Pareto Front Approximation for Multi-Objective Session-Based Recommender Systems

Timo Wilm [1] Philipp Normann [1] Felix Stepprath [1]

[1]OTTO (GmbH & Co KG), Hamburg, Germany

Abstract

This work introduces **MultiTRON**, an approach that adapts Pareto front approximation techniques to multi-objective session-based recommender systems using a transformer neural network. Our approach optimizes trade-offs between key metrics such as click-through and conversion rates by training on sampled preference vectors. A significant advantage is that after training, a single model can access the entire Pareto front, allowing it to be tailored to meet the specific requirements of different stakeholders by adjusting an additional input vector that weights the objectives. We validate the model's performance through extensive offline and online evaluation. For broader application and research, the source code is made available. The results confirm the model's ability to manage multiple recommendation objectives effectively, offering a flexible tool for diverse business needs.

What is MultiTRON?

Building on the Transformer-based TRON model [8], MultiTRON leverages sampled preference vectors and a scalarization approach with a customized regularization term to approximate the entire Pareto front of session-based recommendation tasks.

1. Train just **one** model instead of multiple.
2. Get access to the **entire** Pareto front at inference time
3. **Maintain** competitive performance across multiple objectives

1. Session-Based Recommender Systems

Multi-objective session-based recommender systems predict the next item interaction based on prior user activities. Each user session consists of user-item interactions $s_{raw} = [i_1^{a_1}, i_2^{a_2}, \ldots, i_T^{a_T}]$, where T is the session length, and $i_t^{a_t}$ represents the action taken on item i at time t. Actions include clicking or ordering, typically with orders following clicks. Sessions are modelled as $s := [(c_1, o_1), (c_2, o_2), \ldots, (c_{T-1}, o_{T-1})]$, where c_t is the clicked item at time t, and o_t indicates if the item was ordered up to time T.

1.1 Recommender Model and Loss Functions

Our multi-objective recommender model $\mathcal{R}$ leverages past clicks to predict scores r_t^i for potential item interactions, optimizing the trade-off between click $\mathcal{L}_c(c_t, r_t^i)$ and order $\mathcal{L}_o(o_t, r_t^i)$ losses. Standard scalarization methods [3, 6] use a fixed preference vector $\pi := [\pi_c, \pi_o]$, with $\pi_c + \pi_o = 1$, and minimize:

$$\mathcal{L}(c_t, o_t, \mathcal{R}_t, \pi) = \pi_c \mathcal{L}_c(c_t, \mathcal{R}_t) + \pi_o \mathcal{L}_o(o_t, \mathcal{R}_t). \tag{1}$$

This approach does not scale for large datasets because each point on the Pareto front requires a separate model.

1.2 Pareto Front Approximation

In Pareto front approximation, sampling $\pi \sim Dir(\beta)$ from a Dirichlet distribution with parameter $\beta \in \mathbb{R}^2_{>0}$ during training and adding it to the input yields a model $\mathcal{M}(\cdot, \pi)$ conditioned on π during inference [6, 5, 1, 7]. We adapt this approach to sequential recommender models $\mathcal{R}(\cdot, \pi)$ and conclude from [1] that if $\mathcal{R}^*$ minimizes

$$\mathbb{E}_\pi \mathcal{L}(c_t, o_t, \mathcal{R}_t(\cdot, \pi), \pi) = \mathbb{E}_\pi \left(\sum_{k \in \{c,o\}} \pi_k \mathcal{L}_k(k_t, \mathcal{R}_t(\cdot, \pi)) \right), \tag{2}$$

then $\mathcal{R}^*$ minimizes Equation 1 almost surely w.r.t $\mathbb{P}_\pi$.

1.3 Regularization and Pareto Front Coverage

To address the limitation of narrow Pareto fronts [6], we also leverage the non-uniformity term from [4], defined as:

$$\mathcal{L}_{reg}(\pi) = \mathrm{KL}(g(\hat{\pi}) \mid \mathbf{1}/2), \tag{3}$$

where $\hat{\pi}_k := \frac{\pi_k \mathcal{L}_k}{\pi_c \mathcal{L}_c + \pi_o \mathcal{L}_o}$, $\mathbf{1}/2 = \left[\frac{1}{2}, \frac{1}{2}\right]$, and KL is the Kullback–Leibler divergence. The function g maps $\hat{\pi}$ to a vector of probabilities summing to 1. For instance, g could be chosen as the identity or the softmax function. By adding this regularization term in Equation 3 to the primary loss function in Equation 2, we avoid the need for solving a linear program after each forward pass, thereby maintaining efficient training speeds. This approach yields a Pareto front that approximately intersects with the inverse preference vector $\pi^{-1} = \left[\frac{1}{g(\pi_c)}, \frac{1}{g(\pi_o)}\right]$ at the point $[\mathcal{L}_c^*(\cdot, \pi), \mathcal{L}_o^*(\cdot, \pi)]$ [4].

