DAT410/DIT728 Design of AI systems
Module 2

Assignment 2: Group 54

January 30, 2024

Group members:

Member 1:
Name: Hongguang Chen
Program: MPHPC
Uni email: chenhon@chalmers.se
Time spent: 20 hrs

Member 2:
Name: Sanjaya Thilak Bandara Asurasinghe Rajamantrilage
Program: MPDSC
email: arstbandara@gmail.com
Uni email: sanraj@chalmers.se
Time spent: 20 hrs

The declaration:
"We hereby declare that we have both actively participated in solving every exercise. All solutions are entirely our own work, without having taken part of other solutions."

1 Reading and reflection

Takeaways from the Papers:

Both two paper provided valuable insights into the design features of recommendation systems, offering innovative approaches to address the challenges of predicting user preferences. Two key design features that stood out are the incorporation of diverse models and the utilization of weighted neighborhood information.

Diverse Model Integration:
Both papers emphasized the important of using a diverse set of models to improve recommendation accuracy. The Netflix Prize encouraged the use of models like KNN, latent factor models (SVD, Restricted Boltzmann Machines), and neighborhood-aware factorization methods. This diversity allowed for a more comprehensive understanding of user preferences, capturing both localized relationships and overall structural patterns. The Lessons paper extended this idea by introducing complementary perspectives, such as the binary view, to further improve model performance. This design feature acknowledges that no a single model can do well in all scenarios, and a combination of models can compensate for individual weaknesses.

Weighted Neighborhood Information:
Both papers highlighted the importance of integrating more neighborhood information in recommendation models. The Netflix Prize introduced interpolation weights for neighborhood models, allowing for a more refined representation of how similar items or users contribute to predictions. The second paper delved into refining these weights by addressing issues like the variability among neighbors and the potential lack of informative neighbors. The introduction of adjustable weights and regularization techniques aimed to improve the reliability of neighborhood-based predictions.

Application Differences:
In another application of recommendation systems, such as e-commerce product recommendations, the diversity of models might differ. For example, collaborative filtering and content-based models would be more relevant. The specific characteristics of the

application domain would influence the selection of models, involving emotion and behavior analysis, which might not have been paid enough attention in the Netflix Prize challenge (no relevant data in provided dataset).

In conclusion, the design features of various model integration and weighted neighborhood information in the papers offer valuable lessons for recommendation systems. While these features remain fundamental, their application may vary based on the characteristics and requirements of different recommendation domains, necessitating tailored approaches to achieve optimal results.

2 Implementation

2.1 System design

Firstly, we conduct an analysis of the provided data, which consists of 2000 movies and 600 users. Our examination involves two perspectives: the distribution of user ratings for each movie (refer to Figure 1) and the distribution of movies rated by each user (refer to Figure 2). The observations reveal that, on average, each movie is rated by approximately 5 to 15 users, while each user rates around 10 to 30 movies. These numbers, however, are relatively small compared to the total counts of 2000 movies and 600 users.
Therefore, in the initial phase, we opt to implement a system based on the K-NN model. Subsequently, if feasible, we plan to enhance its performance using the methods outlined in the second paper, "Lessons from the Netflix Prize Challenge."

For the KNN model, we choose $K = 50$, a value within the range of 20 to 50, as suggested in the paper, which has shown to yield optimal results.

To estimate the rating of a user for a movie, we employ the following equation:

$$r^*_{ui} \leftarrow b_{ui} + \frac{\sum_{j\in N(i,u)} s_{ij}(r_{uj} - b_{uj})}{\sum_{j\in N(i,u)} s_{ij}} \tag{1}$$

1.b_{ui} represents the bias rating predicted by the average mean rating of user u.
2.$N_{(i,u)}$ denotes the set of neighbors of movie i, rated by user u, sharing similar features as computed by the aforementioned KNN model.
3. s_{ij} represents the similarity between movie i and j. In our analysis, we have explored two methods for measuring this similarity: the "Pearson correlation coefficient" and the "closely related cosine similarity."
4. For movies that do not have neighboring movies, we utilize b_{ui} as the estimated result.

