NEURAL INFORMATION PROCESSING SYSTEMS

KAIST AI
Kim Jaechul Graduate School

kumo

RDB2G-Bench: A Comprehensive Benchmark for Automatic Graph Modeling of Relational Databases





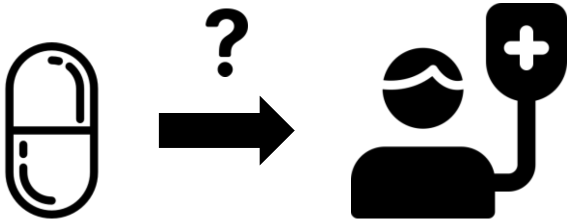



Dongwon Choi | Sunwoo Kim | Juyeon Kim | Kyungho Kim | Geon Lee | Shinhwan Kang | Myunghwan Kim | Kijung Shin

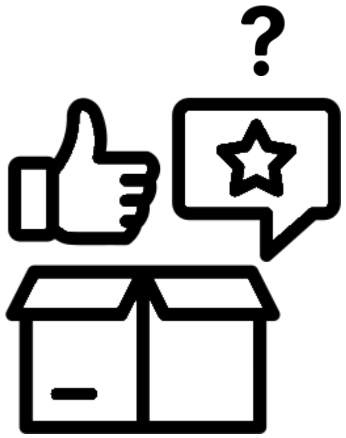
About RDB-to-Graph modeling

- **Relational Databases (RDBs)** have various industrial applications.

Finance



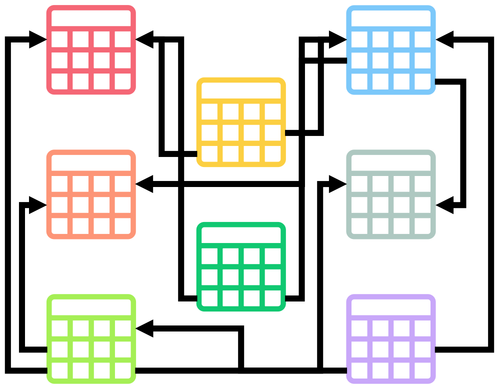
Healthcare



E-commerce

About RDB-to-Graph modeling (cont.)

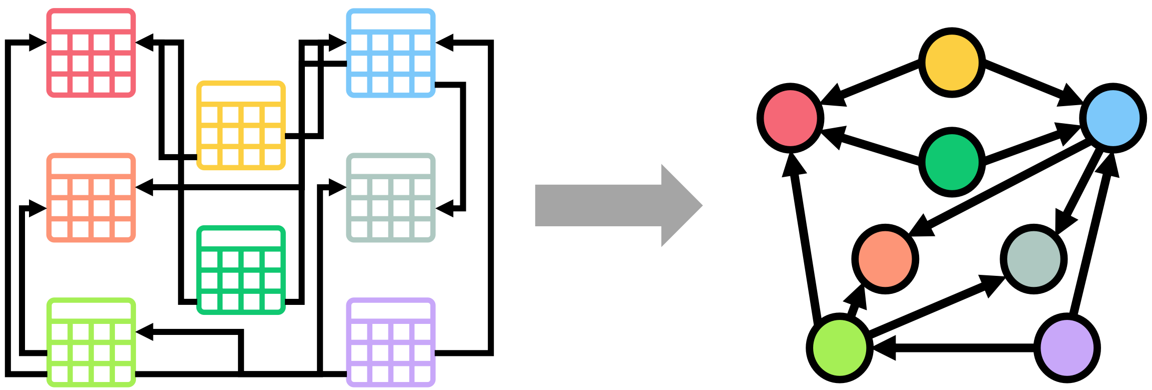
- Recent advances leverage **graph-based ML** for RDB prediction tasks.



RDB schema

About RDB-to-Graph modeling (cont.)

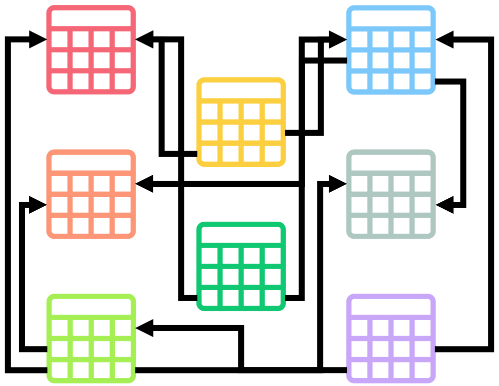
- Recent advances leverage **graph-based ML** for RDB prediction tasks.
  - Table rows as **nodes** and foreign-key (FK) relations as **edges**.



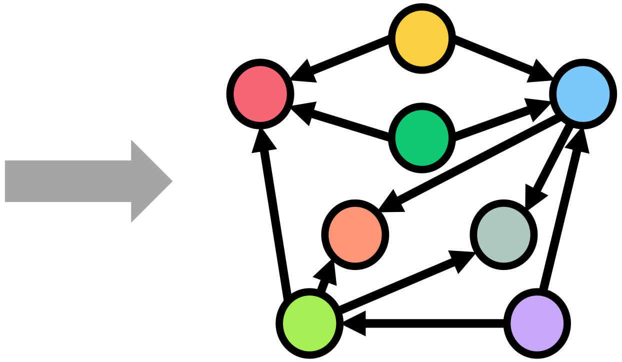
RDB schema **RDB graph**

About RDB-to-Graph modeling (cont.)

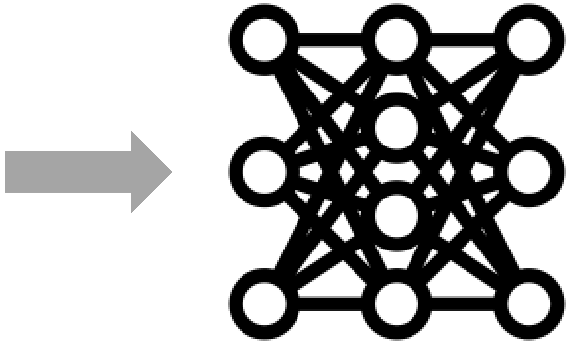
- Recent advances leverage **graph-based ML** for RDB prediction tasks.
  - By using **Graph Neural Networks (GNNs),** we can get a final prediction.



RDB schema



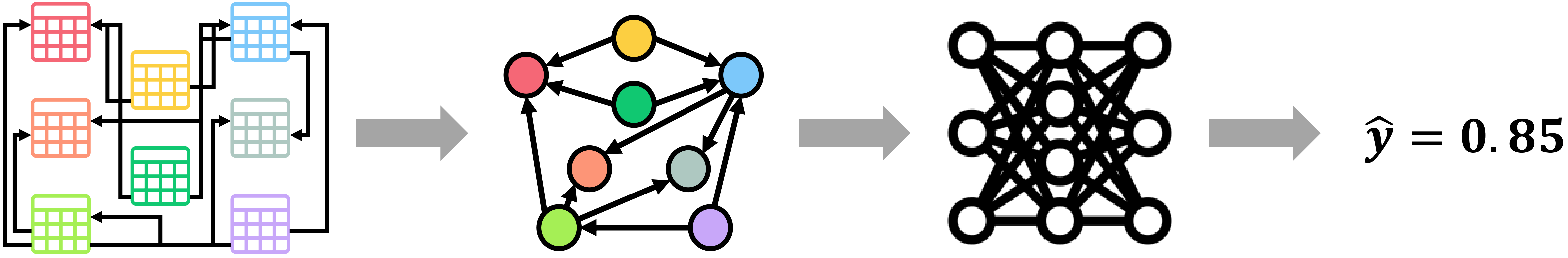
RDB graph



Predictive GNN

About RDB-to-Graph modeling (cont.)

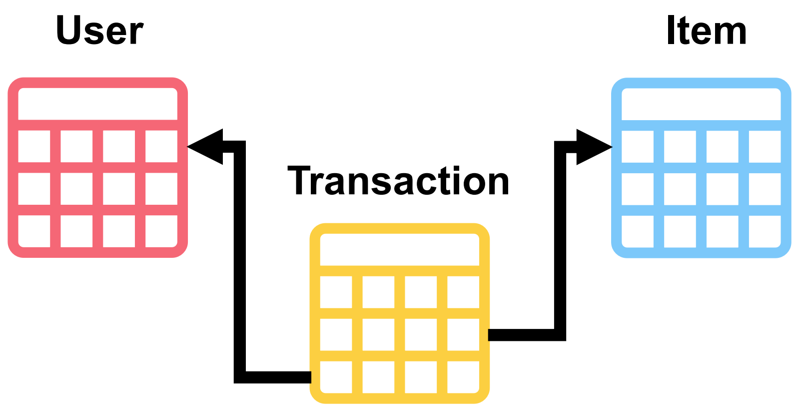
- Recent advances leverage **graph-based ML** for RDB prediction tasks.
  - By using **Graph Neural Networks (GNNs),** we can get a final prediction.



