

NEURAL NETWORK COLLABORATIVE FILTERING

Final Assignment - Sekolah Data Pacmann

Stefanus Yudi Irwan – January 2024

Table of Content



Problem



Dataset



Model NCF



Model Comparison



Conclusion



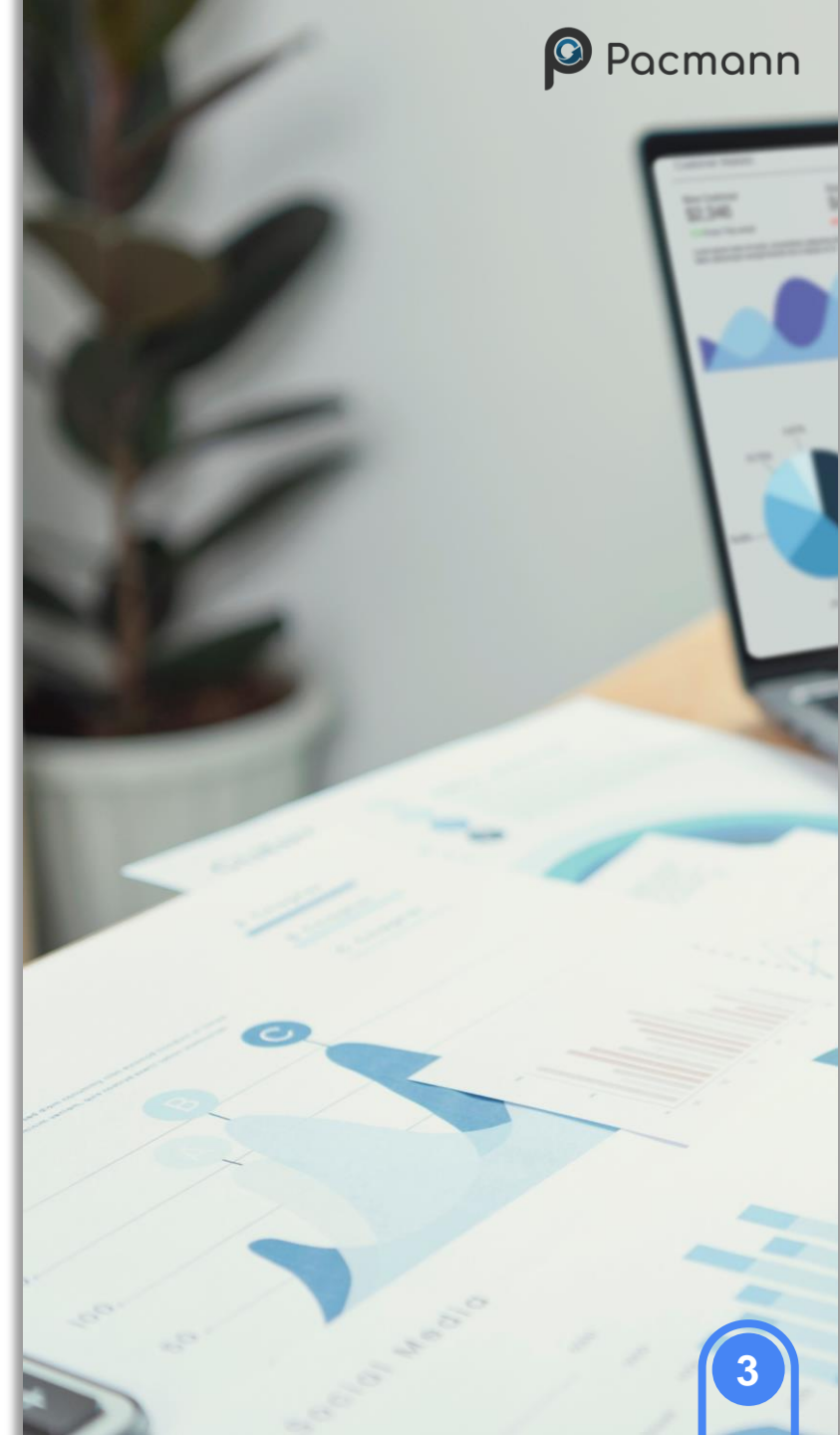
① PROBLEM

Business Problem

- A streaming platform **nonton-yuk.com** is having a problem with its user retention.
- In 3 months, the user retention rate dropped almost 15% which was really affects **nonton-yuk.com** revenues.
- After doing an urgent user research, **nonton-yuk.com** teams found that **users find it difficult** to browse movie in **nonton-yuk.com** which has nearly ~4000 movies