1.4 Overall Loss Function

The overall loss is formulated as:

$$\mathbb{E}_\pi \mathcal{L}(\cdot, \pi, \lambda) = \mathbb{E}_\pi \left(\sum_{k \in \{c,o\}} \pi_k \mathcal{L}_k(k_t, \mathcal{R}_t(\cdot, \pi)) + \lambda \mathcal{L}_{reg}(\pi) \right), \tag{4}$$

with $\lambda \geq 0$ as a regularization parameter. Our model MultiTRON minimizes the loss in Equation 4 to approximate the Pareto front.

2.1 Experimental Study On Public Datasets

We only use click events for our experiments, maintaining a minimum item support of five and a session length of at least two for all datasets [2]. We use a temporal train/test split method, using the last day (Yoochoose) or the last week of data (OTTO) to form the test sets. The remaining data is used for training.

Table 1. Statistics of the datasets used in our experiments.

		Diginetica	Yoochoose	OTTO
Train	sessions	187k	7.9M	12.9M
	click events	906k	31.6M	194.7M
	order events	12k	1.12M	5.1M
Test	sessions	18k	15k	1.6M
	click events	87k	71k	12.3M
	order events	1.1k	1.2k	355k
	Items	43k	37k	1.8M

Table 2. Hypervolumes for $\beta = [\frac{1}{2}, \frac{1}{2}]$ and different values of λ for each dataset. The models are trained on Diginetica for 20 epochs and Yoochoose and OTTO for 10 epochs.

Datasets	$\lambda = 0.02$	$\lambda = 0.2$	$\lambda = 0.5$	$\lambda = 1$
Diginetica	**0.20609**	0.20012	0.20605	0.19296
Yoochoose	0.0806	0.0776	**0.0838**	0.0831
OTTO	1.524	1.537	1.544	**1.546**

Evaluation of MultiTRON

For the **online evaluation**, we utilized a model trained on OTTO's private data collected in May 2024. A live A/B test was conducted the following week, with four groups, each assigned different π values. The test results confirmed that the offline trade-off between $-\mathcal{L}_c$ and $-\mathcal{L}_o$ translates into real-world trade-offs between click-through rates (CTR) and conversion rates (CVR). Specifically, higher $-\mathcal{L}_o$ values correlated with increased CVR, while higher $-\mathcal{L}_c$ values correlated with increased CTR, as detailed in Figure 2.

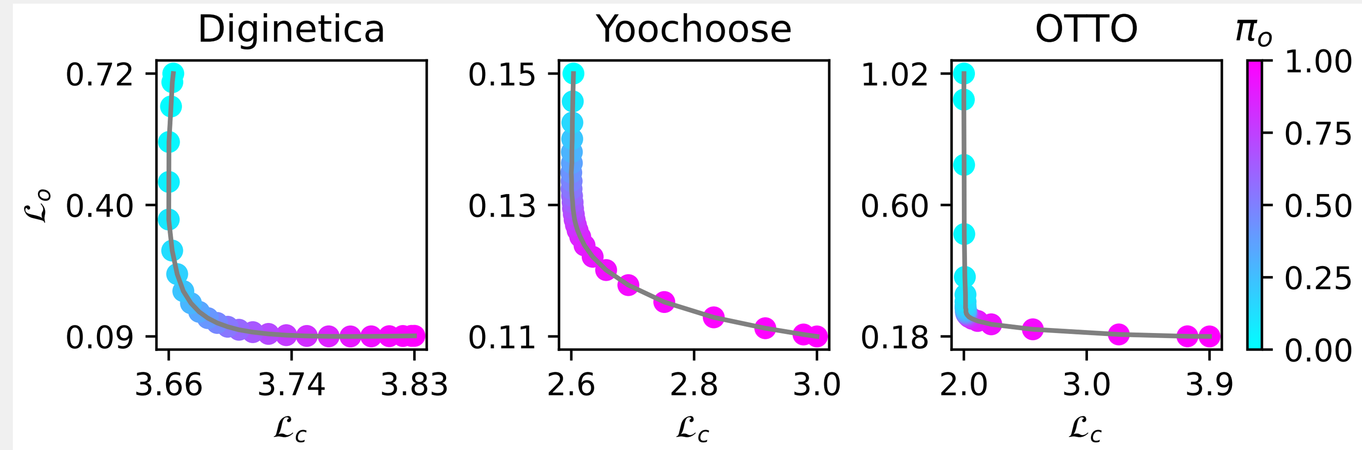

Figure 1. The best performing Pareto fronts from the offline evaluation on all three datasets showing the trade-off between $\mathcal{L}_c$ and $\mathcal{L}_o$ for 26 increasing values of π_o.

