ProtoMF: Prototype-based Matrix Factorization for Effective and Explainable Recommendations

Alessandro B. Melchiorre
alessandro.melchiorre@jku.at
https://karapostk.github.io

Navid Rekab-saz
navid.rekabsaz@jku.at

Christian Ganhör
christian.ganhoer@jku.at

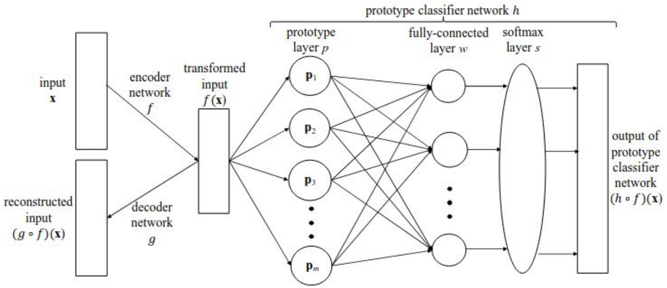
Markus Schedl
markus.schedl@jku.at

JOHANNES KEPLER UNIVERSITY LINZ

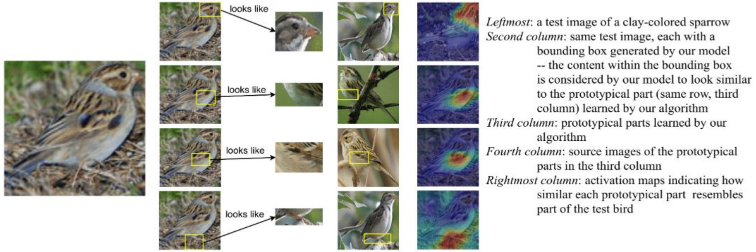
MULTIMEDIA MINING AND SEARCH GROUP

Institute of Computational Perception

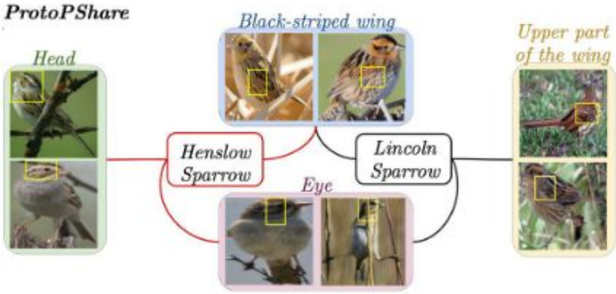
Prototype-based Machine Learning Models



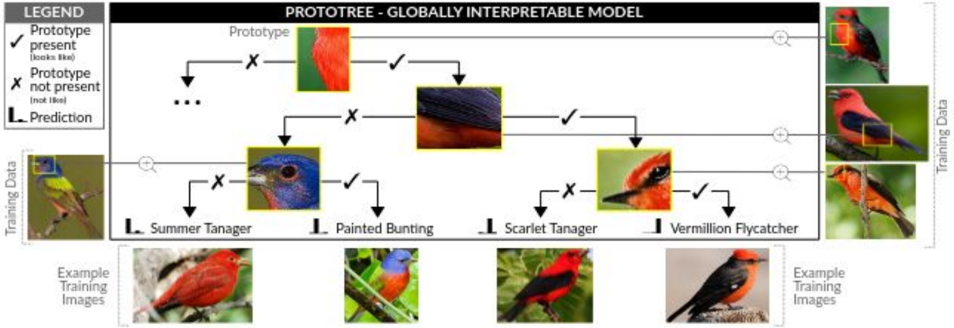
[1] Li, Oscar, et al. "Deep learning for case-based reasoning through prototypes: A neural network that explains its predictions." *AAAI* 2018.



[2] Chen, Chaofan, et al. "This looks like that: deep learning for interpretable image recognition." *NeurIPS* 2019.



[3] Rymarczyk, Dawid, et al. "Protopshare: Prototypical parts sharing for similarity discovery in interpretable image classification." SIGKDD 2021.



