Meet Fresh Recommender System

Group 4
Fenny Jin
Baomei Wen
Qiu Zheng
Yibo Wang
Yunying Liang
Yangfeng Peng

Introduction

- Meet Fresh
  - A Taiwanese dessert restaurant chain
  - Established in 2007
  - Based in New Taipei City
  - Specializes in soft taro balls and delicate herbal jelly
- Project background
  - Objective : build a recommendation system
  - Four different types of recommender algorithms
    - Popularity based
    - Collaborative filtering
    - Deep Learning CDN Retrieval Model
    - Content-based Filtering

Discovery summary

- Reviews observations
- User interview
- Post-event protocols

Brainstorming

- Individual brainstorming
- Group brainstorming

Low fidelity prototyping

Verbral prototype

Paper prototype

Wizard of oz

Evaluation

- Concerns:
    - Privacy issue
    - Time-consuming
    - No “back” or “menu” buttom
- Areas to improve:
    - Efficiency
    - Simplicity

Final product solution design

Research & Consultation

Brainstorming

Prototypes

Evaluation: user interviews

Feedbacks & narrow down → 4 recommendation systems

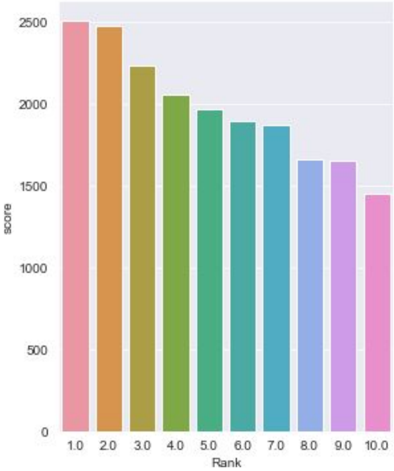
1. **Popularity based recommender system**: not customized, recommends the same items for all the users, works with the trend(highly rated and highly ranked)
2. **Collaborative filtering recommender system**: predictions interests of a user based on their preferences information from other users
3. **Deep Learning Model(Cross-Deep-Net (CDN) Retrieval Model)**: user information & item information
4. **Content Based Filtering Model**: description of the item and a profile of the user's preferences, more information about item, less information about users

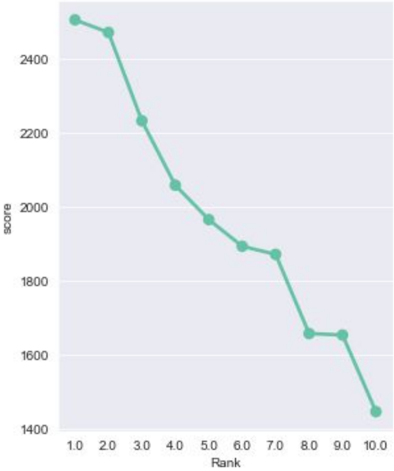
Data Collection

- User: gender, age, ethnicity, and preference for ice or hot drinks
- Item: series name, item name, calories, and contains
- Transaction: record id, transaction id, customer id, item id, quantity
- Review:

Popularity based recommender

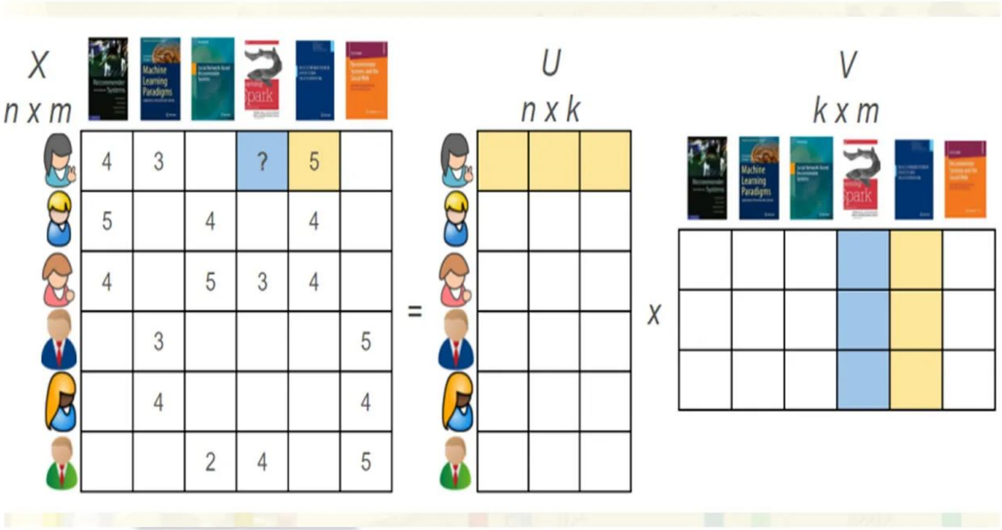
	user_id	Unique_ID	Name	score	Rank
0	6	16-1	Hot Almond Drink	2505	1.0
1	6	5-2	Mini Q & Melon Jelly	2471	2.0
2	6	13-3	Fluffy Oolong Tea	2232	3.0
3	6	15-3	Coffee Smoothie	2058	4.0
4	6	5-1	Creamy Milk	1965	5.0
5	6	12-2	Mixed Fruit Green Tea Slush	1893	6.0
6	6	13-7	Fluffy Mini Q (Mini Taro Ball) Winter Melon Tea	1871	7.0
7	6	16-3	Mini Q (Mini Taro Ball) Hot Almond Drink	1657	8.0
8	6	5-3	Boba & Caramel Pudding	1653	9.0
9	6	15-2	Coffee Milk Tea	1446	10.0