The entire process can be summarized as follows:
1. Choose a user and make predictions for all movies that the user has not watched.
2. Sort the predicted results and select the top five movies.

Clearly, the system is quite simple and straightforward to implement. In fact, the testing results appear reasonable. However, a notable issue arises as depicted in Figure-3, where nearly half of the movies are rated 5. Additionally, the count of neighboring movies for a specific user and a movie is often too low. To address this concern, our intention is to incorporate a latent factor module into the system for optimization and utilizing the information between neighboring movies (as what mentioned in the paper). Unfortunately, due to time constraints, we were unable to implement this enhancement.

Furthermore, as highlighted in the paper, regularization plays a crucial role in enhancing performance. Our system can be further optimized by incorporating regularization techniques, as suggested in the literature.

2.2 Test Result

See table 1 to table 5

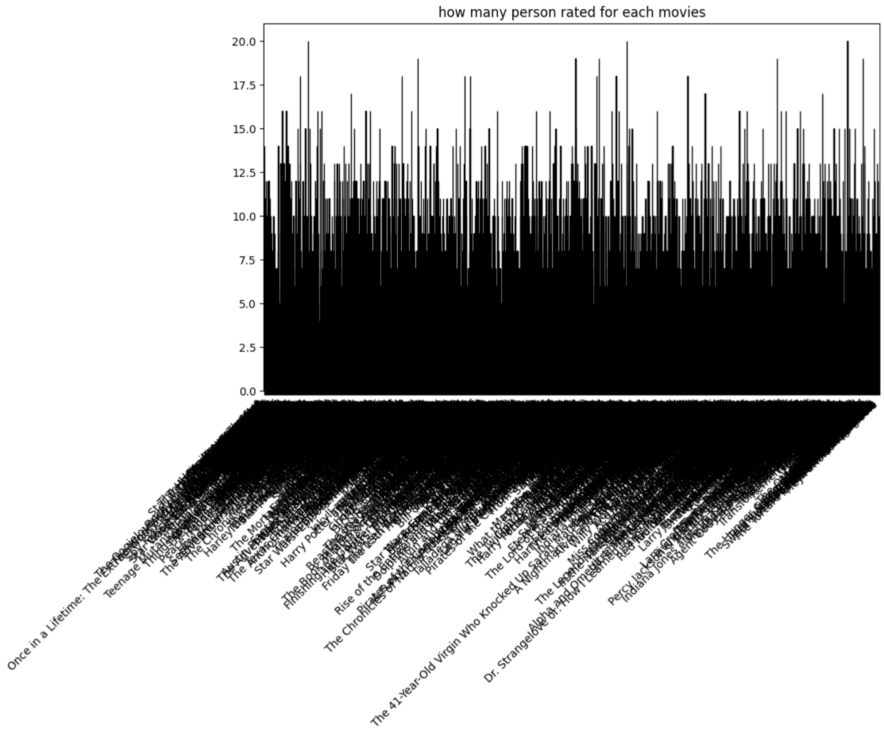

Figure 1: How many user rate for each movie

movies	estimated rating	genres
Prom	5.0	'comedy', 'drama'
The American President	5.0	'comedy', 'drama', 'romance'
Theresa Is a Mother	5.0	'comedy', 'drama', 'music'
Ghost World	5.0	'comedy', 'drama'
The Net	5.0	'action', 'crime', 'drama', 'mystery', 'thriller'

Table 1: Vincent

3 Discussion

Evaluating the quality of a recommendation system is difficult because of several inherent difficulties.

Firstly, there is a lack of a universally metric for measuring recommendation system effectiveness and performance. Different condition have different expectations and criteria, making it hard to build a general evaluation rules.

