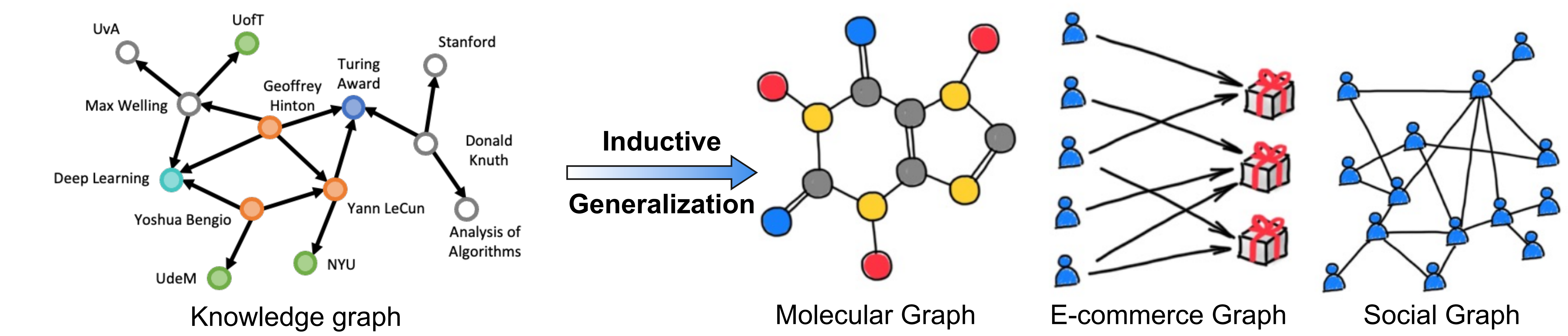
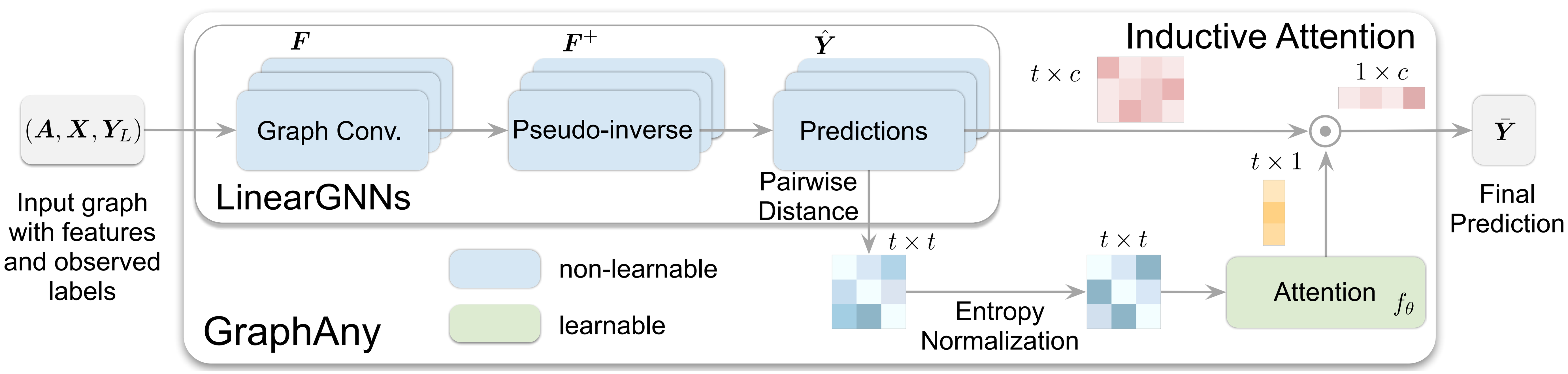
Fully-inductive Node Classification on Arbitrary Graphs

Jianan Zhao, Mikhail Galkin, Hesham Mostafa, Michael Bronstein, Zhaocheng Zhu, Jian Tang

