Personalized Recommendations for Customers in E-commerce Platform Taobao by Bayesian Methods

Presentors: Baihan Liu, Yiwei Mai, Yizhou Zhang

Goal

The overarching goal of this endeavor is to design and implement a robust recommendation system capable of accurately predicting and suggesting items to users based on their historical behaviors.

Data Description

Source: Taobao–an
e-commerce company in China.

Table 1: Descriptions of each column variables

Column Variables	Type	Dimension/Number of Each Variable
User ID	An integer	9820
Item ID	An integer	396855
Category ID	An integer, item category	5843
Behavior type	A string	4
Timestamp	An integer	488038

Table 2: Descriptions of each types of behaviors

Behavior	Explanation	Dimension/Number of Each Behavior
pv	Page view of an item's detail page	882949
cart	Add an item to shopping cart	55229
fav	Favor an item	28407
buy	Purchase an item	19405

Definitions

- User-Item Pair:

  A pair of user and item
- User-item Interactions:

  A behavior (at certain time) for

  a pair of user and item

User-Item Interaction

Data Integration

- No Duplicated and missing value
- Focus on 2017-11-25 to 2017-12-02

```
2017-12-02    139670
2017-12-03    122451
2017-12-01    112001
2017-11-26    105262
2017-11-30    105068
2017-11-25    102365
2017-11-29    100395
2017-11-28     99709
2017-11-27     99069
2017-11-24     12447
2017-11-23        62
2017-11-18        25
2017-11-22        17
2019-11-01        14
2017-11-16         5
2017-11-19         5
2017-11-21         4
2017-11-11         3
2017-11-12         2
2017-11-14         2
2017-11-20         2
2017-10-21         2
2017-10-29         1
2017-11-10         1
2017-11-15         1
2017-11-13         1
2017-11-02         1
2017-11-05         1
Name: Date, dtype: int64
```

Data Transformation

	User_ID	Item_ID	Categorical_ID	Behavior	Timestamp	Date
1	1000061	1288773	1735195	pv	1511576626	2017-11-25
2	1000061	4074215	4684862	pv	1511677113	2017-11-26
3	1000061	1120968	512076	pv	1511685194	2017-11-26
4	1000061	4862918	3164550	cart	1511686985	2017-11-26
5	1000061	3150019	3164550	cart	1511687015	2017-11-26
...	...	...	...	...	...	...
985986	999949	2652397	2355072	pv	1512268408	2017-12-03
985987	999949	694434	2355072	pv	1512268415	2017-12-03
985988	999949	552658	737184	pv	1512268526	2017-12-03
985989	999949	4016076	1521931	pv	1512274117	2017-12-03
985990	999949	424771	375240	pv	1512274628	2017-12-03