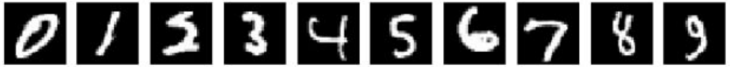
[4] Nauta, Meike, et al. "Neural prototype trees for interpretable fine-grained image recognition." CVPR 2021.

JOHANNES KEPLER UNIVERSITY LINZ

Prototypes

*“A Prototype is an object that is representative of a set of similar instances and is part of the observed points, or **it is an artifact summarizing a subset of them with similar characteristics**”*

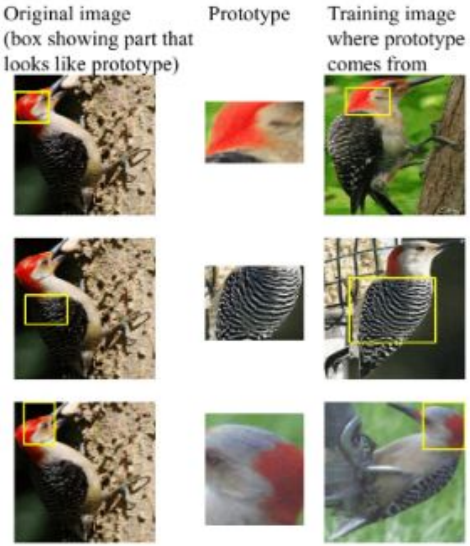
[5] Guidotti, Riccardo, et al. "A survey of methods for explaining black box models." CSUR 2018.



Data points

Prototypes

[1] Li, Oscar, et al. "Deep learning for case-based reasoning through prototypes: A neural network that explains its predictions." *AAAI* 2018.



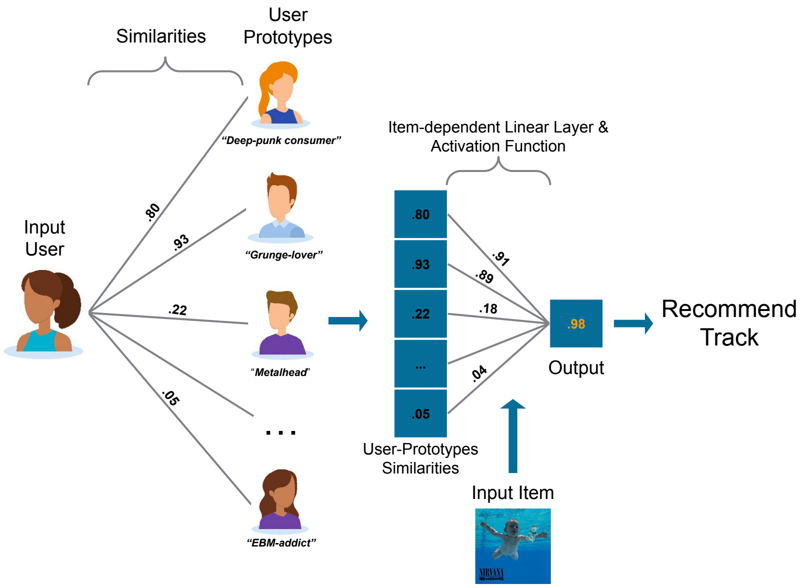
[2] Chen, Chaofan, et al. "This looks like that: deep learning for interpretable image recognition." *NeurIPS* 2019.

Aim:

Explain recommendations of matrix factorization approaches using item and user prototypes.

Prototype-based Recommender Systems

Prediction

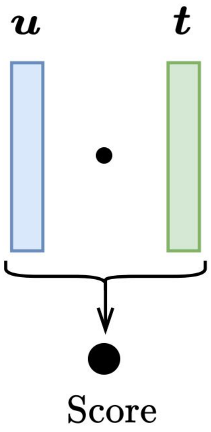


Sample Textual Explanation

The input user and input item are a good match since the input user is very similar to the user prototypes and which are compatible with the input item.

