Relational Deep Learning in Practice

From Enterprise Databases to Predictive Models

Mahannop Thabua[2] · Monthol Charattrakool[3] · Tutanon Sinthuprasith[2]
Peem Puangmaha[1] · Vijay Prakash Dwivedi[1] · Saksorn Ruangtanusak[3]

Stanford Data-Centric AI Workshop
Foundations, Agents, and Scientific Discovery

Research Question:* *What makes relational deep learning work on enterprise data?

Enterprise data → predictive tasks

Impressions
Impression_ID
User_ID
Ad_ID
Timestamp

Users
User_ID
Age_group
Region
User_segment

Ads
Ad_ID
Ad_category
Ad_format
Campaign_ID

→

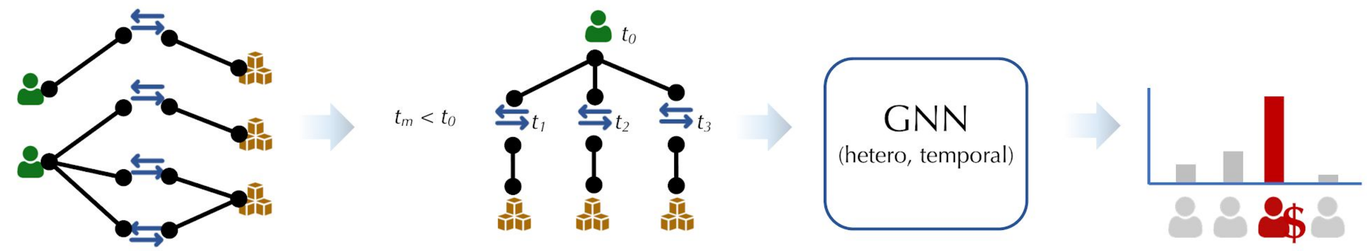
Training table

Will this impression receive a click?

Impression_ID
Seed time
Clicked (0/1)

Advertising schema · Impression-level click prediction.

From database to prediction



Sample history at t → encode columns → learn representations → predict

Classification · Regression · Link prediction; foundation models extend reuse via pretraining.

Three model routes

Engineered features → GBDT / TFM
RDBLearn features → TFM
Relational graph + encoders → GNN / RelGT

From task to deployment

Past 90 days → Seed time t → Next 90 days

Active customer — No purchase = churn

Audit leakage · Tune tabular baselines · Match serving latency

Impression-level winners vary

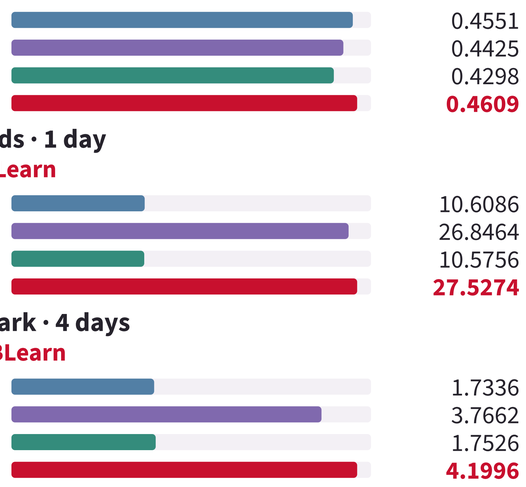
Logged-impression click prediction · AUPRC (%)

Model	SCBx	Alimama	rel-Avito
Random	0.3549	5.0390	0.6725
XGBoost	0.6596	6.3718	1.6331
LightGBM	**0.9747**	6.3749	1.6517
CatBoost 48K	0.7206	5.9705	1.2121
CatBoost 1M	0.8628	7.0755	1.5484
GraphSAGE	0.5364	6.5352	1.5813
RelGNN	0.2915	5.0390	0.6726
RelGT	0.5306	6.3939	**1.9197**
TabICL v2 48K	0.5397	5.3313	0.9971
TabPFN-3 1M	0.9014	5.8718	1.6100
TabICL v2 48K + RDBLearn	0.6634	6.7510	0.8282
TabPFN-3 1M + RDBLearn	0.6094	**7.2506**	1.7699

Teal cells mark each dataset's highest reported AUPRC.

Future user AUPRC by family

Augmented TFMs lead all 3 · AUPRC (%) · Separate scales



RDBLearn and user AUPRC

TabPFN-3 improves on all three datasets; TabICL v2 declines on SCBx.

TabPFN-3 1M

User AUPRC (%)	Base	+ RDBLearn
SCBx	0.3930	**0.4609**
Alimama	10.5756	**27.5274**
rel-Avito	1.7526	**4.0290**

TabICL v2 48K

User AUPRC (%)	Base	+ RDBLearn
SCBx	0.4298	**0.2244**
Alimama	10.4640	**26.3961**
rel-Avito	1.7351	**4.1996**

Teal: improvement · Red: decline
User AUPRC (%), higher is better.

User AUROC leaders: RelGT (SCBx, Alimama); TabPFN-3 + RDBLearn (rel-Avito). Evaluated tasks: clicks and user engagement; recommendation links are broader context.