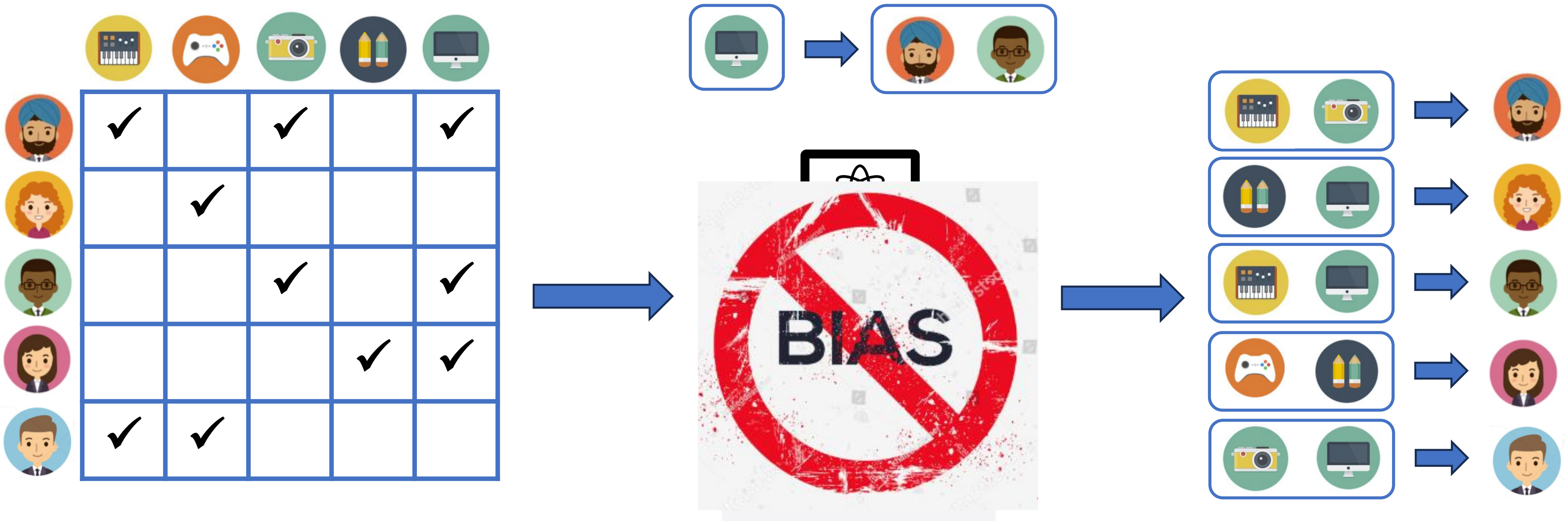
Countering Mainstream Bias via End-to-End Adaptive Local Learning

Jinhao Pan[1], Ziwei Zhu[2], Jianling Wang[1], Allen Lin[1], James Caverlee[1]

