Assignment 3

Collaborative Filtering - Slope One Scheme

S3961136 - Benjaporn Wongmayura

1. Slope One Scheme

The Slope One algorithm is a collaborative filtering technique that generates **item-based recommendations**. It operates by **computing average rating differences between pairs of items**, which are then utilized to **predict missing ratings**.

$$P^{S1}(u)_j = \bar{u} + \frac{1}{card(R_j)} \sum_{i \in R_j} \mathrm{dev}_{j,i}.$$

	Item 1	Item 8
Alice	3	?
Bob	4	5

How does it work?

Take the average of these differences of the co-ratings to make the prediction

$$P(\text{Alice}, \text{Item8}) = 3 + (5 - 4) = 4$$

Steps

1. **Data Preparation:** Start with a dataset of user ratings represented as a matrix.
2. **Rating Difference Computation:** Calculate the rating differences between items based on the ratings of users who have rated both items.
3. **Prediction Generation:** For a target user, predict ratings for items they haven't rated using the average rating differences associated with the items they have rated.
4. **Aggregation and Recommendation:** Aggregate the predicted ratings for unrated items and generate a list of recommended items sorted by predicted ratings.

2. Weighted - Slope One Scheme

The Weighted Slope One Scheme was developed to overcome a limitation of the original Slope One Scheme, which did not consider the reliability or significance of rating differences between items. In the original scheme, the average rating difference between items was calculated based on all available ratings, without taking into account the number of users who had rated those items.

$$P^{wS1}(u)_j = \frac{\sum_{i\in S(u)-\{j\}}(\text{dev}_{j,i}+u_i)c_{j,i}}{\sum_{i\in S(u)-\{j\}} c_{j,i}}$$

where $c_{j,i} = card(S_{j,i}(\chi))$.

Whereas the Weighted Slope One Scheme assigns weights to rating differences based on the number of users who have rated both items. This means that **rating differences derived from a larger number of users are given more importance** or significance during the weighted averaging process.

3. How the parameter λ affects the performance of the personalised weighted Slope One method?

The higher the Lambda(λ), the less MAE, the better the prediction result.

```
print(evaluate(test_ds, np_predictions))
```

```
(0.8097533038472071, 1.0268274969416544)
```

(a) $\lambda = 0$

```
print(evaluate(test_ds, np_predictions))
```

```
(0.7573963167317079, 0.9643263570586944)
```

(b) $\lambda = 0.5$

```
print(evaluate(test_ds, np_predictions))
```

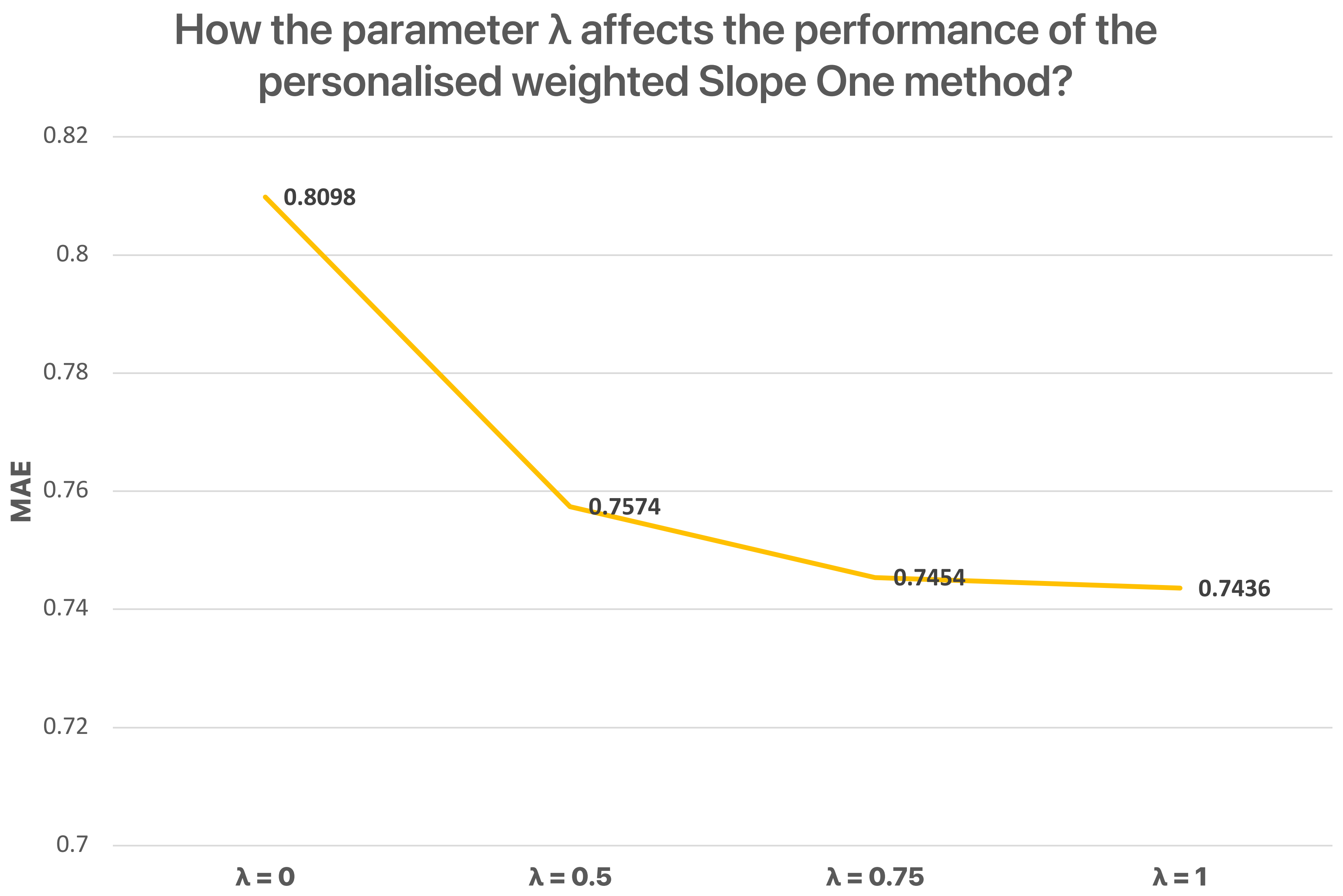
```
(0.7454351672626034, 0.9514827186178385)
```