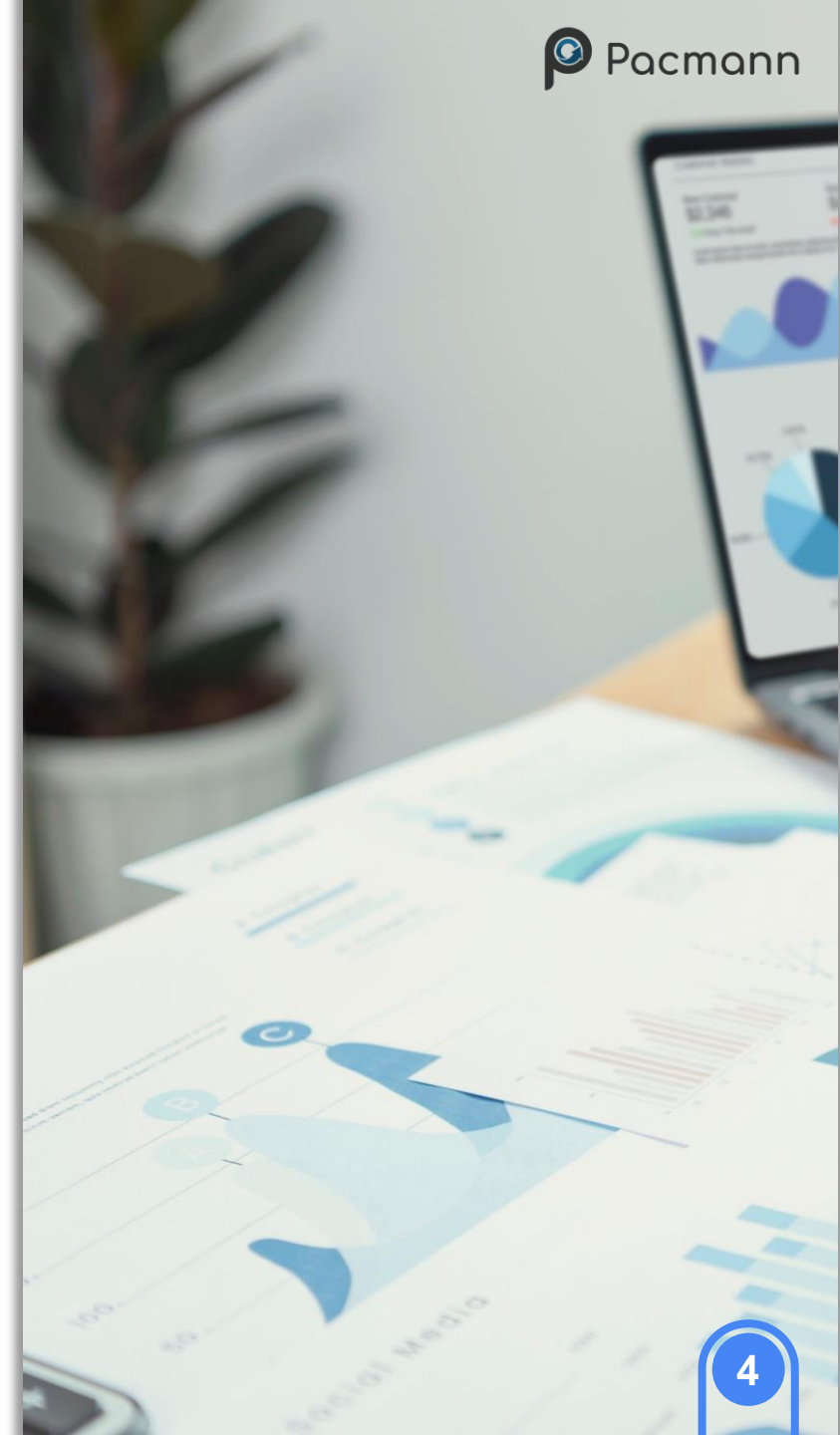


Retention rate drop ~15%



Business Objective & Solution

- Our business objective would be **increasing user retention** by **15%** within 3 months.
- **Movie recommendation** will be created to help **users browse** the movie **easily** --> remove the users difficulty in using **nonton-yuk.com** platform.



②

DATASET

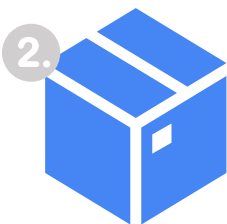


Dataset Overview

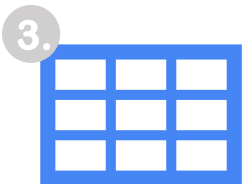
grouplens



