Toronto Restaurant Recommender System

Presented by Jiangning Luo and Hang Gong

Contents

PART 1

Introduction

Introduction

Motivation

There are thousands of restaurants in Toronto City. Sometimes, it is hard for consumers to decide which one they should visit.

Goal

Build a personalized recommendation system for the Yelp users in Toronto, recommend the most suited restaurants for the yelp users.

Method

Modelling users' performance on items based on their past interactions. Using the neural collaborative filtering, compare it with classical matrix factorization.

PART 2

Data Pipeline

Data Pipeline

Raw Data: .json files

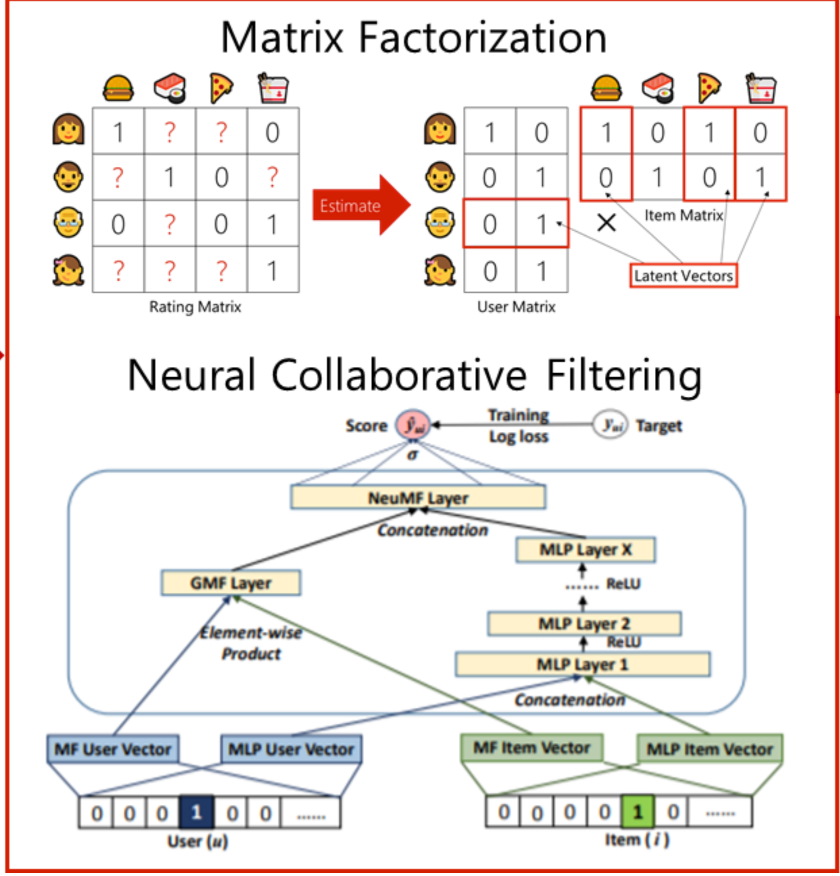
Business
0.2 Million Rows

Users
1.96 Million Rows

Reviews
8.02 Million Rows

Data Cleaning

Data Models



Evaluation

Data Cleaning

- Keeping the users with only a few reviews leads to a noisy training set and therefore leads to high bias.
- Eliminating too much users leads to a small training set and therefore leads to high variance.

The challenge here is deciding the users with less than what number of reviews need to be eliminated.

After the experiments, we decide to eliminate the users with less than 5 reviews.

number of items: 5417
number of users: 90665
number of reviews: 338328
sparsity: 99.93%

Original

Eliminate users with less than 5 reviews

number of items: 5417
number of users: 13651
number of reviews: 210467
sparsity: 99.55%

Cleaned

PART 3

Data Models

Interaction Matrix

We define the interaction matrix from the interactions between the users and the restaurants

- 1 if the user is interested in the restaurant
- 0 if the user have no interest in the restaurant

