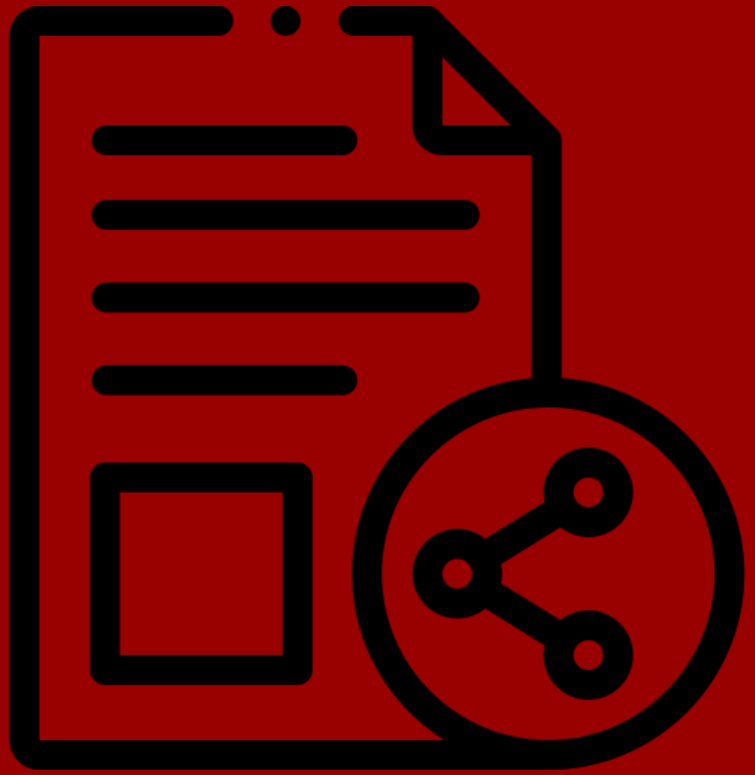
Big Data Analytics and Text Mining Project

Testing GraphAny on a RelBench task

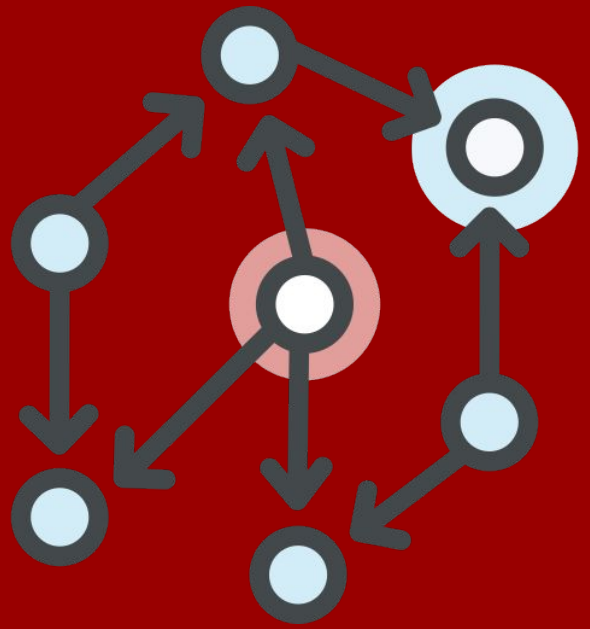
Riccardo Romeo
University of Bologna

01
The Context



GraphAny

A fully-inductive Node Classification on Arbitrary Graphs



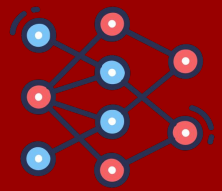
It is trained to perform node classification on:

- Any graph.
- With any feature.
- With any number of labels.

Totally Inductive Node Classification!

Some other fancy aspects

GraphAny is a young model, (june 2024 on Arxiv)!



Built on only 5 LinearGNNs!

Tested on 27 datasets.

And the number of parameters?

4 different checkpoints available.

RelBench

Open benchmark for machine learning over relational databases

Full of interesting tasks to perform on relational graphs

More importantly: we have some Binary Node Classification Tasks!

Project Idea

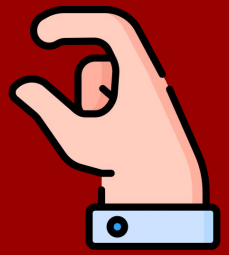
Import a task from RelBench and test GraphAny on it!