985990 rows × 6 columns

Pivoting

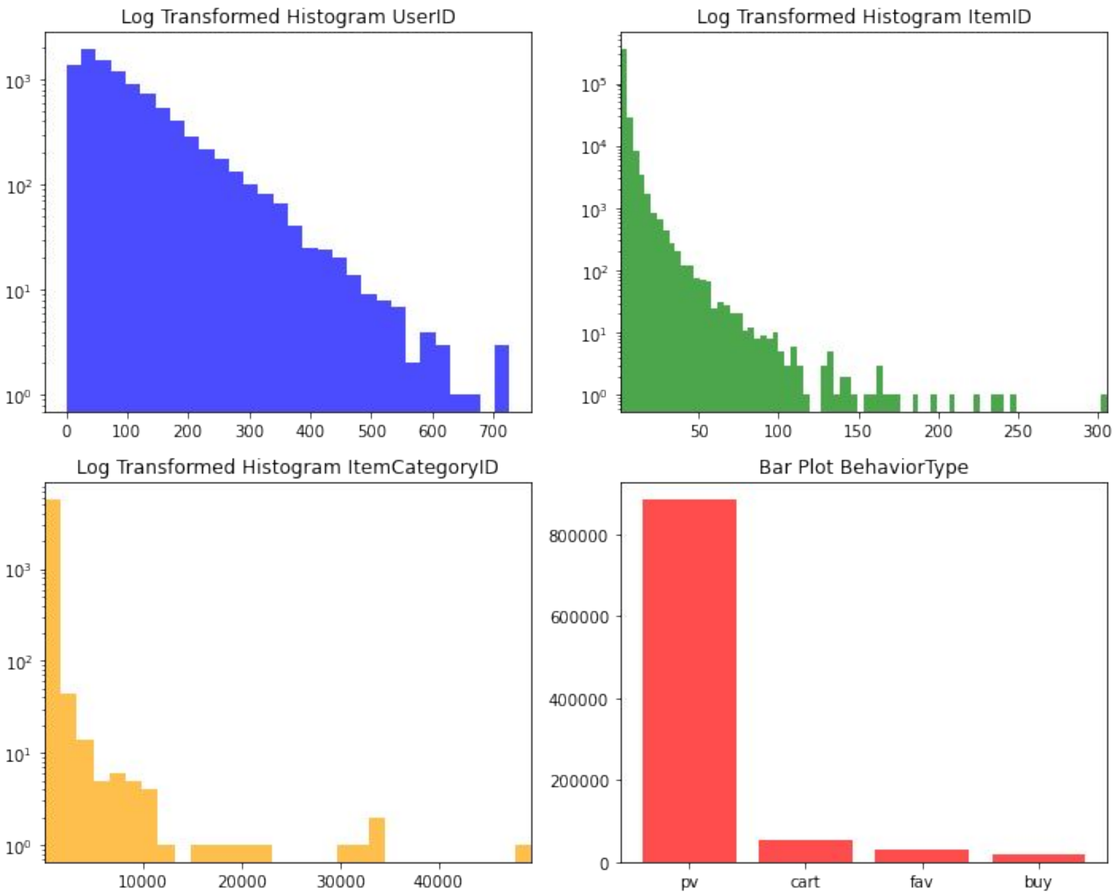
BehaviorType	UserID	ItemID	ItemCategoryID	buy	cart	fav	pv
0	253	137988	2409937	0	0	0	1
1	253	257634	2885642	0	0	0	1
2	253	500819	2409937	0	0	0	1
3	253	543072	2885642	0	0	0	1
4	253	937072	2885642	0	0	0	1
...	...	...	...	...	...	...	...
746367	1017952	4011217	3422001	0	0	0	1
746368	1017952	4183961	4159072	0	0	0	1
746369	1017952	4490910	411153	0	0	0	1
746370	1017952	4501732	1573426	0	0	0	1
746371	1017952	4814776	2885642	0	0	0	1

