Taking a Step Back

- We have looked at GLADC’s approach closely.
- Perhaps it’s a good time to take a step back and look at the problem we are dealing with.

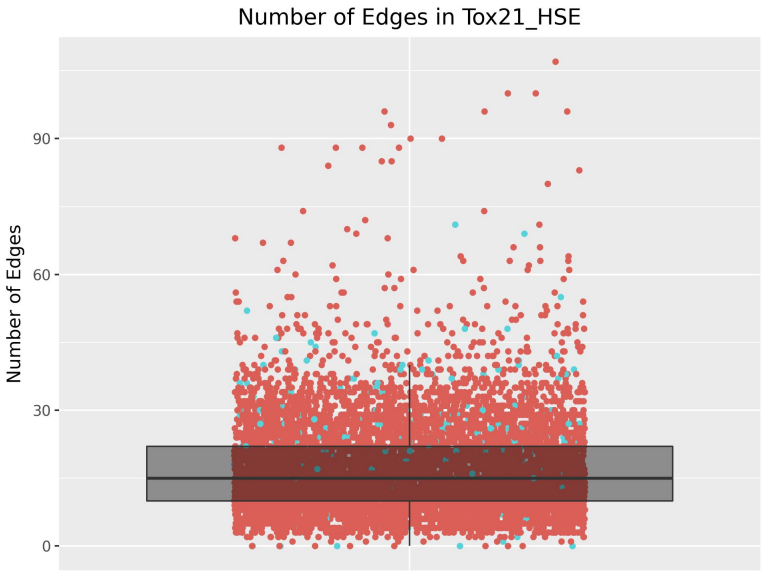
Classifier Base Comparison

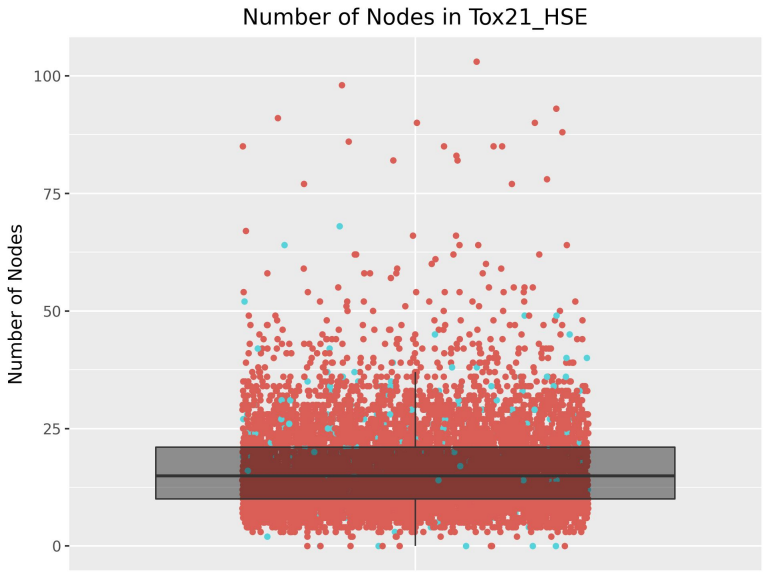
Datasets	GLADC	Dummy	Naive
MMP	**0.696 ± 0.042**	0.5 ± 0.0	0.5 ± 0.0
HSE	**0.618 ± 0.110**	0.5 ± 0.0	0.5 ± 0.0
p53	**0.649 ± 0.216**	0.497 ± 0.02	0.498 ± 0.0
BZR	**0.715 ± 0.067**		
DHFR	**0.612 ± 0.041**		
COX2	**0.615 ± 0.044**		
ENZYMES	**0.583 ± 0.035**		
IMDB	**0.656 ± 0.023**		
AIDS	0.993 ± 0.005		
NCI1	0.683 ± 0.011	0.5 ± 0.0	0.501 ± 0.001

Classifier Base Comparison (cont.)