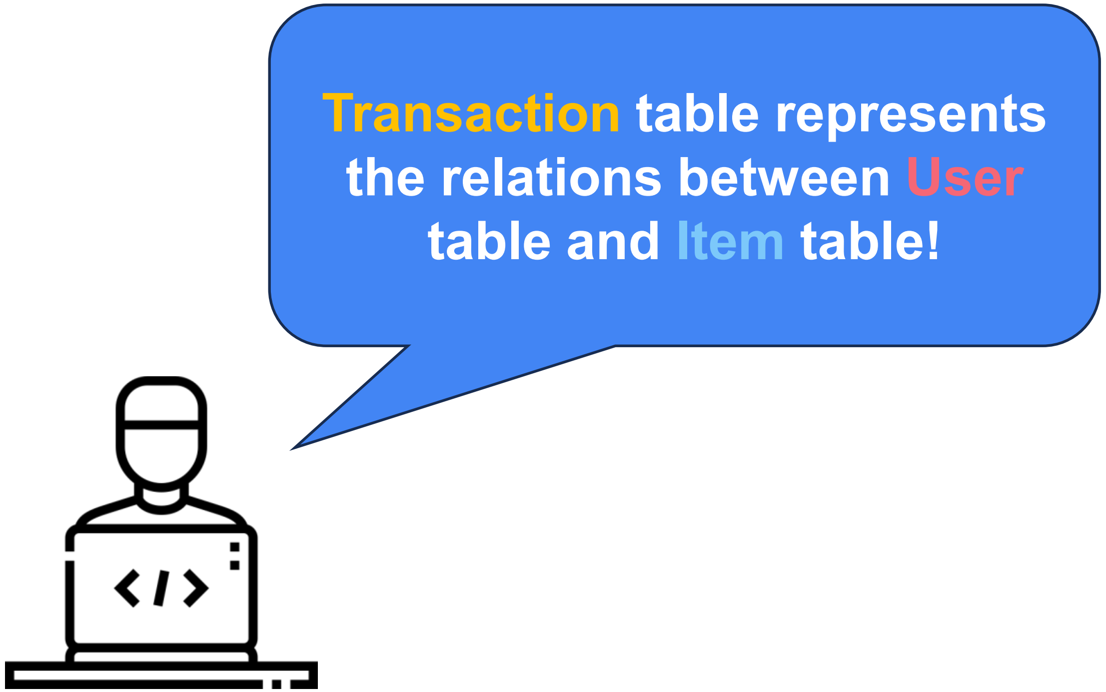
RDB schema **RDB graph** **Predictive GNN** **Prediction**

About RDB-to-Graph modeling (cont.)

- Option 1. Some table rows can be modeled as **edges** between tables.

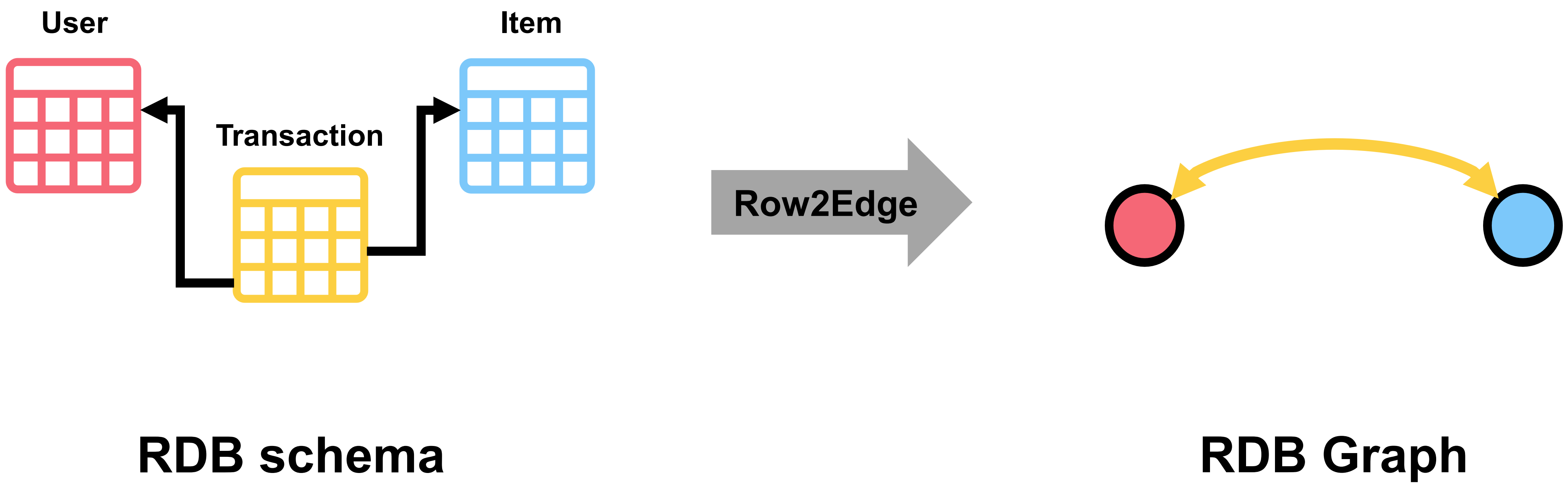


RDB schema



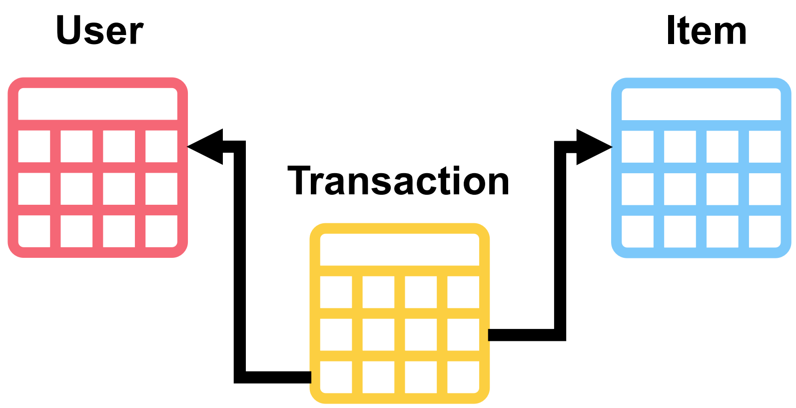
About RDB-to-Graph modeling (cont.)

- Option 1. Some table rows can be modeled as **edges** between tables.

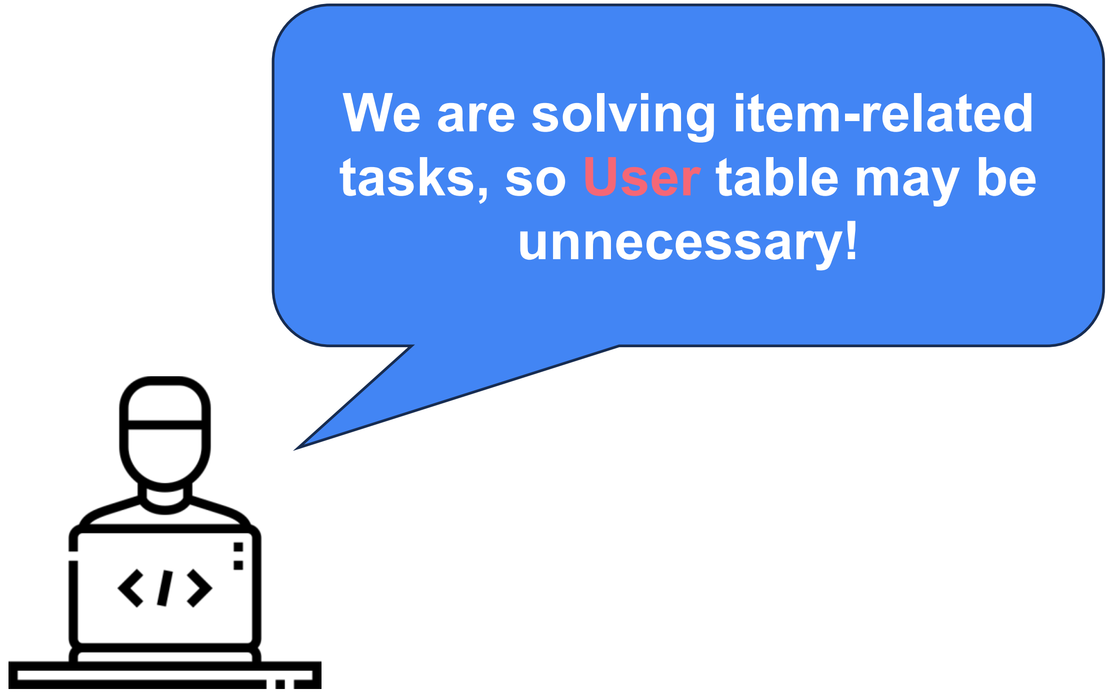


About RDB-to-Graph modeling (cont.)

- Option 2. We can choose only a **subset** of tables and FK relations.

