A Lightweight Method for Modeling Confidence in Recommendations with Learned Beta Distributions (LBD)

Cristobal Vasquez, Rodrigo Pozo,
Luis Miranda, Oscar Loch

Professor : Andrés Francisco Carvallo de Ferari

Departamento de Ciencias de la Computacion

Pontificia Universidad Católica de Chile

November 7 of 2023

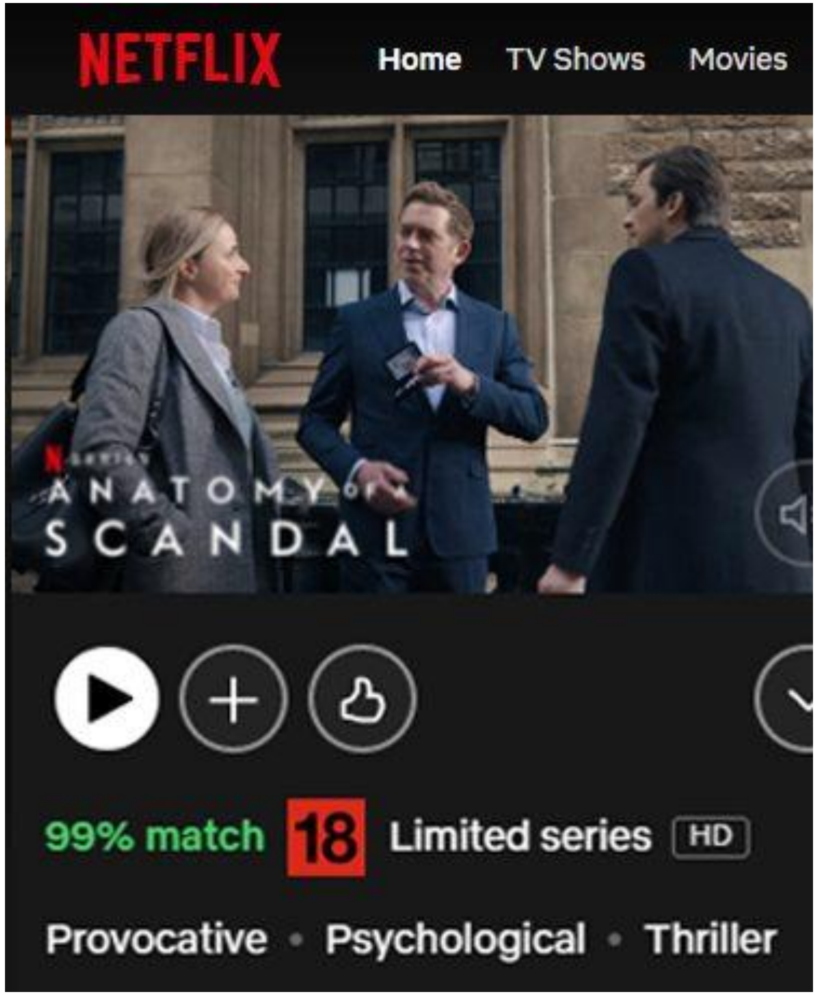
Outline

Outline

Why have a confidence level?