[1]Texas A&M University, [2]George Mason University

Recommenders Link Users with Items



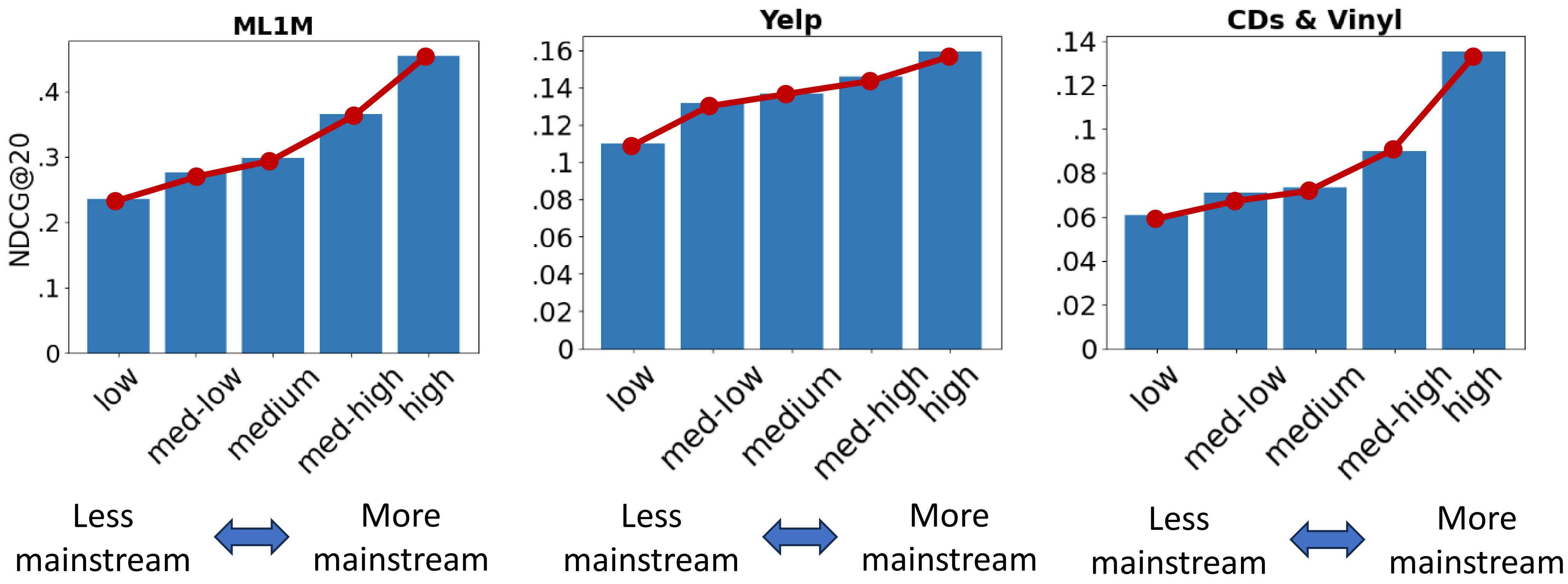
User-item interaction Data

Collaborative Filtering (CF) based Recommendation Model

Recommendation Results

Mainstream Bias is a Critical Problem in RecSys [Zhu, WSDM 2022]

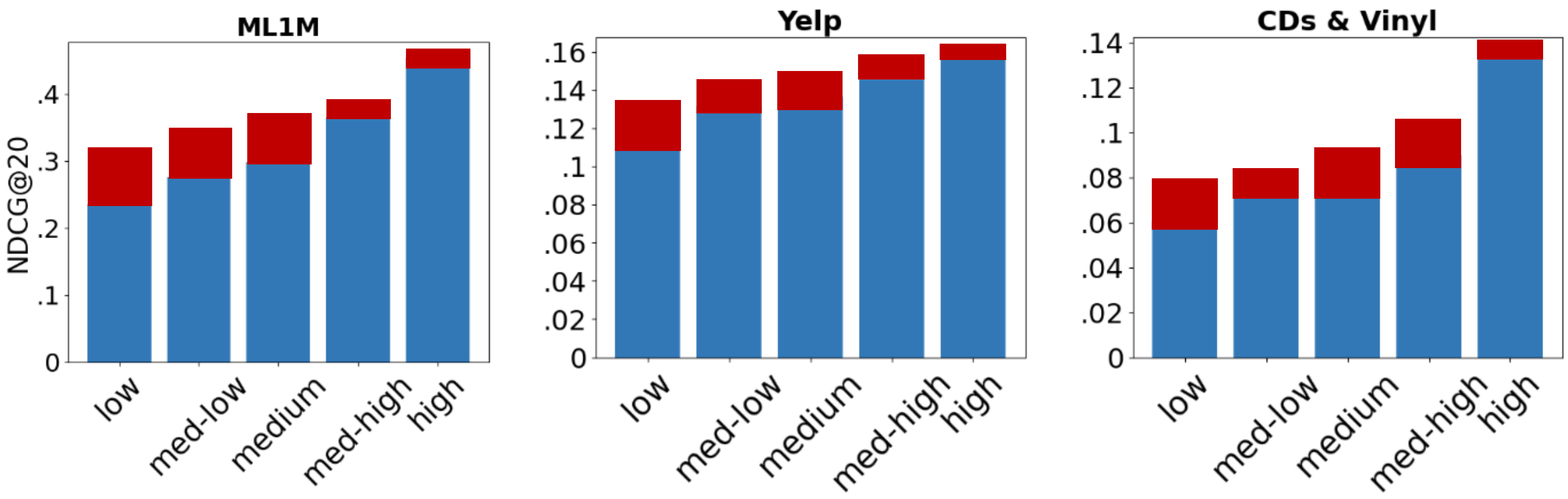
A CF-based algorithm delivers recommendations of higher utility to users with mainstream interests at the cost of poor recommendation performance for users with niche or minority interests.