Raw Data comprises of **3 txt files, user data, movie data, ratings data.**



Size of the raw data is **24 Mega-Byte**



Raw Data comprises of **6040 User, 3952 Movie, and 1 Million interaction**

Raw Data Composition

1

User Data

- UserID
- Gender
- Age
- Occupation
- Zip Code

2

Movie Data

- MovieID
- Title
- Genres

3

Ratings Data

- UserID
- MovieID
- Rating
- Timestamp



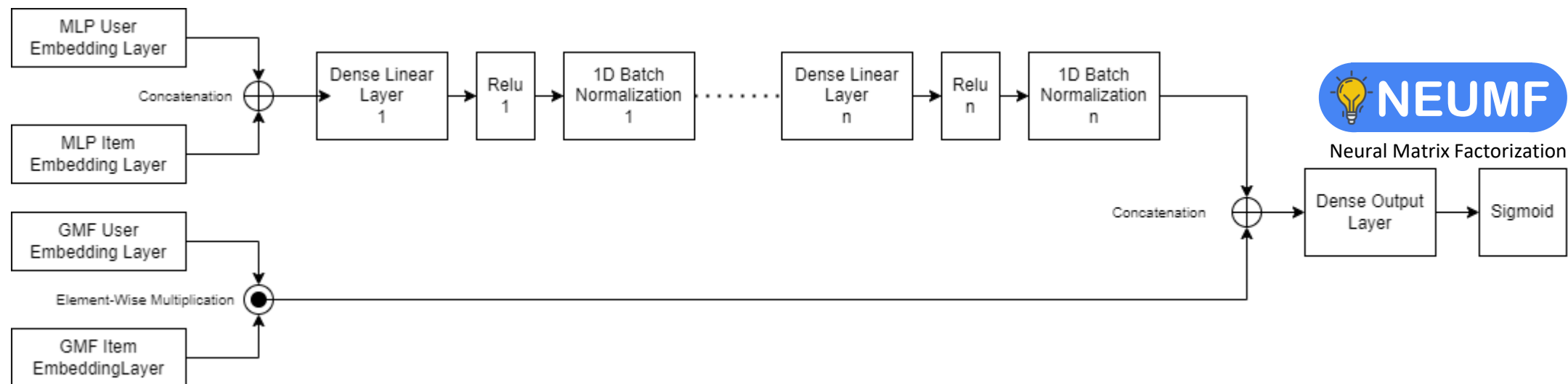
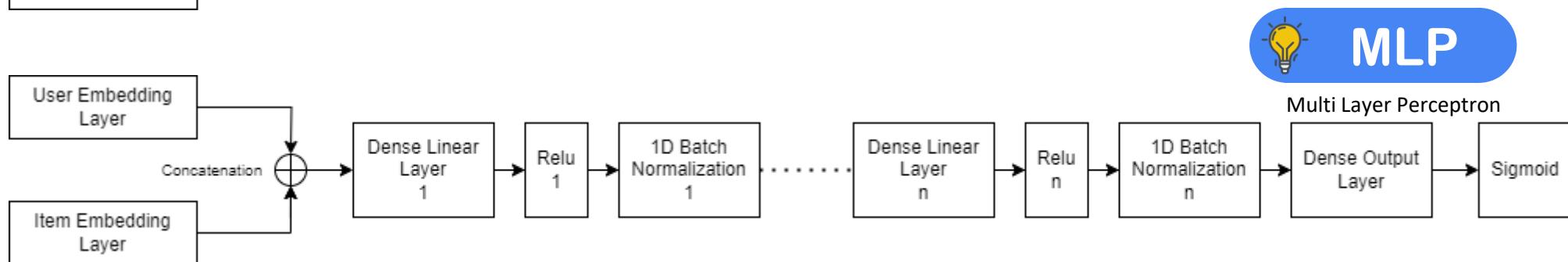
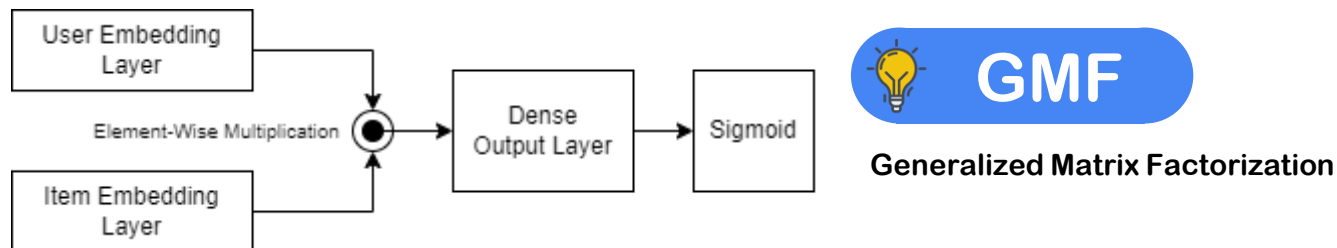
③