- It provide insights into the reliability of recommendations

- Personalization



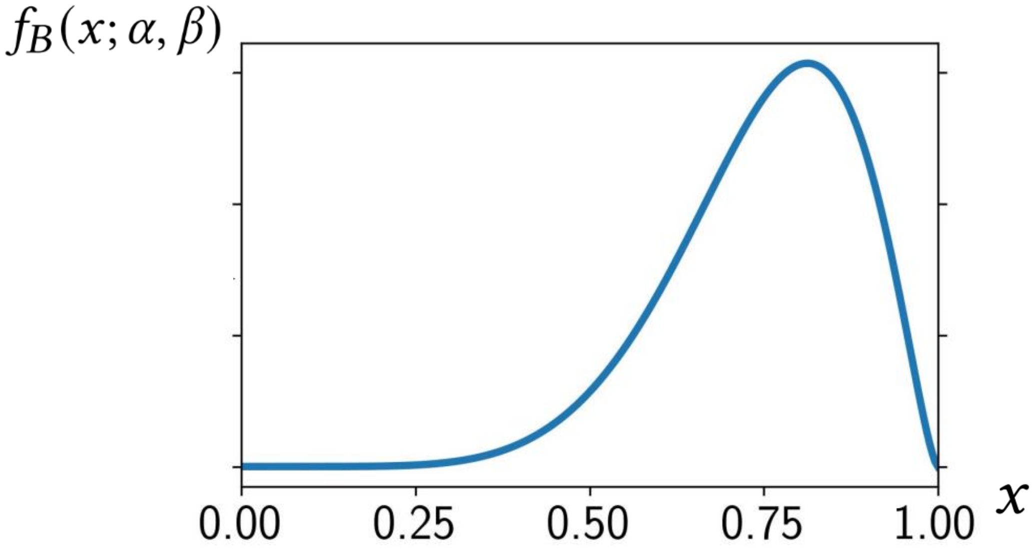
Related work

Method	Confidence	Distribution	Ranking bounding
CMF	Yes	Gaussian	No
OrdRec	Yes	Logistic	Yes
MF	No	N/A	Yes
LBD	**Yes**	**Beta (discrete)**	**Yes**

Outline

From Beta Distribution to a Discrete Rating Distribution

Beta Distribution



$$f_B(x;\alpha,\beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$$

- Parameters: $\alpha, \beta > 0$. Normalization factor: $B(\alpha,\beta) = \int_0^1 f_B(t;\alpha,\beta)dt$

Parametrization by mean μ and sample size ν

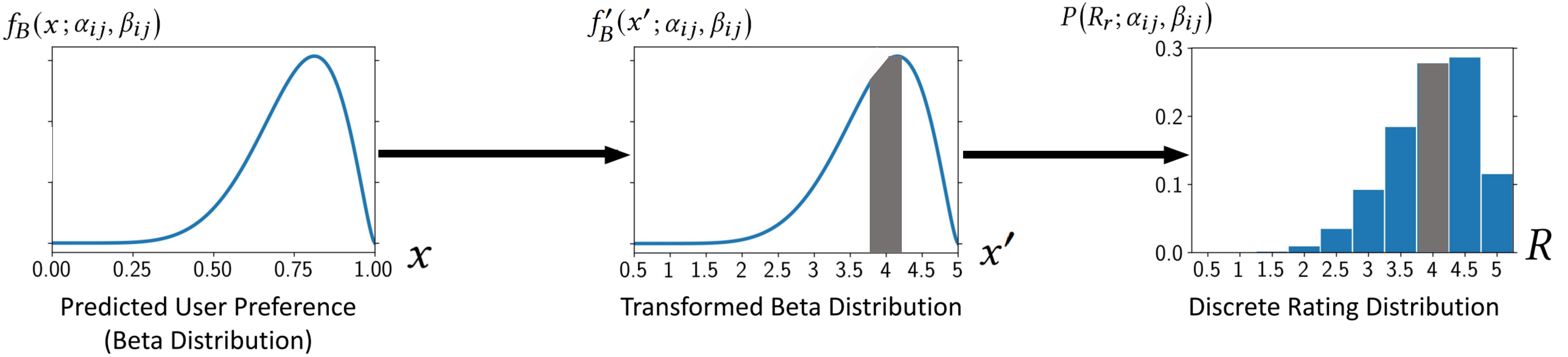
$$\mu = \alpha/(\alpha+\beta) \qquad \mu \in (0,1)$$

$$\nu = \alpha+\beta \qquad \nu \in (0,\infty)$$

- ν as “confidence”

From Beta Distribution to a Discrete Rating Distribution

- Assuming that user i provides ratings: $R \in \{R_1, R_2, \ldots, R_n\}$ for an item j:



$$x' = x(R^{\max} - R^{\min}) + R^{\min}$$

Modeling Means and Confidences

Mean and confidence parameters proposed

$$\mu_{ij} = \frac{1}{2} + \frac{1}{2}\text{Cosine-similarity}(U_i, V_j) = \frac{1}{2} + \frac{1}{2}\frac{U_i^\top V_j}{||U_i||\,||V_j||}$$

$$v_{ij}^{\text{norm}} = ||U_i||\,||V_j||, \quad v_{ij}^{\text{sum}} = ||U_i + V_j||, \quad v_{ij}^{\text{dot}} = |U_i^\top V_j|$$