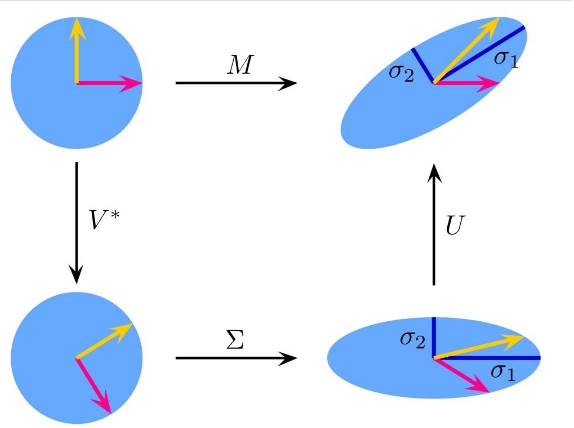




Popularity based recommender system is simple and very straightforward. But the model is not actually personalized it simply recommends to a user the most popular items that the user has not previously consumed. It recommends the same items for all users.

Collaborative Filtering:





$$M = U \cdot \Sigma \cdot V^*$$

Singular Value Decomposition (SVD):

It is popular latent factor model which compress user-item matrix into a low-dimensional representation in terms of latent factors. More factors there's more specific the model is, it may result some overfitting if there's to much factors.

Collaborative Filtering

	user_id	Unique_ID	Name	Rating
0	8	7-5	Oolong Tea	3.810004
1	8	3-5	Purple Rice Soup Combo C	3.809523
2	8	18-2	Red Bean Almond Pudding	3.808935
3	8	8-4	Fresh Milk Herbal Tea	3.808759
4	8	3-6	Hot Almond Purple Rice Soup	3.804683
5	8	10-8	Fresh Milk Jin Xuan Tea	3.799658
6	8	6-1	Plain Tofu Pudding	3.799541
7	8	3-13	Purple Rice Drink with Boba & Fresh Milk	3.798839
8	8	2-2	Icy Grass Jelly Signature	3.796997
9	8	9-8	Jin Xuan Oolong Milk Tea	3.794788

Collaborative Filtering provides strong predictive power for recommender systems, and requires the least information at the same time. Some of the challenges that are faced:

Con:

- **Cold start issue**
- **Data sparsity**
- **Data Scaling**
- **Popularity bias**

Model Evaluation:

Root Mean Squared Error (RMSE)

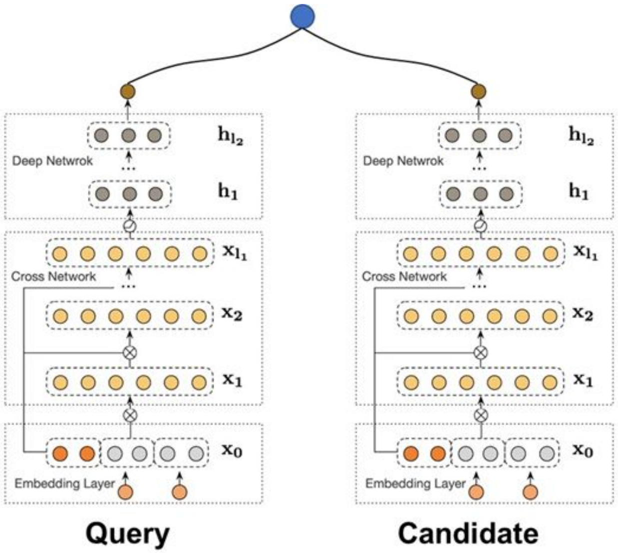
- **0.42**

Recall:

- **80%@10**

	Query Model	Candidate Model
Embedding dimension	385	577
Cross Net projection dimension	100	150
Dropout layer rate	0.1	0.1
Deep net dimensions	256, 128	512, 256, 128
L2 regularization layer dimension	64	64

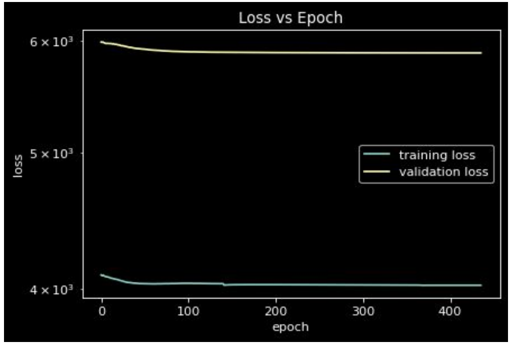
Cross Deep Net (CDN) Retrieval Model

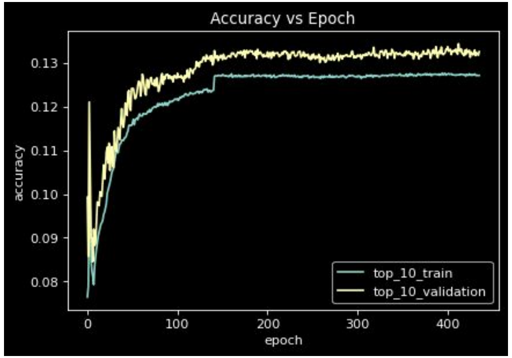


	Features	Embedding methods				
		String Lookup	Vectorization	Integer Lookup	Normalization	Discretization
Query Model	User ID	√				
	Age			√	√	√
	Gender	√				
	Ethnicity	√				
	Preference	√				
Candidate Model	Item Id	√				
	Ice/Hot	√				
	Series	√	√			
	Name	√	√			
	Contains	√	√			
	Calories				√	√

Top-K Categorical Accuracy	Training	Validation	Testing	Random Selection
K = 1	1.2%	1.4%	1.2%	0.6%
K = 5	6.8%	7.0%	6.6%	3.2%
K = 10	12.7%	13.3%	12.6%	6.4%
K = 50	51.2%	51.3%	50.7%	32.2%
K = 100	84.5%	84.6%	84.3%	64.5%

CDN Retrieval Model (Backtesting)





Actual Purchase Item	Mango Green Tea Slush	Mini Q Tofu Pudding
Top-10 Recommendation	Fresh Milk Black Tea	Rice Ball Tofu Pudding
	Icy Taro Ball Combo C	Mini Q Tofu Pudding
	Mango & Passion Fruit Green Tea	Potaro Ball Tofu Pudding
	Red Bean Almond Pudding	Iced Fruit Black Tea
	Herbal Tea	Taro Paste Volcano Shaved Ice
	Orange & Passion Fruit Green Tea	Fluffy Winter Melon Tea
	Fresh Milk Green Tea	Potaro Ball Tofu Pudding
	Potaro Ball Tofu Pudding	Win Win (Boba & Lychee) Milk Tea
	Fresh Milk Green Tea	Red Bean Tofu Pudding
	Jin Xuan Oolong Tea	Jin Xuan Oolong Tea

Content-based Filtering

	Name	Unique-ID	correlation	avg_rating	num_of_ratings
0	Icy Taro Ball Combo B	1–2	0.891326	3.170792	404
1	Icy Taro Ball Combo C	1–3	0.788955	3.143296	977
2	Icy Grass Jelly Combo A	1–4	0.757855	3.033333	270
3	Hot Red Bean Soup Combo B	1–11	0.689812	3.279446	866
4	Hot Red Bean Soup Combo C	1–12	0.652847	3.040724	221