The "cold start" problem further complicates estimation, because new items or users lack sufficient historical data for predicting.

The dynamic nature of user preferences and evolving content presents another challenging. User preference change over time, and new items continuously enter the system, making it hard to create a static evaluation framework that remains high accuracy.

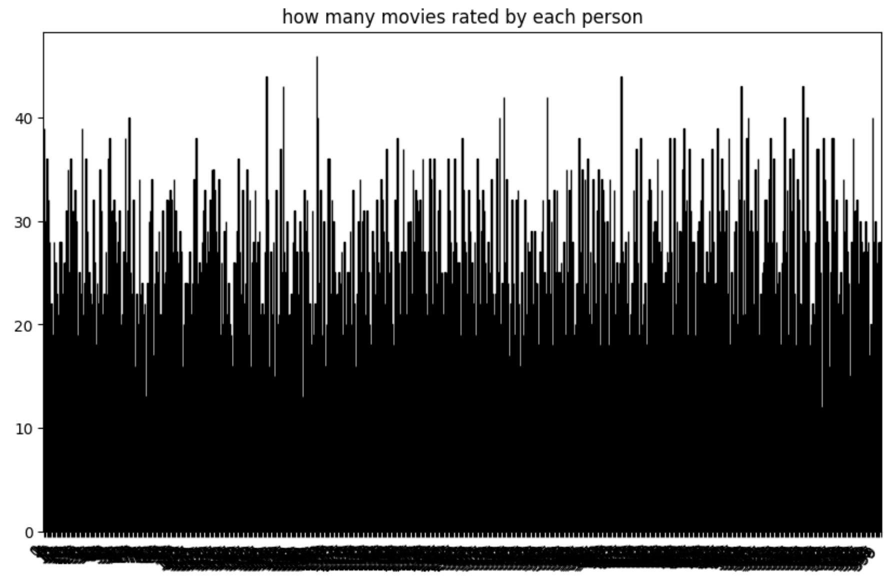

Figure 2: How many movies rated by each user

movies	estimated rating	genres
Running Scared	5.0	'action', 'crime', 'drama', 'thriller'
Face/Off	5.0	'action', 'crime', 'sci-fi', 'thriller'
Safe Haven	5.0	'drama', 'romance', 'thriller'
The Descendants	5.0	'comedy', 'drama'
The Net	5.0	'action', 'crime', 'drama', 'mystery', 'thriller'

Table 2: Edgar

movies	estimated rating	genres
This Is England	5.0	'crime', 'drama'
Butterfly	5.0	'crime', 'drama'
Sonny with a Chance	5.0	'comedy', 'family', 'romance'
Sahara	5.0	'action', 'adventure', 'comedy', 'thriller'
Piranha 3D	5.0	comedy', 'horror'

Table 3: Addilyn

movies	estimated rating	genres
N-Secure	5.0	'crime', 'drama', 'thriller'
Veronica Guerin	5.0	'biography', 'crime', 'drama', 'thriller'
The Reef	5.0	'horror', 'thriller'
Transsiberian	5.0	'crime', 'drama', 'mystery', 'thriller'
Freaky Deaky	5.0	'comedy', 'crime', 'thriller

Table 4: Marlee

movies	estimated rating	genres
TMNT	5.0	'action', 'adventure', 'animation', 'comedy', ...
Yu-Gi-Oh! Duel Monsters	5.0	'action', 'adventure', 'animation', ...
Alpha and Omega 4	5.0	'action', 'adventure', 'animation', 'comedy', 'drama', ...
Mulan	5.0	'adventure', 'animation', 'family', 'fantasy', 'musical', 'war'
The Princess and the Frog	5.0	'animation', 'family', 'fantasy', 'musical', 'romance'