Ziwei Zhu and James Caverlee. 2022. Fighting mainstream bias in recommender systems via local fine tuning. In Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining. 1497–1506.

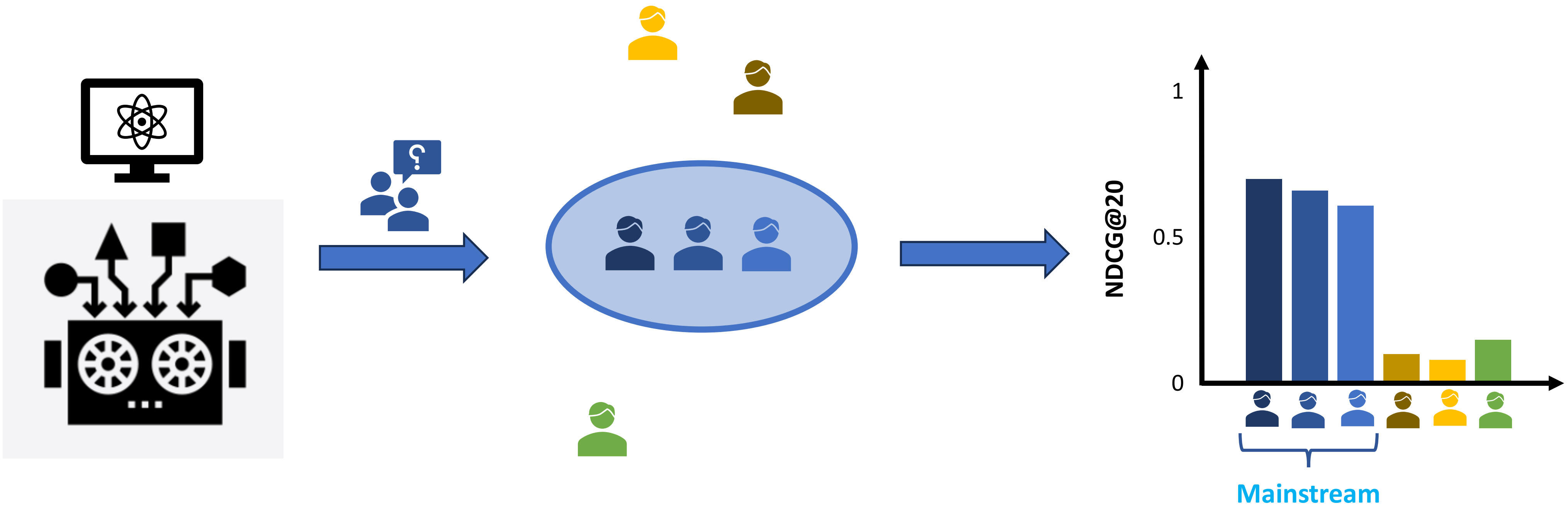
Goals on Alleviating Mainstream Bias [Rawls, 1958]

- Rawlsian Max-Min fairness principle of distribute justice
- To promote the average NDCG@20 for subgroups with low mainstream scores
- To preserve or even improve the utility for subgroups with high mainstream scores at the same time
- Anticipate an increase in the overall NDCG@20