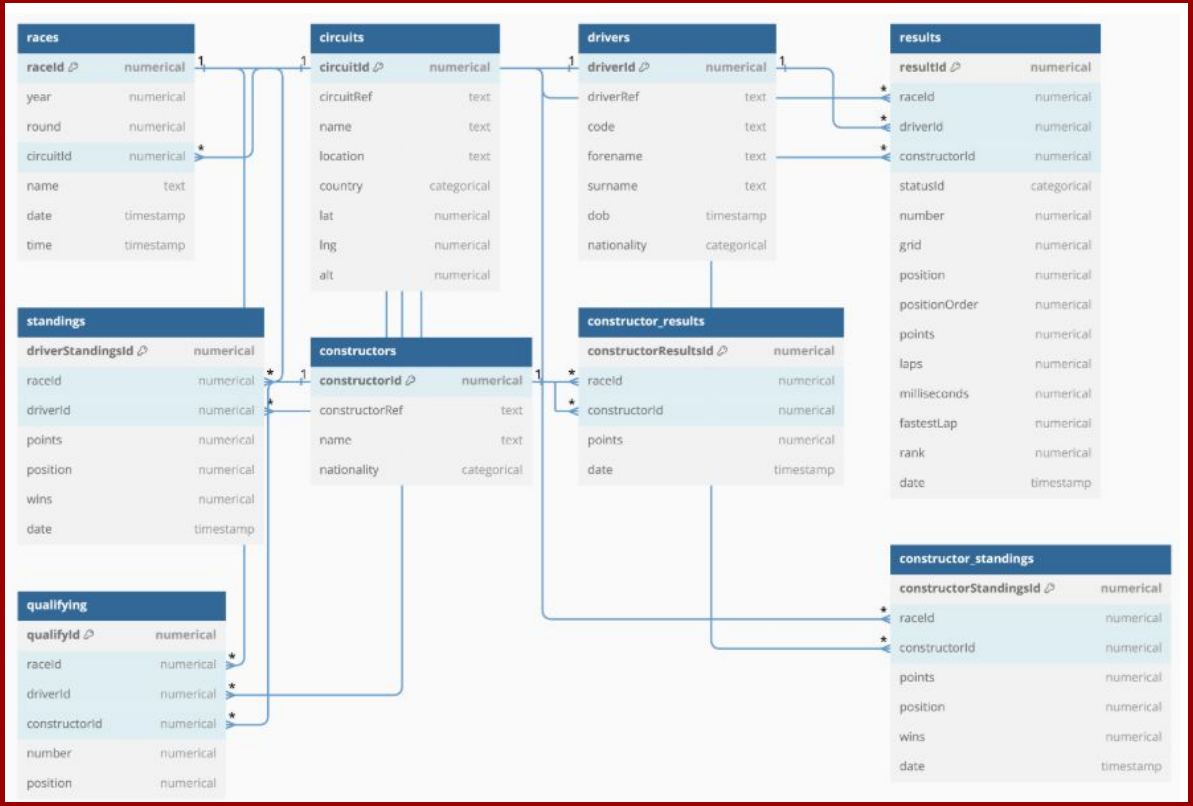
But How?

02
The Approach

The choice: F1 Database

The smallest dataset with a node classification task.





The task: For each driver predict if they will DNF (did not finish) a race in the next 1 month

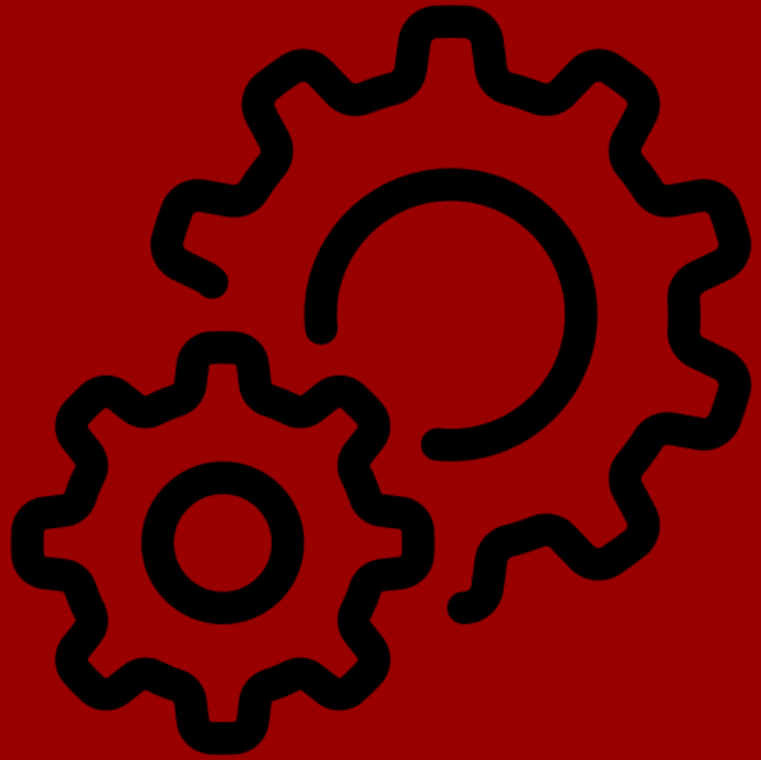
The HeteroGraph problem

Different nodes with different features...

