SIGIR 2020
XI'AN·CHINA

Neural Interactive Collaborative Filtering

Lixin Zou[1,5], Long Xia[2], Yulong Gu[4],
Xiangyu Zhao[3], Weidong Liu[1], Jimmy Huang[2], Dawei Yin[5]
[1]Tsinghua University, [2]York University, [3]Michigan State University
[4]JD Data Science Lab, [5]Baidu Inc.

43rd International ACM SIGIR Conference on Research and Development in Information Retrieval
July 25-30, 2020 (Virtual Event, China)

Outline

- Introduction
- Related work
- Method
- Experiments

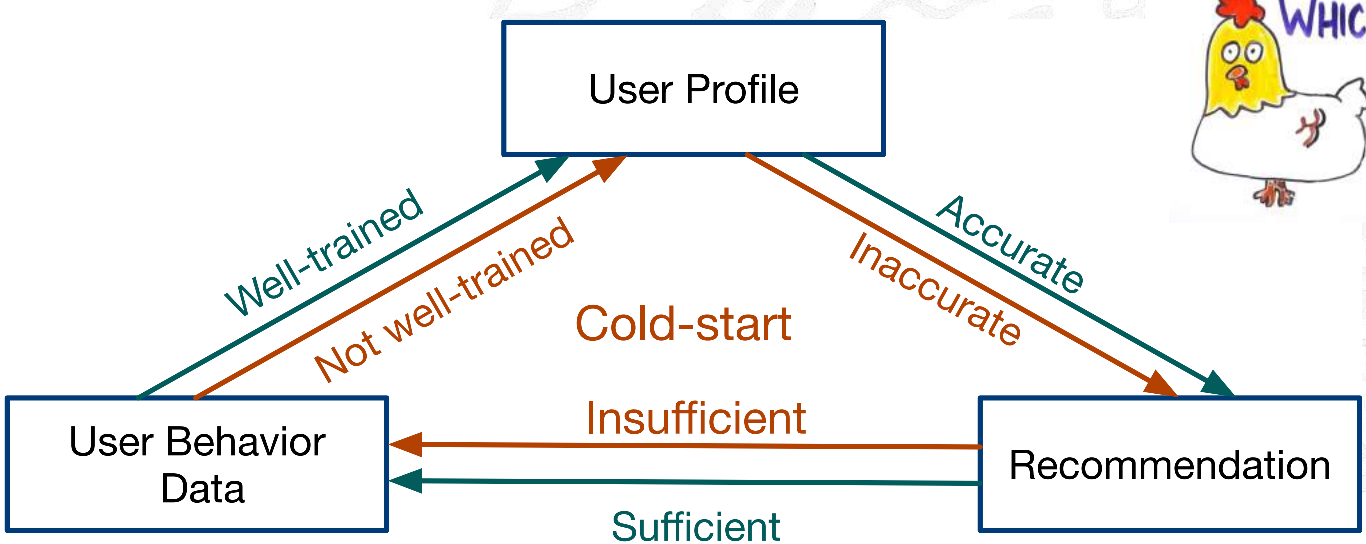
Introduction

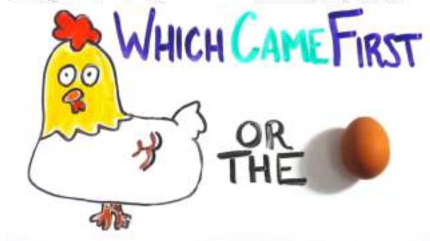
- Interactive recommender system

Introduction

- Cold start

The challenge in the interactive recommendation is to suggest items with insufficient observations.





Related work

- Problem formulation

For any user u (can be viewed as a Bandit) , our goal is to design a policy π so that the expected total payoff $G_\pi(T)$ is maximized,

$$G_\pi(T) = \mathbb{E}_{i_t \sim \pi(s_t)} \left[\sum_{t=1}^{T} r_{u,i_t} \right].$$

Here, r_{u,i_t} is user's rating on item i_t .

