ADM - RecSys 2025

Karol Ochman-Milarski   Dawid Kubkowski

June 30, 2025

Challenge Overview

- Goal: Build **Universal Behavioral Profiles** (user representations).
- Input data: *user events* (e.g., purchases, cart events, page views, searches).
- Submissions: **user embeddings only**.
- Evaluation: Performed on multiple **downstream tasks**.

Key Points

- One set of embeddings $\Rightarrow$ used for all tasks.
- Models trained separately per task using submitted embeddings.
- Final score $=$ sum of performances across all tasks.
- Both **open** and **hidden tasks** used in evaluation.

Open Tasks

- **Churn Prediction:** Binary classification into 1: user will churn or 0: user will not churn. Churn task is performed on a subset of active users with at least one `product_buy` event in history.
  Task name: `churn`

- **Categories Propensity:** Multi-label classification into one of 100 possible labels. The labels represent the 100 most often purchase product categories.
 Task name: `propensity_category`

- **Product Propensity:** Multi-label classification into one of 100 possible labels. The labels represent the 100 most often purchase products in train target data.
  Task name: `propensity_sku`
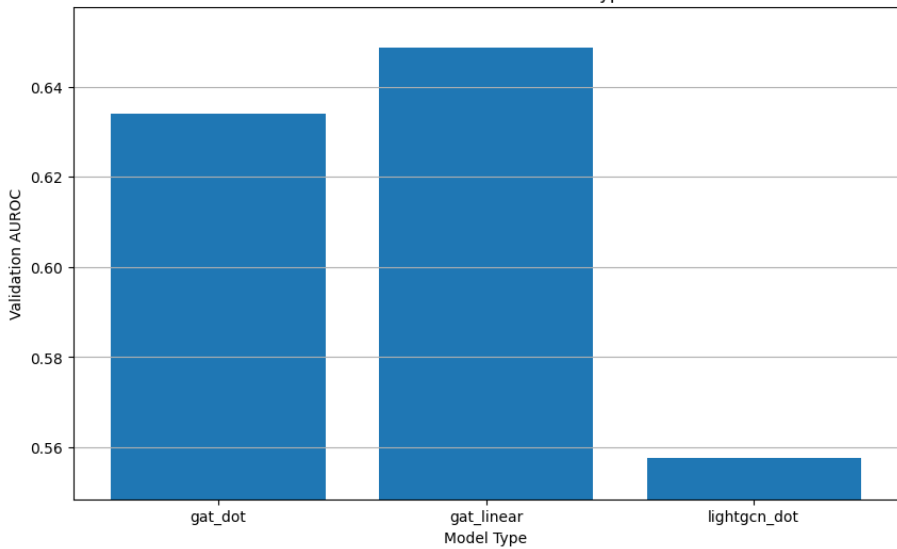
Evaluation Metric

$$\text{Score} = 0.8 \times \text{AUROC} + 0.1 \times \text{Novelty} + 0.1 \times \text{Diversity}$$

- The metric is computed on the validation set consisting the last 14 days of the dataset.
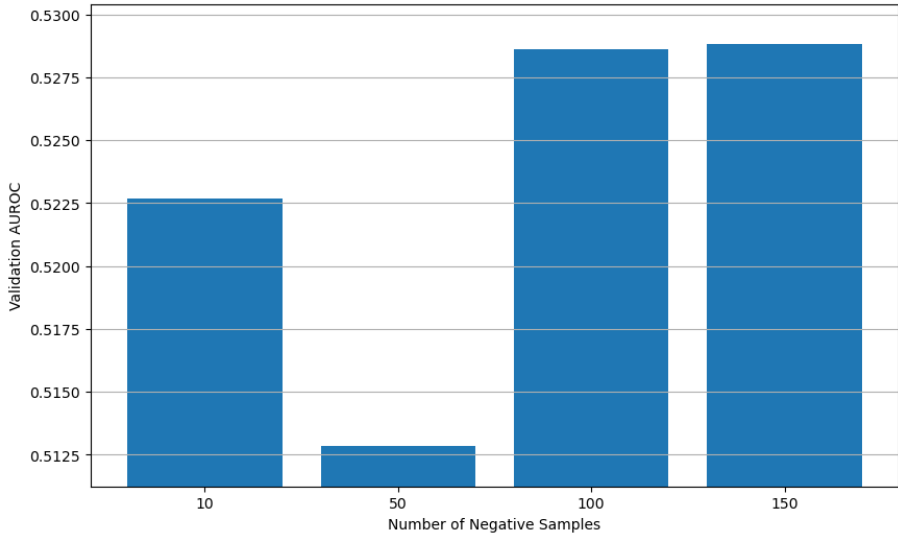
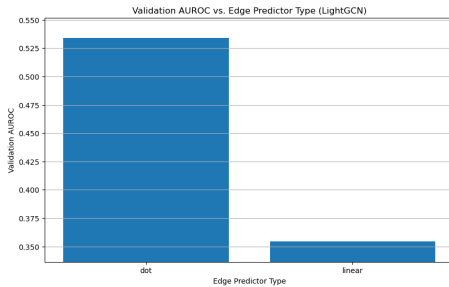FULL DATASET: Distribution of Interactions per User
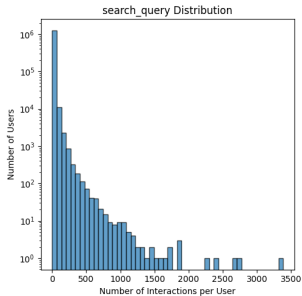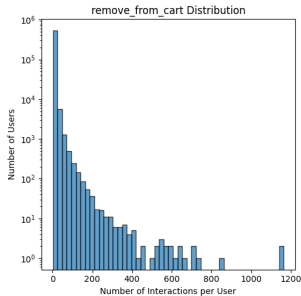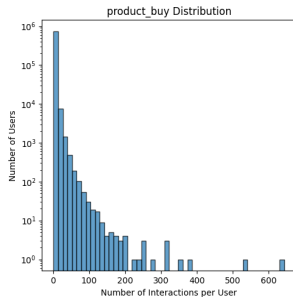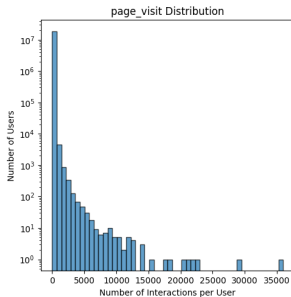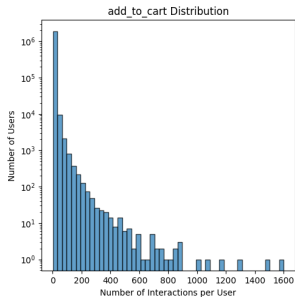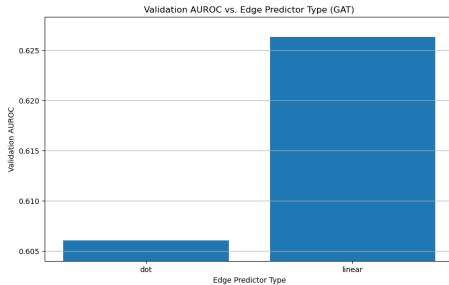
Validation AUROC vs. Model Type

