# The Missing Structure
# When Graph Representations Outperform Tabular Models

Tamara Cucumides and Floris Geerts
**University of Antwerp**

## Why tabular models cannot handle cross-row tasks

Task **T**: *"the row has a unique value in A"*

| id | A | X | T |
|---|---|---|---|
| r1 | a | y | 1 |
| r2 | b | x | 0 |
| r3 | b | z | 0 |
| r4 | b | y | 0 |

add r5

| id | A | X | T |
|---|---|---|---|
| r1 | a | y | 0 |
| r2 | b | x | 0 |
| r3 | b | z | 0 |
| r4 | b | y | 0 |
| r5 | a | z | 0 |

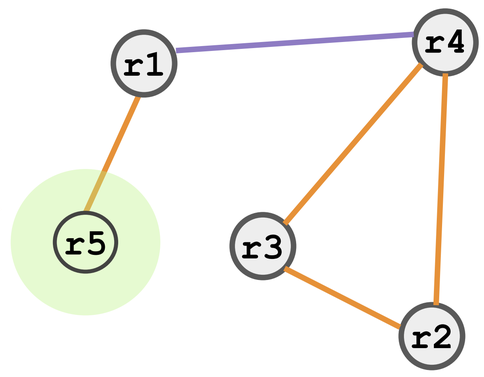
Tabular models are **row-local:** their predictions depend only on the row's own features and **cannot change when other rows are added or removed.**

## Conceptual gap

**Proposition:** Row-local tabular models cannot capture uniqueness, but message passing on a graph built from shared attribute values can

| id | A | X | T |
|---|---|---|---|
| r1 | a | y | 0 |
| r2 | b | x | 0 |
| r3 | b | z | 0 |
| r4 | b | y | 0 |
| r5 | a | z | 0 |

table to graph




A row has a **unique value in column *A*** exactly when its node has *no neighbors* via an orange edge.

## Method: Three Levels of Structure

### Original tabular data

| id | A | X |
|---|---|---|
| r1 | a | y |
| r2 | b | x |
| r3 | b | z |
| r4 | b | y |

Row-local model (*F*) (MLP, GBDT)

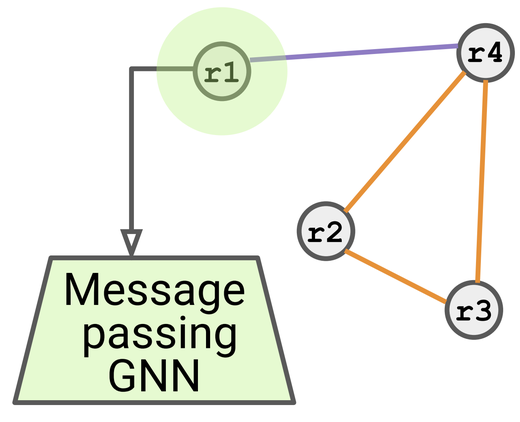
### Tabular data with aggregated one-hop graph features

| id | A | X | A=a | A=b | X=x | X=y | X=z |
|---|---|---|---|---|---|---|---|
| r1 | a | y | 0 | 1 | 0 | 1 | 0 |
| r2 | b | x | 0 | 2 | 1 | 1 | 1 |
| r3 | b | z | 0 | 2 | 1 | 1 | 1 |
| r4 | b | y | 1 | 2 | 1 | 1 | 1 |

one-hop neighbor aggregated features from graph $G_T$

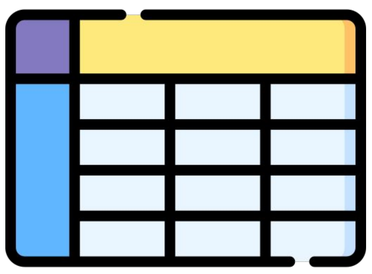
Row-local model (*F*) (MLP, GBDT)

### Graph




**Our framework progressively adds structural access: from row-local features, to static one-hop graph feature aggregates, to dynamic message passing. This isolates how much cross-row reasoning each level can express.**

## Experiments and results



- **10k synthetic tabular dataset** with 3 categorical and 2 numerical features.
- **Uniqueness and counting** (equal and greater than equal) tasks
- **80/20/20** train, validation, test split

| Model | Uniqueness Acc | Uniqueness ROC-AUC | Counting (eq, $k=3$) Acc | Counting (eq, $k=3$) ROC-AUC | Counting (geq, $k=3$) Acc | Counting (geq, $k=3$) ROC-AUC |
|---|---|---|---|---|---|---|
| RealMLP (row-local) | 0.770 | 0.540 | 0.318 | 0.371 | 0.374 | 0.382 |
| LightGBM (row-local) | 0.774 | 0.514 | 0.506 | 0.485 | 0.498 | 0.500 |
| RealMLP + NFA | 0.770 | 0.451 | 0.452 | 0.300 | 0.380 | 0.387 |
| LightGBM + NFA | 0.778 | 0.529 | 0.536 | 0.452 | 0.420 | 0.487 |
| TabPFN | 0.786 | 0.453 | 0.566 | 0.474 | 0.692 | 0.516 |
| MPGNN (2 L) | **0.811** | **0.738** | **0.806** | **0.771** | **0.998** | **0.998** |

- **Row-local models (realMLP and LightGBM)** near chance performance
- **Neighbor feature aggregation (NFA)** small, inconsistent gains
- **TabPFN** above baseline, still limited vs MPGNN
- **MPGNN** strong, reliable performance across tasks

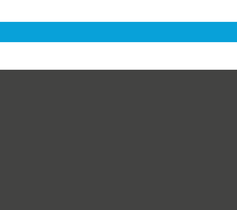
## Interpretation and practical recipe

**Tabular models**
- completely misses structure

**Neighbor feature aggregation**
- static, limited reach

**TabPFN (cross-row attention)**
- attention helps, structure wins

**Message passing**
- propagation works, logic recovered



- Try row-local lightweight model first
  - add NFA if needed ....
    - if it's still not enough, try message passing

## Open challenges and future work

(1) **Real data with cross-row dependencies**
→ Need curated datasets that expose when structural reasoning truly matters.

(2) **Robustness to noisy structure**
→ Methods should handle uncertain, partial, or noisy relational signals.

(3) **Clarify when neighbor feature aggregations are enough**
→ We need criteria or diagnostics to decide when NFA suffices and when message passing is required.