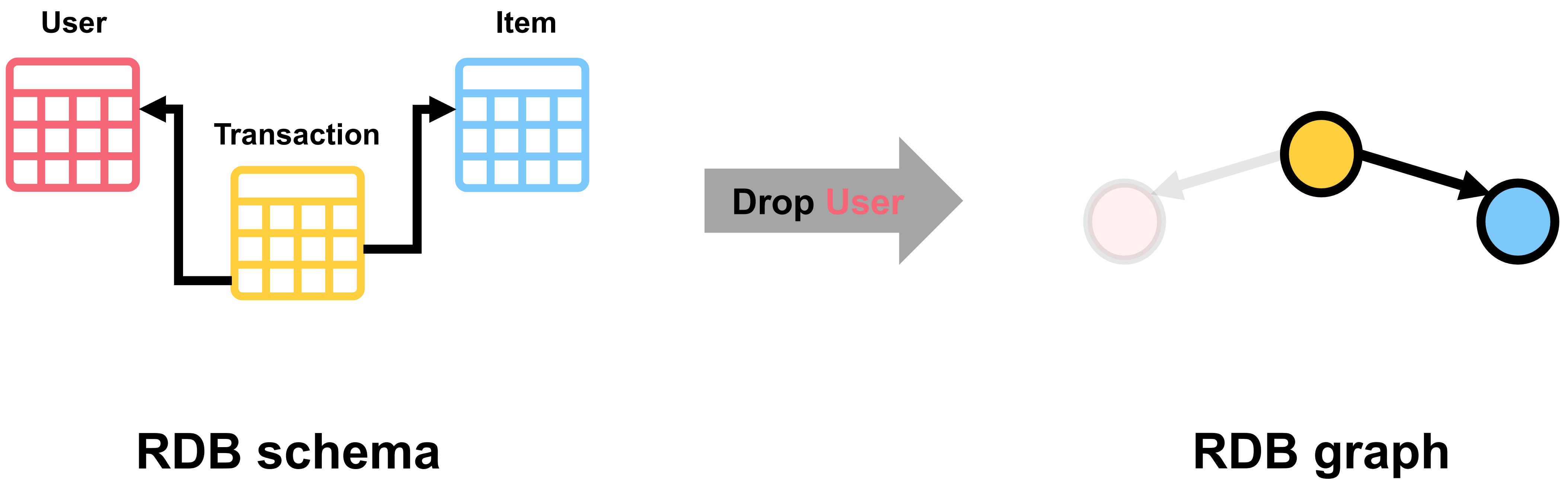


RDB schema



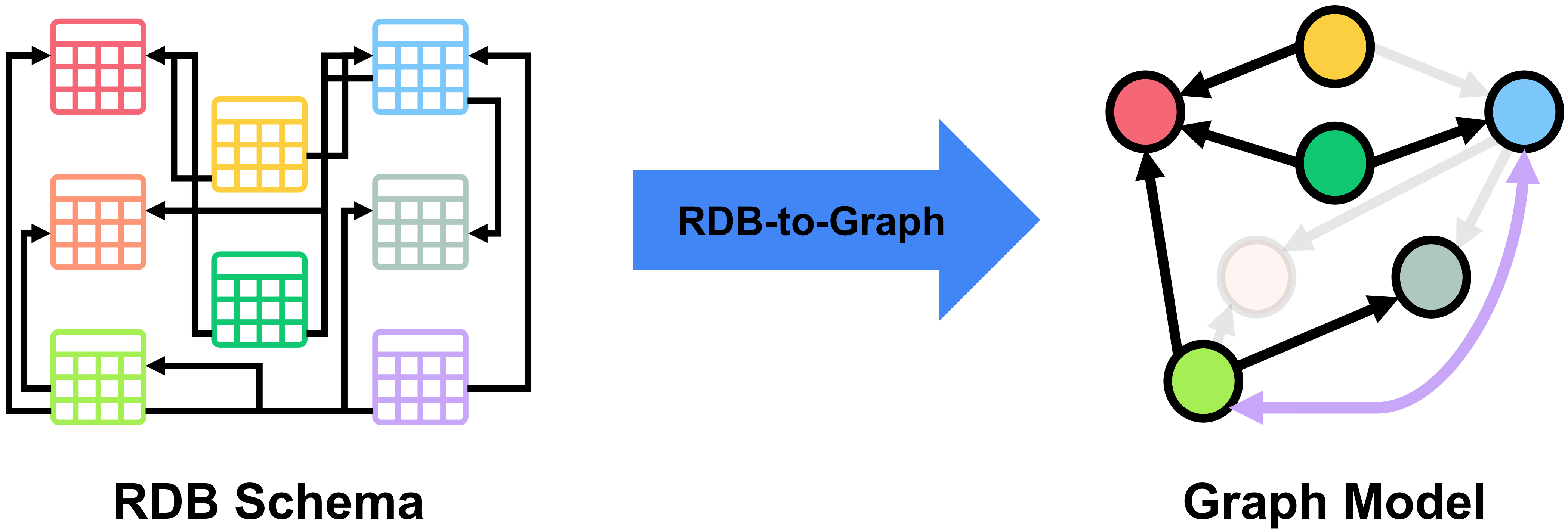
About RDB-to-Graph modeling (cont.)

- Option 2. We can choose only a **subset** of tables and FK relations.



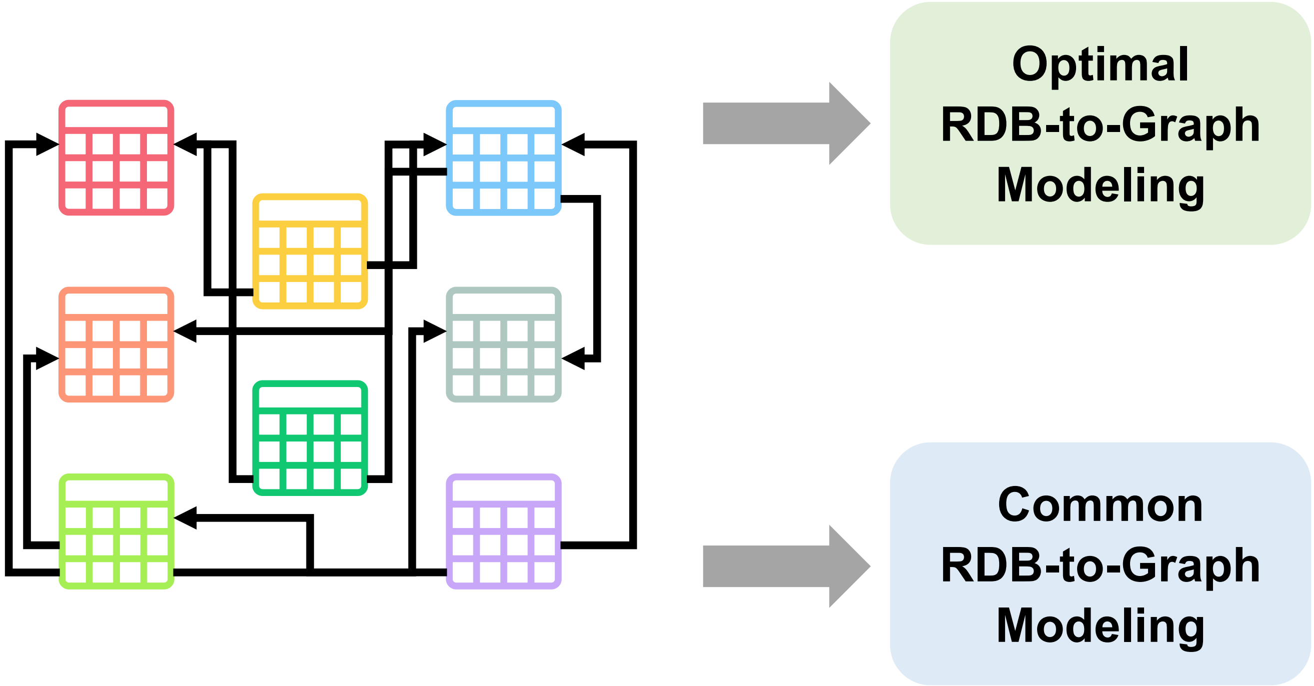
About RDB-to-Graph modeling (cont.)

- **RDB-to-Graph** modeling aims to find a **graph model** (i.e., graph representation of RDB) to maximize the performance of the predictive GNN on the given task.



About RDB-to-Graph modeling (cont.)

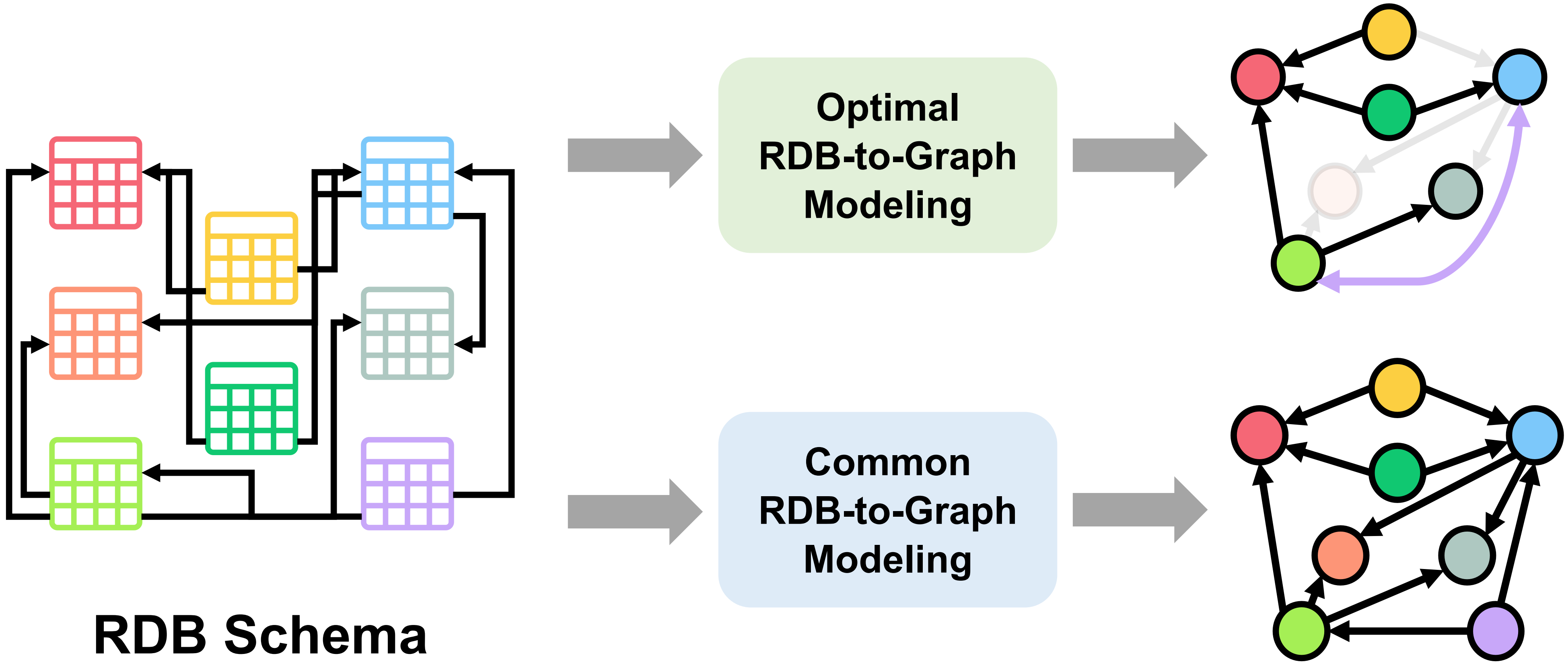
- Careful **RDB-to-Graph** modeling is crucial in practice.