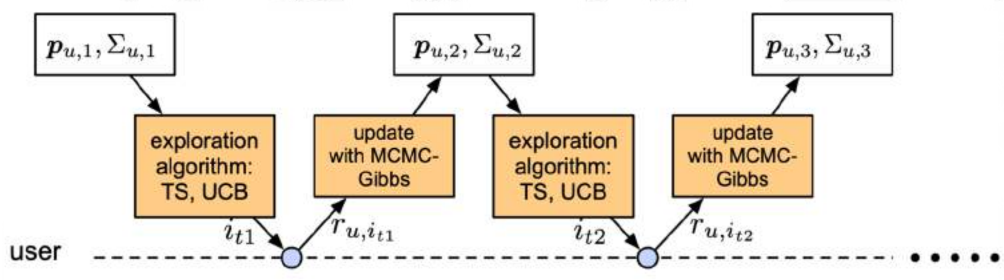
Related work

- Multi-Armed Bandit (optimize in face of uncertainty)

Assume $Pr(r_{u,i}|\boldsymbol{p}_u^\top \boldsymbol{q}_i, \sigma^2) = \mathcal{N}(r_{u,i}|\boldsymbol{p}_u^\top \boldsymbol{q}_i, \sigma^2)$. $\boldsymbol{p}_u$, $\boldsymbol{q}_i$ are the latent factors for users and items (Probabilistic Matrix Factorization).

(1) Obtaining the posterior distributions of the user and item feature vectors after the (t − 1)-th interaction.

(2) Heuristically select the item for the t-th recommendation with the aim of maximizing the cumulative reward.



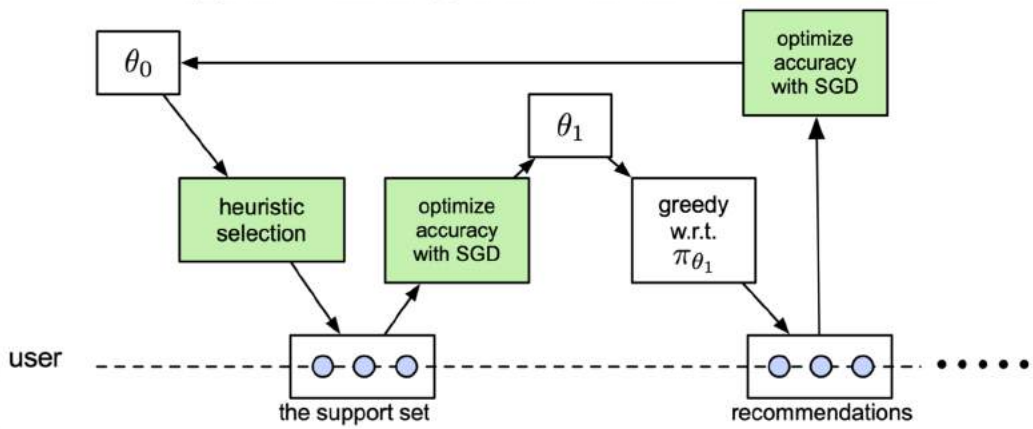
(a) The MAB based approaches for cold-start recommendation.

Related work

- Meta-learning

It aims to learn an initialization θ_0 that can identify users' interests after updating θ_1 with small support set D.

$$\begin{aligned} \theta_1 &= \theta_0 - \alpha \ell(\pi_{\theta_0}, D) \\ \theta_0 &\leftarrow \theta_0 - \alpha \ell(\pi_{\theta_1}, D), \end{aligned}$$