Solution: Linearization and Glove embedding!

```
HeteroData(
  constructor_standings={
    tf=TensorFrame([13051, 4]),
    time=[13051],
  },
  results={
    tf=TensorFrame([26080, 11]),
    time=[26080],
  },
  circuits={ tf=TensorFrame([77, 7]) },
  drivers={ tf=TensorFrame([857, 6]) },
  races={
    tf=TensorFrame([1101, 5]),
    time=[1101],
  },
  standings={
    tf=TensorFrame([34124, 4]),
    time=[34124],
  },
  qualifying={
    tf=TensorFrame([9815, 3]),
    time=[9815],
  },
  constructors={
tf=TensorFrame([211, 3]) },
  constructor_results={
    tf=TensorFrame([12290, 2]),
    time=[12290],
  }
```

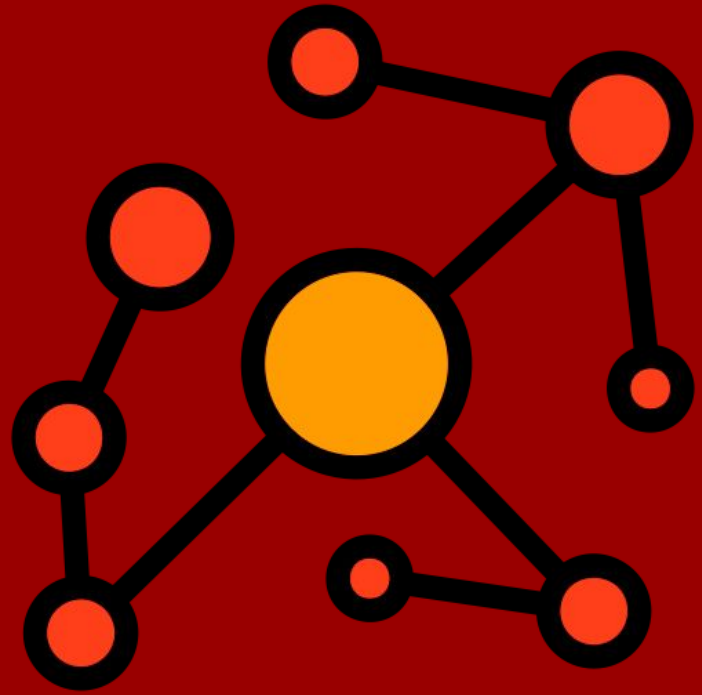
03 Experimental Setup



Our Dataset

A HUGE Knowledge Graph:

- 97605 nodes.
- 455432 edges.
- but only 3 labels.