	🍔	🍣	🍕	🥡
👩	1	0	1	0
👨	0	1	0	0
👴	0	0	0	1
👧	0	1	0	1

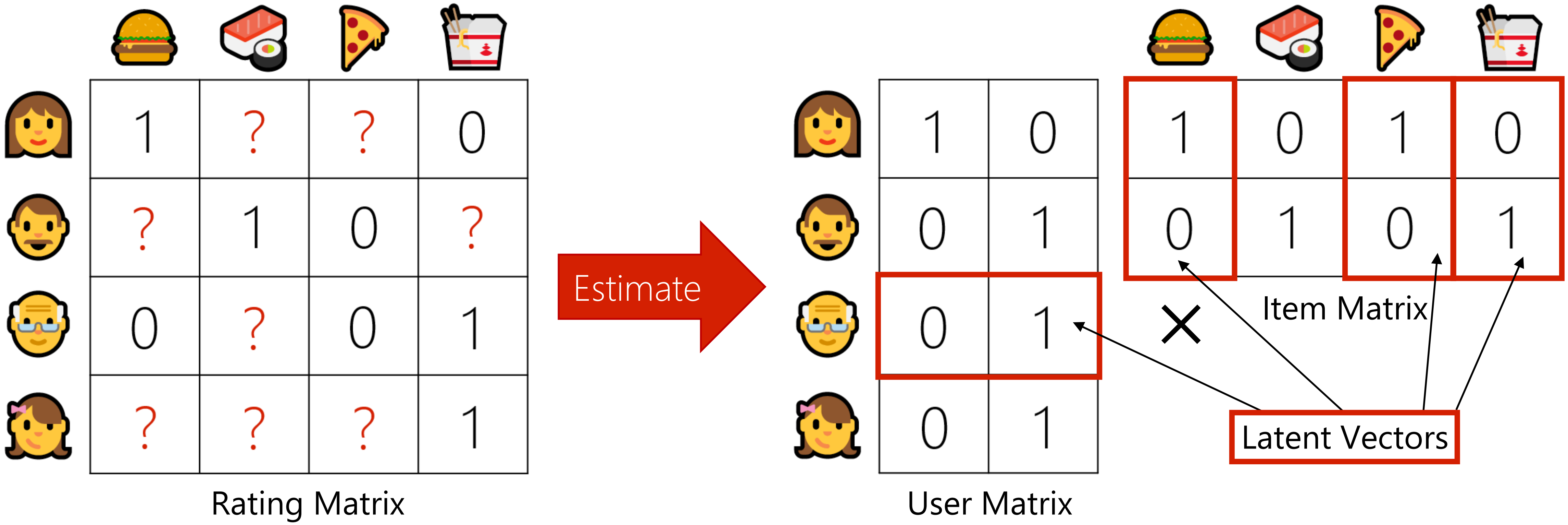
Goal of Our Recommender System

	🍔	🍣	🍕	🥡
👩	1	?	?	0
👨	?	1	0	?
👴	0	?	0	1
👧	?	?	?	1

Predict

	🍔	🍣	🍕	🥡
👩	1	0	1	0
👨	0	1	0	0
👴	0	0	0	1
👧	0	1	0	1

Matrix Factorization



Matrix Factorization

User Matrix

👩	1	0
👨	0	1
👴	0	1
👧	0	1

×

Item Matrix

🍔	🍣	🍕	🥡
1	0	1	0
0	1	0	1

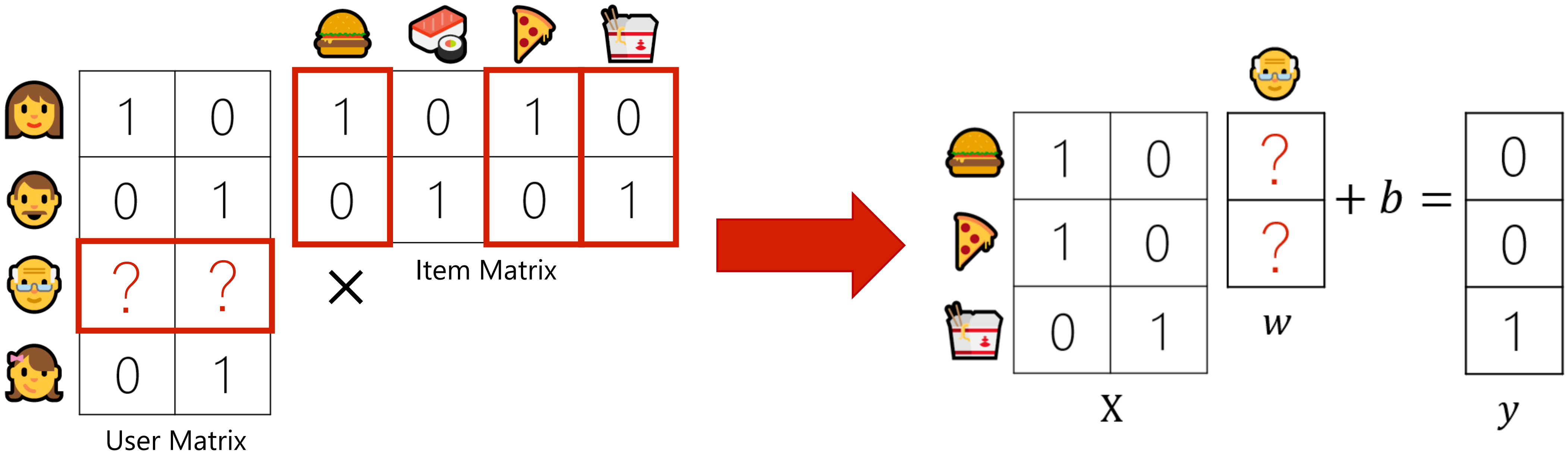
Predict →

Predicted Rating Matrix

	🍔	🍣	🍕	🥡
👩	1	0	1	0
👨	0	1	0	1
👴	0	1	0	1
👧	0	1	0	1

Matrix Factorization

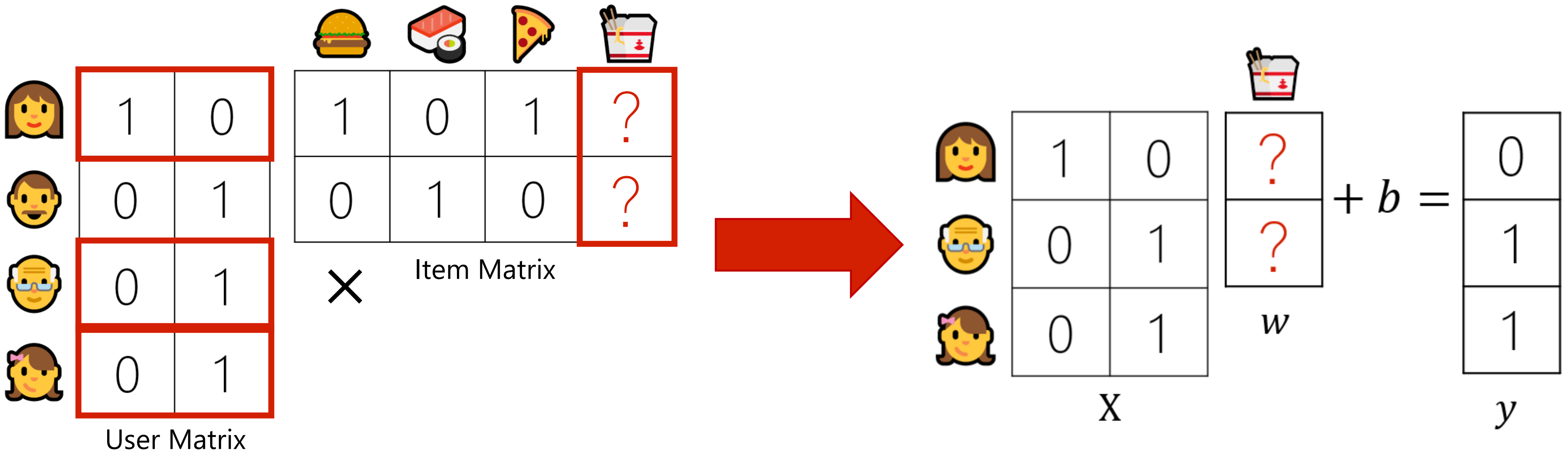
The training of matrix factorization is actually training a bunch of linear regression models.