746372 rows × 7 columns

User-Item Interactions: 985990

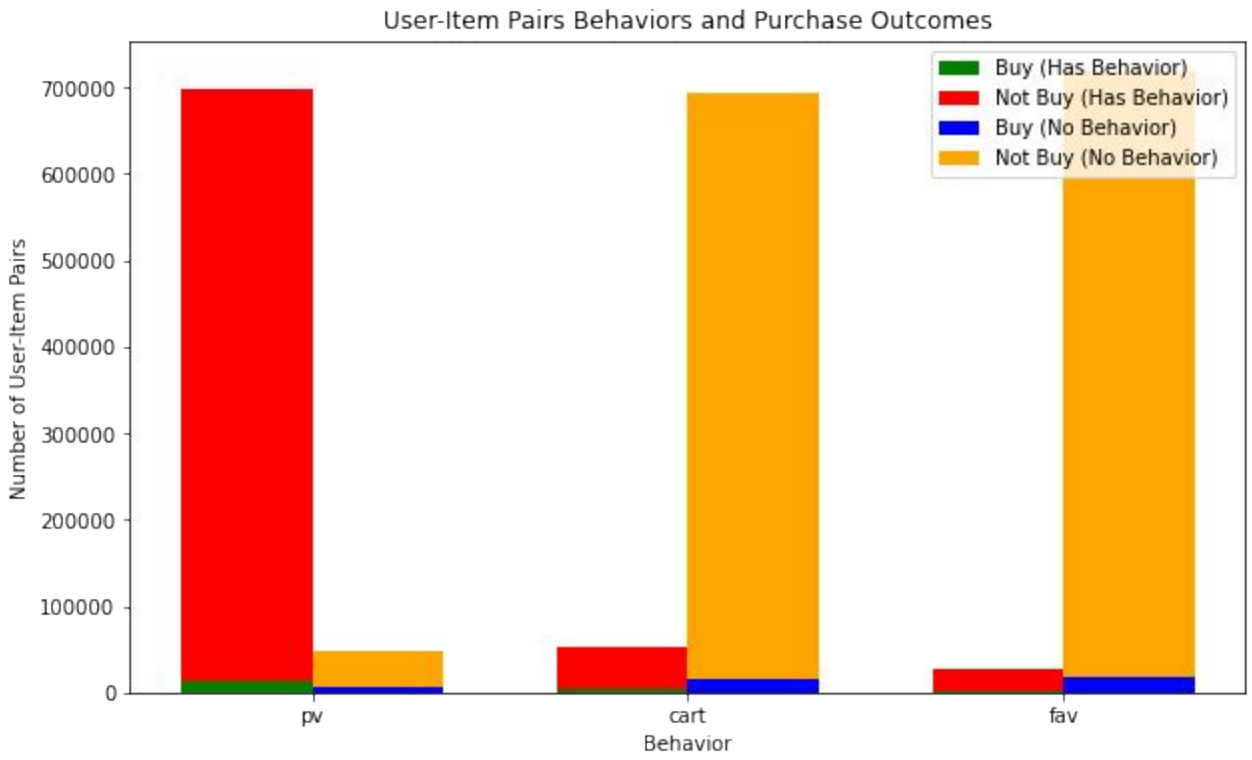
User-Item Pairs: 746372

Variable Distribution



Consider
user-itemcategory
pair in BPR model

Variable Relationship



Simpson's Paradox

Chi-Square Test and Column Ratios

Contingency Tables

buy pv	0	True
0	42481	5972
True	685293	12626
	χ^2: 20618	p=0

buy cart	0	True
0	677288	15284
True	50486	3314
	χ^2: 3209	p=0

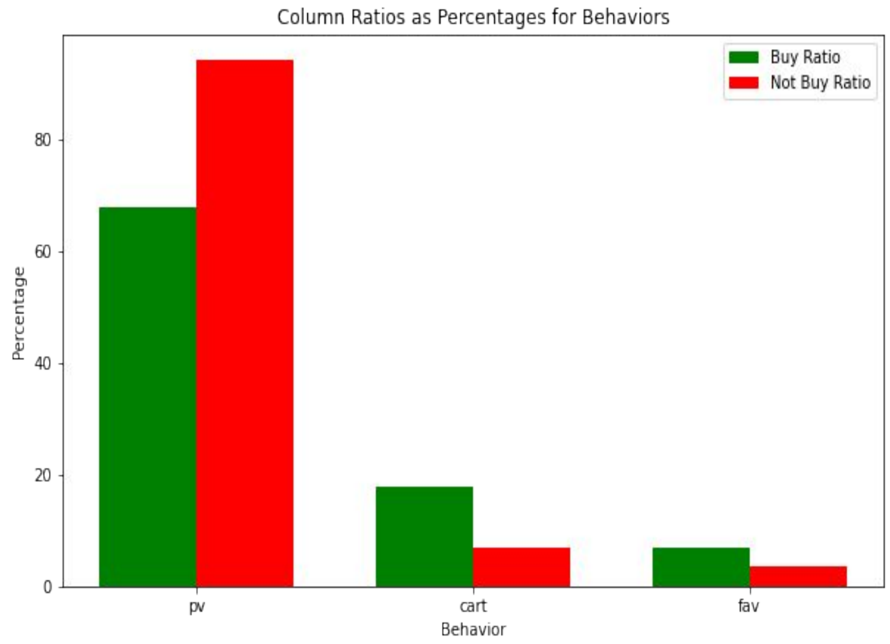
buy fav	0	True
0	700803	17308
True	26971	1290
	χ^2: 519	p=0

Column Ratios

buy pv	0	True
0	0.058371	0.32111
True	0.941629	0.67889

buy cart	0	True
0	0.93063	0.821809
True	0.06937	0.178191

buy fav	0	True
0	0.96294	0.930638
True	0.03706	0.069362



Consider pv, cart, fav as predictors to predict buying decision

Potential Defects

- Lack demographic variables for customers, such as gender, region and economic status
- Lack information about commodity, such as price
- Limited Time span

Bayesian Logistic Regression
via Metropolis-Hastings Algorithm

- Task: predicting whether a customer would buy an item or not according to their shopping behaviors for unpopular categories.

