start scenarios
- Personalization and combined signals lead to better predictions.

Conclusion

- Accuracy improves as methods shift from global assumptions to personalized, similarity-aware predictions.
- Method 5 outperforms by integrating both user- and item-level insights.

THANK YOU