- Where $U_i \in \mathbb{R}^D$ and $V_j \in \mathbb{R}^D$ are embeddings for each user i and item j

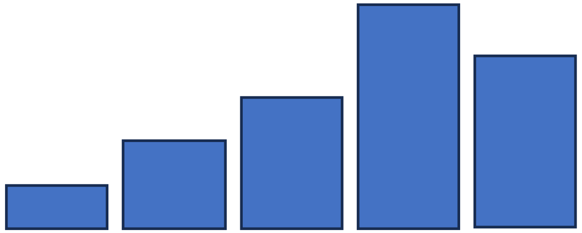
Modeling biases and discretization strategies

Modeling User and items biases

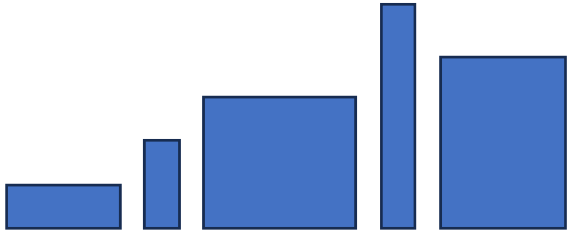
- Weights for α_{ij}, β_{ij} : $\longrightarrow$ $f_B(a_0 + a_i + a_j + \alpha_{ij}, b_0 + b_i + b_j + \beta_{ij})$
- Weights for μ_{ij} ν_{ij} : $\longrightarrow$ $u_0\ u_i\ u_j$ $v_0\ v_i\ v_j$

Modeling static and adaptive discretization strategies

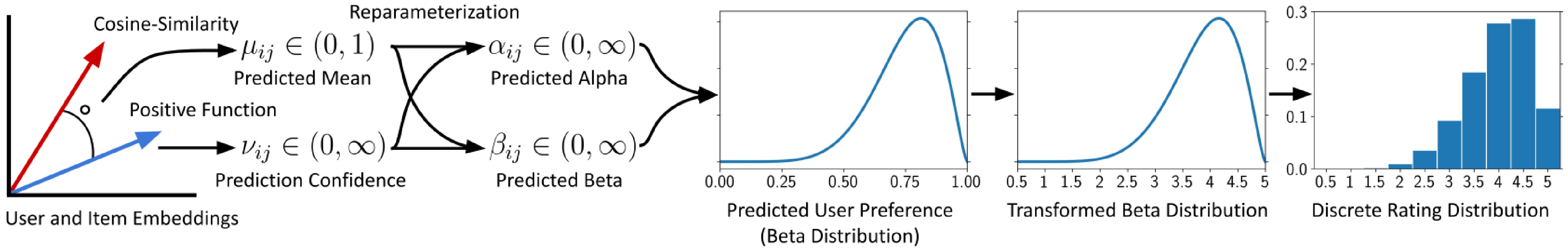
- Static Strategy (LBD-S) makes each bin equisized $\longrightarrow$



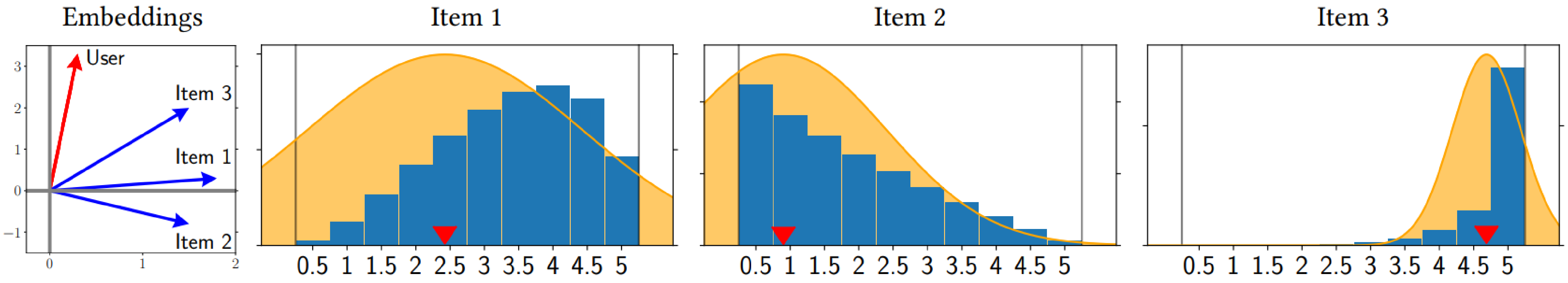
- Adaptative Strategy (LBD-A) varies the bin widths per rating, user and item. $\longrightarrow$



Overview of the LBD Method



Comparison with existing methods



- The prediction of MF is pointwise; CMF is a Gaussian distribution; and LBD is a discrete rating prediction.