RDB Schema

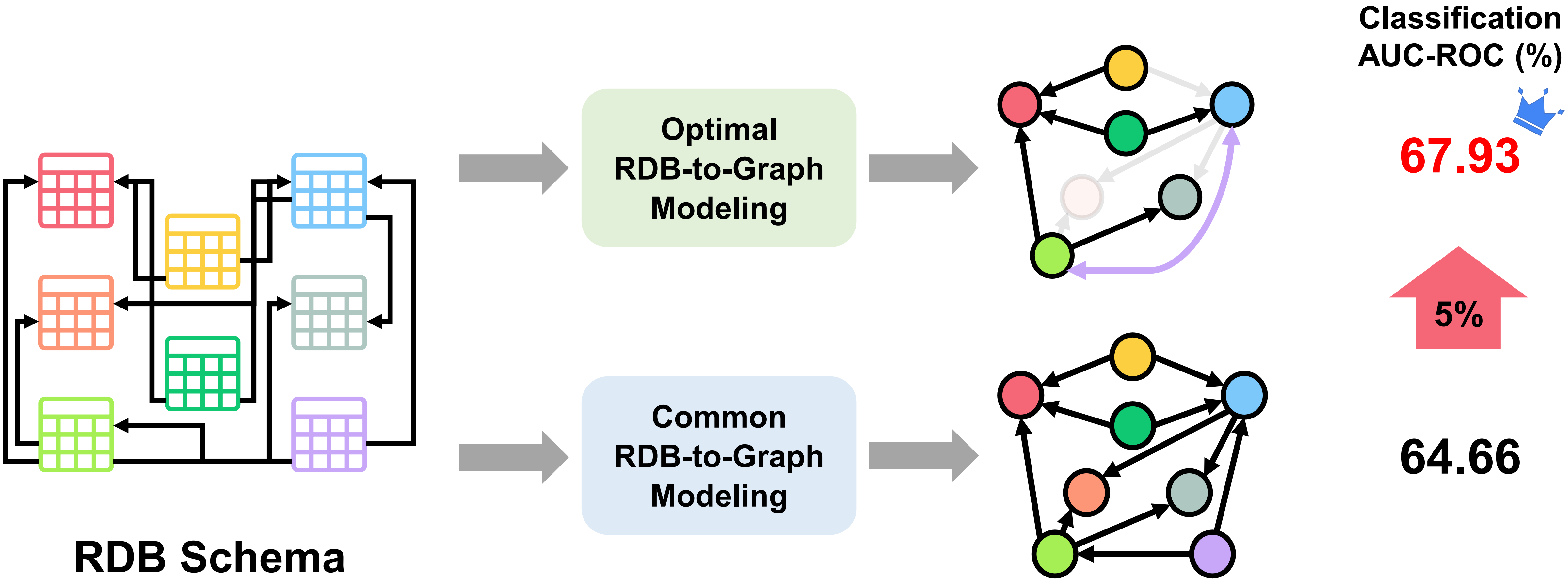
About RDB-to-Graph modeling (cont.)

- Careful **RDB-to-Graph** modeling is crucial in practice.



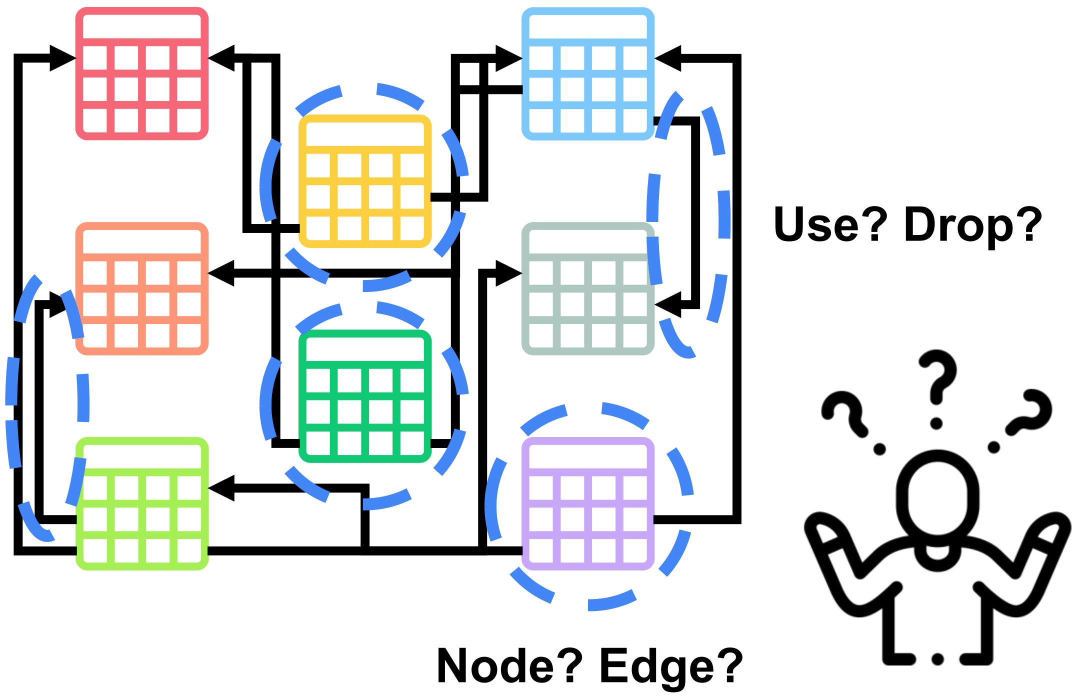
About RDB-to-Graph modeling (cont.)

- Careful **RDB-to-Graph** modeling is crucial in practice.



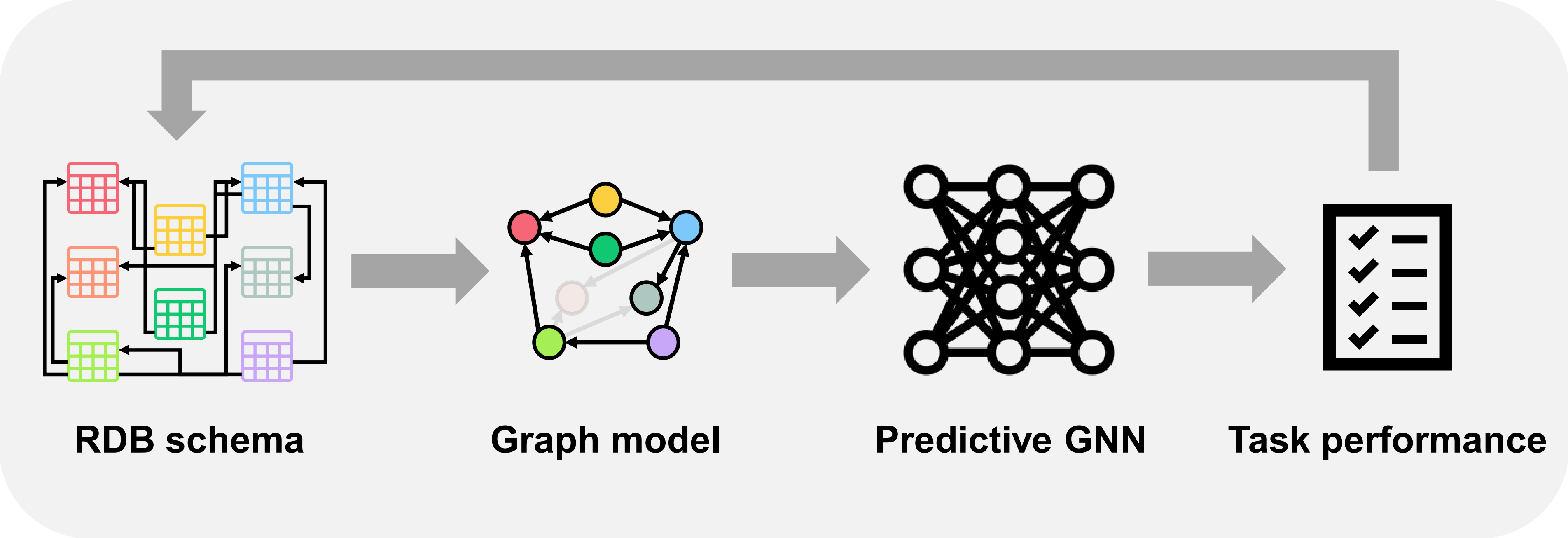
Challenges in RDB-to-Graph Modeling

- However, there are numerous possible graph models.

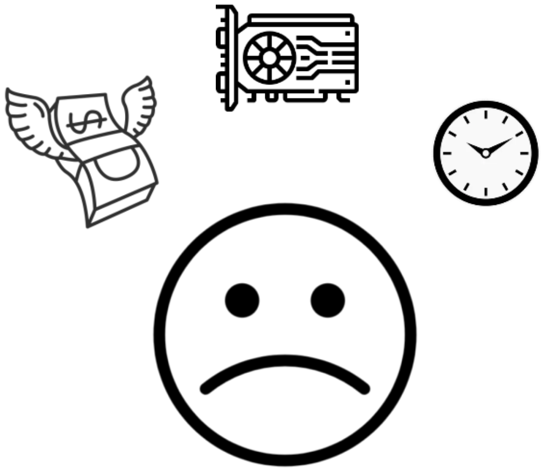


Challenges in RDB-to-Graph Modeling (cont.)

- Evaluating RDB-to-graph modeling is computationally costly.



RDB-to-Graph evaluation



Challenges in RDB-to-Graph Modeling (cont.)

- E… graph modeling is computationally costly.

RDB-to-Graph evaluation

In our paper, we introduce RDB2G-Bench, the first RDB-to-graph modeling benchmark.