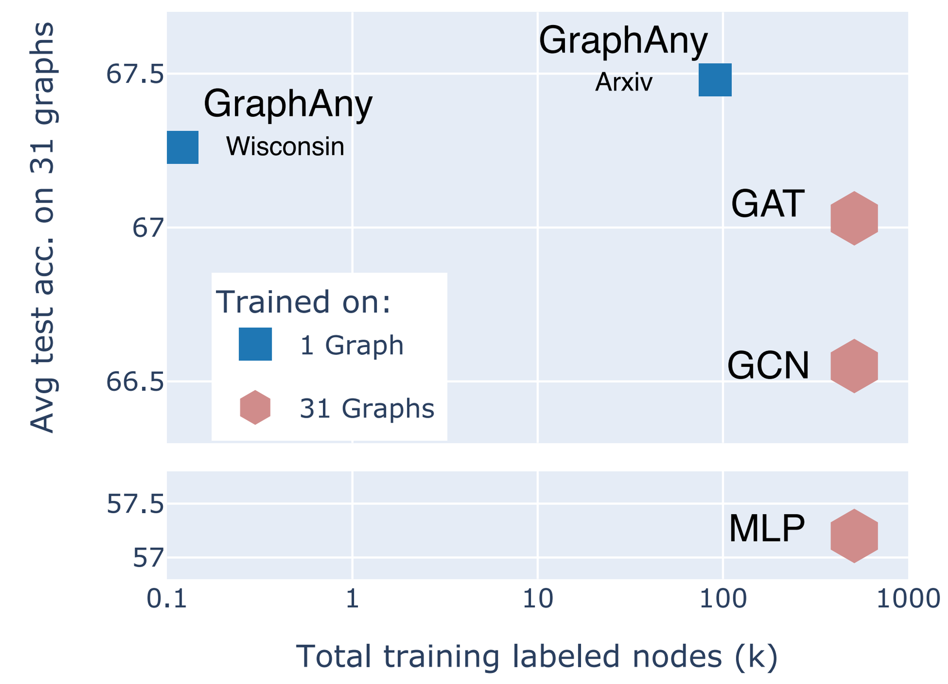
Mila · intel labs · Université de Montréal · University of Oxford

Fully Inductive Generalization: Inference on ***any graph*** with ***unseen structure, feature, and label*** spaces.





The GraphAny Architecture



Node Classification Performance

LinearGNN

Training-free inference on ***any graph*** with an ***analytical solution***.

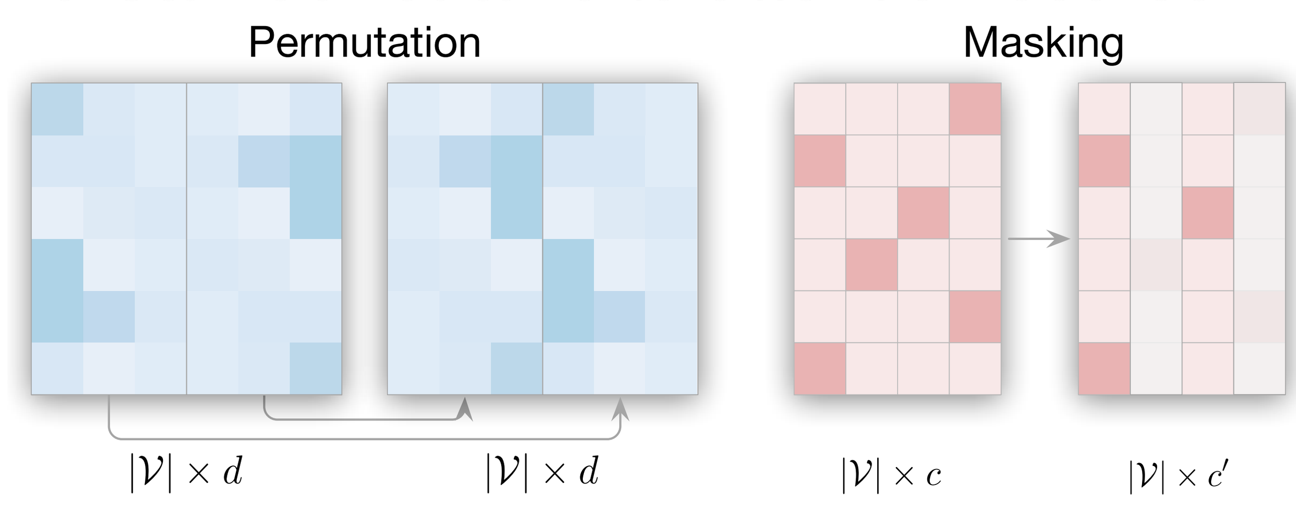
$$\hat{Y} = FW, \quad W^* = \arg\min_{W} \left\| \hat{Y}_L - Y_L \right\|^2,$$

$$W^* = F_L^+ Y_L, \quad \hat{Y} = F F_L^+ Y_L.$$

- Precomputed propagated feature F, e.g. AX, A^2X.
- Weights W^* are given in analytical form, optimizing the mean-squared error (MSE) loss.
- Solving W^* takes $O(|V_L|)$ complexity.
- Provides fast and *training-free inductive inference* with reasonably good performance.