Presented by Jiangning Luo and Hang Gong

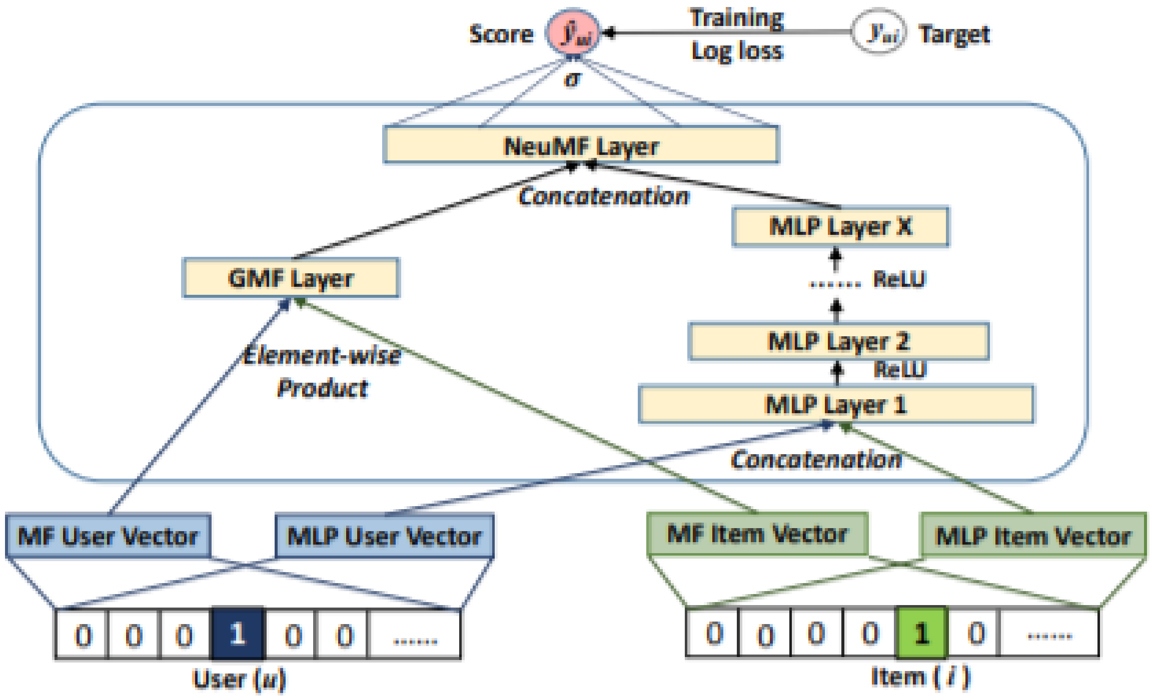
Matrix Factorization

Actually, Matrix Factorization problem is a bunch of linear regression problems.



Neural Collaborative Filtering

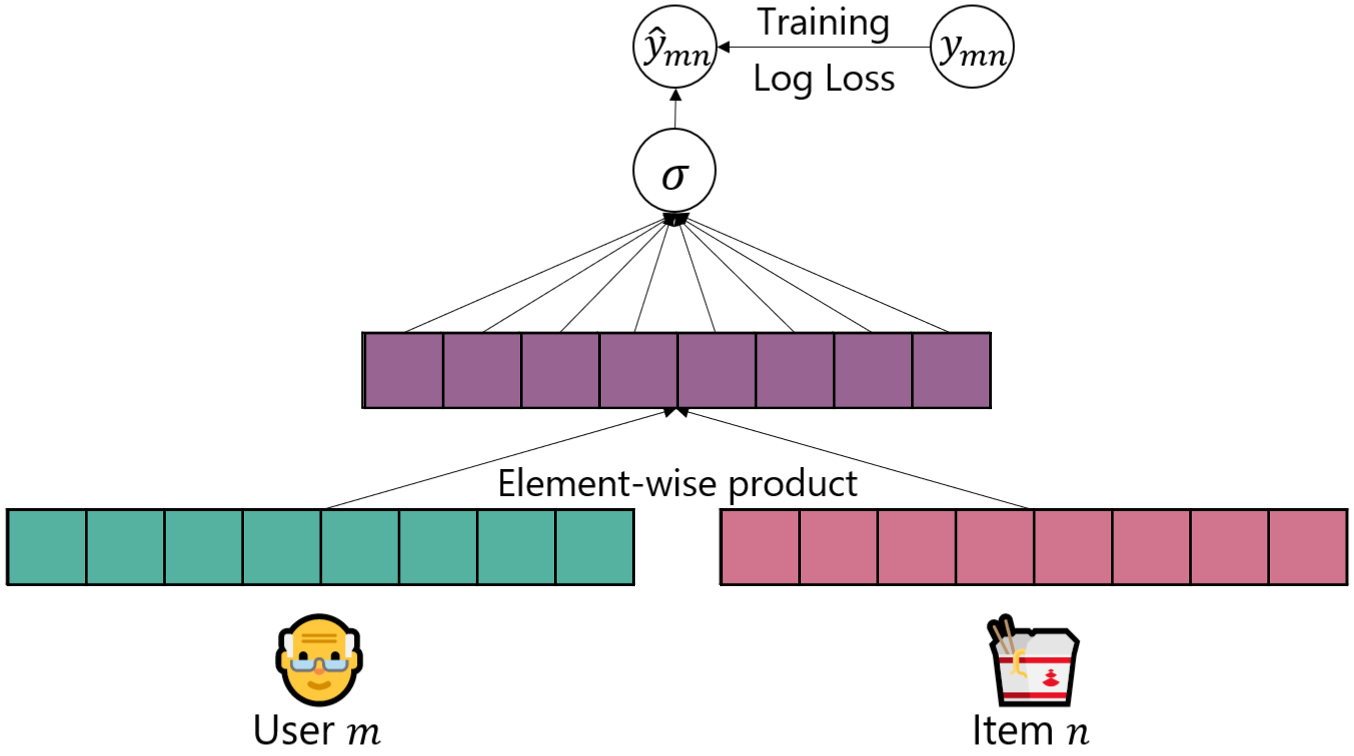
Fusion of generalized matrix factorization (GMF) and multi-layer perceptron (MLP) collaborative filtering.



* Image from Neural Collaborative Filtering by He et. al.

