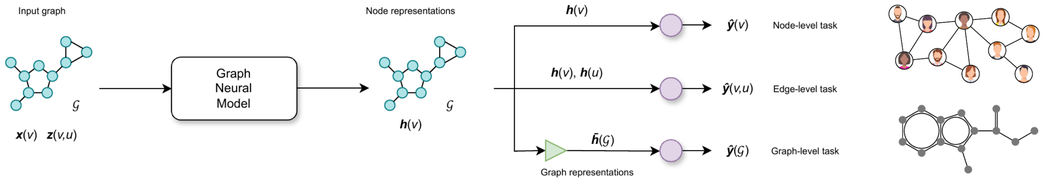
University of Pisa
Department of Computer Science

Efficient Deep Learning on Graphs

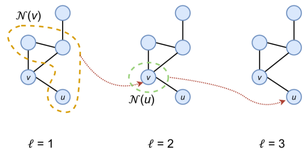
Computational Intelligence and Machine Learning Group

Domenico Tortorella (✉ domenico.tortorella@phd.unipi.it)

Neural Networks for Graphs



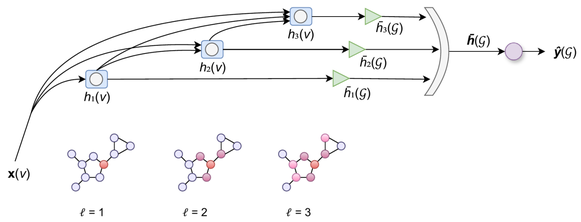
- Message-passing leverages locality to build a hierarchy of representations



Issues and challenges

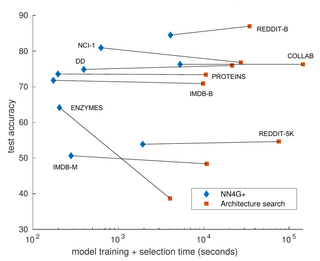
- Gradient back-propagation and other training issues (over-smoothing/over-squashing, …)
- Model selection on increasingly large hyper-parameter spaces
- Scaling to larger graphs and larger datasets → efficiency

Automatic Construction of Deep Neural Networks – NN4G+

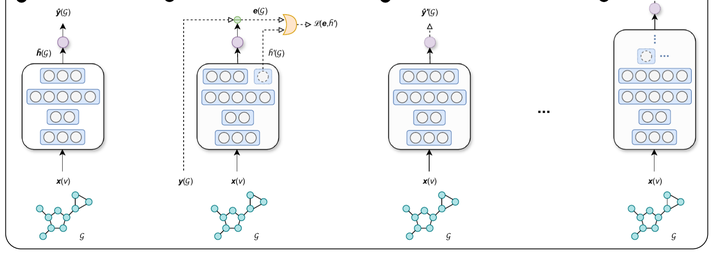


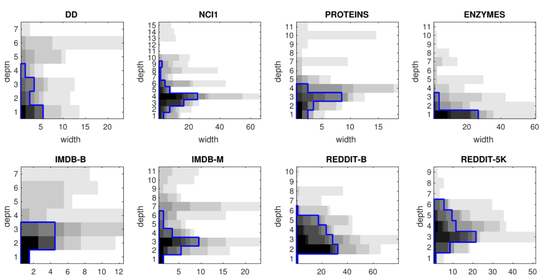
- Build the DNN architecture one unit at a time
- Greedy exploration:
  - ❖ Keep expanding the current layer
  - ❖ Then move on to the next deeper one
- Advantages:
  - ✓ No deep gradient back-propagation issues
  - ✓ No explicit architecture selection

- Fast, effective learning



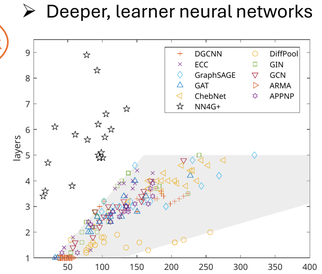
How does it work?