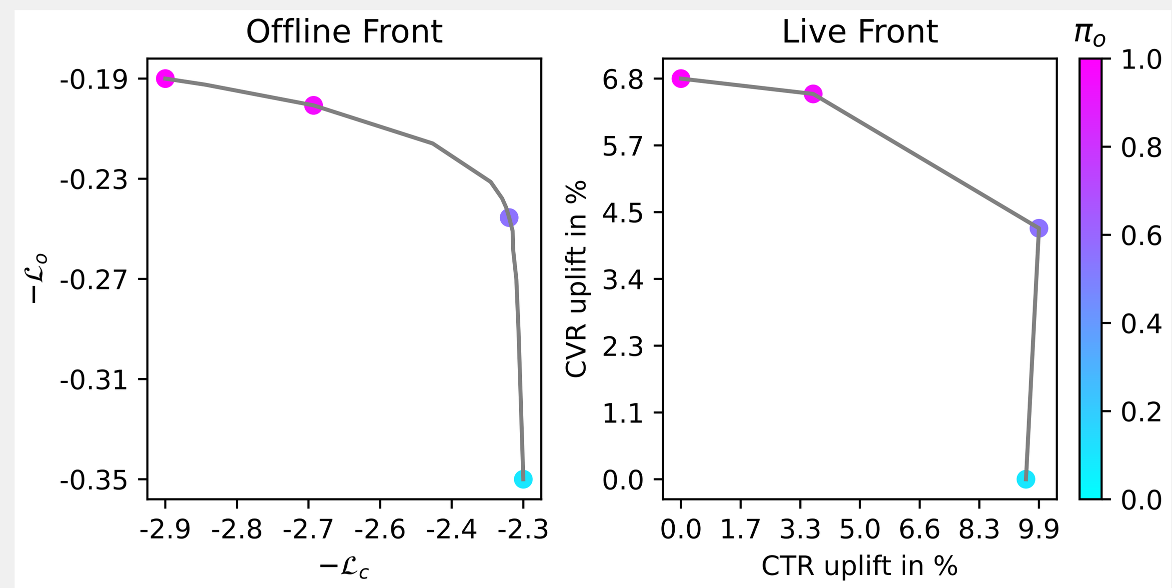

Figure 2. Evaluation results from the offline dataset (left) and live A/B test (right). It demonstrates that points on the predicted offline front translate into real-world trade-offs of CTR and CVR. The colored points correspond to differing values of π_o, representing the four different A/B test groups.

References

[1] Alexey Dosovitskiy and Josip Djolonga. You Only Train Once: Loss-Conditional Training of Deep Networks. In *International Conference on Learning Representations*, 2020.

[2] Balázs Hidasi and Alexandros Karatzoglou. Recurrent Neural Networks with Top-k Gains for Session-based Recommendations. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 843–852, October 2018.

[3] Xi Lin, Hui-Ling Zhen, Zhenhua Li, Qingfu Zhang, and Sam Kwong. Pareto Multi-Task Learning. In *Thirty-third Conference on Neural Information Processing Systems (NeurIPS)*, pages 12037–12047, 2019.

[4] Debabrata Mahapatra and Vaibhav Rajan. Multi-Task Learning with User Preferences: Gradient Descent with Controlled Ascent in Pareto Optimization. In Hal Daumé III and Aarti Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 6597–6607. PMLR, July 2020.

[5] Aviv Navon, Aviv Shamsian, Gal Chechik, and Ethan Fetaya. Learning the Pareto Front with Hypernetworks. In *International Conference on Learning Representations*, 2021.

[6] Michael Ruchte and Josif Grabocka. Scalable Pareto Front Approximation for Deep Multi-Objective Learning. In *2021 IEEE International Conference on Data Mining (ICDM)*, pages 1306–1311, Auckland, New Zealand, December 2021. IEEE.

[7] Tran Anh Tuan, Long P. Hoang, Dung D. Le, and Tran Ngoc Thang. A framework for controllable Pareto front learning with completed scalarization functions and its applications. *Neural Networks*, 169:257–273, January 2024.

[8] Timo Wilm, Philipp Normann, Sophie Baumeister, and Paul-Vincent Kobow. Scaling Session-Based Transformer Recommendations using Optimized Negative Sampling and Loss Functions. In *Proceedings of the 17th ACM Conference on Recommender Systems*, pages 1023–1026, Singapore Singapore, September 2023. ACM.

https://doi.org/10.1145/3640457.3688048

Eighteenth ACM Conference on Recommender Systems (RecSys '24), October 14–18, 2024, Bari, Italy

timo.wilm@otto.de, philipp.normann@otto.de