Prototyped-based Matrix Factorization (ProtoMF)

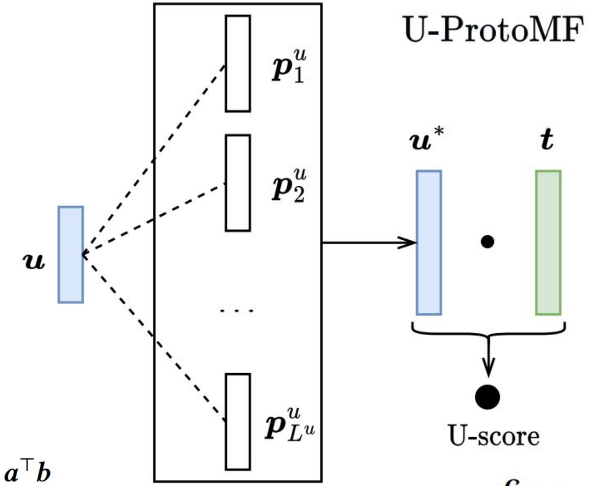
Matrix Factorization



$$\mathcal{L}_{rec}(\mathcal{I}, \Theta) = -\sum_{(u,t) \in \mathcal{I}} \ln p(t|u) + \lambda_{L2} \|\Theta\|, \quad p(t|u) = \frac{e^{\mathrm{Score}(u,t)}}{\sum_{j=1}^{M} e^{\mathrm{Score}(u,t_j)}}$$

[6] Koren, Yehuda, et al.. "Matrix factorization techniques for recommender systems." Computer 2009.
[7] Rendle, Steffen, et al. "BPR: Bayesian personalized ranking from implicit feedback." UAI 2009.

User Prototype Matrix Factorization (U-ProtoMF)



$$\boldsymbol{u}^* = \begin{bmatrix} \mathrm{sim}(\boldsymbol{u}, \boldsymbol{p}_1^u) \\ \dots \\ \mathrm{sim}(\boldsymbol{u}, \boldsymbol{p}_{L^u}^u) \end{bmatrix} \in \mathbb{R}^{L^u}, \qquad \mathrm{sim}(\boldsymbol{a}, \boldsymbol{b}) = 1 + \frac{\boldsymbol{a}^\top \boldsymbol{b}}{\|\boldsymbol{a}\| \cdot \|\boldsymbol{b}\|}$$

$$\text{U-score}(u, t) = \sum_{l=1}^{L^u} s_l^{\text{user}}, \qquad \boldsymbol{s}^{\text{user}} = \boldsymbol{u}^* \odot \boldsymbol{t}$$

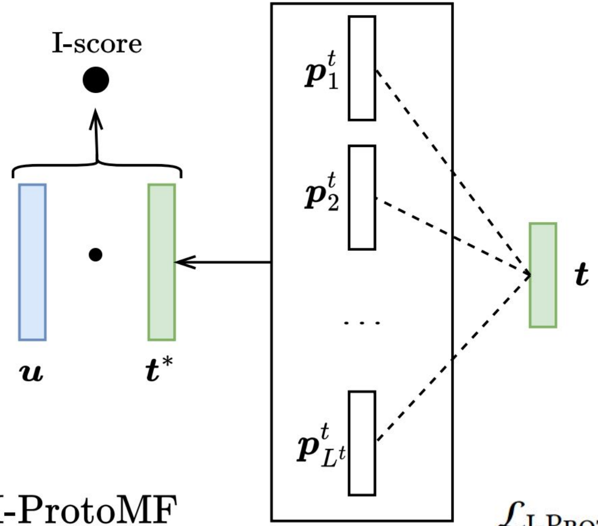
$$\mathcal{L}_{\text{U-PROTO}} = \mathcal{L}_{rec} + \lambda_1 R_{\{\mathcal{P}^u \to \mathcal{U}\}} + \lambda_2 R_{\{\mathcal{U} \to \mathcal{P}^u\}}$$