Models tailored to the learning task

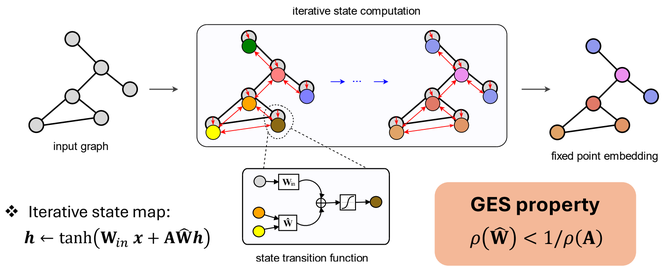
- Deeper, learner neural networks



Reservoir Computing for Graphs – Graph Echo State Networks

Reservoir Computing

- Inputs are encoded as states of a dynamical system
- Weights are randomly initialized, only readout is trained
- Conditions on weight initialization ensure effective embeddings



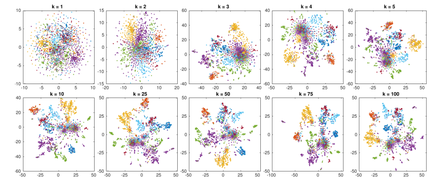
❖ Iterative state map:
$\boldsymbol{h} \leftarrow \tanh(\mathbf{W}_{in}\,\boldsymbol{x} + \mathbf{A}\widehat{\mathbf{W}}\boldsymbol{h})$

GES property
$\rho(\widehat{\mathbf{W}}) < 1/\rho(\mathbf{A})$

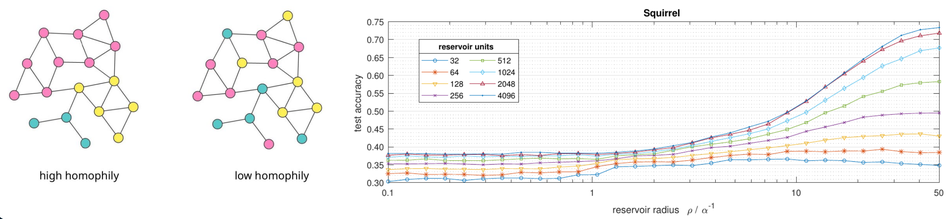
MY RESEARCH

Addressing Heterophily – GESN for Node Classification

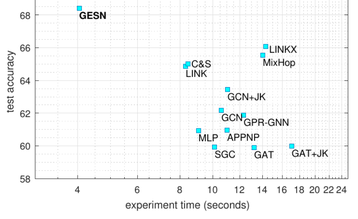
- ❑ **Homophily**
  - ➢ Prevalence of intra-class links
  - ➢ Local information is reliable for classification
- ❑ **Heterophily**
  - ➢ Significant fraction of inter-class links
  - ➢ Short-range information no longer good



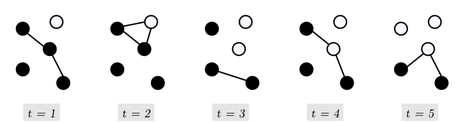
Beyond GES $\rho(\widehat{\mathbf{W}}) \gg 1/\rho(\mathbf{A})$



- **Node-level GESN**
  - ✓ No over-smoothing/squashing issues
  - ✓ State-of-the-art accuracy on several heterophilous node classification task
  - ✓ More efficient, effective when scaling to large-size graphs (>100K nodes)



Learning on Temporal Graphs – Dynamic Graph Echo State Networks



- Temporal graphs represent relationships that evolve through time
- Challenges for Temporal Graph Nets
  - ❖ All issues of training in DNNs
  - ❖ *plus* Memory fading, BPTT, …
- **Dynamic Graph Echo State Networks**
  - ❖ Reservoir Computing approach
  - ❖ Spatio-temporal graph convolution

- **Extreme efficiency**
  - ✓ 10× faster inference vs TGNs
  - ✓ 100× faster training vs TGNs
- **Mostly better accuracy than TGNs**

Echo State Property
$\|\widehat{\mathbf{W}}\| < 1 \,/\, \lceil \alpha_t \rceil,\ \alpha_t = \|\mathbf{A}_t\|$

- Representation over different time-scales