FCN

Neural Collaborative Filtering

 [NCF Research Paper](#)

 [NCF Repository Reference](#)

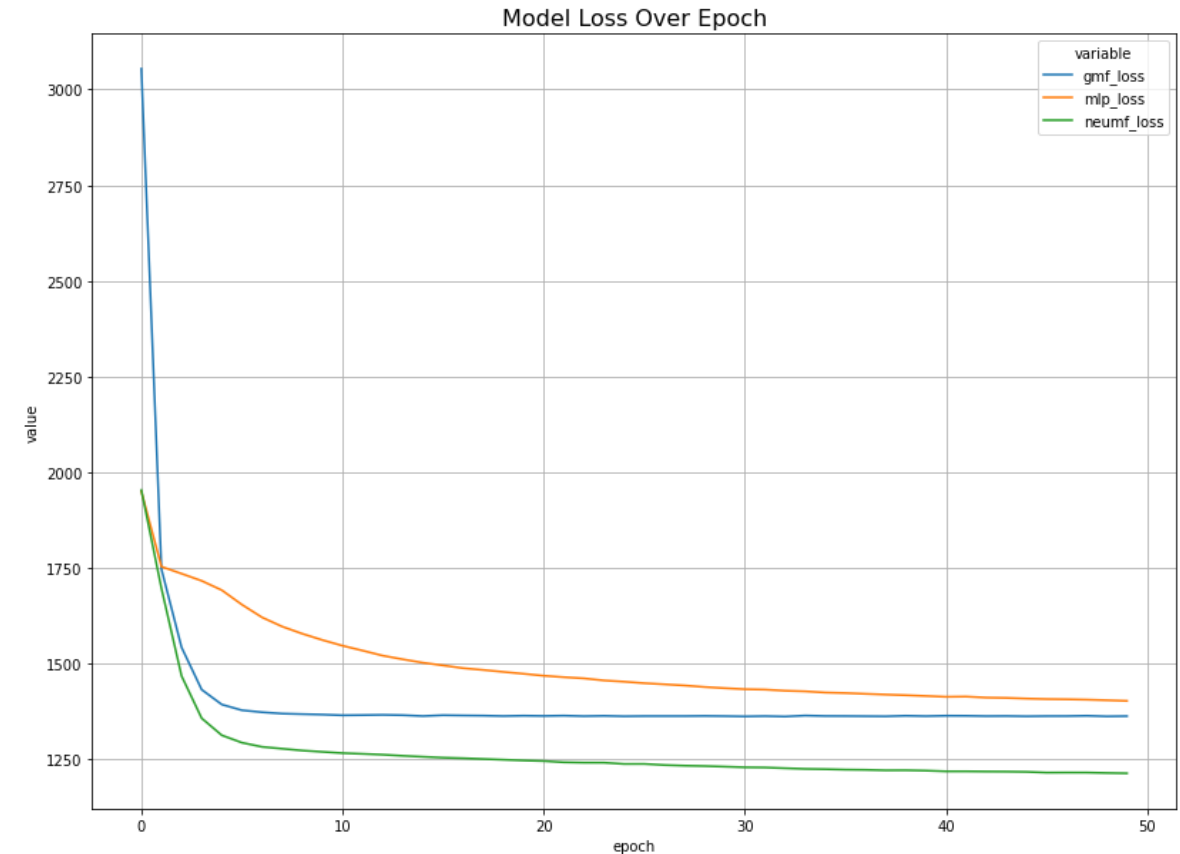
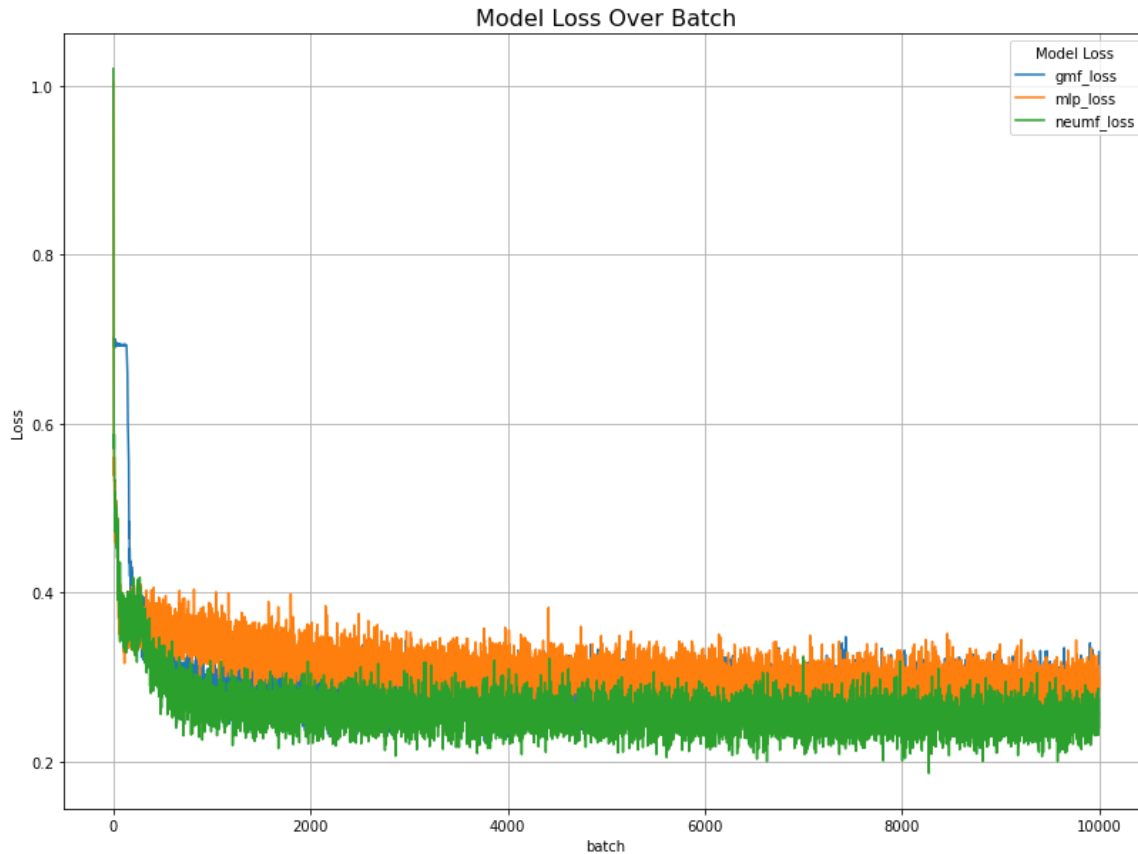