Similar food for 1-13 Hot Almond Soup Combo A

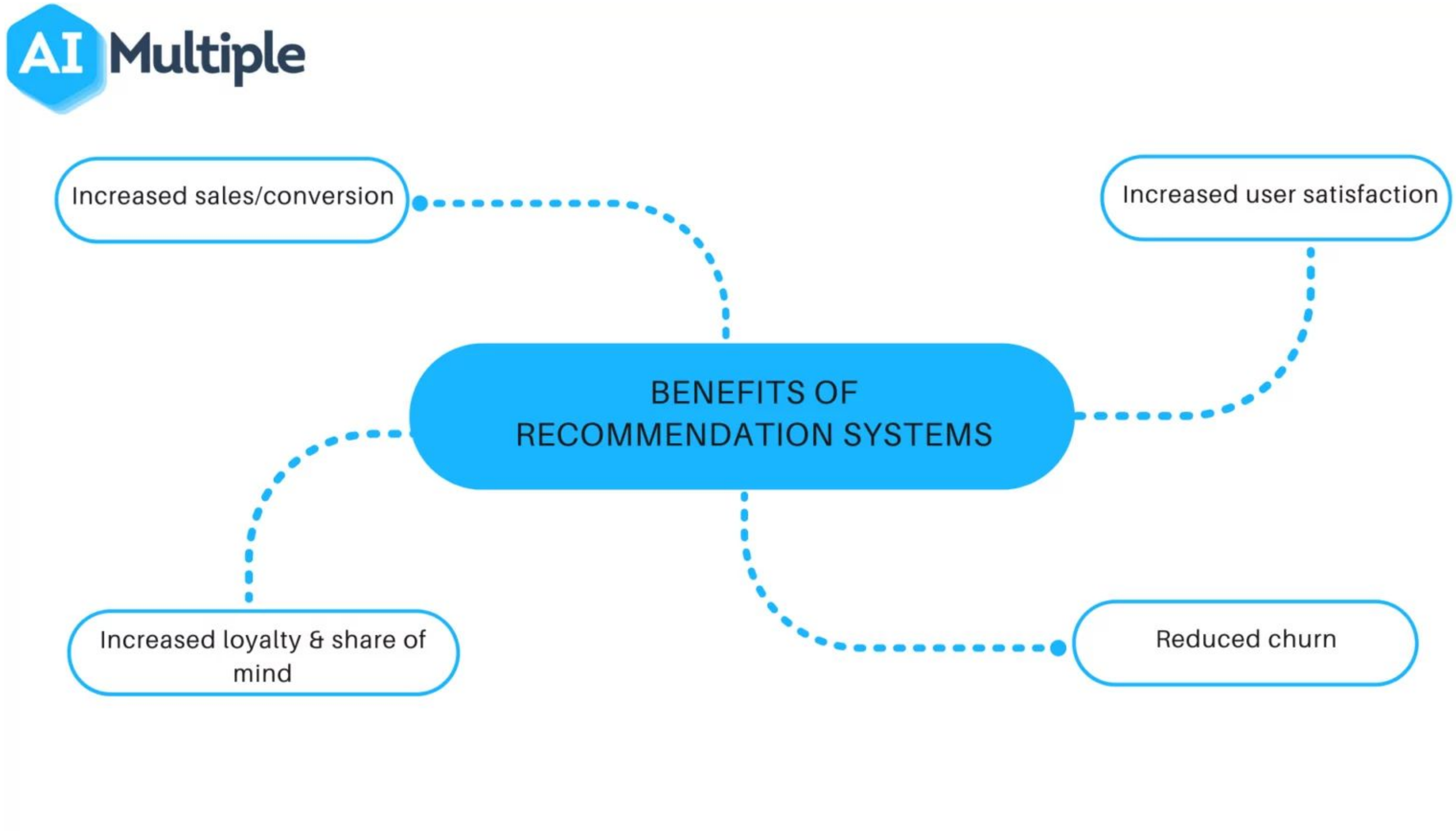
Pro:
Good for limited user info data

Recommend based on user's favourite features such as ingredients, ice/hot, combo, calories

Con:
Recommended only by previous purchases, not useful for new users

Did not consider user info and popularity, may not as accurate as other models

Business implications



Next steps & Final thought

- More relevant recommendations
- Reserach customers through multiple channels
- Adding weather and geological locations
- Breath of the data & depth of the data

Thank you!

Any Questions?