Generalized Matrix Factorization

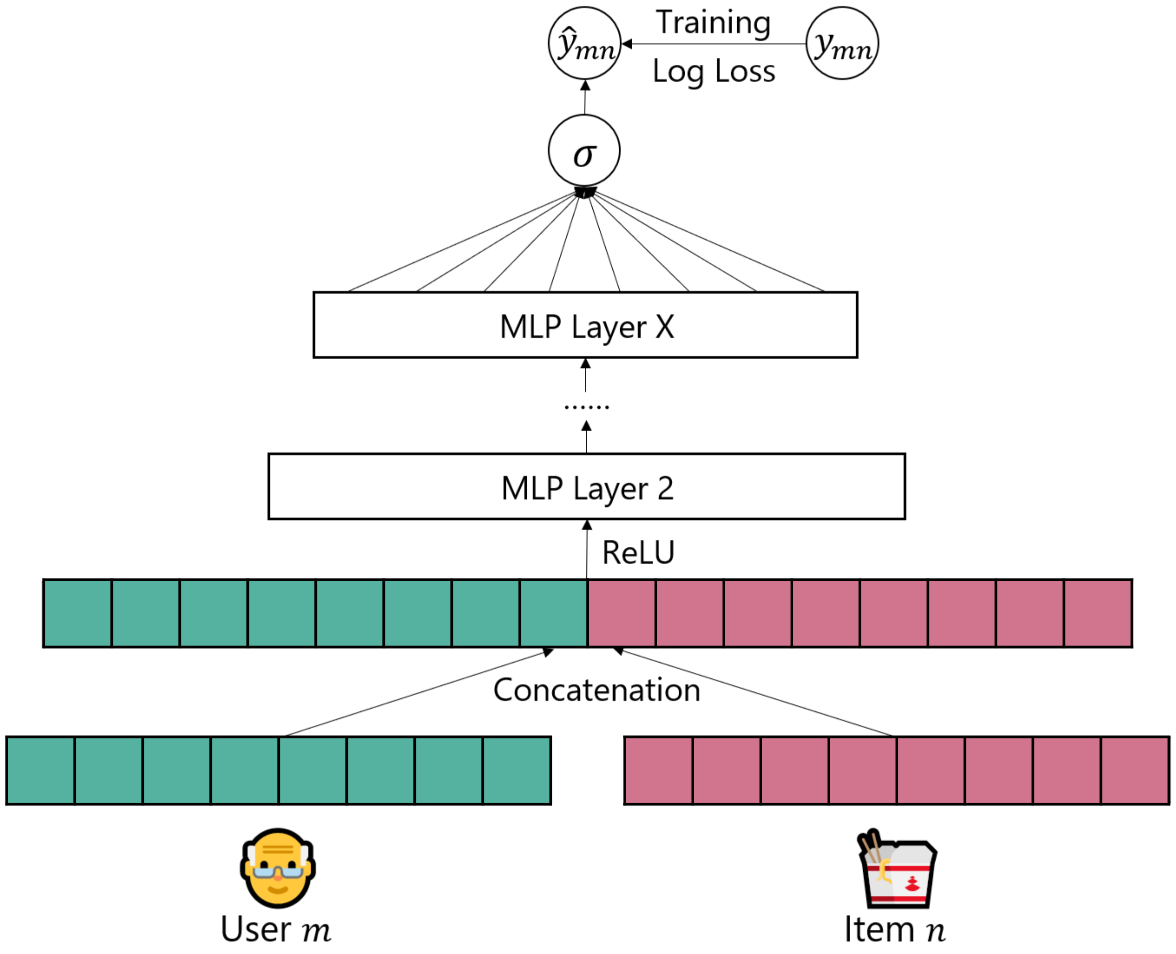
Matrix factorization, but this time it is training a bunch of logistic regressions instead of training linear regressions.



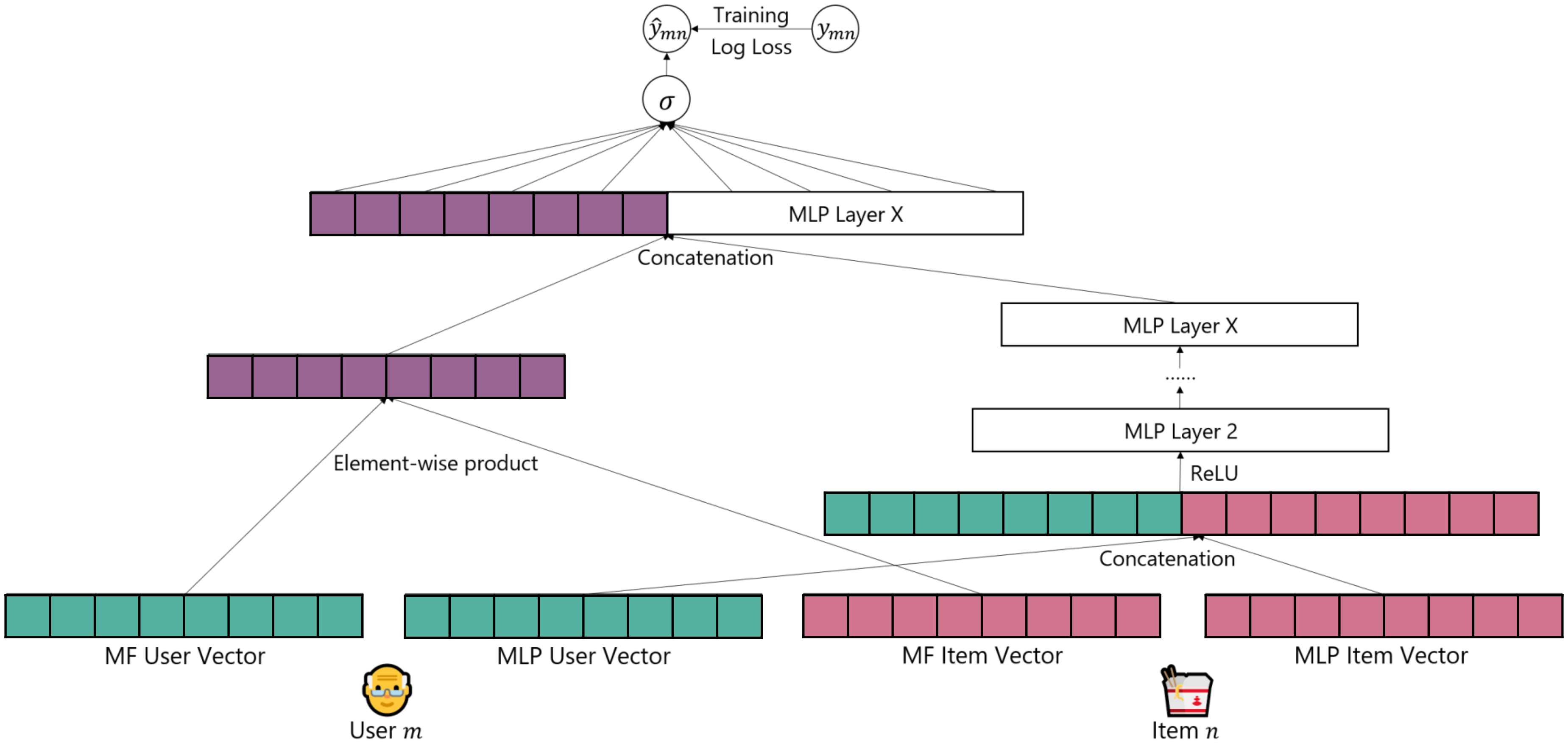
Multi-Layer Perceptron

Combine the features of a user and an item by concatenating them, then use multi-layer perceptron to learn the interaction between user and item latent features.