- Data: customer behavior records on some very unpopular item categories for which we have fewer than 300 pieces of data.

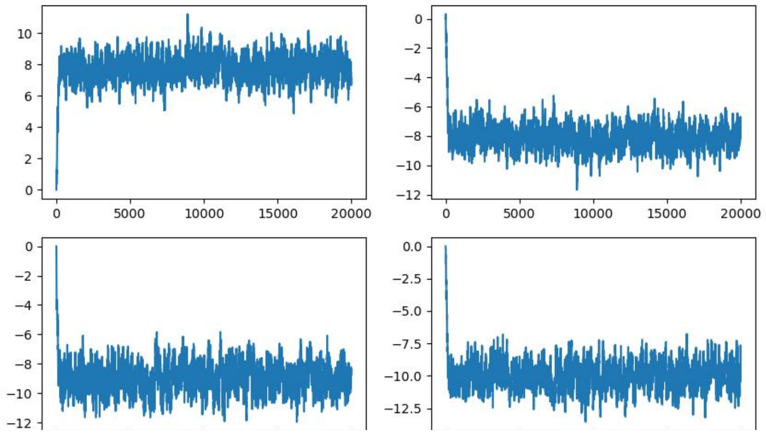
- Features: 'pv', 'cart' and 'fav'

  Target: 'buy'

Experiment

- Settings:
  - The prior mean is (10, -10, -10, -10), the variance of proposal distribution is 0.01.
  - Iterating 20000 times.
  - Experiment respectively on data of 5 unpopular categories.

- Results:
  - The sampling distributions converged well. The posterior mean is (7.77,-8.06,-8.93,-8.95) for category 359388
  - The sampling distributions are influenced by the prior mean to a great extent



Model Performance Comparison

- Bayesian Logistic Regression performs the best in such small sample situation.

- With tiny amount and imbalanced data, non-Bayesian approach failed to categorize any positive instances correctly while Bayesian approach correctly categorized all of them.

Model	Precision	Recall
Non-Bayesian LR (without cross-terms)	0.0	0.0
Non-Bayesian LR (with cross-terms)	0.0	0.0
Bayesian LR	**1.0**	**1.0**

- Similar results found for experiments on the other four categories.

Discussions

- Recall the prior mean and posterior mean:

  Prior mean: (10, -10, -10, -10)

  Posterior mean: (7.77,-8.06,-8.93,-8.95)

- Although the absolute values do not tell us much since they vary a lot with the prior mean, but we can interpret from their relative size that the influences of 'fav' ≈ 'cart' > 'pv'. Such relationships also hold for experiments on other categories.

- This method can be useful under situations where we have tiny amount of data due to product features but we have some prior knowledge.

BPR: Bayesian Personalized Ranking

Algorithm Description:

- **Pairwise**: BPR will learn from every user's preference between any two items.
- **Bayesian**: BPR uses a bayesian approach to find the optimal value for its parameters.

Objective:

Considering the sparsity of User-Item interactions, we decided to use User-ItemCategory interactions. Thus, we aim to build a recommender system that provides accurate ranking of ItemCategory for each user.

BPR: Bayesian Personalized Ranking

Transform Data:

- We believe that the four kinds of users' behavior may have different importances which will affect the preference structure for each user.

- We tried three different ways to assign a value to each kind of behavior, and set the interaction value between User u and ItemCategory i as the sum of all behaviors' importances between u and i.

BPR: Bayesian Personalized Ranking

Train/test split:

- We split 80% of the non-zero interactions between user and ItemCategory as the training data, and let the rest be test data. We will examine whether the model will rank correctly on test data.