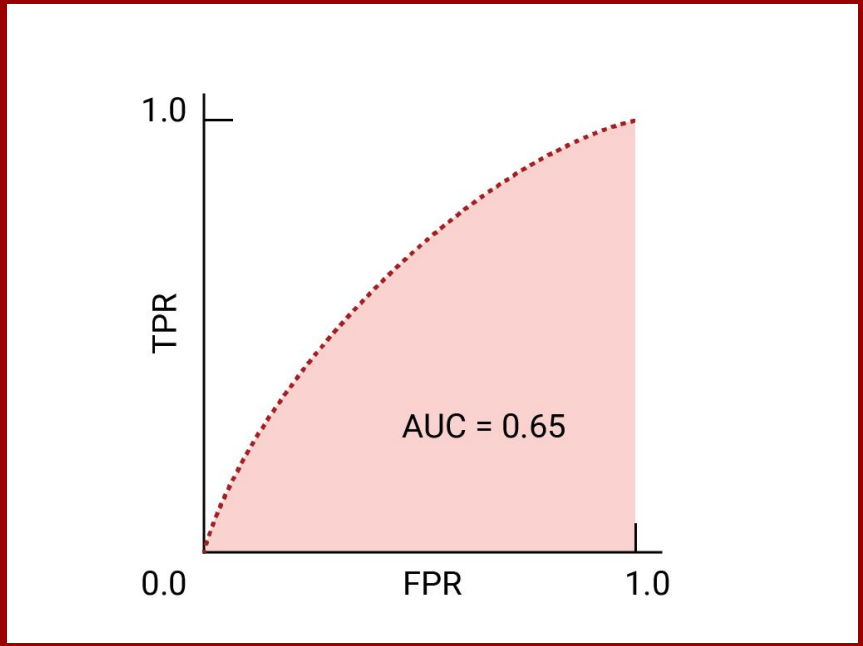


The Metrics

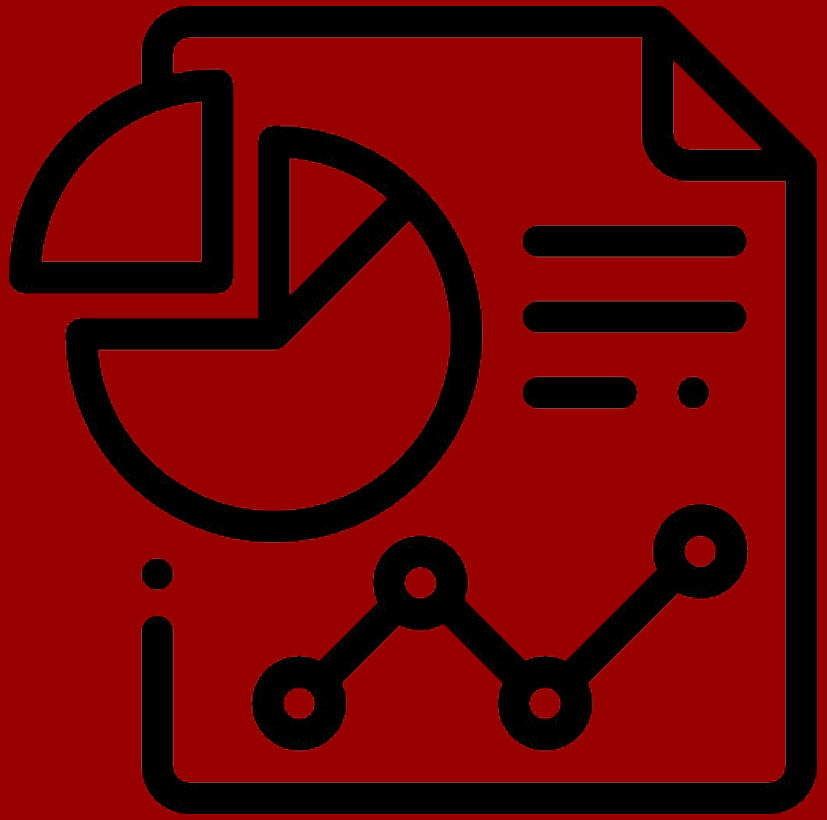
1. F1-Score

2. Binary ROC AUC

3. One-vs-rest ROC AUC



04
The Results



GraphAny performances on Task

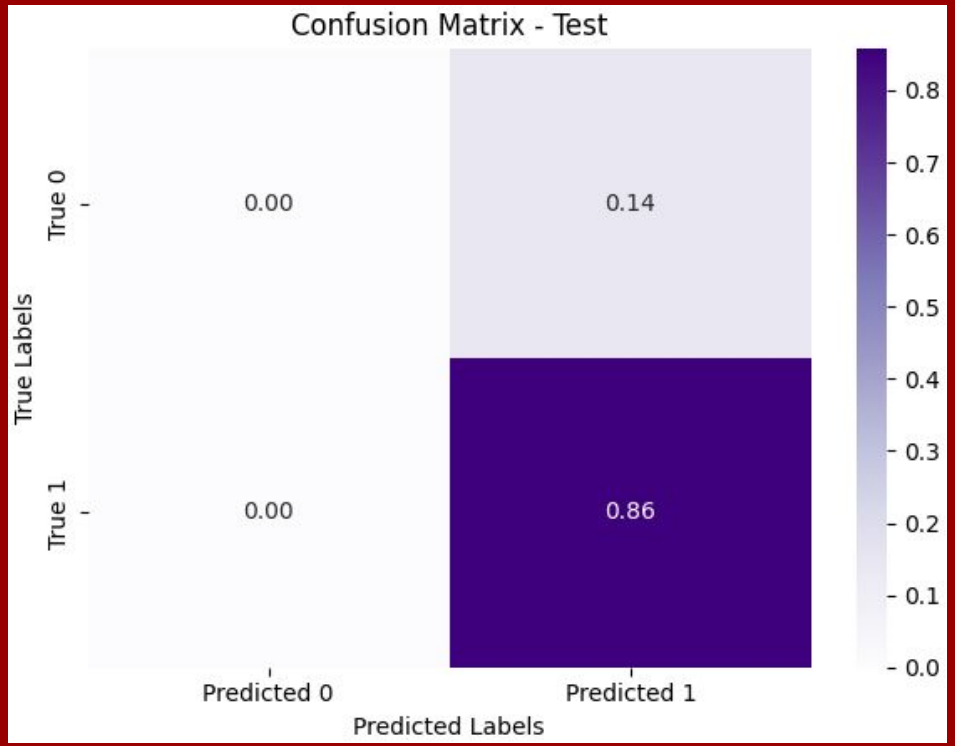
Precision, Recall and F1 per class.

