Ahsanullah University of Science and Technology

Department of Computer Science and Engineering

Soft Computing Lab

CSE 4238

Assignment 01

Submitted To

Nibir Chandra Mandal
Lecturer, CSE, AUST

Sanzana Karim Lora
Lecturer, CSE, AUST

Submitted By

Fairuz Nawar

17.01.04.122

July 25, 2021

Contents

1. Loss Curves

1.1 Optimal Hyperparameters

Iterations :

For K = 700, a global minimum was found at iteration 4. So, instead of breaking the loop, the model was run for a total of 400 iterations. Among the iterations, there were numerous local minimum found with values around 4.74 for 2 decimal places. The global maximum of the 400 iterations was achieved at iteration 285 which is 4.7391, The difference is approximately 0.01 between the 44th and 285th iterations.
Hence, Around 50-75 iterations were used for tuning K as it requires less time than 285 iterations.
The Figure 2 demonstrates the values MSE with respect to of Iterations.

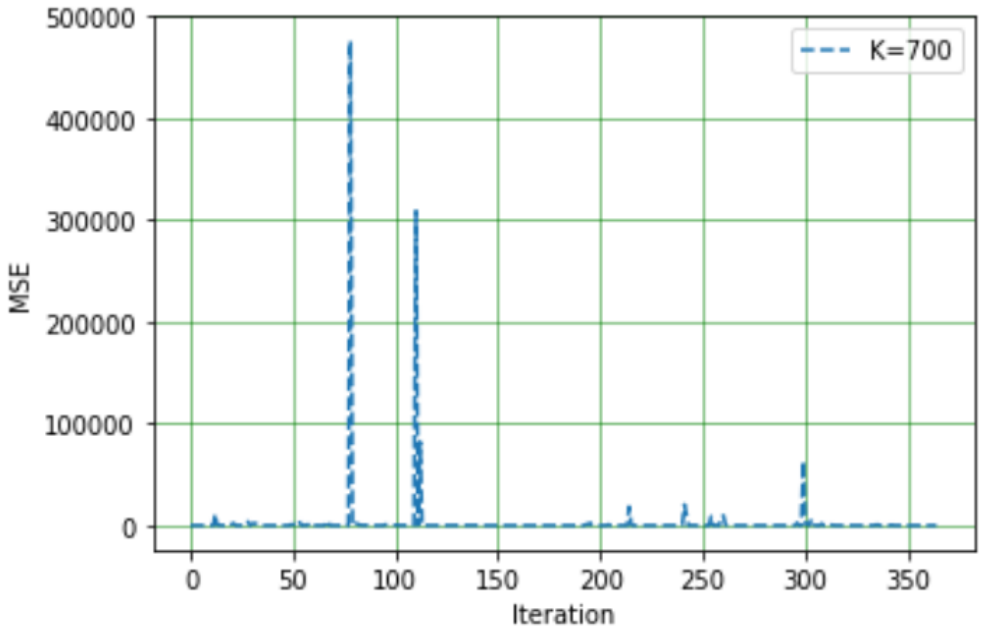
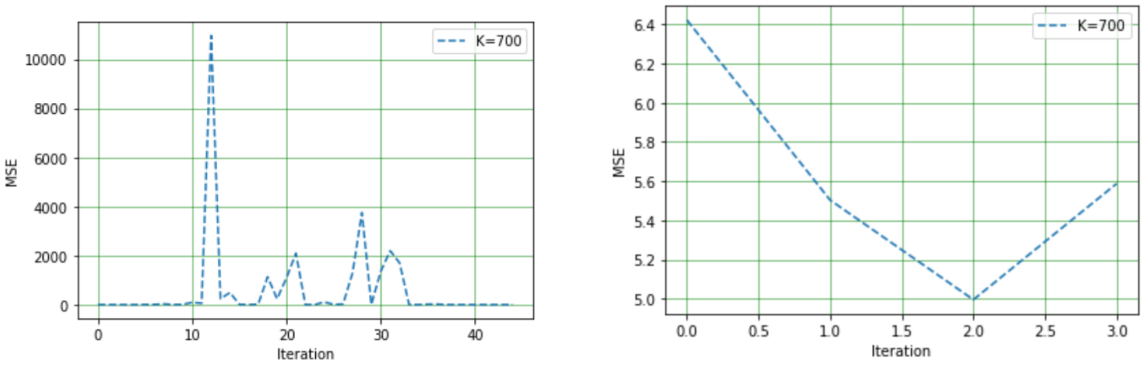


Figure 1: MSE vs Iterations



(a) (b)

Figure 2: MSE Vs (a) 45 iterations and (b) 4 iterations

Iteration	MSE
4	**4.99573**
24	4.91118
44	**4.74989**
69	5.59200
86	11.49184
94	4.75862
100	7.67668
124	4.77789
144	4.76050
177	4.77363
186	4.74734
188	4.74549
218	4.74325
221	4.74324
228	4.74108
238	4.74097
271	4.73961
277	4.73930
281	4.73926
285	**4.73913**
365	4.74977

Figure 3: Local minimum MSE for different Iterations

K :

The following table lists the result of the experiments done for tuning the value of K. Higher values of K lead to better performance of the model.