④

MODEL COMPARISONS



Analysis : Loss Over Training Attempt

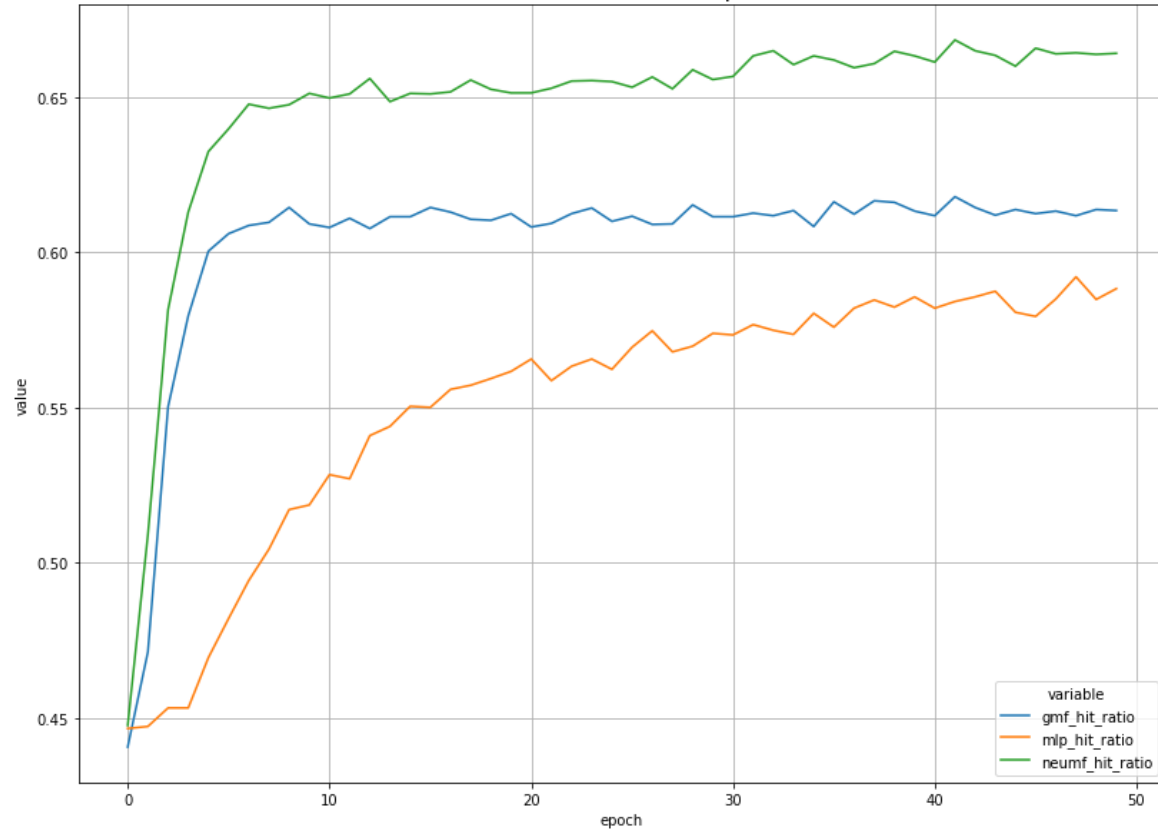


Insight.