Roadmap

- Overview
- **Dataset for RDB2G-Bench**
- Benchmark for RDB2G-Bench
- Conclusions

Dataset designs of RDB2G-Bench

- Our **RDB2G-Bench** datasets were created based on 5 real-world RDBs and 12 predictive tasks provided in **RelBench.**

RELBENCH
RELATIONAL DEEP LEARNING BENCHMARK

Sports

Formula 1

E-commerce

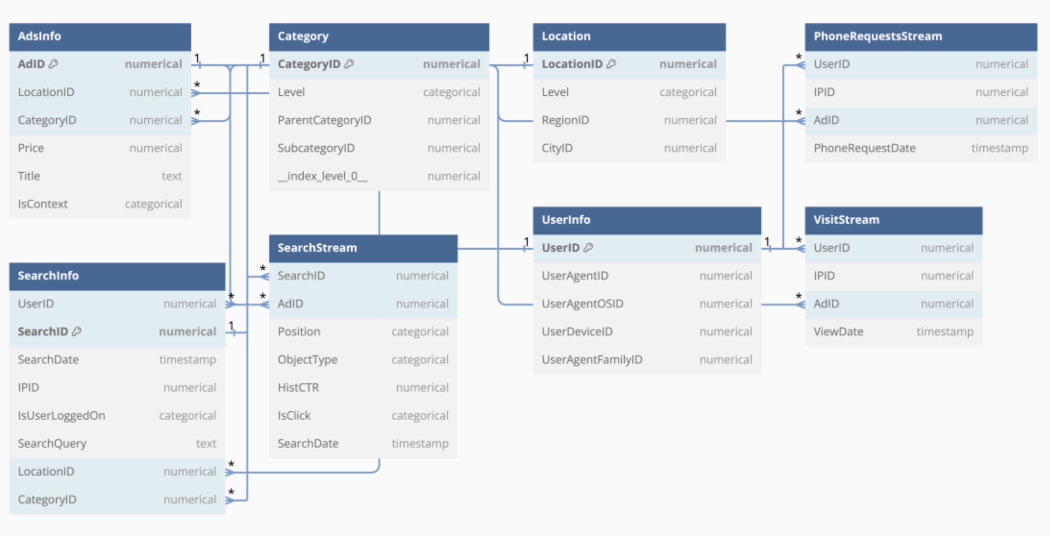
Avito

Social

stackoverflow

Healthcare

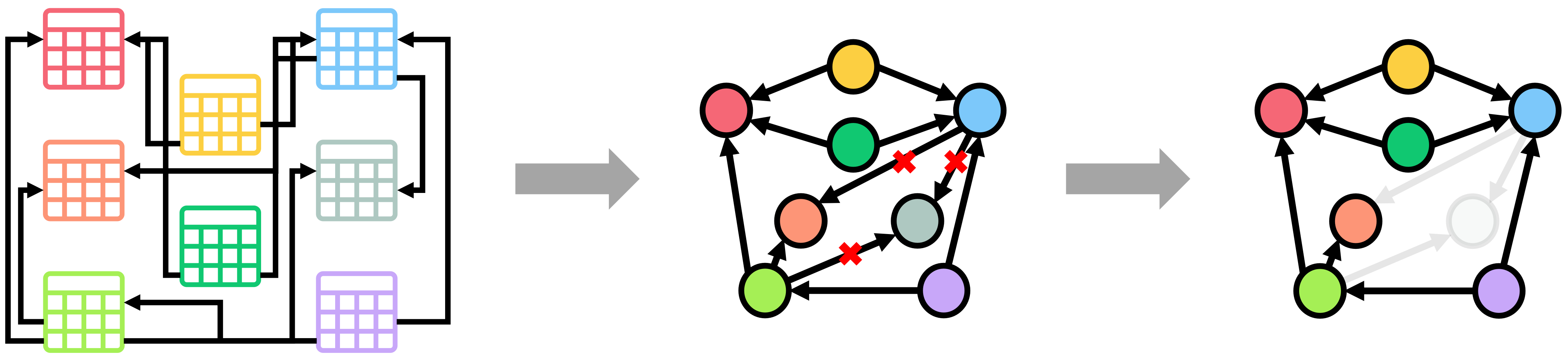
NIH National Library of Medicine
National Center for Biotechnology Information



Example schema (rel-avito)

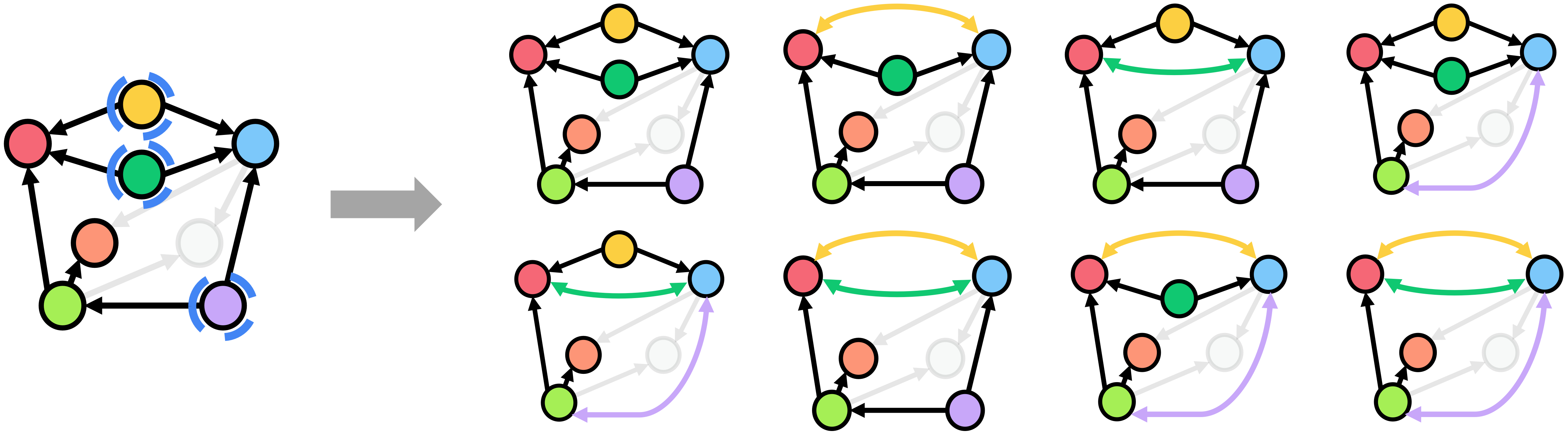
Dataset designs of RDB2G-Bench (cont.)

- Step 1. Selecting which tables and FK relationships to include in the graph models.



Dataset designs of RDB2G-Bench (cont.)

- Step 2. Selecting how to represent the rows of each table, as either nodes or edges, in the graph.



Dataset designs of RDB2G-Bench (cont.)

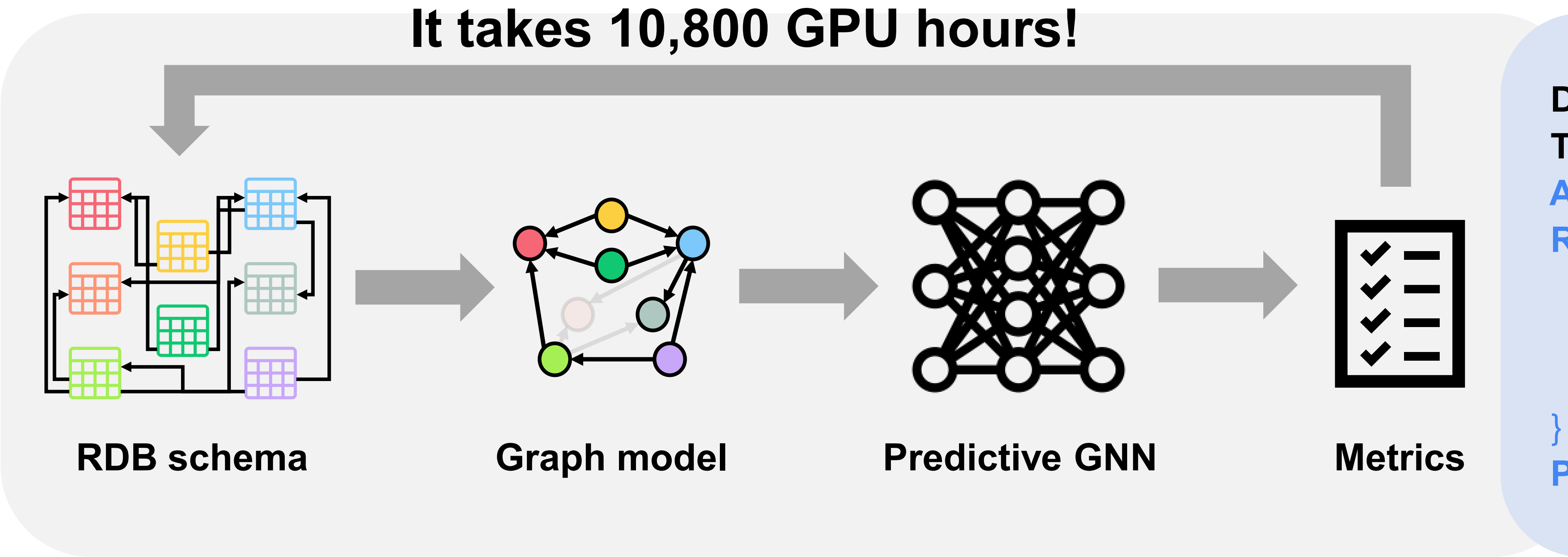
- By doing so, we constructed about 50k distinct graph models.

RDB	Task Name	Type	# Tables	# Graph Models	Performance Statistics Best	AR2N [29]	Worst
rel-avito	user-clicks (UC)	classification	8	944	67.93	64.66	60.89
	user-visits (UV)	classification		944	66.33	65.97	59.83
	ad-ctr (AC)	regression		1304	0.039	0.040	0.044
	user-ad-visit (UAV)	recommendation		909	3.682	3.661	0.159
rel-event	user-repeat (UR)	classification	5	214	82.29	77.65	63.96
	user-ignore (UI)	classification		214	82.82	82.22	74.29
	user-attendance (UA)	regression		214	0.237	0.244	0.266
rel-f1	driver-dnf (DD)	classification	9	722	74.56	73.14	67.40
	driver-top3 (DT)	classification		722	81.88	78.11	75.37
	driver-position (DP)	regression		722	3.831	3.913	4.171
rel-stack	post-post-related (PPL)	recommendation	7	7979	12.04	10.82	0.006
rel-trial	study-outcome (SO)	classification	15	36863	70.91	68.09	62.85

≈**50k!**

Dataset designs of RDB2G-Bench (cont.)