Table 5: Javier

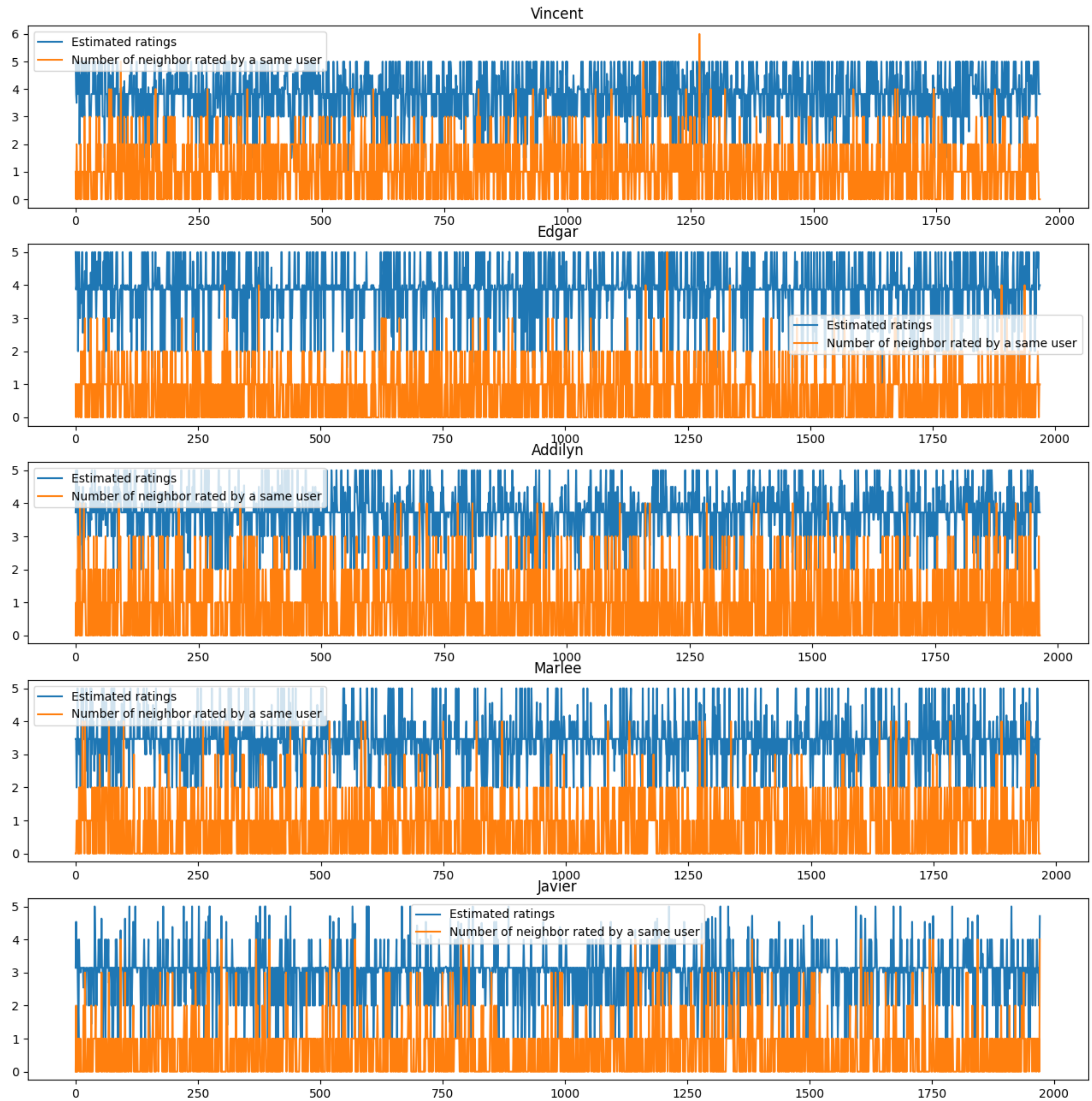

Figure 3: Five user, estimated rating for all un-watched movies