Outline

Research questions

- What modeling of the parameters of LBD results in the highest predictive performance for rating prediction?
- Does our LBD approach provide rating prediction performance that is competitive with MF, CMF and OrdRec?
- Is there a stronger correlation between the confidence and accuracy of predictions by LBD than for CMF and OrdRec?
- Does the confidence modeling of LBD translate to higher performance compared to MF, CMF and OrdRec in a highprecision targeted recommendation task?

Baselines, metrics and parameter tuning

Baseline and metrics

Baselines
- MF
- CMF
- OrdRec

New Methods
- LBD-S
- LBD-A

Metrics

RMSE
Accuracy
NDCG@k

Peformance

- Classification
- Regression
- Ranking

Parameter tuning

- Extensive random search
- Different embedding sizes
- RMSE on validation test

Outline

Parametrization of the LBD Model

- What modeling of the parameters of LBD results in the highest predictive performance for rating prediction?

Size	$\nu^{(\cdot)}$	Bias	RMSE	MAE	Accuracy	Ave. Log-L.	NDCG@3	NDCG@10
512	*sum*	-	0.7932 (0.0005)	0.6130 (0.0004)	0.3016 (0.0005)	-1.769 (0.001)	0.9330 (0.004)	0.9568 (0.0002)
512	*sum*	μ, ν	0.7864 (0.0005)	0.6058 (0.0003)	0.3092 (0.0003)	-1.760 (0.001)	0.9339 (0.0003)	0.9568 (0.0001)
512	*sum*	α, β	0.7761 (0.0005)	0.5895 (0.0005)	0.3139 (0.0003)	-1.755 (0.003)	**0.9357** (0.0003)	**0.9583** (0.0001)
512	*norm*	α, β	0.7852 (0.0007)	0.5942 (0.0005)	0.3140 (0.0005)	-1.818 (0.004)	0.9336 (0.0002)	0.9572 (0.0002)
512	*dot*	α, β	0.8387 (0.0005)	0.6391 (0.0006)	0.2972 (0.0005)	-1.865 (0.005)	0.9330 (0.0004)	0.9495 (0.0001)
256, 256	*sum*	α, β	0.7850 (0.0006)	0.5964 (0.0006)	0.3133 (0.0006)	-1.774 (0.002)	0.9331 (0.0003)	0.9570 (0.0001)
512 + 10	*sum*	α, β	**0.7759** (0.0004)	**0.5875** (0.0004)	**0.4356** (0.0006)	**-1.432** (0.002)	0.9351 (0.0003)	0.9579 (0.0001)

Best Results: confidence function ν^{sum}, α and β bias terms, both with and without dynamic binning