Endow the model a large level of flexibility and non-linearity to learn the interactions between the users and the items, instead of only using a fixed elementwise product in GMF



Fusion of GMF and MLP

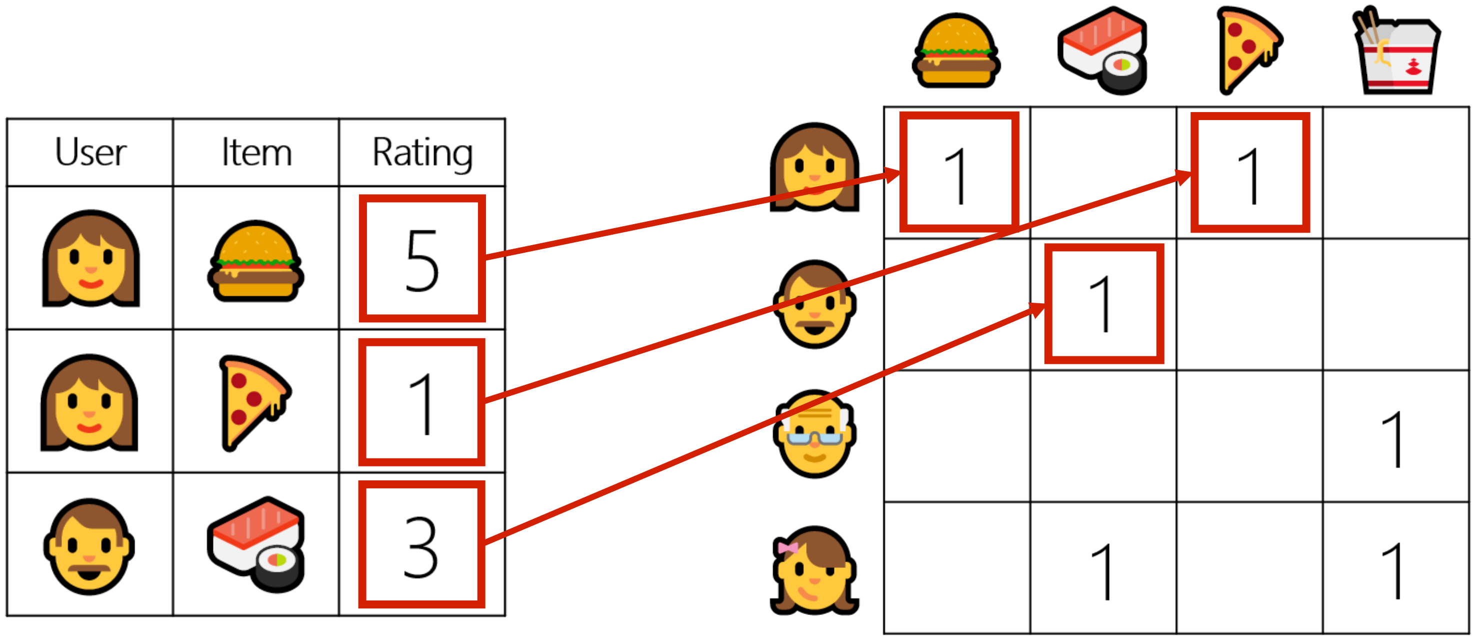


PART 4

Evaluation

Interested vs Not Interested

- If a user gave a restaurant a review, we assume the user is interested in the restaurant.
- If a user haven’t reviewed a restaurant, we assume the user is either not interested in that restaurant or have not aware of that restaurant.



1

I like burgers

There's a burger restaurant! I'll have a try!

The burger from that restaurant is disgusting!

I'll post a 1 star review for that restaurant on Yelp.