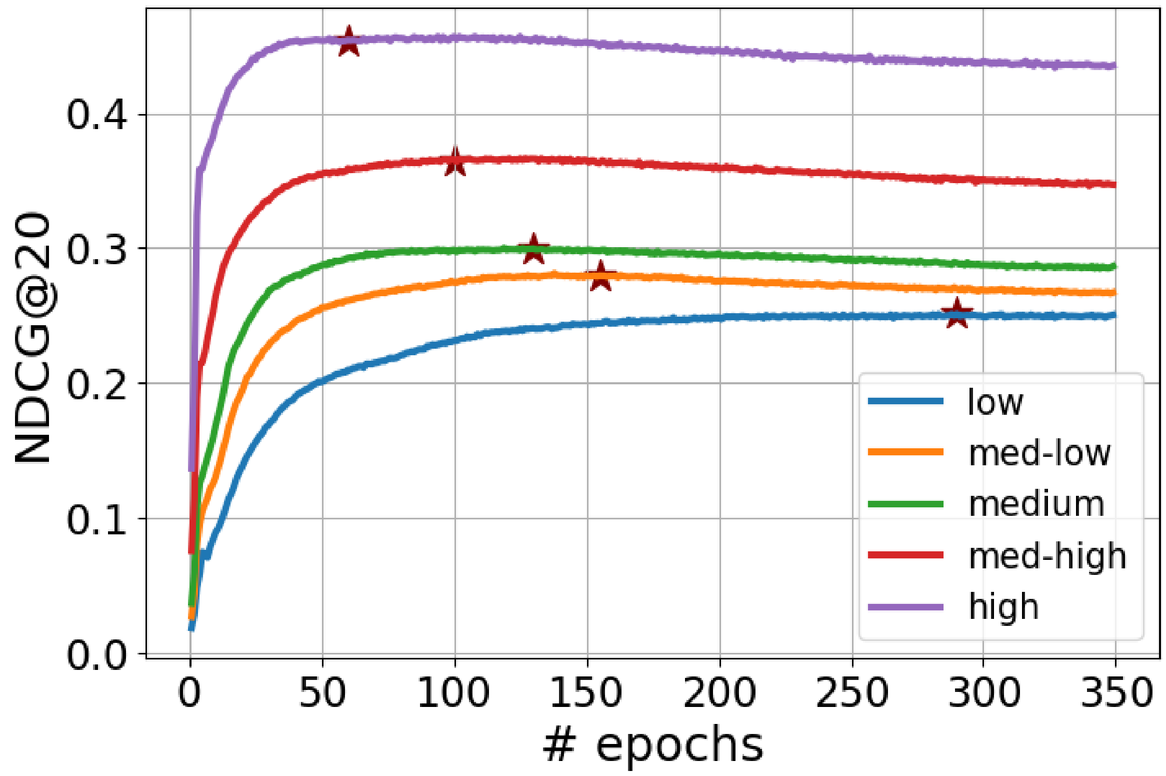
Challenges: Discrepancy Modeling

- Niche users have very different preference patterns from the majority, a model influenced by the discrepant mainstream users will result in poor recommendations for these niche users
- Find an adaptive method to build locally customized models for different types of users in an end-to-end fashion



Challenges: Unsynchronized Learning

- Mainstream users who have sufficient training signals can reach peak learning performance earlier than niche users
- Seek a solution to synchronize the learning paces for different types of users