Recommendation Performance

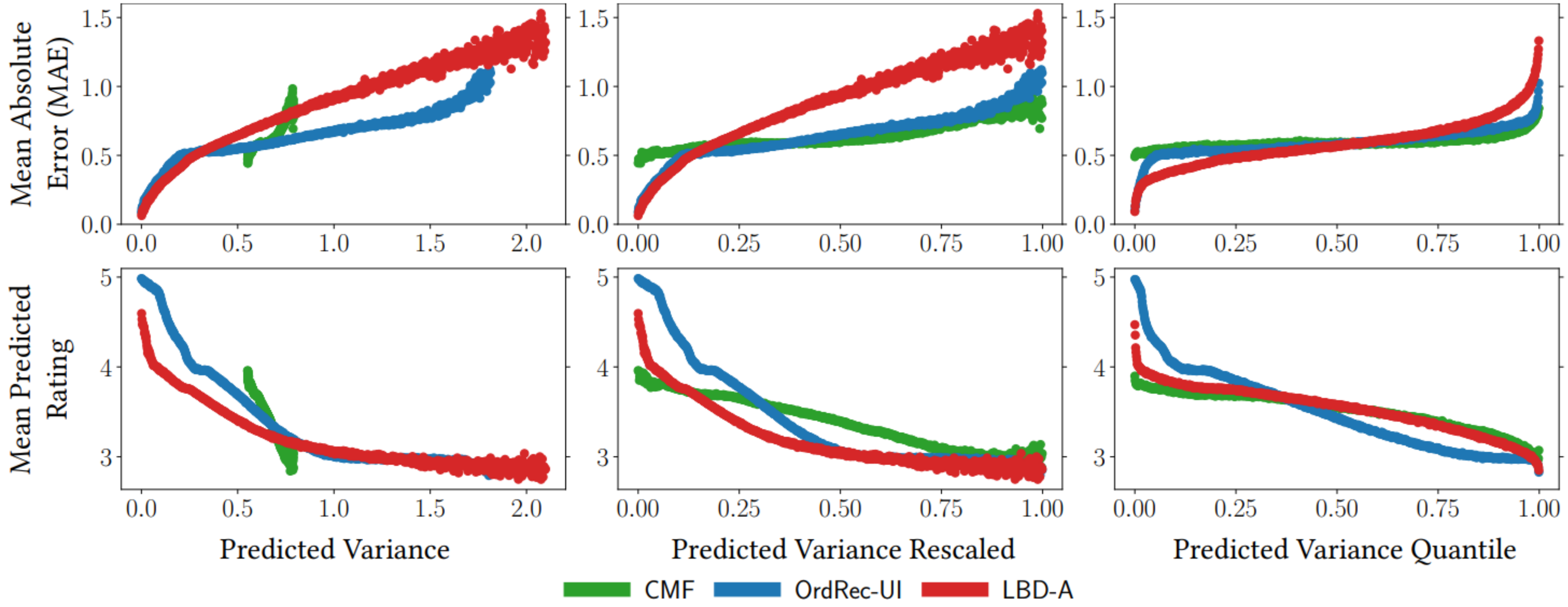
- Does our LBD approach provide rating prediction performance that is competitive with MF, CMF and OrdRec?

Model	RMSE	MAE	Accuracy	Ave. Log-L.	NDCG@3	NDCG@10
MF (128)	0.7802 (0.0006)	0.5984 (0.0005)	0.2923 (0.0005)	-2.004 (0.002)	0.9334 (0.0002)	0.9570 (0.0001)
MF (512)	0.7883 (0.0006)	0.6022 (0.0005)	0.2934 (0.0004)	-2.022 (0.002)	0.9321 (0.0002)	0.9560 (0.0001)
CMF (128)	0.7760 (0.0007)	0.5936 (0.0004)	0.3226 (0.0004) $^{\triangledown}$	-1.777 (0.001)	0.9338 (0.0002)	0.9573 (0.0001)
CMF (512)	0.7820 (0.0005)	0.5967 (0.0003)	0.2849 (0.0015)	-1.801 (0.001)	0.9329 (0.0003)	0.9567 (0.0001)
OrdRec-U (512)	0.7821 (0.0006)	0.6043 (0.0004)	0.2322 (0.0007)	-1.881 (0.001)	0.9343 (0.0003)	0.9575 (0.0001)
OrdRec-UI (512)	0.7765 (0.0006)	0.5896 (0.0005)	0.4187 (0.0006)$^{\triangledown}$	-1.569 (0.001)$^{\triangledown}$	0.9349 (0.0003)	0.9578 (0.0001)
LBD-S (512)	0.7761 (0.0005)	0.5895 (0.0005)	0.3139 (0.0003)	-1.755 (0.003)	**0.9357** (0.0003) $^{\triangle}$	**0.9583** (0.0001)$^{\triangle}$
LBD-A (512)	**0.7759** (0.0004)	**0.5875** (0.0004) $^{\triangle}$	**0.4356** (0.0006) $^{\triangle}$	**-1.432** (0.002) $^{\triangle}$	0.9351 (0.0003)	0.9579 (0.0001)

Evaluating prediction Confidence

- Is there a stronger correlation between the confidence and accuracy of predictions by LBD than for CMF and OrdRec?