$$R_{\{\mathcal{U} \to \mathcal{P}^u\}} = -\frac{1}{N} \sum_{i=1}^{N} \max_{l \in [1,..,L^u]} \mathrm{sim}(\boldsymbol{u}_i, \boldsymbol{p}_l^u)$$

$$R_{\{\mathcal{P}^u \to \mathcal{U}\}} = -\frac{1}{L^u} \sum_{l=1}^{L^u} \max_{i \in [1,..,N]} \mathrm{sim}(\boldsymbol{u}_i, \boldsymbol{p}_l^u)$$

JOHANNES KEPLER UNIVERSITY LINZ

Item Prototype Matrix Factorization (I-ProtoMF)



I-ProtoMF

$$\boldsymbol{t}^* = \begin{bmatrix} \mathrm{sim}(\boldsymbol{t}, \boldsymbol{p}_1^t) \\ \dots \\ \mathrm{sim}(\boldsymbol{t}, \boldsymbol{p}_{L^t}^t) \end{bmatrix} \in \mathbb{R}^{L^t}$$

$$\text{I-score}(u, t) = \sum_{l=1}^{L^t} s_l^{\text{item}}, \quad \boldsymbol{s}^{\text{item}} = \boldsymbol{t}^* \odot \boldsymbol{u}$$

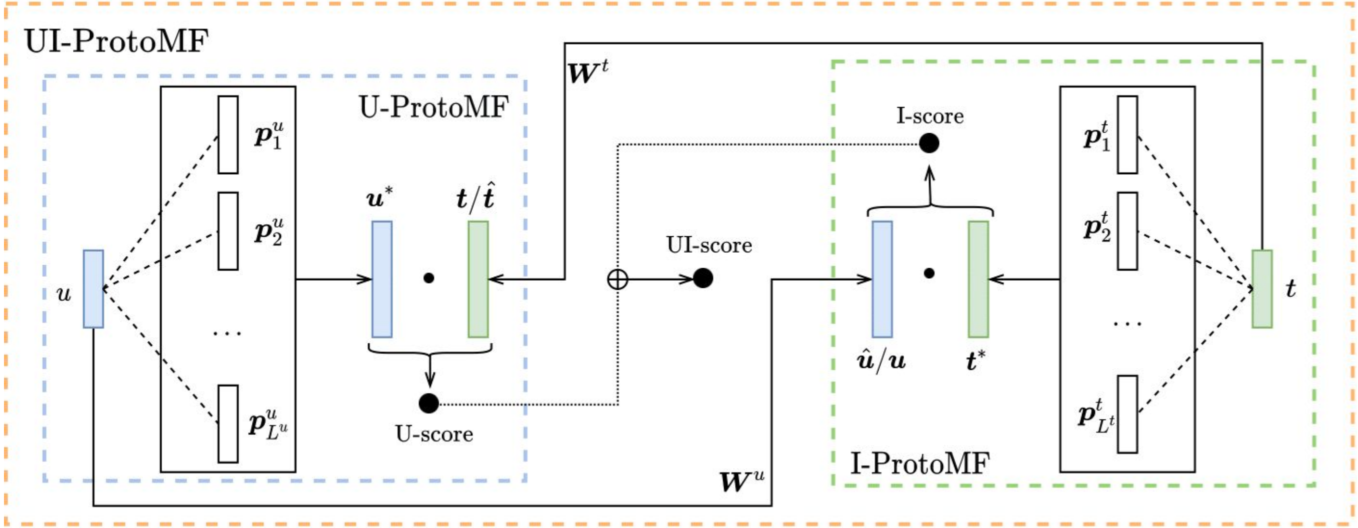
$$\mathcal{L}_{\text{I-PROTO}} = \mathcal{L}_{rec} + \lambda_3 R_{\{\mathcal{P}^t \to \mathcal{T}\}} + \lambda_4 R_{\{\mathcal{T} \to \mathcal{P}^t\}}$$