Checkpoint	Precision	Recall	F1 Class 0	F1 Class 1
Wisconsin	0.86	1.00	0.00	0.92
Cora	0.86	1.00	0.00	0.92
Arxiv	0.86	1.00	0.00	0.92

Table 1: GraphAny Performances on Formula 1 Test set RelBench Knowledge Graph in terms of Precision, Recall and F1-Score on both classes.

GraphAny performances on Task

Confusion Matrix.



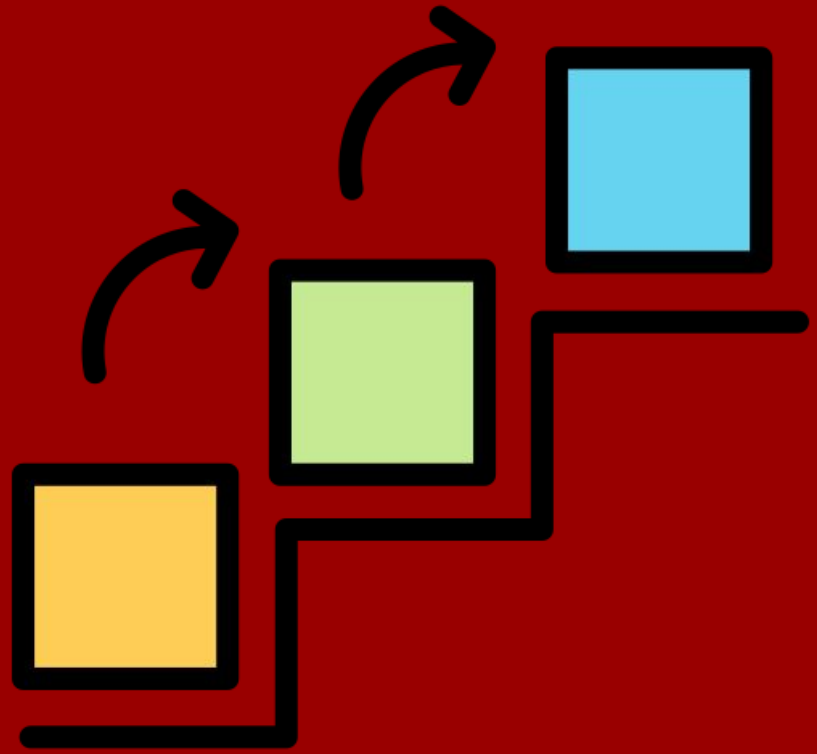
GraphAny performances on Task

Cohen's Kappa Score, Binary ROC-AUC and One-vs-Rest ROC-AUC.

Checkpoint	Cohen's kappa	Bin-AUROC	OvR-AUROC
Wisconsin	0.00	58.80	53.70
Cora	0.00	60.19	56.25
Arxiv	0.00	56.94	53.47

Table 2: GraphAny Performances on Formula 1 Test set RelBench Knowledge Graph in terms of Cohen's Kappa (computed between a random classifier and GA classifier), Binary ROC-AUC and One-vs-Rest ROC-AUC.

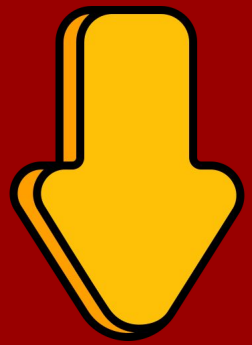
05 Future Project Improvements



Some critical aspects

- We tested the model on only one task.

- We tested the model on only one dataset.



Our metrics could be biased by the task choice!

Thank you for your attention!