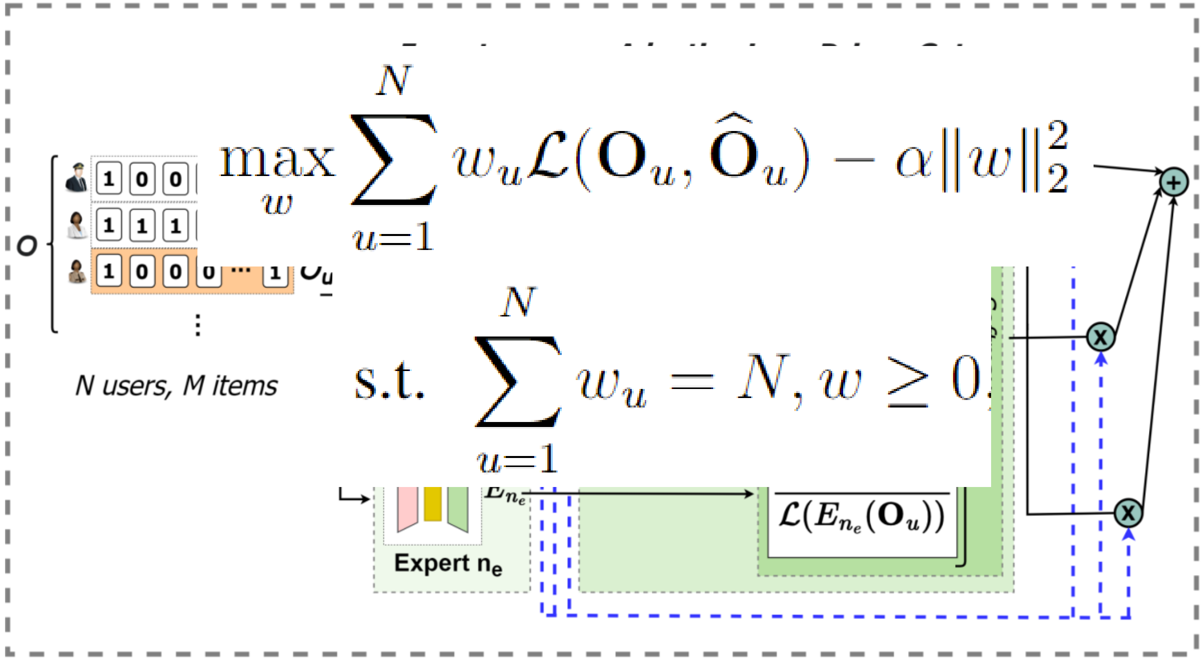
Proposed Method: TALL

- Proposed an end-To-end Adaptive Local Learning (TALL) to address mainstream bias

- Design centered around two challenges of this bias:
  - Discrepancy modeling
  - Unsynchronized learning

Proposed Debiasing Solution: Model Framework

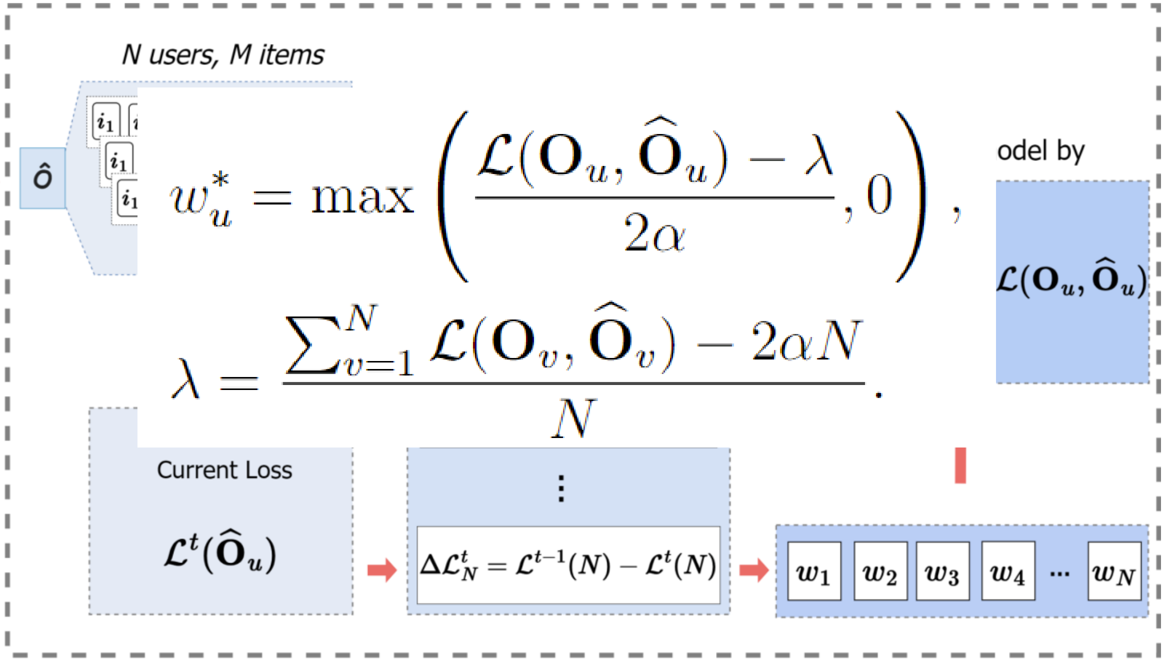
- end-**T**o-end **A**daptive **L**ocal **L**earning (TALL)
- Two main components fight against Two root causes of mainstream bias:
  - Loss-Driven MoE vs. Discrepancy Modeling
  - Adaptive Weight vs. Unsynchronized Learning



$$\max_{w} \sum_{u=1}^{N} w_u \mathcal{L}(\mathbf{O}_u, \widehat{\mathbf{O}}_u) - \alpha \|w\|_2^2$$

$$\text{s.t.} \sum_{u=1}^{N} w_u = N, w \geq 0$$

Loss-Driven Mixture-of-Experts



$$w_u^* = \max\left(\frac{\mathcal{L}(\mathbf{O}_u, \widehat{\mathbf{O}}_u) - \lambda}{2\alpha}, 0\right),$$

$$\lambda = \frac{\sum_{v=1}^{N} \mathcal{L}(\mathbf{O}_v, \widehat{\mathbf{O}}_v) - 2\alpha N}{N}.$$