$$R_{\{\mathcal{P}^t \to \mathcal{T}\}} = -\frac{1}{L^t} \sum_{l=1}^{L^t} \max_{j \in [1,..,M]} \mathrm{sim}(\boldsymbol{t}_j, \boldsymbol{p}_l^t)$$

$$R_{\{\mathcal{T} \to \mathcal{P}^t\}} = -\frac{1}{M} \sum_{i=1}^{M} \max_{l \in [1,..,L^t]} \mathrm{sim}(\boldsymbol{t}_i, \boldsymbol{p}_l^t)$$

JOHANNES KEPLER
UNIVERSITY LINZ

User-Item Prototype Matrix Factorization (UI-ProtoMF)



$$\text{UI-score}(u,t) = \text{U-score}(u,t) + \text{I-score}(u,t)$$

$$\hat{\boldsymbol{u}} = \boldsymbol{W}^u \boldsymbol{u} \in \mathbb{R}^{L^t}, \quad \hat{\boldsymbol{t}} = \boldsymbol{W}^t \boldsymbol{t} \in \mathbb{R}^{L^u}$$

$$\mathcal{L}_{\text{UI-PROTO}} = \mathcal{L}_{\text{U-PROTO}} + \mathcal{L}_{\text{I-PROTO}}$$

JOHANNES KEPLER UNIVERSITY LINZ

Experiment Setup

- Datasets:
  - *MovieLens-1M* [8]
  - *LFM2B-1Month* [9]
  - *Amazon Video Games* [10]
- Baselines:
  - *Matrix Factorization* [6]
  - *Representative-based Matrix Factorization* [11]
  - *Anchor-based Collaborative Filtering* [12]
- *NDCG* and *Hit Ratio* at 5, 10, 50.
- Statistical tests assessing the improvements.
- Extensive hyperparameter optimization.

	ML-1M	LFM2B-1MON	AMAZONVID
# Users	6,034	3,555	6,950
# Males/Females	4,326/1,708	2,965/590	-
	(72%/28%)	(83%/17%)	-
# Items	3,125	77,985	14,494
# Interactions	574,376	877,365	132,209

Table 1: Statistics of the datasets after filtering.

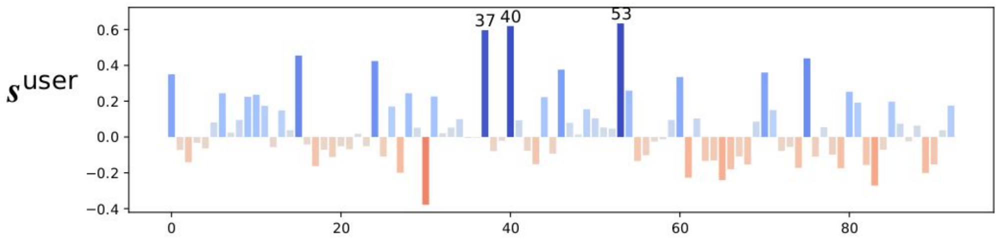
[8] Harper, F. Maxwell, and Joseph A. Konstan. "The Movielens Datasets: History and Context." TIIS 2015.
[9] Schedl, Markus, et al. "LFM-2b: A Dataset of Enriched Music Listening Events for Recommender Systems Research and Fairness Analysis." CHIIR 2022.
[10] McAuley, Julian, et al. "Image-based Recommendations on Styles and Substitutes." SIGIR 2015.
[6] Koren, Yehuda, et al.. "Matrix factorization techniques for recommender systems." Computer 2009.
[11] Liu, Nathan N., et al. "Wisdom of the better few: cold start recommendation via representative based rating elicitation." RecSys 2011.
[12] Barkan, Oren, et al. "Anchor-based collaborative filtering." CIKM 2021.