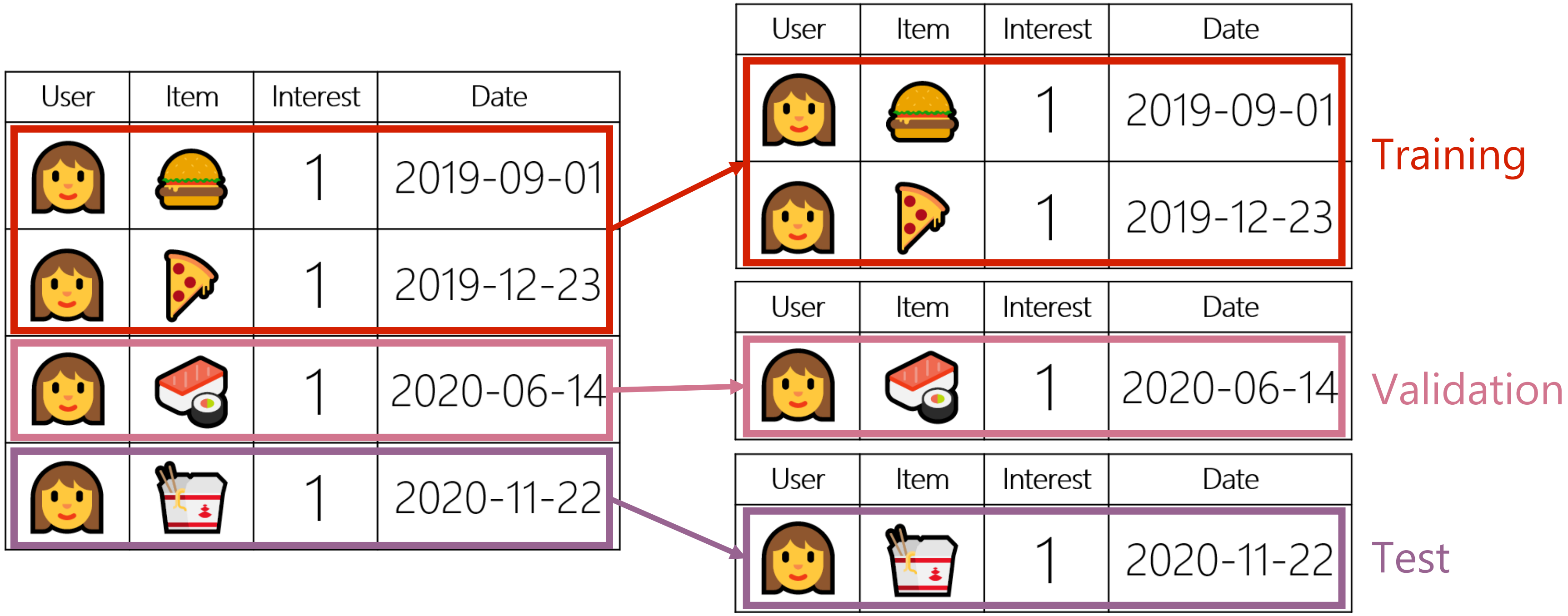
0

I hate Chinese food

There's a Chinese restaurant, I definitely don't want to have a try.

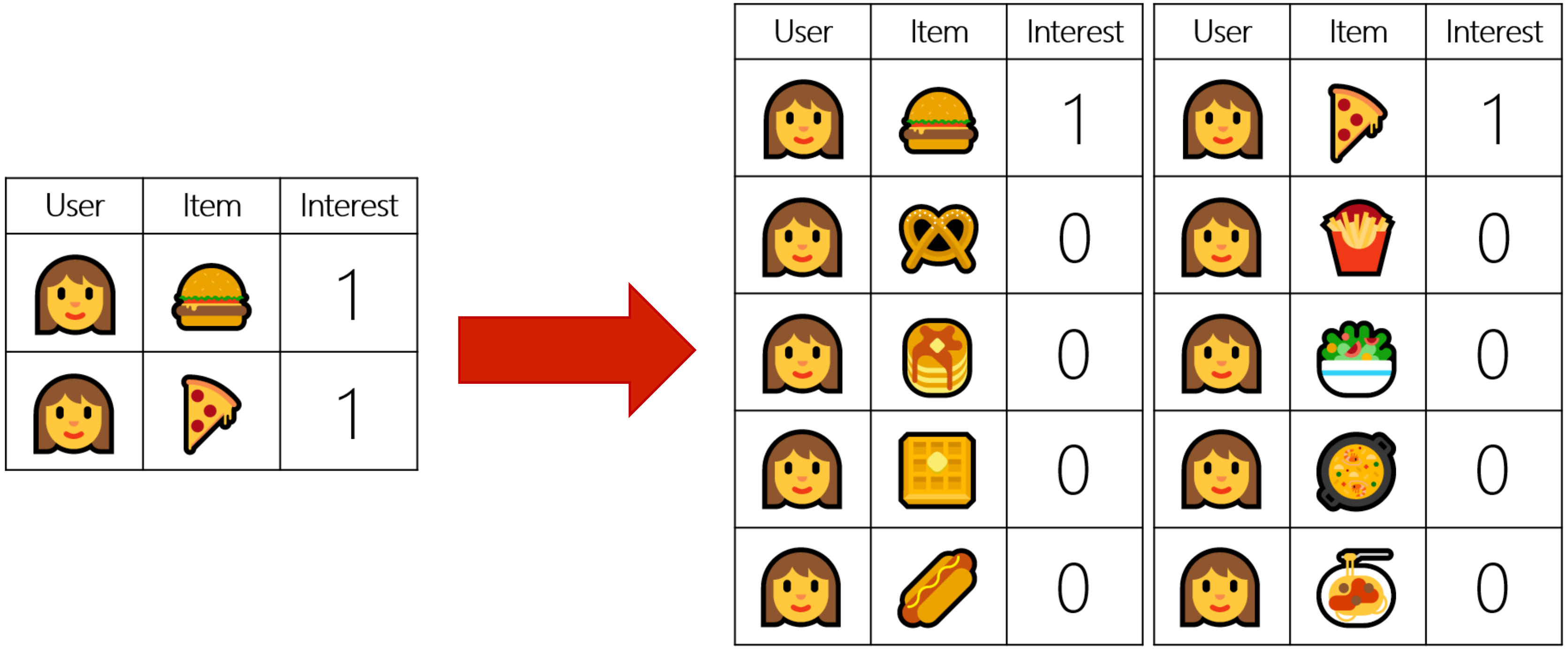
Train-Validation-Test Split

Put the latest review of each user into the test set, put the second latest review of each user into the validation set, the other reviews are in the training set.