- For each constructed graph model, we collected **predictive performance, runtime,** and **parameter size**



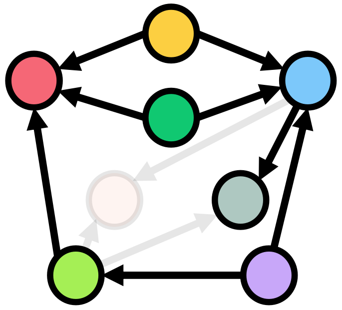
Dataset: rel-avito
Task: user-clicks
AUC-ROC (%): 65.73
Runtime: {
  Train: 2.23s
  Valid: 0.65s
  Test: 0.61s
}
Parameters: 1.78M

RDB-to-Graph evaluation

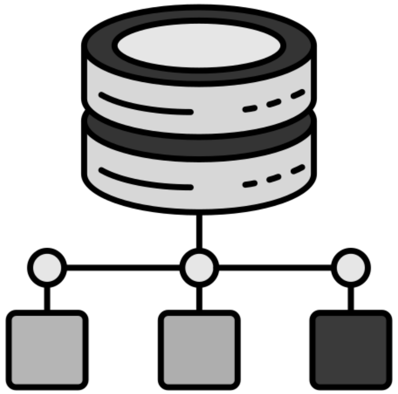
Performance metrics

Dataset designs of RDB2G-Bench (cont.)

- We can get the given graph model's performance immediately by looking up **RDB2G-Bench** datasets, without costly evaluations.



Graph model



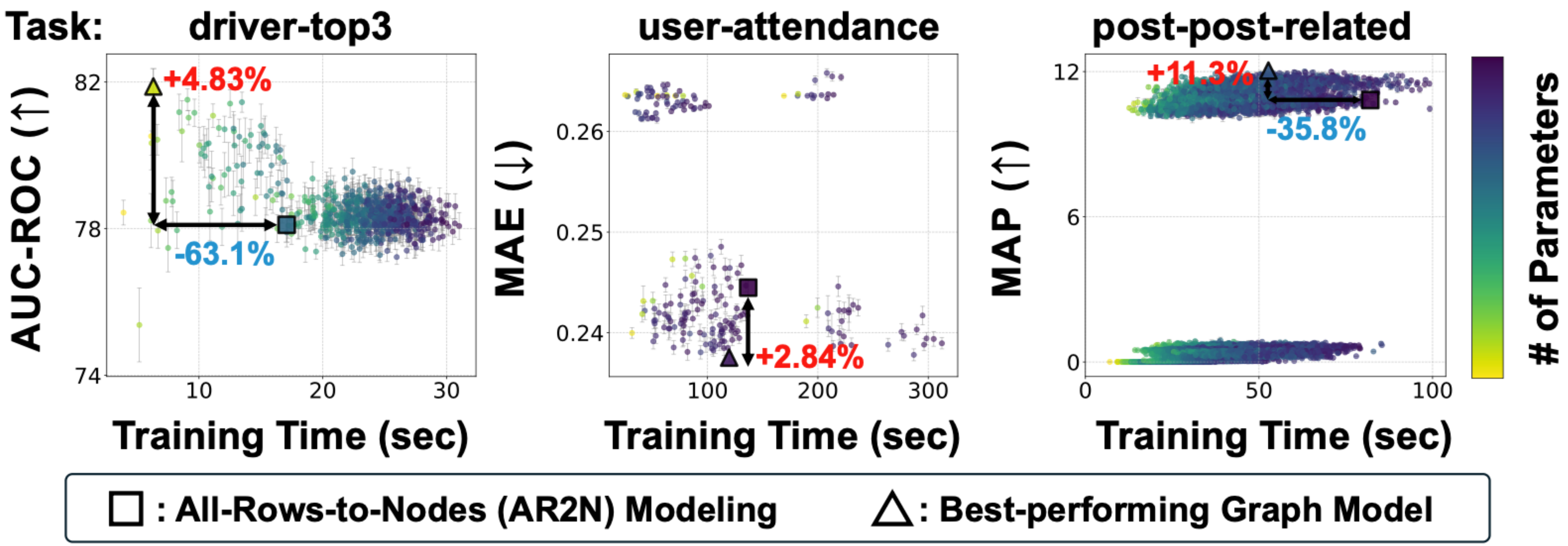
RDB2G-Bench

Dataset: rel-avito
Task: user-clicks
AUC-ROC (%): 65.73
Runtime: {
  Train: 2.23s
  Valid: 0.65s
  Test: 0.61s
}
Parameters: 1.78M

Performance metrics

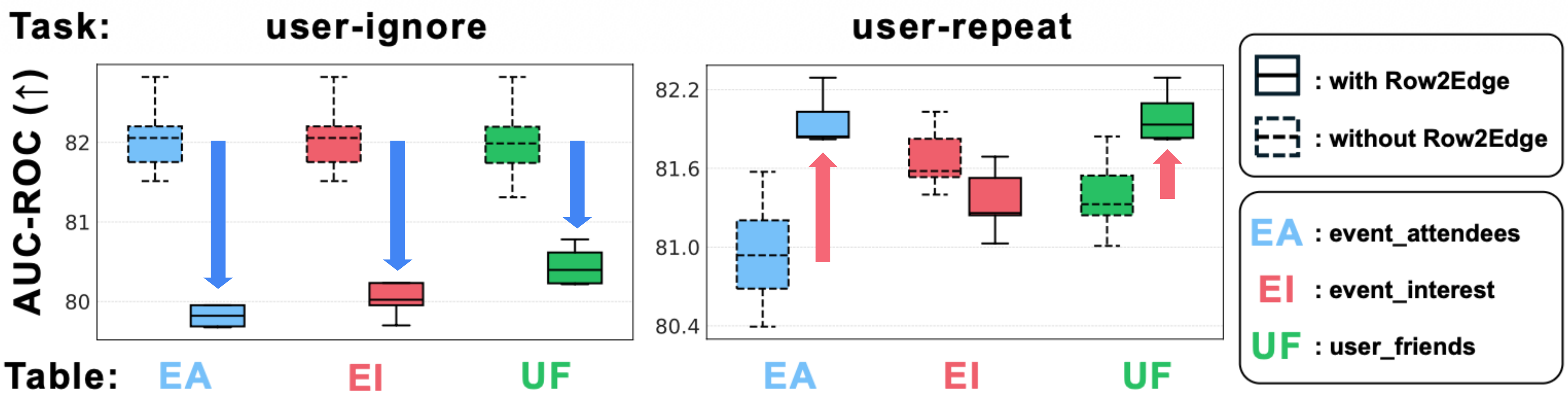
Observations from datasets

- Observation 1. Finding the best graph model is **worthwhile.**



Observations from datasets (cont.)

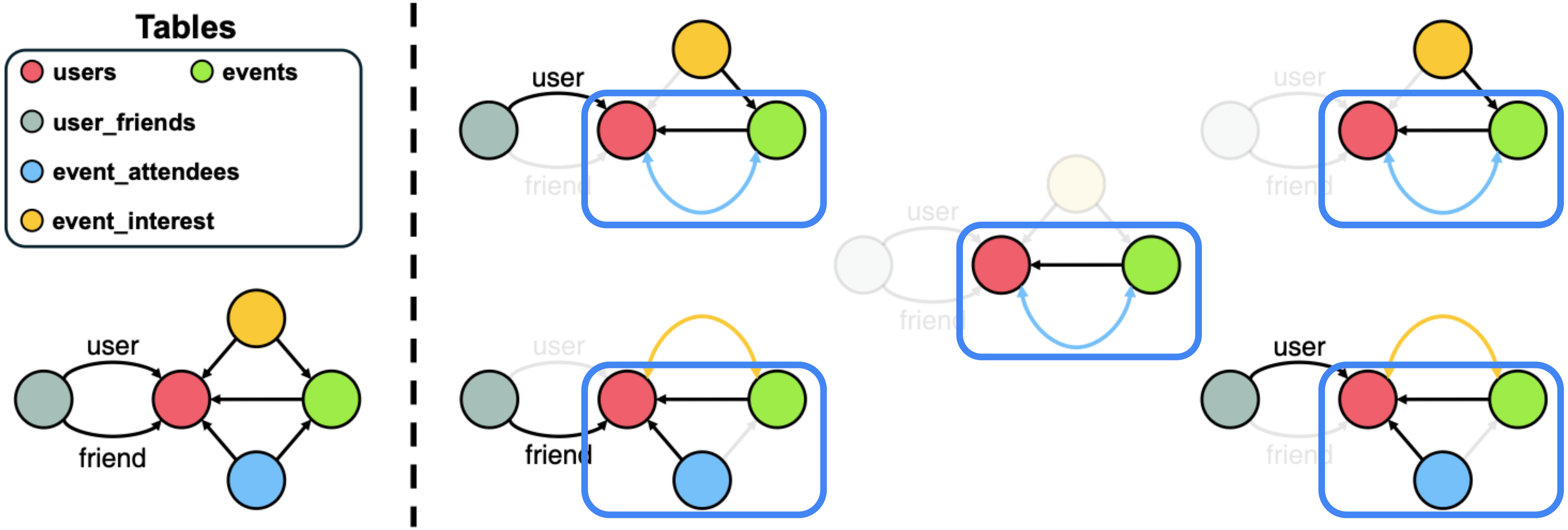
- Observation 2. Modeling **rows as edges (Row2Edge)** can be crucial.



It depends on the task, even on the same RDB!