K	MSE	Iterations
100	47.80791	5
500	5.26483	4
500	5.12035	50
1000	5.70539	2
1000	4.26205	26
2000	2.80037	37
4000	0.71699	61
5000	**0.25643**	19

1.2 Cosine Similarity

From sklearn library, cosine_similarity function was used to generate both the cosine similarity of Users and Movies. But the process was not successful as the 12GB RAM of Google Colab was unable to allocate all information, hence the function crashed.

1.3 Movie Suggestion

For a particular User, such as 1050, the following top 5 movies were suggested, which have not yet been rated by the user.

- 1065 [Predicted Rating = 6.522]
- 549 [Predicted Rating = 6.497]
- 4866 [Predicted Rating = 6.446]
- 4965 [Predicted Rating = 6.431]
- 3242 [Predicted Rating = 6.723]

1.4 Performance Analysis

The lowest MSE value of 0.25 was generated for K = 5000 for 19 iterations. Following is the loss curve of the optimal output.

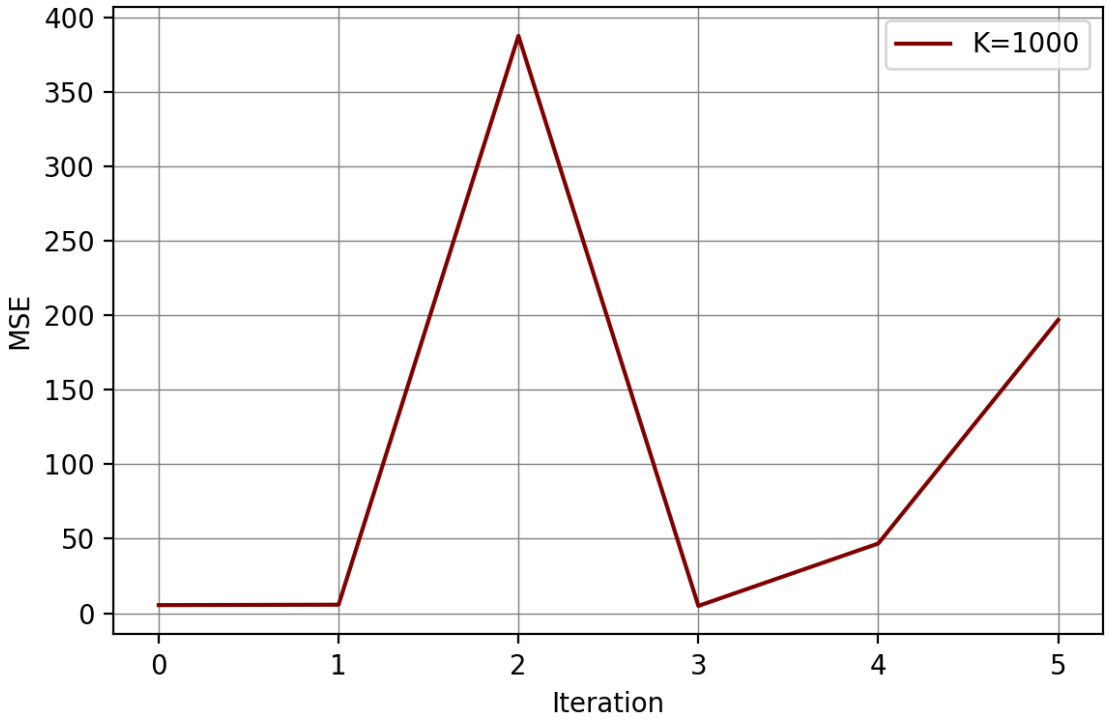


Figure 4: Loss curve for Optimal Output

1.5 Applications of the study

Matrix Factorization can be used for a number of applications, mostly recommection systems. Some applications are -

- Recommentation System of Streaming services - Netflix, Disney+, HBO Max etc
- Music, Book, Game Genre classification
- Recommendation system for music - Spotify, Youtube etc
- Product recommendation, filtering, searching
- Content discovery platforms
- Decision Support System
- Text mining