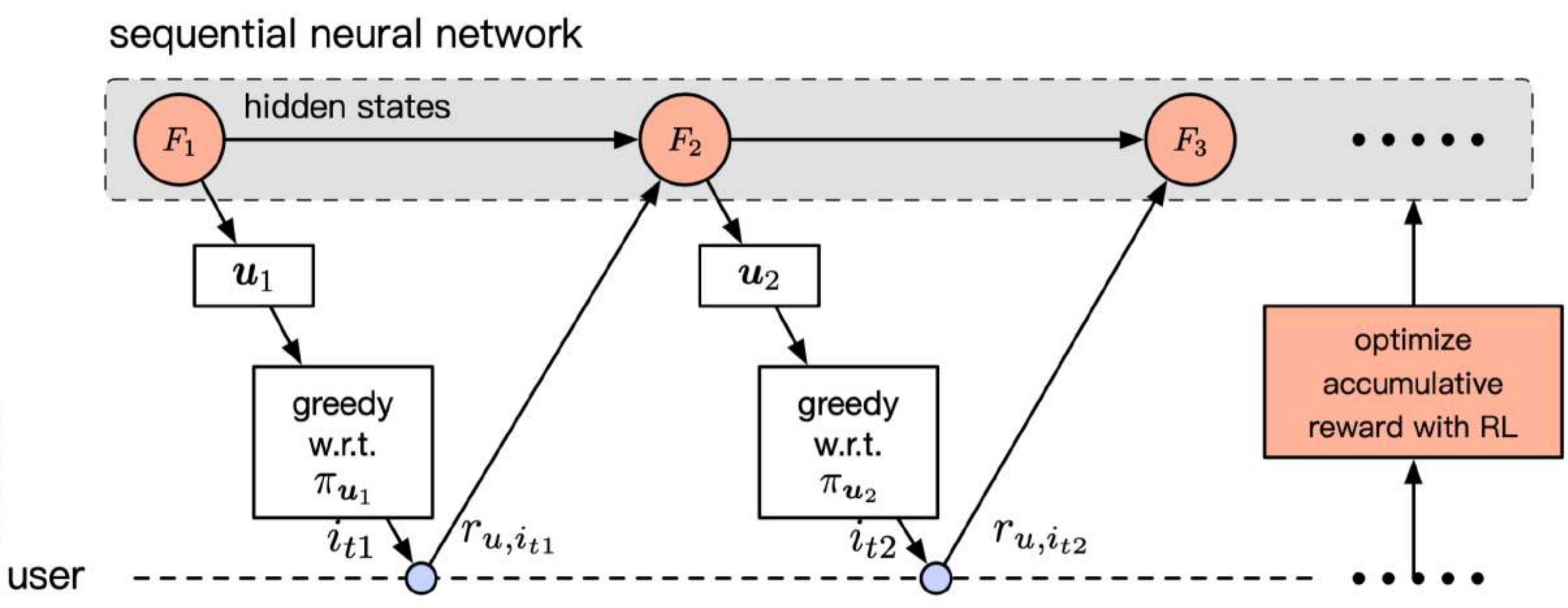


(b) Model agnostic meta-learning based approach for cold-start recommendation.

sequential neural network

Method

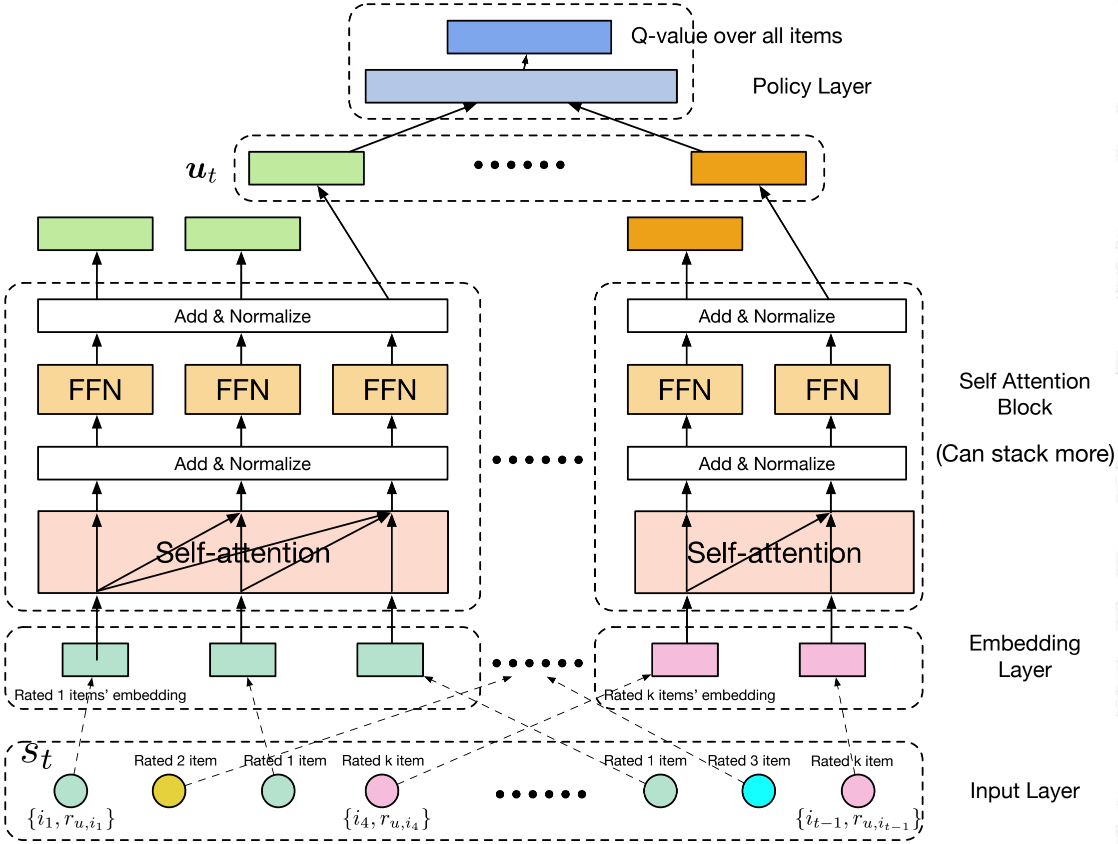
- Learn a neural network based exploration strategy whereby the recommender agent can capture users' interests rapidly for different users.
- From meta-learning, the objective is to learn a learning algorithm that can take as the input of the user's history s_t and will output a model (policy function) that can be applied to new items.
- From RL, the users' preferences gathered by exploration recommendations can be viewed as the delayed reward for the recommendations.