Validation AUROC vs. Number of Negative Samples (LightGCN)

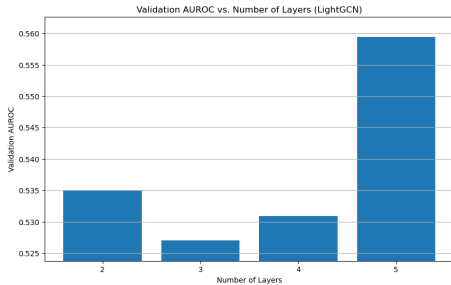LightGCN – Edge Type

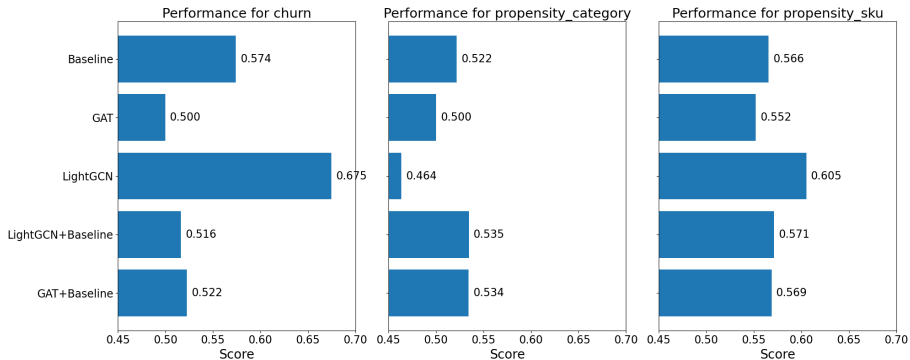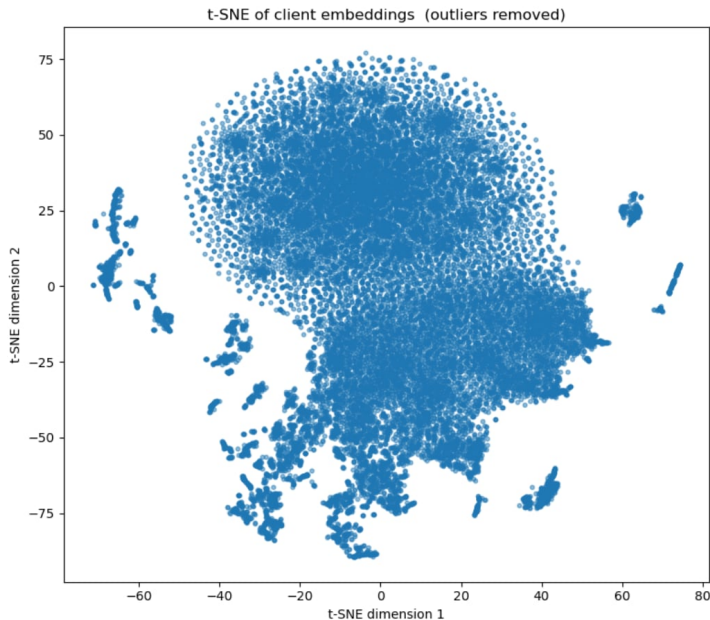GAT – Edge Type

Validation AUROC vs. Number of Layers (LightGCN)

LightGCN – Number of Layers

Validation AUROC vs. Number of Layers (GAT)

GAT – Number of Layers

Performance for churn

Model	Score
Baseline	0.574
GAT	0.500
LightGCN	0.675
LightGCN+Baseline	0.516
GAT+Baseline	0.522

Performance for propensity_category

Model	Score
Baseline	0.522
GAT	0.500
LightGCN	0.464
LightGCN+Baseline	0.535
GAT+Baseline	0.534

Performance for propensity_sku

Model	Score
Baseline	0.566
GAT	0.552
LightGCN	0.605
LightGCN+Baseline	0.571
GAT+Baseline	0.569

LightGCN embeddings

t-SNE of client embeddings (outliers removed)

GAT embeddings

t-SNE of my_user_emb (outliers removed)

Baseline embeddings

t-SNE of client embeddings (outliers removed)