Adaptive Weight

Debiasing Experiments and Results: Experimental Setup

- Three public datasets:
  - ML1M
  - Yelp
  - Amazon CDs & Vinyl
- Compare the average NDCG@20 for each subgroup
- To promote the average NDCG@20 for subgroups with low mainstream scores while preserving or even improving the utility for subgroups with high mainstream scores at the same time. We should also anticipate an increase in the overall NDCG@20 of the model

	#users	#items	density
ML1M	6,040	3,706	4.46%
Yelp	20,001	7,643	0.32%
CDs & Vinyl	12,023	8,050	0.32%

Debiasing Experiments and Results: Debiasing Performance [Choi, WSDM 2021][ZHU, WSDM 2022]

- **MultVAE –** Vanilla recommendation model without debiasing
- **WL –** Global model, puts more weights on niche users' loss function
- **LOCA –** Local Collaborative Autoencoder, offers predication for a target user by aggregating different anchor models
- **LFT –** SOTA local learning model, which fine-tunes a global model with a small collection of local data for each user so that each user will be predicted by a customized local model
- **EnLFT –** ensembled version of LFT, which is similar to LOCA but trains the anchor models by the approach of LFT

MultVAE: Dawen Liang, Rahul G Krishnan, Matthew D Hoffman, and Tony Jebara. 2018. Variational autoencoders for collaborative filtering. In Proceedings of the 2018 world wide web conference. 689–698.
LOCA: Minjin Choi, Yoonki Jeong, Joonseok Lee, and Jongwuk Lee. 2021. Local Collaborative Autoencoders. In Proceedings of the 14th ACM International Conference on Web Search and Data Mining. 734–742.
WL, EnLFT, LFT: Ziwei Zhu and James Caverlee. 2022. Fighting mainstream bias in recommender systems via local fine tuning. In Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining. 1497–1506.

Debiasing Experiments and Results: Debiasing Performance

	ML1M						Yelp						CDs & Vinyl					
	NDCG	Subgroups of mainstream levels					NDCG	Subgroups of mainstream levels					NDCG	Subgroups of mainstream levels				
	@20	L	ML	M	MH	H	@20	L	ML	M	MH	H	@20	L	ML	M	MH	H
MultVAE	.3260	.2354	.2764	.2986	.3652	.4546	.0877	.0686	.0710	.0733	.0901	.1355	.1367	.1100	.1316	.1366	.1457	.1596
WL	.3278	.2448	.2801	.2970	.3639	.4532	.0870	.0700	.0708	.0720	.0888	.1332	.1361	.1133	.1334	.1361	.1441	.1534
EnLFT	.3341	.2586	.2875	.3025	.3661	.4556	.0887	.0697	.0715	.0740	.0915	.1369	.1453	.1230	.1387	.1423	.1561	.1666
LOCA	.3308	.2551	.2780	.2972	.3622	.4617	.0942	.0723	.0758	.0764	.0970	.1494	.1573	.1341	.1510	.1556	.1665	.1795
LFT	.3416	.2707	**.2918**	.3072	.3727	.4657	.0927	.0740	.0738	.0768	.0956	.1432	.1557	.1343	.1481	.1515	.1678	.1770
TALL	**.3456**	**.2746**	.2903	**.3112**	**.3784**	**.4734**	**.0992**	**.0772**	**.0803**	**.0826**	**.1056**	**.1505**	**.1700**	**.1392**	**.1599**	**.1652**	**.1844**	**.2013**
$\Delta_{MultVAE}(\%)$	6.01	16.65	5.03	4.22	3.61	4.14	13.11	12.54	13.1	12.69	17.2	11.07	24.36	26.54	21.5	20.94	26.56	26.13
$\Delta_{LFT}(\%)$	1.17	1.44	-0.51	1.30	1.53	1.65	7.01	4.32	8.81	7.55	10.46	5.10	9.18	3.65	7.97	9.04	9.89	13.73
$\Delta_{LOCA}(\%)$	4.47	7.64	4.42	4.71	4.47	2.53	5.31	6.78	5.94	8.12	8.87	0.74	8.07	3.80	5.89	6.17	10.75	12.14

L: low, ML: med-low, M: medium, MH: med-high, H: high

Preserve utility for mainstream users at the same time.

Debiasing Experiments and Results: Debiasing Performance

More experiment details and results in paper!

Thank you!
Q&A