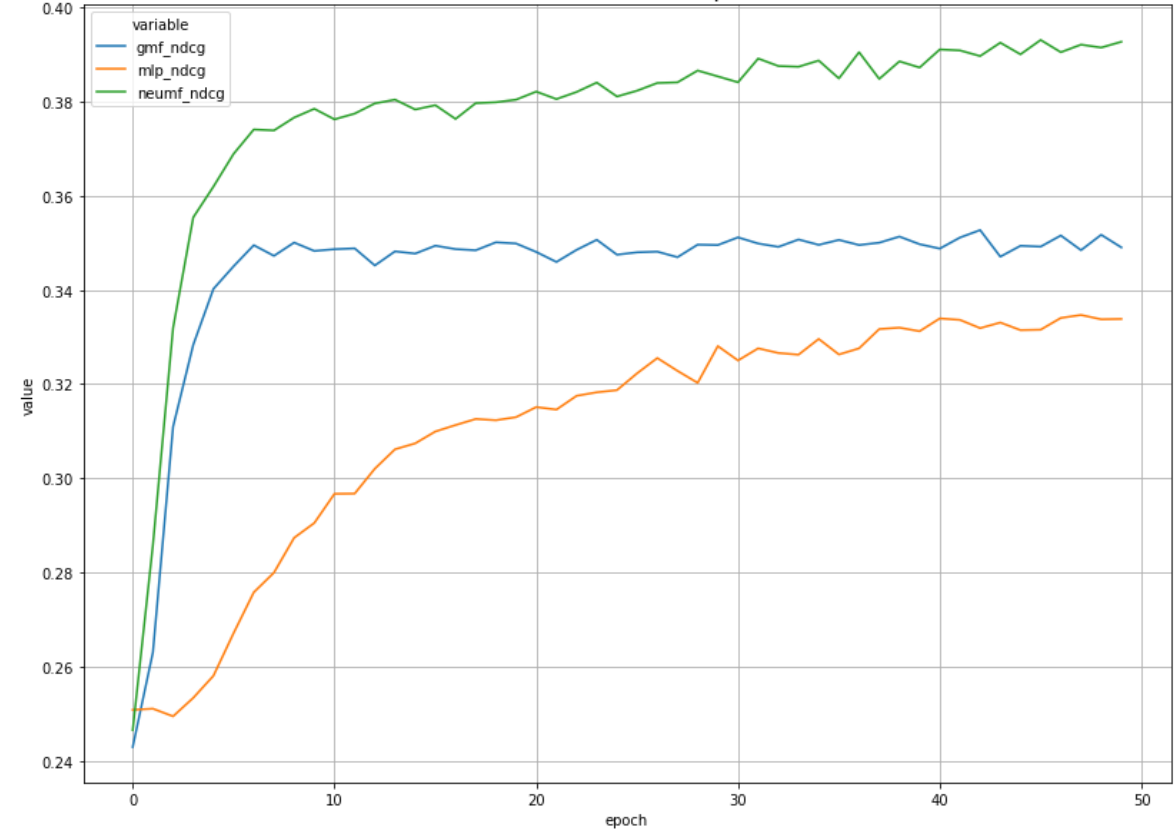
- Both loss over batch as well as epoch, is decline towards minimum value. With final value of NEUMF model losses is the smallest, and MLP model losses is the largest
- Initial total loss over epoch of GMF model is the highest, compared to MLP or NEUMF, this is probably be the effect of a fewer layer of GMF than MLP or NEUMF

Analysis : Recommendation System Evaluation

Model Hit Ratio Over Epoch



Model NDCG Over Epoch



Insight.

- NEUMF model has the highest value of Hit Ratio, just above 65% compared to GMF at slightly above 60% and MLP just under 60%,
- NEUMF model also has the highest value of NDCG, just under 40%, whereas GMF at around 35%, and MLP a little of under 34%

⑤ CONCLUSIONS



Conclusion

- NEUMF model is better compared to GMF and MLP, yet the model architecture is the most complicated compared to MLP and GMF. This would require a lot of training time and memory to create recommendation system, hereby the cost is the most expensive
- On the other hand, although GMF is the simplest one, but the performance is better than MLP. For this reason, GMF model is a better choice if the budget is strict and training time becoming concern.

Thank You

