ML Homework 5

3.c)

Text

Description automatically generated

3.d)

The best value would be related to Item based CF. But the value can change depending on the similarity distance being used. For example, when I tested the User based and Item based with cosine distance then the best model was User based and not Item based. The above code was done using the default MSD similarities.

3.e)

Chart, line chart

Description automatically generated

Chart, line chart

Description automatically generated

As we can observe it from the graph being plotted, the values are being consistent with the impact of the similarity metrics. Item based and User based show similar trends with each other while using different similarity trends.

3.f)

A picture containing graphical user interface

Description automatically generated

A picture containing shape

Description automatically generated

3.g)

From understanding of the above images, we can say that the best range for K would be as follows: -

1. User based CF: - The best k would be in the range of ‘8-10’ (The best in my case is 9). The value would be in the range of 0.9608. (The same has been solved using the code. It is present in the file)
2. Item based CF: - The best k would be in the range of ’19-21’ (The best in my case is 20). The value would be in the range of 0.934. (The same has been solved using the code. It is present in the file)

Code Upload Link: -

<https://drive.google.com/drive/folders/14ksTYGLYqOG8fkymfgv4xbt0uQ5c2dXP?usp=sharing>