(c) The neural interactive collaborative filtering framework.

Method

- Self-Attentive Neural Policy



Method

- Policy Learning

We improve the value function by adjusting θ to minimize the mean-square loss function, defined as follows:

$$\ell\left(\theta_Q\right)=\mathbb{E}_{\left(s_t, i_t, r_t, s_{t+1}\right) \sim \mathcal{M}}\left[\left(y_t-Q\left(s_t, i_t ; \theta_Q\right)\right)^2\right]$$

$$y_t=r_t+\gamma \max _{i_{t+1} \in \mathcal{I}} Q\left(s_{t+1}, i_{t+1} ; \theta_Q\right)$$

By differentiating the loss function w.r.t. θ_Q , we arrive at the following gradient:

$$\begin{aligned} \nabla_{\theta_Q} \ell\left(\theta_Q\right) \quad = \quad & \mathbb{E}_{\left(s_t, i_t, r_t, s_{t+1}\right) \sim \mathcal{M}}\left[\left(r+\gamma \max _{i_{t+1}} Q\left(s_{t+1}, i_{t+1} ; \theta_Q\right)\right.\right. \\ & \left.\left.-Q\left(s_t, i_t ; \theta_Q\right)\right) \nabla_{\theta_Q} Q\left(s_t, i_t ; \theta_Q\right)\right] \end{aligned}$$

Trick: We adapt a constantly increased γ during the training as $\gamma_e=\frac{1}{1+(E-e)^\eta}$. The increasing γ_e can be treated as an increasingly difficult curriculum, which gradually guides the learning agent from 1-horizon (greedy solution), 2-horizon, …. , to overall optimal solutions .

Experiments

- Dataset

Dataset	# User	# Item	# Interactions	# Interactions/ Per User	# Interactions/ Per Item
MovieLens 1M	6,040	3,706	1,000,209	165.60	269.89
EachMovie	1,623	61,265	2,811,718	1,732.42	45.89
Netflix	480,189	17,770	100,480,507	209.25	5,654.50

- Setting

Assume that the ratings recorded in the datasets are users' instinctive actions, not biased by the recommendations provided by the system. In this way, the records can be treated as unbiased to represent the feedback in an interactive setting. Additionally, we assume that the rating is no less than 4 is the satisfied recommendation, otherwise dissatisfied.