Evaluation metrics:

- ROC-AUC score
- Precision at k: the percentage of top k ranked items that are known positives.

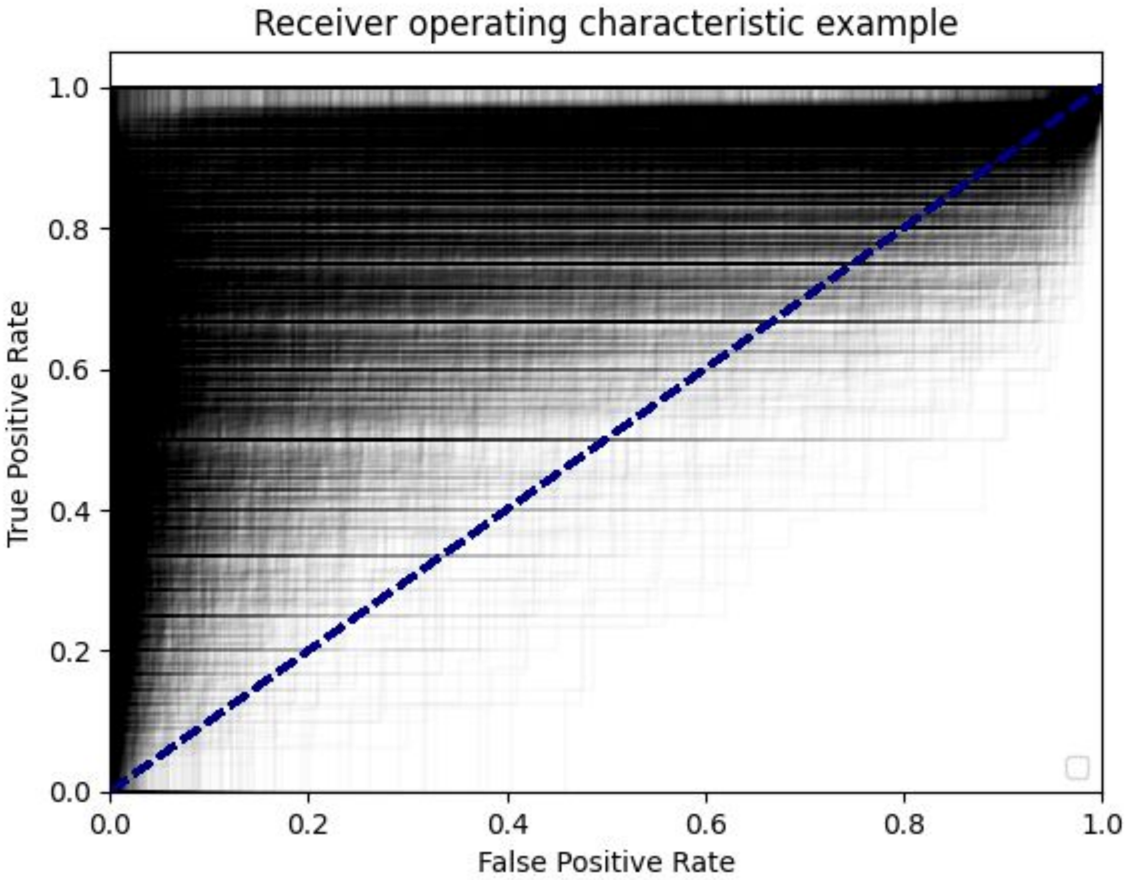
BPR: Bayesian Personalized Ranking

Results:

Method	Importance (pv, fav, cart, buy)	AUC Score	Precision at 5	Precision at 20
BPR	(1,1,1,1)	0.705	0.037	0.036
BPR	(1,2,3,4)	**0.709**	**0.041**	**0.038**
BPR	(1,10,100,1000)	0.706	**0.041**	**0.038**

The second model has a slightly better result than the others.

Changing importance for each behavior does not make a big difference in the ranking.


Receiver operating characteristic example
1.0
0.8
0.6
0.4
0.2
0.0
True Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
False Positive Rate

Thank you