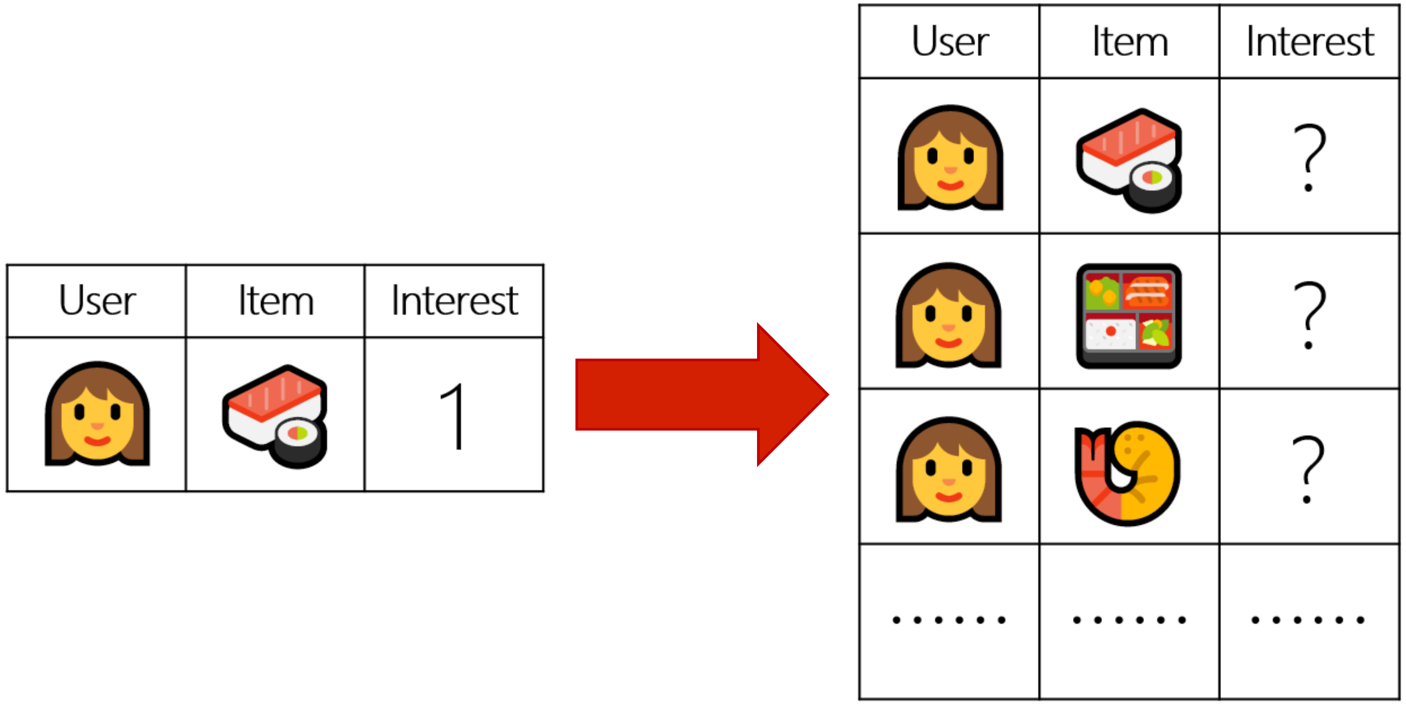
Training Set

For each review in the training set, randomly sample 4 restaurants that have not been reviewed by the user as not interested.



Validation and Test Set

For each user in the validation and test set, randomly sample 99 restaurants that have not been reviewed by the user as unknown interactions.



Evaluation Metric

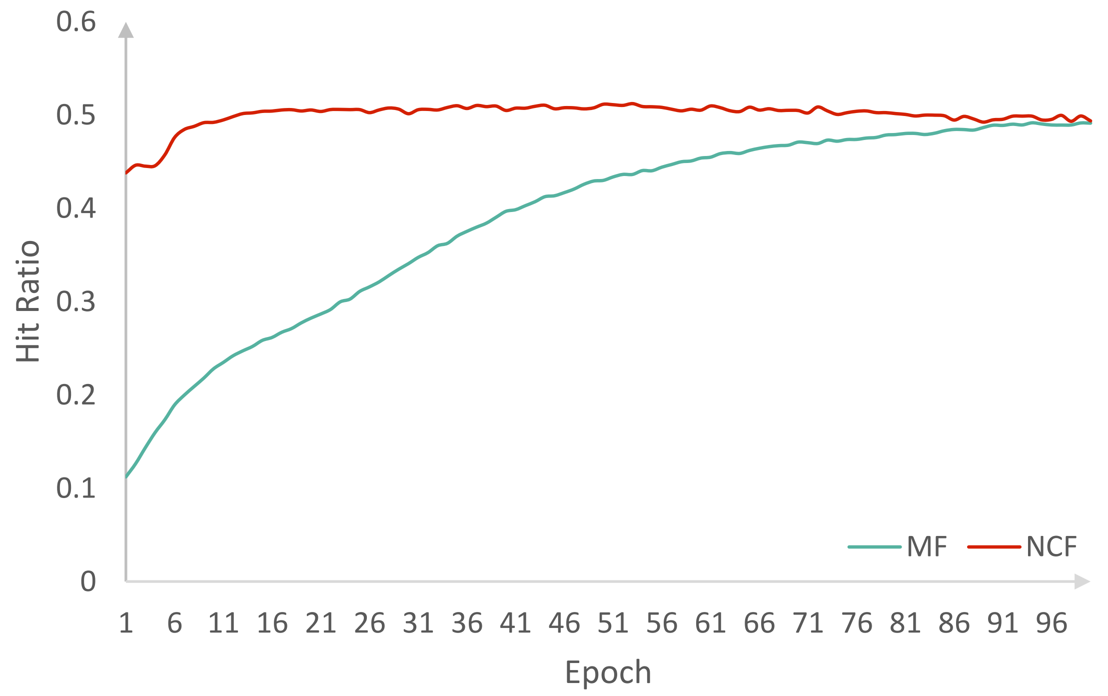
During the evaluation, for each user, recommend 10 items that he/she will be most interested in from the validation set and the test set.

The performance is judged by Hit Ratio (HR).

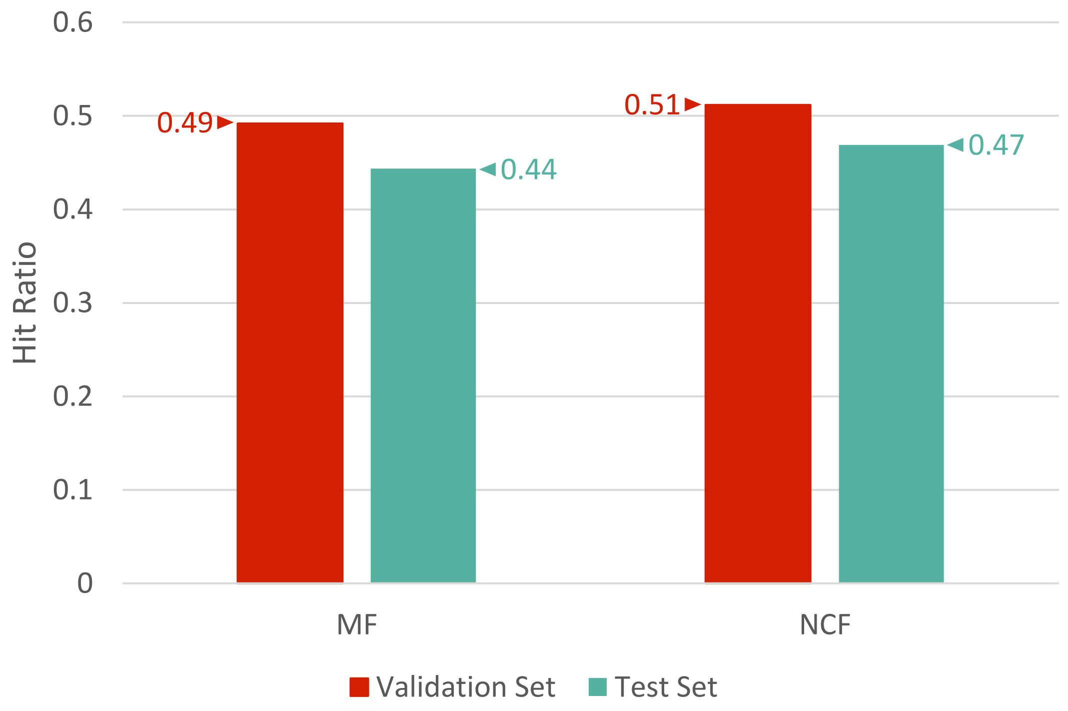
The HR is the percentage of users that have the reviewed restaurant in their top 10 recommendations.