(c) $\lambda = 0.75$

```
print(evaluate(test_ds, np_predictions))
```

```
(0.7435957831804696, 0.9516747501357116)
```

(d) $\lambda = 1$



As the value of Lambda (λ) increases, the MAE decreases, indicating improved prediction performance.

Why the personalised weighted Slope One method is better than normal weighted Slope One method?

The personalized weighted Slope One approach exhibits superior performance in comparison to the conventional weighted Slope One method proposed by Daniel Lemire and Anna Maclachlan. This is substantiated by the observation of lower MAE.

Scheme	EachMovie	Movielens
BI-POLAR SLOPE ONE	0.194	0.188
WEIGHTED SLOPE ONE	0.198	0.188
SLOPE ONE	0.200	0.188

```
print(evaluate(test_ds, np_predictions))
```

```
(0.7435957831804696, 0.9516747501357116)
```

(d) λ = 1

```
print(evaluate(test_ds, np_predictions))
```

```
(0.7454351672626034, 0.9514827186178385)
```

(c) λ = 0.75

< 0.188*4 =0.752 → Better Prediction

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”:

Appendix

Collaborative Filtering (CF)

- Widely used technique in recommendation systems to offer personalized recommendations to users.

- Users who share similar preferences or behaviours can **benefit from recommendations** made by **other like-minded users**.

- Unlike other recommendation methods, collaborative filtering **does not rely on** explicit knowledge about the **recommended items** but instead taps into the collective knowledge of the **user community**.

Why do we prefer Collaborative Filtering to other recommendation techniques?

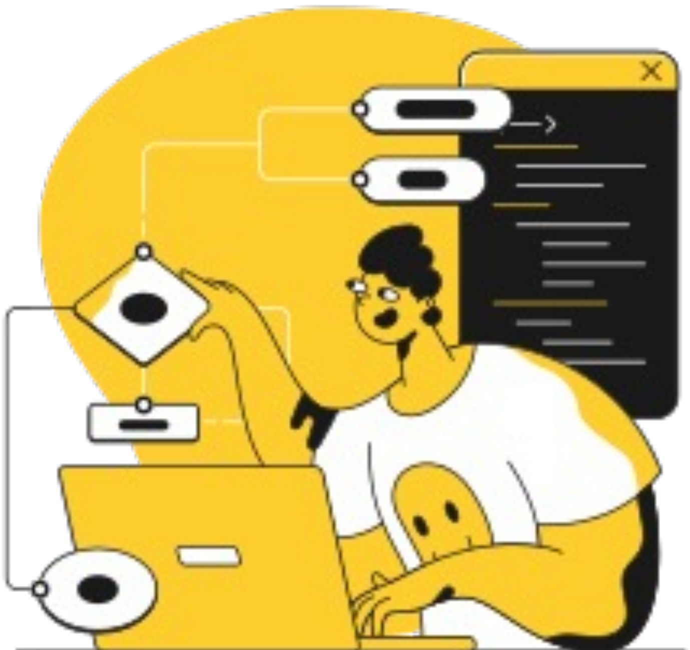
- Collaborative filtering is employed to address the limitations of alternative recommendation approaches, like content-based filtering. **Content-based methods rely on item attributes or user profiles** to generate recommendations, which can be restrictive when there is inadequate information about the items or users.

- In contrast, collaborative filtering **exploits user preferences and behaviour to provide recommendations,** making it more effective when item attributes are unavailable or challenging to analyze.

1.User-Centric Approach: Collaborative filtering focuses on the preferences and behaviours of users themselves rather than relying solely on item attributes or user profiles.

2.No Dependency on Item Attributes: Collaborative filtering can work effectively even when item characteristics are unknown or difficult to obtain.

3.Serendipitous Recommendations: Collaborative filtering can provide serendipitous recommendations by suggesting items that users may not have discovered on their own.



4. **Scalability and Adaptability:** Collaborative filtering techniques, such as memory-based approaches, are scalable and can handle large datasets with a large number of users and items. Additionally, collaborative filtering can adapt to changes in user preferences and behaviors over time, allowing the recommendations to remain relevant and up-to-date.

5. **Cold Start Problem Mitigation:** Collaborative filtering can handle the "cold start" problem, which refers to situations where there is limited or no information available for new users or items.

6. **Diversity of Recommendations:** Collaborative filtering has the potential to offer diverse recommendations by considering the preferences of multiple users. It can provide a broader range of options, reducing the risk of presenting users with repetitive or redundant recommendations.