Evaluation Results (@10)

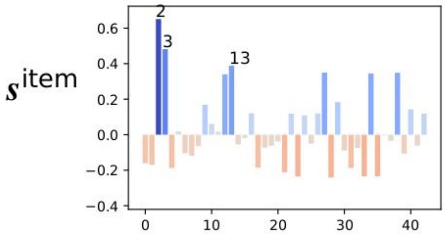
Model	ML-1M		AmazonVid		LFM2B-1Mon	
	NDCG	HitRatio	NDCG	HitRatio	NDCG	HitRatio
MF	.326	.571	.140	.255	.118	.215
RBMF	.282	.505	.093	.166	.279†	.384†
ACF	.335	.597†	.202†	.392†	.291†	.517†
U-ProtoMF	.333	.583	.152†	.276†	.179†	.322†
I-ProtoMF	.303	.544	.194†	.371†	.251†	.457†
UI-ProtoMF	**.383†‡**	**.657†‡**	**.220†‡**	**.401†**	**.347†‡**	**.579†‡**

Table 2: Evaluation results w.r.t. accuracy metrics at cutoff 10. The sign † indicates significant improvement over MF while ‡ indicates significant improvement over ACF.

Explaining UI-ProtoMF Recommendations



(a) U-ProtoMF



(b) I-ProtoMF

User Prototype 53	User Prototype 40	User Prototype 37	Item Prototype 3	Item Prototype 2	Item Prototype 13
Roman Holiday *Comedy|Romance*	**Runaway Bride** *Comedy|Romance*	**Cinderella** *Anim.|Children's|Musical*	**City of Angels** *Romance*	**Broadway Melody, The** *Musical*	**Chambermaid on the Titanic, The** *Romance*
To Catch a Thief *Com.|Romance|Thriller*	**She's All That** *Comedy|Romance*	**Little Mermaid, The** *Anim.|Child.|Com.|Musical*	**It Could Happen to You** *Drama|Romance*	**Slipper and the Rose, The** *Adventure|Musical|Romance*	**Dreaming of Joseph Lees** *Romance*
Sabrina *Comedy|Romance*	**Affair to Remember, An** *Romance*	**Sleeping Beauty** *Anim.|Children's|Musical*	**Walk in the Clouds, A** *Drama|Romance*	**Penny Serenade** *Drama|Romance*	**Passion of Mind** *Romance|Thriller*
Sleepless in Seattle *Comedy|Romance*	**Double Jeopardy** *Action|Thriller*	**She's All That** *Comedy|Romance*	**One Fine Day** *Drama|Romance*	**Perils of Pauline, The** *Comedy*	**Golden Bowl, The** *Drama*
While You Were Sleeping *Comedy|Romance*	**Ever After: A Cinderella Story** *Drama|Romance*	**101 Dalmatians** *Animation|Children's*	**Sommersby** *Drama|Mystery|Romance*	**Damsel in Distress, A** *Comedy|Musical|Romance*	**Up at the Villa** *Drama*

JOHANNES KEPLER UNIVERSITY LINZ

Thank you for your attention!

Code, dataset splits, and more are hosted on
https://github.com/hcai-mms/ProtoMF

ProtoMF: Prototype-based Matrix Factorization for Effective and Explainable Recommendations

Alessandro B. Melchiorre
Navid Rekabsaz
alessandro.melchiorre@jku.at
navid.rekabsaz@jku.at
Institute of Computational Perception,
Johannes Kepler University Linz and
Human-centered AI group, Linz
Institute of Technology
Linz, Austria

Christian Ganhör
christian.ganhoer@jku.at
Institute of Computational Perception,
Johannes Kepler University Linz
Linz, Austria

Markus Schedl
markus.schedl@jku.at
Institute of Computational Perception,
Johannes Kepler University Linz and
Human-centered AI group, Linz
Institute of Technology
Linz, Austria

MULTIMEDIA MINING AND SEARCH GROUP

Institute of Computational Perception