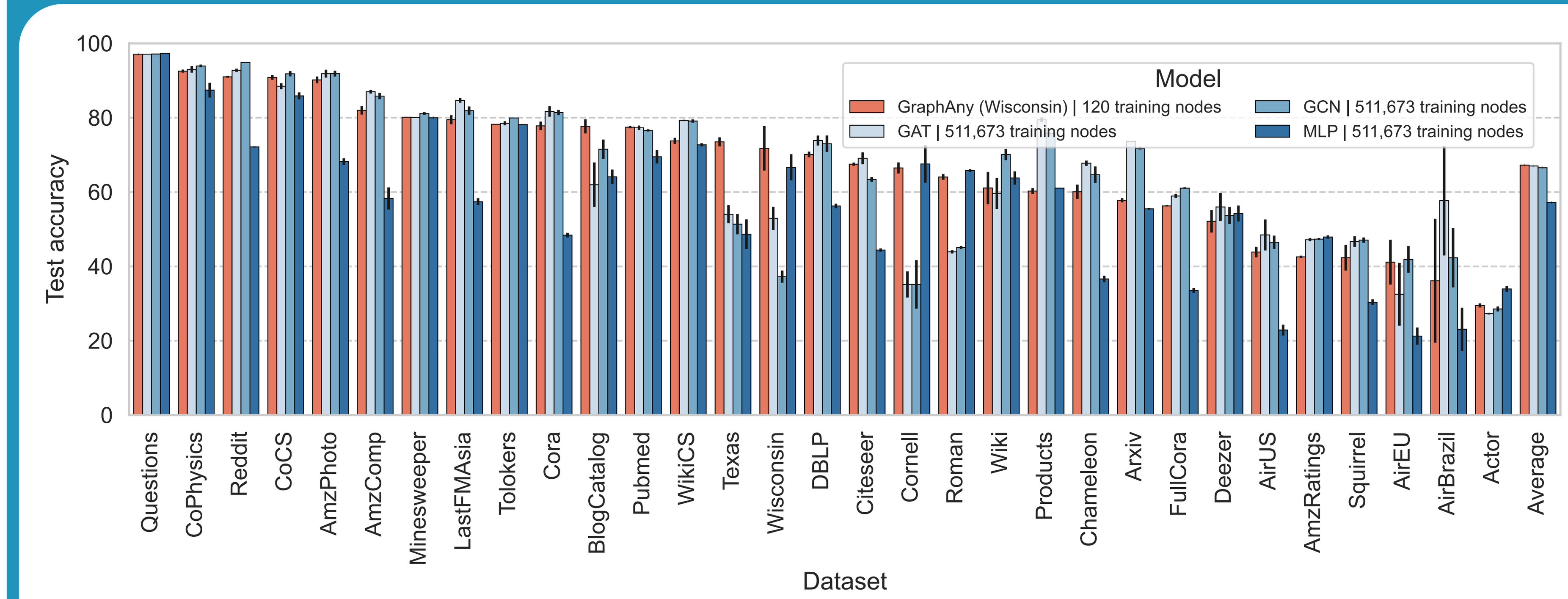
Inductive Attention

An inductive function that maps ***entropy-normed distance feature*** to attention scores .