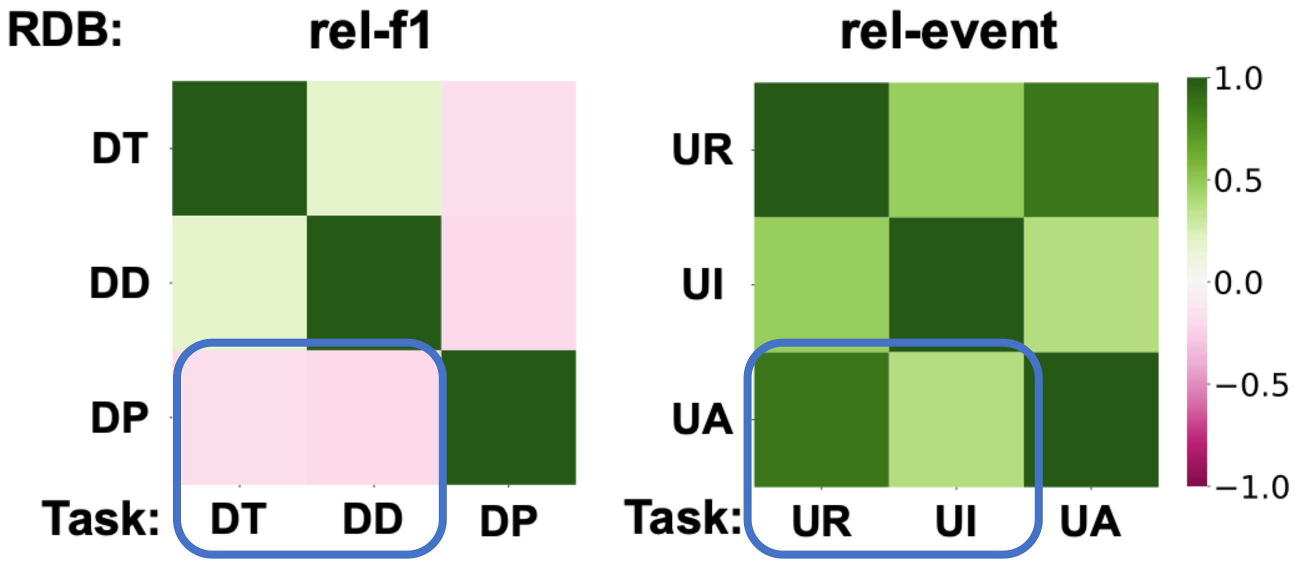
Observations from datasets (cont.)

- Observation 3. Top-performing graph models share **common substructures.**



Observations from datasets (cont.)

- Observation 4. Different tasks may require **different graph models**, even on the same RDB.

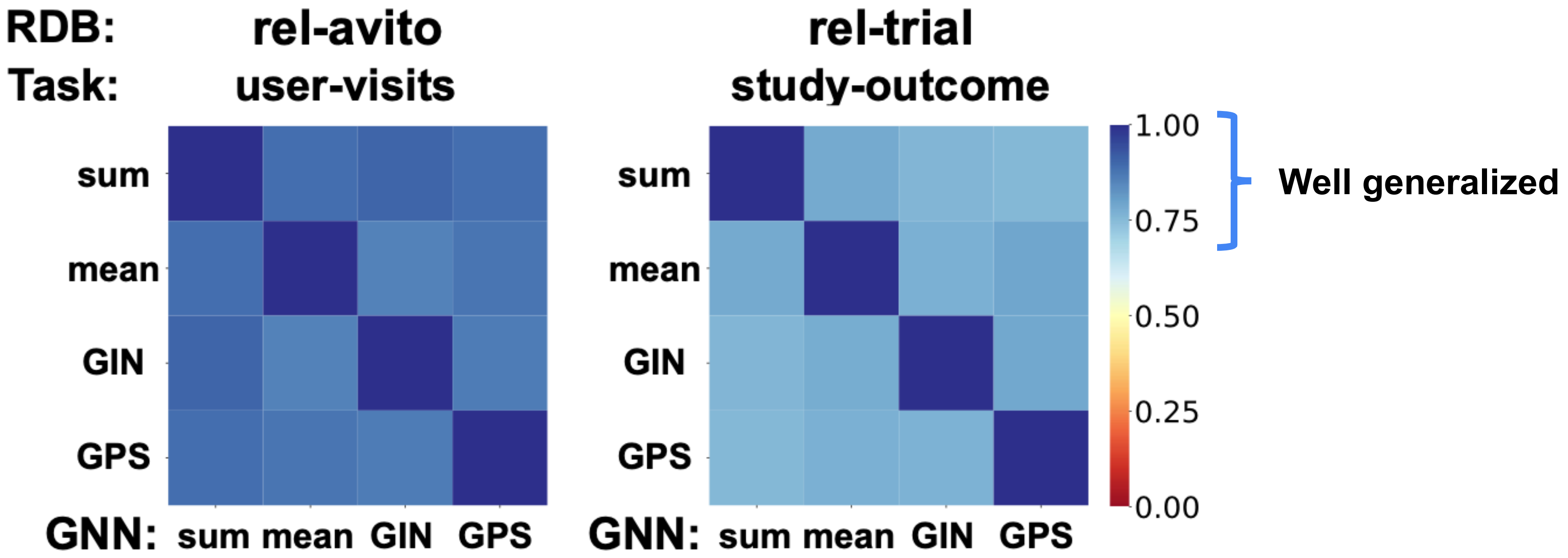


Not generalized

Generalized

Observations from datasets (cont.)

- Observation 5. Effectiveness of graph models **generalizes** across predictive GNNs.

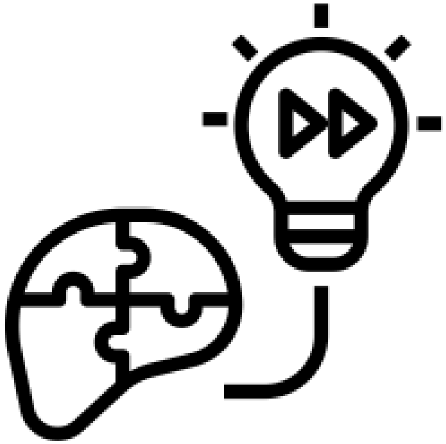


Roadmap

- Overview
- Dataset for RDB2G-Bench
- **Benchmark for RDB2G-Bench**
- Conclusions

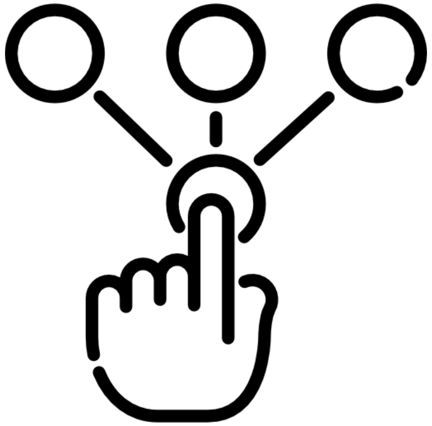
Benchmark settings (cont.)

- We benchmark 10 **RDB-to-Graph** modeling methods.
  - Each method tries to find effective graph models under constrained budgets.



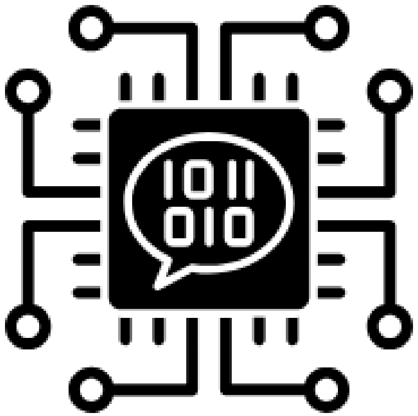
Heuristic-based

(e.g., All-Rows-to-Nodes)



Action-based search

(e.g., Greedy Search)

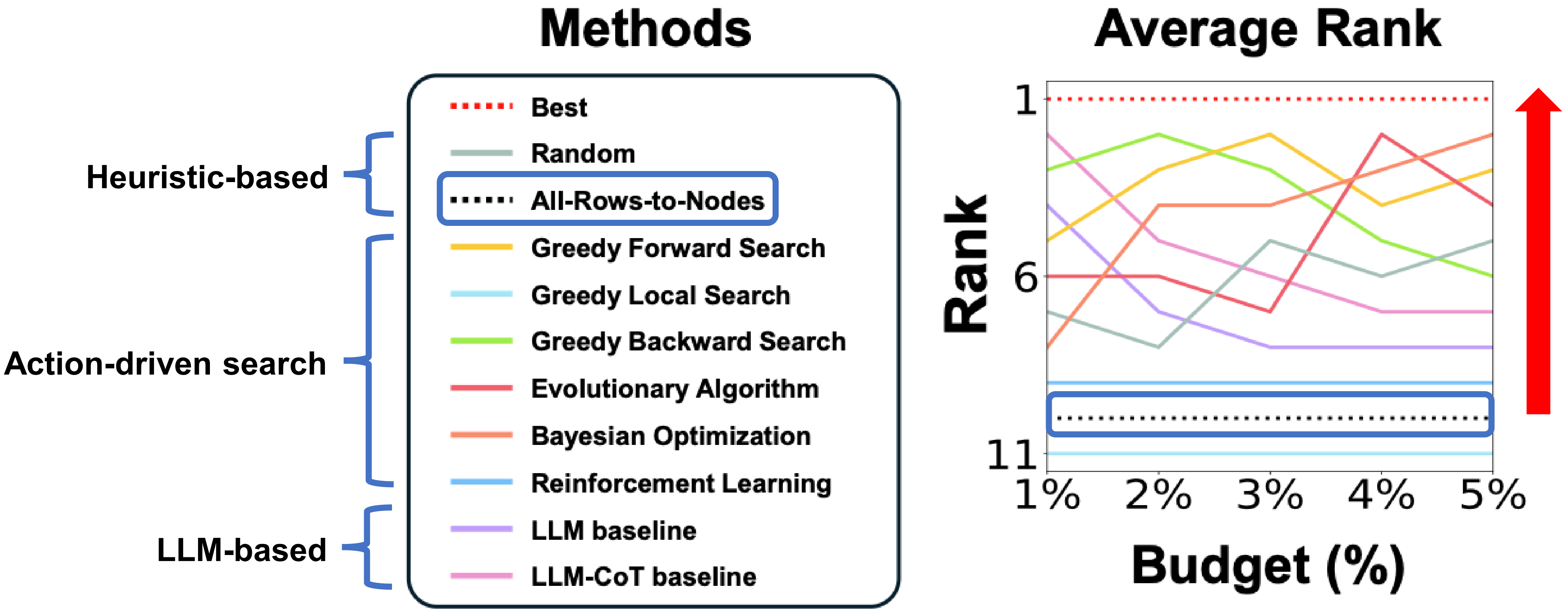


LLM-based

(e.g., Chain-of-Thought)