Experiments

- Baseline

Baseline	Illustration
Random	The random policy is executed in every recommendation, which is a baseline used to estimate the worst performance that should be obtained.
Pop	It ranks the items according to their popularity measured by the number of being rated. This is a widely used simple baseline. Although it is not personalized, it is surprisingly competitive in evaluation, as users tend to consume popular items.
Matrix Factorization	It suggests recommendations based on the ratings of other users who have similar ratings as the target user. For cold-start recommendation, we always greedy w.r.t. the estimated scores and update users' latent factor after every interaction.
MLP (SIGIR17)	Multi-layer perceptron has been a common practice for non-linear collaborative filtering due to its superiority. We deploy a MLP based recommender agent.
BPR (ICML09)	It optimizes the MF model with a pairwise ranking loss, which is a state-of-the-art model for item recommendation.
ICF (CIKM12)	Interactive collaborative filtering combined the probabilistic matrix factorization with different exploration techniques for recommender system, including GLM-UCB (generalized LinUCB), TS and ε-Greedy, which are strong baselines for handling exploration/exploitation dilemma in recommender system.
MeLU (SIGKDD19)	MeLU is a state-of-the-art method, which adapted MAML for solving the cold start problem by treating it as a few-shot task.
NICF	Our proposed approach for learning to explore in cold-start or warm-start recommendation.

Experiments

- Evaluation metrics

$$\text{precision@}T = \frac{1}{\#\ \text{users}} \sum_{\text{users}} \sum_{t=1}^{T} b_t .$$

$$\text{recall@}T = \frac{1}{\#\ \text{users}} \sum_{\text{users}} \sum_{t=1}^{T} \frac{b_t}{\#\ \text{satisfied items}} .$$

Here, $b_t = 1$ if $r_{u,i_t} >= 4$ and 0 otherwise.

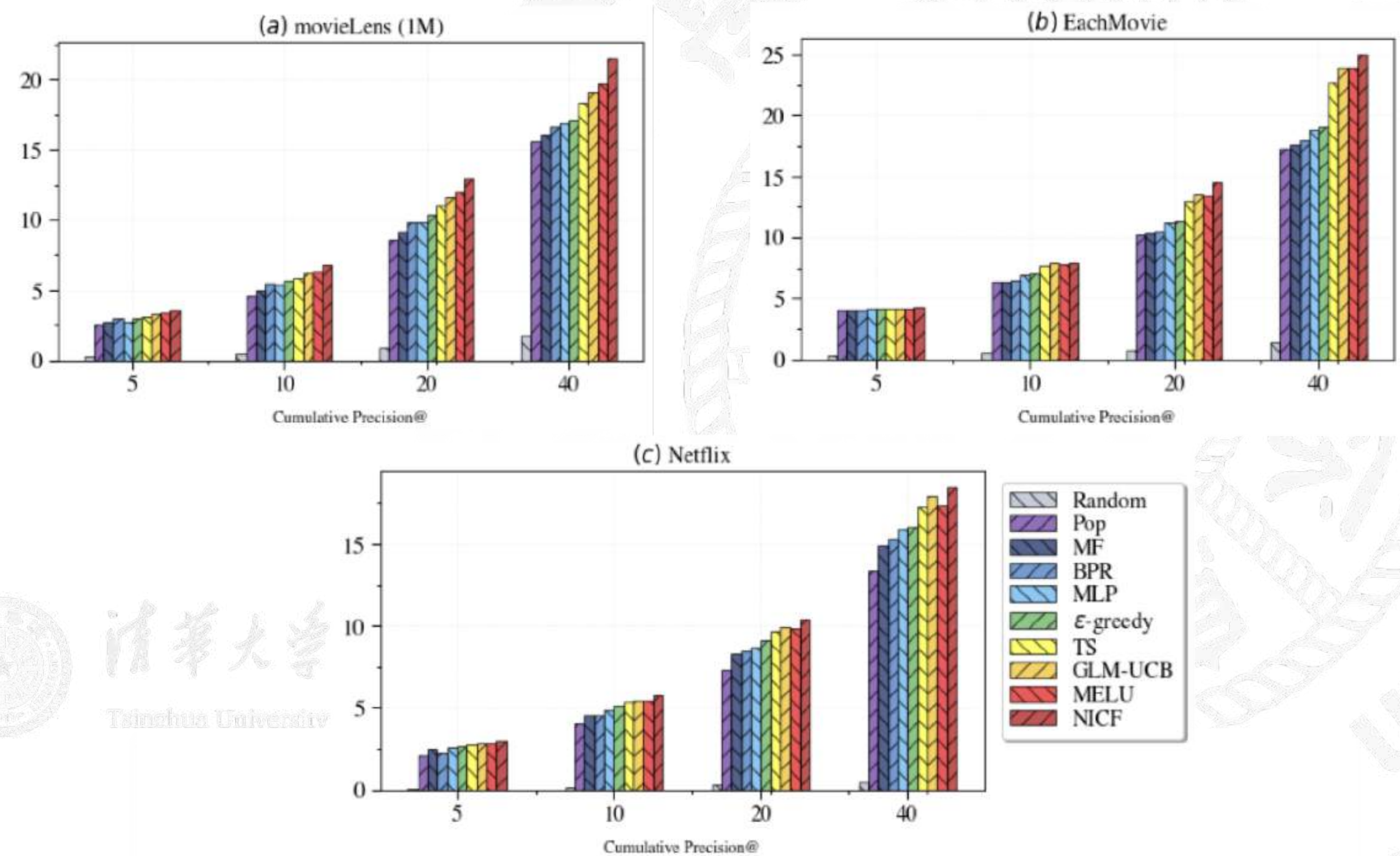
$$\alpha\text{-}NDCG@T = \frac{1}{\mathcal{Z}} \sum_{t=1}^{T} \frac{G@t}{\log(1+t)} .$$

