Node Degree for Attributed Graphs

- **Justification:** Node degree provides information about the node's connectivity, which can complement existing attributes.
- This was added as an additional item in the feature vector of each node.
- The original (paper) model configuration was used to rule out any other factors that may introduce variability in the AUC scores.
- Let’s take a look at the results…

Node Degree for Attributed Graphs (cont.)

Datasets	GLADC	Node Features	Node Features + Node Degrees
MMP	**0.696 ± 0.042**		
HSE	**0.618 ± 0.110**		
p53	**0.649 ± 0.216**		
BZR	**0.715 ± 0.067**	**0.683 ± 0.045**	**0.661 ± 0.078**
DHFR	**0.612 ± 0.041**	**0.560 ± 0.053**	**0.553 ± 0.050**
COX2	**0.615 ± 0.044**	**0.595 ± 0.093**	**0.632 ± 0.053**
ENZYMES	**0.583 ± 0.035**	**0.529 ± 0.071**	**0.547 ± 0.080**
IMDB	**0.656 ± 0.023**		
AIDS	0.993 ± 0.005	**0.993 ± 0.005**	**0.992 ± 0.008**
NCI1	0.683 ± 0.011		

Node Degree | Attributed Graphs | Graph Convolutions

- Main takeaway: adding node degree does not seem to improve AUC scores significantly (or at all).
- From our previous discoveries… Graph convolutions outperformed linear layers when dealing with node degrees.
- Let’s see if we can translate that knowledge.

Node Degree | Attributed Graphs | Graph Convolutions (cont.)

Datasets	GLADC	Node Features	Node Features + Node Degrees	Node Features + Node Degrees + GC
MMP	**0.696 ± 0.042**			
HSE	**0.618 ± 0.110**			
p53	**0.649 ± 0.216**			
BZR	**0.715 ± 0.067**	**0.683 ± 0.045**	**0.661 ± 0.078**	**0.478 ± 0.074**
DHFR	**0.612 ± 0.041**	**0.560 ± 0.053**	**0.553 ± 0.050**	**0.574 ± 0.048**
COX2	**0.615 ± 0.044**	**0.595 ± 0.093**	**0.632 ± 0.053**	**0.465 ± 0.038**
ENZYMES	**0.583 ± 0.035**	**0.529 ± 0.071**	**0.547 ± 0.080**	**0.494 ± 0.038**
IMDB	**0.656 ± 0.023**			
AIDS	0.993 ± 0.005	**0.993 ± 0.005**	**0.992 ± 0.008**	**0.276 ± 0.363**
NCI1	0.683 ± 0.011			

Node Degree | Attributed Graphs | Graph Convolutions (cont.)

- So no improvement…
- Perhaps GC work better overall with bigger datasets.
  - Note the attributed datasets are considerably smaller than the plain ones
    - The biggest attributed (AIDS): 2000 graphs.
    - The smallest plain (NCI1): 4110 graphs.

Loss Function Reconfiguration

- We have played around with removing L3 from LG.
- Let’s now look at the effects of removing the contrastive loss (L2) to see how it impacts the AUC scores.

Node Degree for Attributed Graphs (cont.)

Datasets	GLADC	LG	LG - L2
MMP	**0.696 ± 0.042**	0.508 ± 0.138	0.507 ± 0.139
HSE	**0.618 ± 0.110**	0.551 ± 0.070	0.553 ± 0.070
p53	**0.649 ± 0.216**	0.497 ± 0.122	0.496 ± 0.121
BZR	**0.715 ± 0.067**	0.683 ± 0.045	0.682 ± 0.078
DHFR	**0.612 ± 0.041**	0.560 ± 0.053	0.560 ± 0.050
COX2	**0.615 ± 0.044**	0.595 ± 0.093	0.595 ± 0.093
ENZYMES	**0.583 ± 0.035**	0.529 ± 0.071	0.529 ± 0.071
IMDB	**0.656 ± 0.023**		
AIDS	0.993 ± 0.005	0.993 ± 0.005	0.992 ± 0.003
NCI1	0.683 ± 0.011	0.668 ± 0.034	0.667 ± 0.034