

Datasets

Datasets	Users	Items	Ratings
Instant Video	426922	23965	583933
Musical Instrument	339231	83046	500176

Table 1: Statistics of the datasets

Datasets

	user	item	rating	timestamp
0	A1YS9MDZP93857	0006428320	3.0	1394496000
1	A3TS466QBAWB9D	0014072149	5.0	1370476800
2	A3BUDYITWUSIS7	0041291905	5.0	1381708800
3	A19K10Z0D2NTZK	0041913574	5.0	1285200000
4	A14X336IB4JD89	0201891859	1.0	1350432000

Figure 1: Samples from Amazon review datasets

Datasets

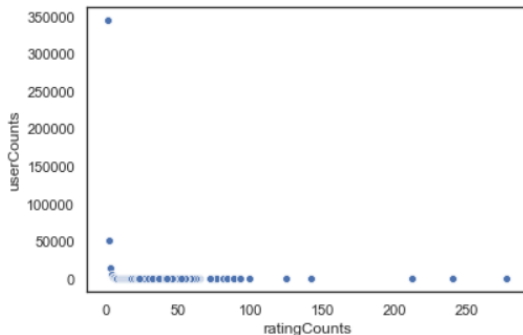


Figure 2: Statistics of users' ratings

Datasets

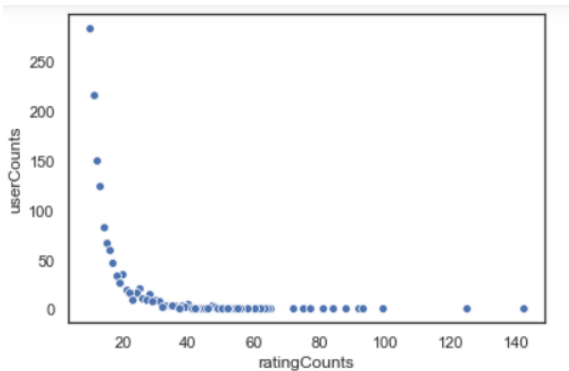


Figure 3: Statistics of users' ratings

Datasets

Datasets	Users	Items	Ratings
Instant Video	426922	23965	583933
Musical Instrument	339231	83046	500176

Table 2: Statistics of the raw datasets

Datasets	Users	Items	Ratings
Instant Video	1372	7957	23181
Musical Instrument	2270	21464	38404

Table 3: Statistics of the preprocessed datasets

Evaluation

Datasets	Instant Video			Musical Instrument		
Measures@10(%)	P	R	F1	P	R	F1
SVD	0.984	2.065	1.285			
kNN	0.496	1.183	0.681			

Table 4: The performance of baselines

$$P@N = \frac{1}{M} \sum_u P_u@N = \frac{1}{M} \sum_u \frac{|R_u \cap T_u|}{|R_u|}$$

$$R@N = \frac{1}{M} \sum_u R_u@N = \frac{1}{M} \sum_u \frac{|R_u \cap T_u|}{|T_u|}$$

$$F_1@N = \frac{1}{M} \sum_u F_{1u}@N = \frac{1}{M} \sum_u \frac{2 \cdot P_u@N \cdot R_u@N}{P_u@N + R_u@N}$$

Our model

For the reinforcement learning stage,

- ▶ **Observation** The prespecified user u and all users' purchasing history.
- ▶ **Action** Output K neighbors of the user u .
- ▶ **Policy** *Upper Confidence Bounds* methods.

$$\bar{x}_{u'}(t) + \sqrt{\frac{2 \ln t}{T_{u',t}}}$$

Our model

For the superposed Hawkes process learning,

- ▶ **Merge** Combining the purchasing history of the K neighbors and the prespecified user u .
- ▶ **MLE**

$$\min -\log \mathcal{L}(\mathcal{H}_{train}^{super}; \{\lambda_c^u\}_{c \in C})$$

- ▶ **Reward**

$$-\sum_{c \in C} \left(N_c^u(t_{test}) - \int_{t_{test}} \lambda_c^u(s) ds \right)^2$$