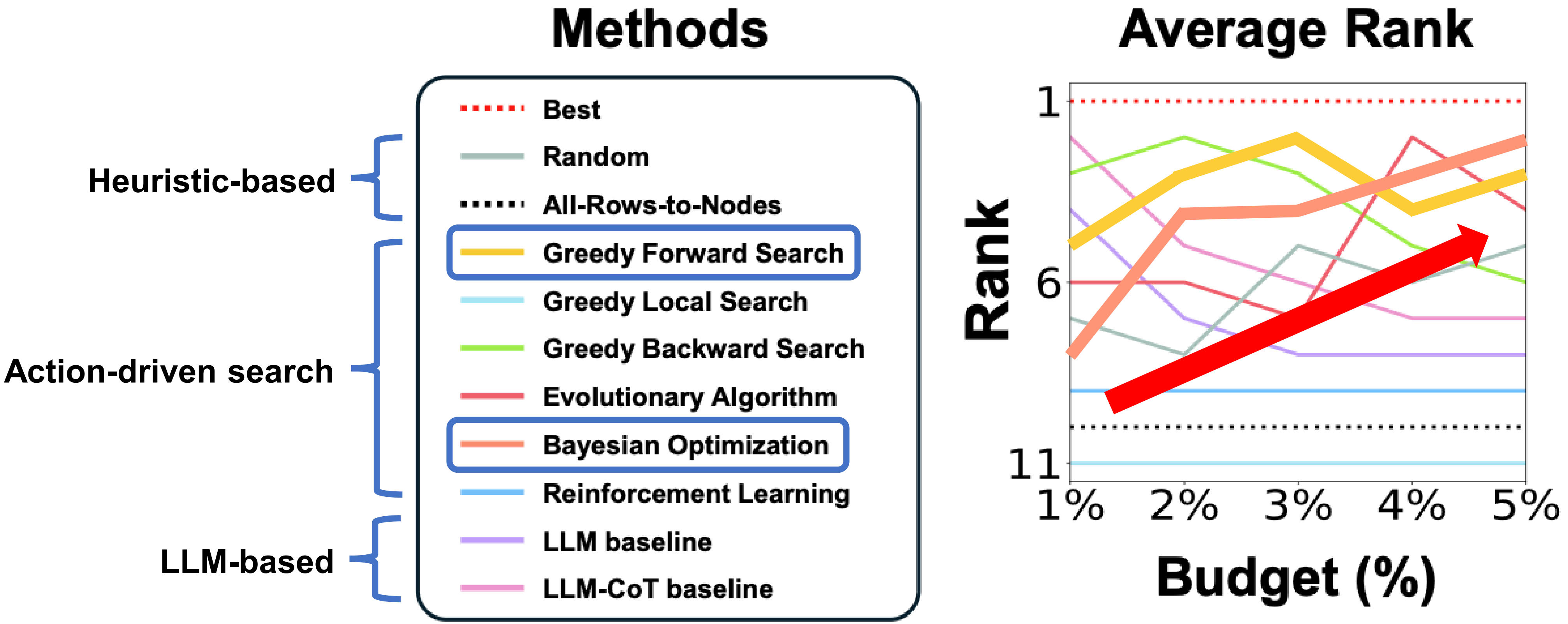
Benchmark results

- Result 1. Most baselines outperform a widely-adopted heuristic (All Rows to Nodes) with minimal explorations.



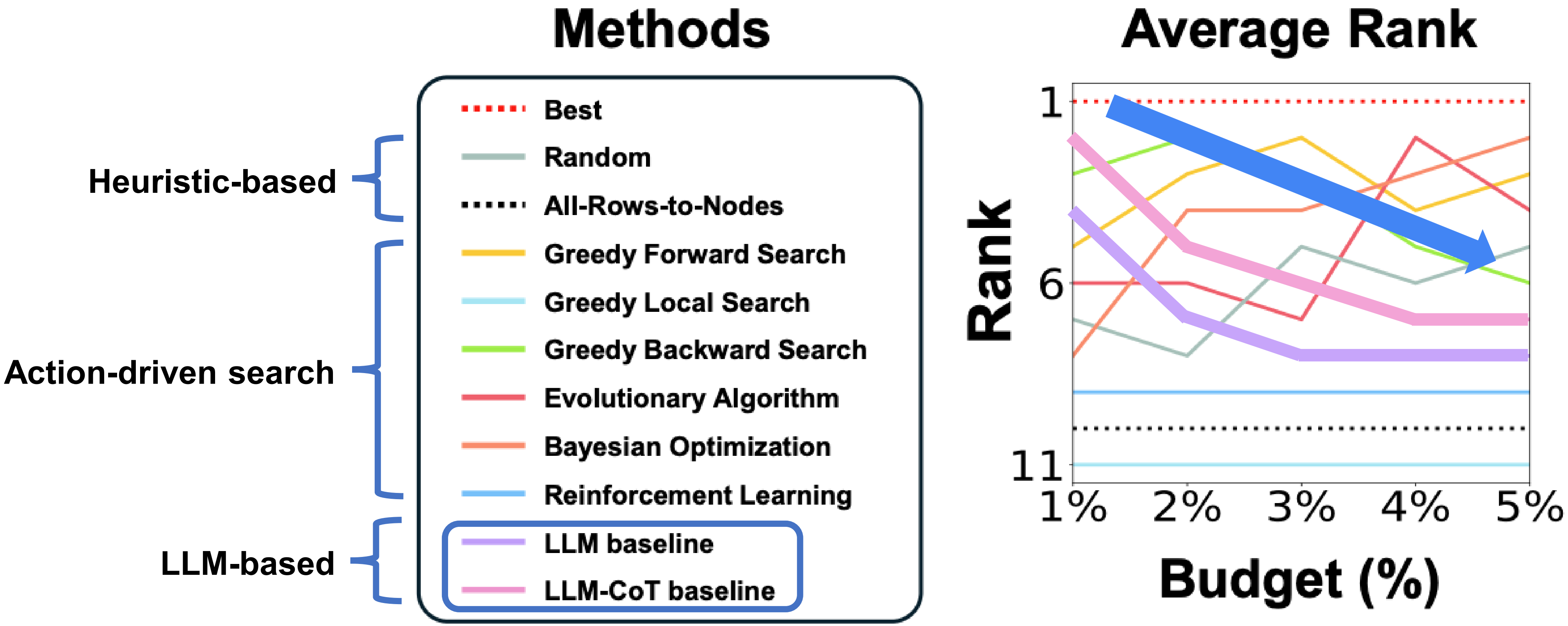
Benchmark results (cont.)

- Result 2. Value-based search algorithms, typically improve as performance feedback accumulates.



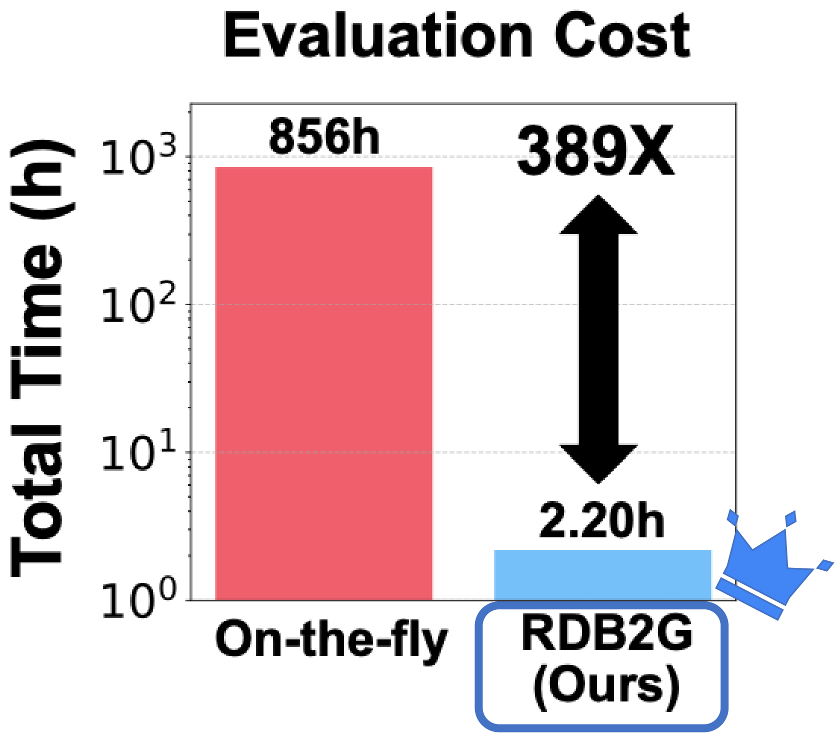
Benchmark results (cont.)

- Result 3. LLM baselines perform well under small budgets but struggles as the search space expands.



Benchmark results (cont.)

- Result 4. **RDB2G-Bench** enables evaluation up to 389x faster than on-the-flying results.



Roadmap

- Overview
- Dataset for RDB2G-Bench
- Benchmark for RDB2G-Bench
- **Conclusions**

Conclusions

Topic: The first benchmark for automatic RDB-to-Graph modeling

- **Datasets:** 50k precomputed graph models from 5 real-world RDBs.
- **Benchmarks:** Enables 380x faster evaluation of modeling methods.
- **Observations:** Provides practical insights for graph modeling.

NEURAL INFORMATION PROCESSING SYSTEMS

KAIST AI
Kim Jaechul Graduate School

kumo

RDB2G-Bench: A Comprehensive Benchmark for Automatic Graph Modeling of Relational Databases








Dongwon Choi   Sunwoo Kim   Juyeon Kim   Kyungho Kim   Geon Lee   Shinhwan Kang   Myunghwan Kim   Kijung Shin