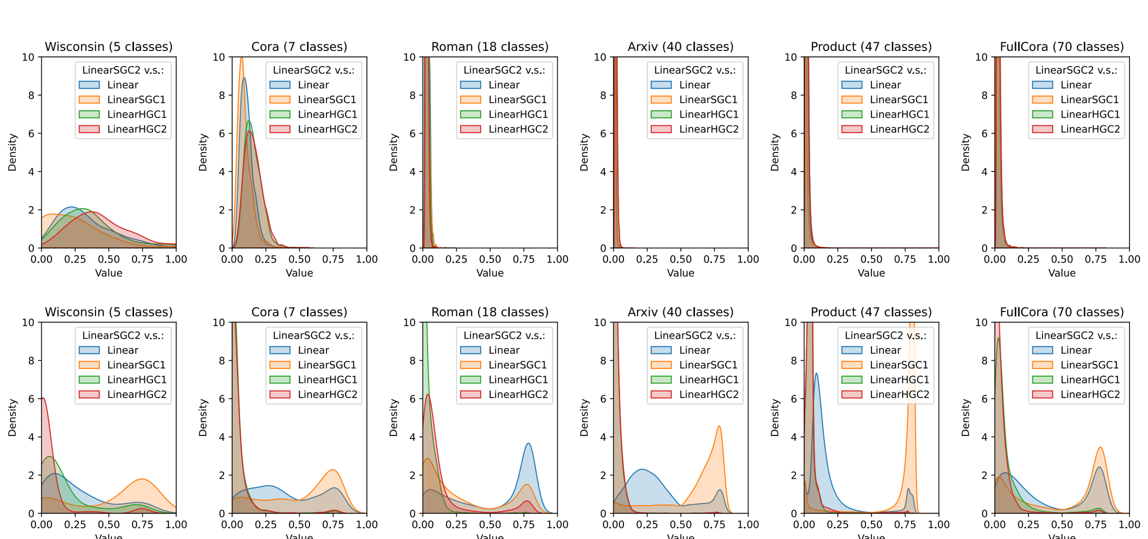
- A fully-inductive inductive function should satisfy: ***Permutation Invariance*** and ***Dimensional Robustness*** across diverse feature/label dimensions.
- We propose ***entropy-normalized distance features*** to compute attention scores that adaptively fuse LinearGNN predictions.

Experiments

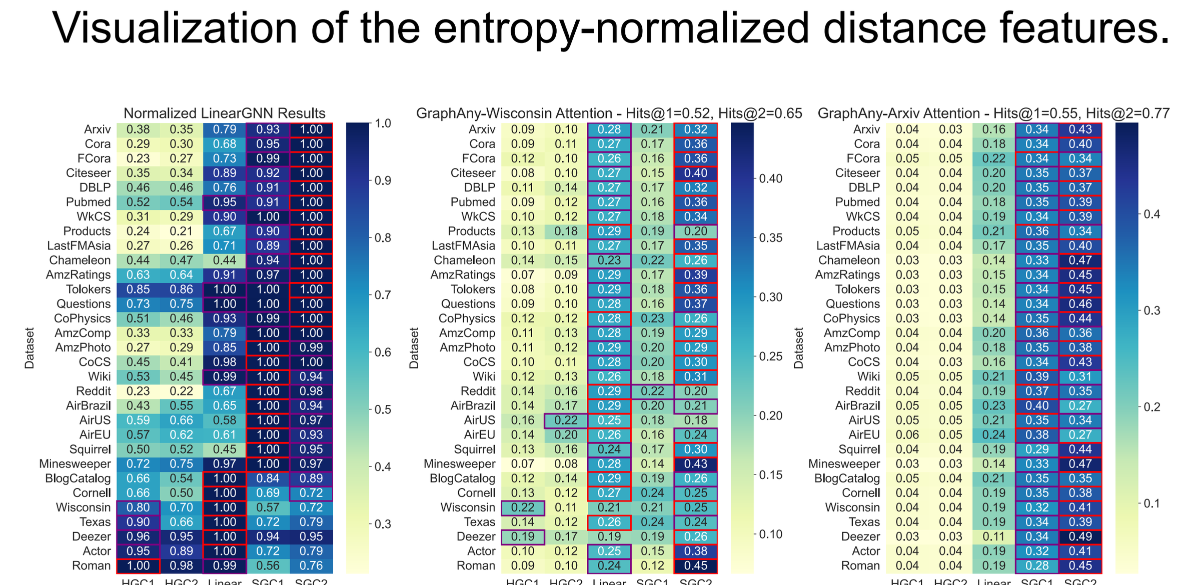


- ***Inference on any graph***: GraphAny trained on one graph, e.g. Wisconsin with 120 labeled nodes, generalizes to 30 new graphs and ***outperforms transductive baselines***.
- ***Effectiveness***: Training with 2.95x speed up on a single V100 or even on your laptop!

Model	Pre-processing	Optimization	Inference	Total Wall Time (31 graphs)
GCN	0	$O(\|\mathcal{E}\|)$	$O(\|\mathcal{E}\|)$	18.80 min
LinearGNN	$O(\|\mathcal{E}\|)$	$O(\|\mathcal{V}_L\|)$	$O(\|\mathcal{V}\|)$	1.25 min (15.04×)
GraphAny	$O(\|\mathcal{E}\|)$	$O(\|\mathcal{V}_L\|)$	$O(\|\mathcal{V}\|)$	6.37 min (2.95×)



Visualization of the entropy-normalized distance features.



Visualization of the learned inductive attention scores.

Paper

Code