Performances

Performances on the validation set



Performances on the validation set and the test set



PART 5

Showcase & Summary

Showcase

Last reviewed restaurants of a user

	item_id	name	categories
0	18	Assembly Chef's Hall	Japanese, Pizza, Coffee & Tea, Ramen, Barbeque, Food, Cafes, Food Court, Restaurants, Nightlife, Public Markets, Shopping, Vegan, Diners, Sushi Bars, Breakfast & Brunch, Bars
1	108	Joy Thai Restaurant	Seafood, Restaurants, Thai, Noodles
2	106	Rudy	Fast Food, Sandwiches, Burgers, Restaurants
3	151	Little Sister	Restaurants, Nightlife, Indonesian, Bars
4	156	GB Hand-pulled Noodles	Noodles, Chinese, Restaurants
5	100	Siam Square Hut	Thai, Asian Fusion, Caterers, Event Planning & Services, Restaurants
6	155	The Parkdale Drink	Nightlife, Lounges, Restaurants, Bars, Sushi Bars
7	160	The Dime	Nightlife, Burgers, Pubs, Restaurants, American (Traditional), Comfort Food, Bars
8	96	Pho Cuu Long Mien Tay	Restaurants, Vietnamese
9	97	Ramen Isshin	Noodles, Tapas/Small Plates, Food, Ramen, Japanese, Restaurants

Top 10 recommendation from NCF to the user

	item_id	name	categories	score
0	18	Assembly Chef's Hall	Japanese, Pizza, Coffee & Tea, Ramen, Barbeque, Food, Cafes, Food Court, Restaurants, Nightlife, Public Markets, Shopping, Vegan, Diners, Sushi Bars, Breakfast & Brunch, Bars	0.830849
1	251	Salad King Restaurant	Thai, Restaurants	0.778795
2	1866	Tabülè Middle Eastern Cuisine	Middle Eastern, Restaurants	0.685031
3	1018	The Good Son	Restaurants, Pizza, Bars, Canadian (New), Cocktail Bars, Nightlife	0.648689
4	1024	Luma	Cocktail Bars, Religious Organizations, Lounges, Bars, Nightlife, American (New), Restaurants, Canadian (New), Wine Bars	0.632588
5	1088	Sushi On Bloor	Restaurants, Sushi Bars	0.586555
6	444	Aji Sai Japanese Restaurant	Japanese, Restaurants, Food, Buffets	0.543902
7	731	Peter Pan Bistro	Mediterranean, Cocktail Bars, American (New), Canadian (New), French, Restaurants, Wine Bars, Nightlife, Bars, Bistros	0.526528
8	206	Rodney's Oyster House	Seafood, Restaurants	0.515795
9	647	St Lawrence Market	Farmers Market, Sandwiches, Restaurants, Grocery, Food	0.515549

Showcase

Last reviewed restaurants of a user

	item_id	name	categories
0	18	Assembly Chef's Hall	Japanese, Pizza, Coffee & Tea, Ramen, Barbeque, Food, Cafes, Food Court, Restaurants, Nightlife, Public Markets, Shopping, Vegan, Diners, Sushi Bars, Breakfast & Brunch, Bars
1	108	Joy Thai Restaurant	Seafood, Restaurants, Thai, Noodles
2	106	Rudy	Fast Food, Sandwiches, Burgers, Restaurants
3	151	Little Sister	Restaurants, Nightlife, Indonesian, Bars
4	156	GB Hand-pulled Noodles	Noodles, Chinese, Restaurants
5	100	Siam Square Hut	Thai, Asian Fusion, Caterers, Event Planning & Services, Restaurants
6	155	The Parkdale Drink	Nightlife, Lounges, Restaurants, Bars, Sushi Bars
7	160	The Dime	Nightlife, Burgers, Pubs, Restaurants, American (Traditional), Comfort Food, Bars
8	96	Pho Cuu Long Mien Tay	Restaurants, Vietnamese
9	97	Ramen Isshin	Noodles, Tapas/Small Plates, Food, Ramen, Japanese, Restaurants

Top 10 recommendation from MF to the user

	item_id	name	categories	score
0	251	Salad King Restaurant	Thai, Restaurants	0.999953
1	18	Assembly Chef's Hall	Japanese, Pizza, Coffee & Tea, Ramen, Barbeque, Food, Cafes, Food Court, Restaurants, Nightlife, Public Markets, Shopping, Vegan, Diners, Sushi Bars, Breakfast & Brunch, Bars	0.944914
2	1088	Sushi On Bloor	Restaurants, Sushi Bars	0.914624
3	292	Gdous Juicy Chicken House	Bubble Tea, Restaurants, Food, Chicken Wings, Chicken Shop	0.700383
4	731	Peter Pan Bistro	Mediterranean, Cocktail Bars, American (New), Canadian (New), French, Restaurants, Wine Bars, Nightlife, Bars, Bistros	0.655764
5	444	Aji Sai Japanese Restaurant	Japanese, Restaurants, Food, Buffets	0.634856
6	1400	Kenzo Ramen	Japanese, Food, Restaurants	0.517021
7	1018	The Good Son	Restaurants, Pizza, Bars, Canadian (New), Cocktail Bars, Nightlife	0.502444
8	206	Rodney's Oyster House	Seafood, Restaurants	0.436970
9	1024	Luma	Cocktail Bars, Religious Organizations, Lounges, Bars, Nightlife, American (New), Restaurants, Canadian (New), Wine Bars	0.378411

Pros and Cons of NCF

Pros:

- Large level of flexibility and non-linearity
- Converges faster
- Better accuracy

Cons:

- Higher computational cost
- Slower implementation (60% more running time per epoch on GTX1060)

Thank You!