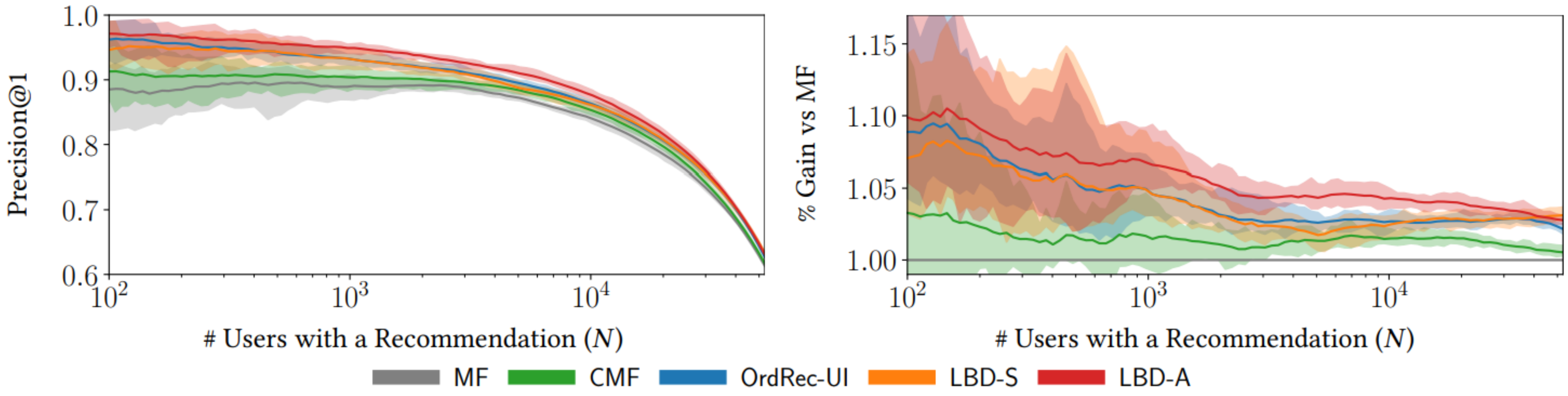
	CMF	OrdRec-U	OrdRec-UI	LBD-S	LBD-A
Pearson's r (Linear Correlation)	0.079 (0.001)	0.157 (0.001)	0.175 (0.001)	0.304 (0.002)$^{\Delta}$	**0.329** (0.002)$^{\Delta}$
Kendall's τ (Rank Correlation)	0.043 (0.002)	0.094 (0.001)	0.112 (0.001)	0.194 (0.001)$^{\Delta}$	**0.216** (0.001)$^{\Delta}$



High precision targeted Recommendation

- Does the confidence modeling of LBD translate to higher performance compared to MF, CMF and OrdRec in a highprecision targeted recommendation task?

Model	Precision@1 for top N users					
	$N = 100$	$N = 320$	$N = 1000$	$N = 3200$	$N = 10000$	$N = 32000$
MF	0.885 (0.034)	0.893 (0.019)	0.889 (0.011)	0.886 (0.004)	0.841 (0.004)	0.726 (0.002)
CMF	0.913 (0.030)	0.907 (0.016)	0.904 (0.013)	0.895 (0.005)	0.853 (0.005)	0.733 (0.002)
OrdRec-UI	0.962 (0.016)	0.949 (0.007)	0.932 (0.006)	0.910 (0.003)	0.863 (0.003)	0.747 (0.002)
LBD-S	0.947 (0.019)	0.944 (0.013)	0.932 (0.009)	0.908 (0.006)	0.861 (0.004)	0.747 (0.003)
LBD-A	**0.971** (0.012)	**0.962** (0.011)$^{\Delta}$	**0.949** (0.005)$^{\Delta}$	**0.925** (0.004)$^{\Delta}$	**0.877** (0.004)$^{\Delta}$	**0.751** (0.003)$^{\Delta}$



Outline

Conclusions and Future Work

Conclusions:

- Low model complexity
- Strong correlation between correlation and accuracy
- Higher precision in recommendation task

Future work:

- Understanding User Preferences and Dataset Uncertainty using LBD
- Confidence Modeling for Implicit Feedback Signals

Gracias 😎