Here, $G@t = \sum_{\forall i \in C} (1-\alpha)^{c_{i,t}-1}$ with $c_{i,t}$ as the number of times that topic i has appeared in the ranking of the recommendation list up to (and including) the t-th position.

Experiments

- Comparison on cold-start cases

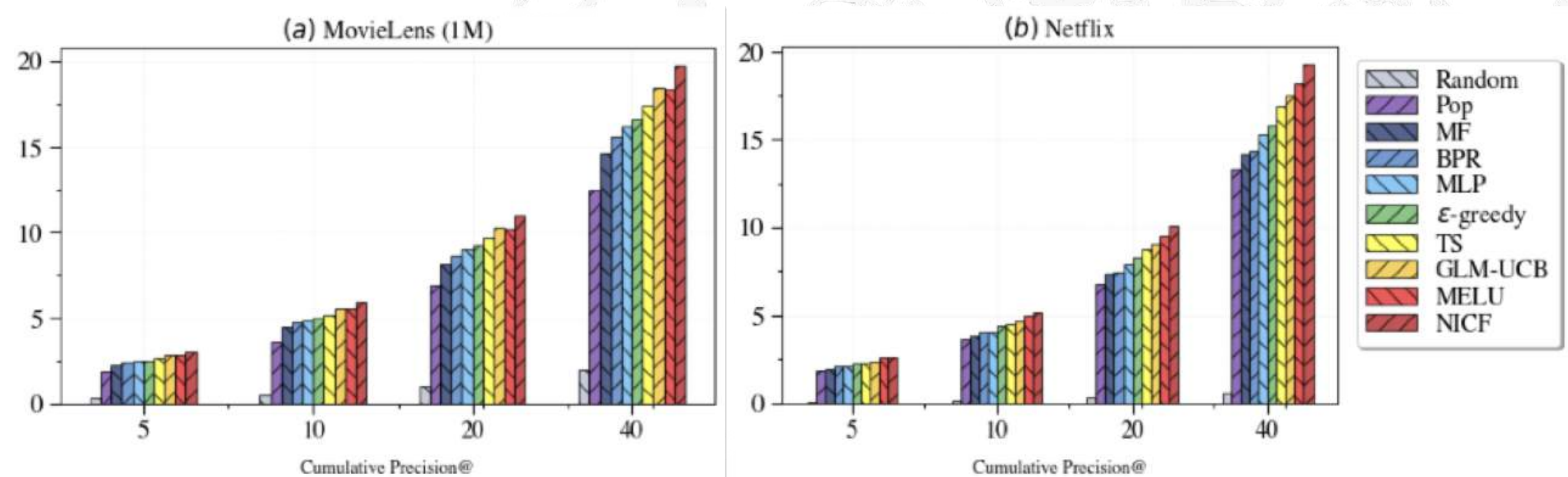
  Cold-start recommendation performance of different models on MovieLens (1M), EachMovie and Netflix Dataset.



Experiments

- Comparison on warm-start cases with taste drift

We respectively selected 4,600 users and 96,037 users whose rated items' topics have significantly changed from MovieLens (1M) and Netflix datasets to evaluate the performance on warm-start cases with taste drift.



Experiments

- Case study

The sequential decision tree learned by NICF without using the genre information of movies.

Thanks for your patience !

Any questions please contact

Lixin Zou
zoulixin15@gmail.com