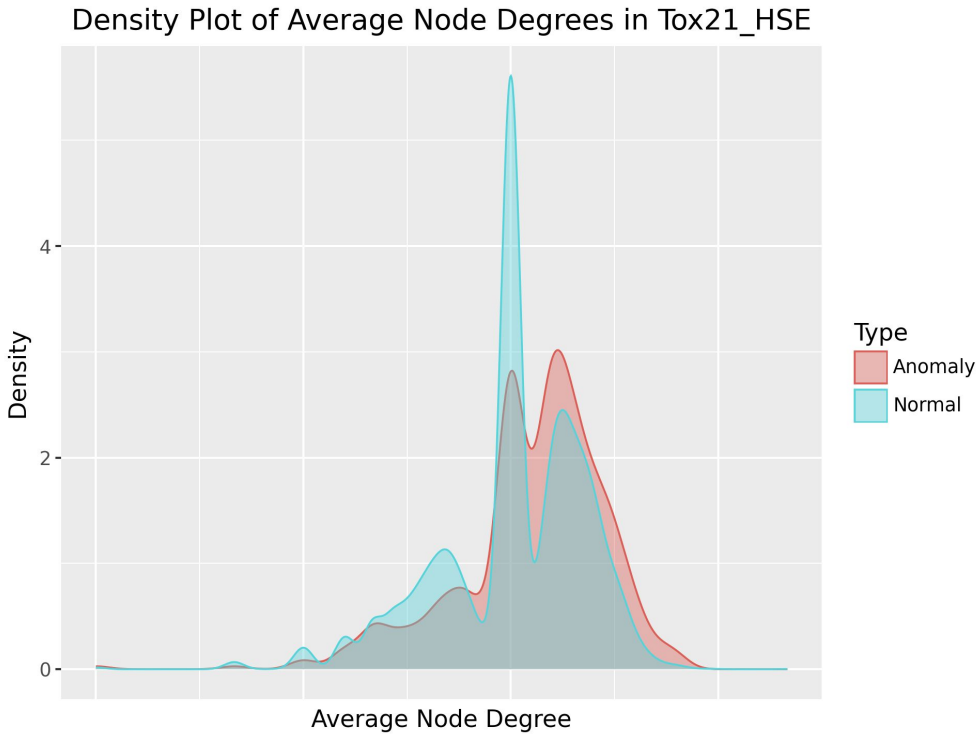
- So is GLADC better than flipping a coin…
  - Yes.
- Both the dummy and naive classifiers have a comparable performance.
- Let’s take a look at the datasets.
  - This may give us better insight on how to approach the problem.

HSE

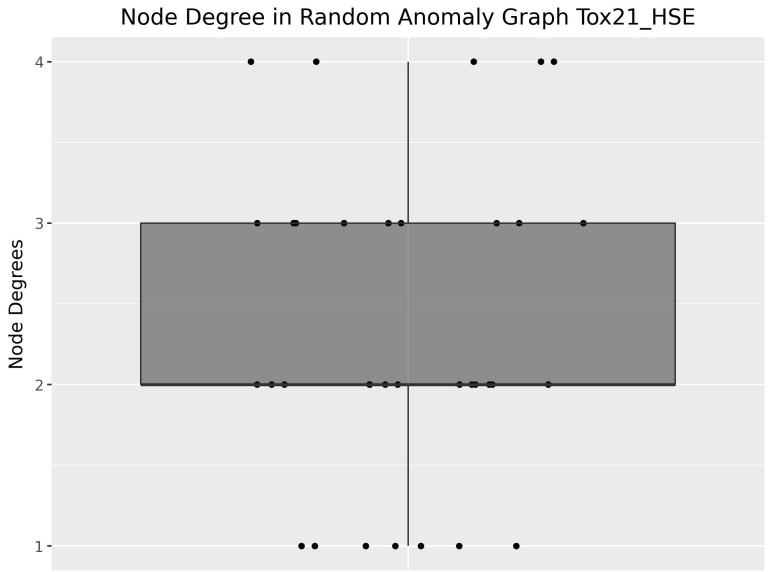

Number of Edges in Tox21_HSE
Number of Edges
90
60
30
0

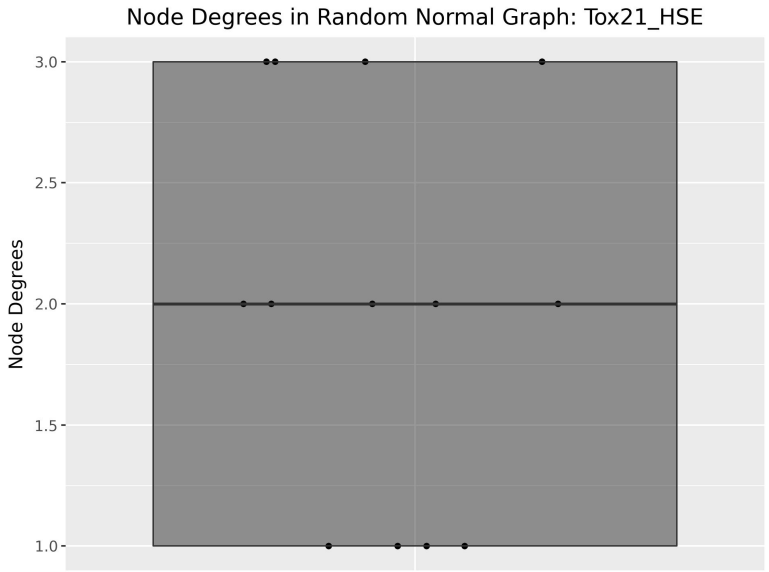

Number of Nodes in Tox21_HSE
Number of Nodes
100
75
50
25
0

HSE (cont.)

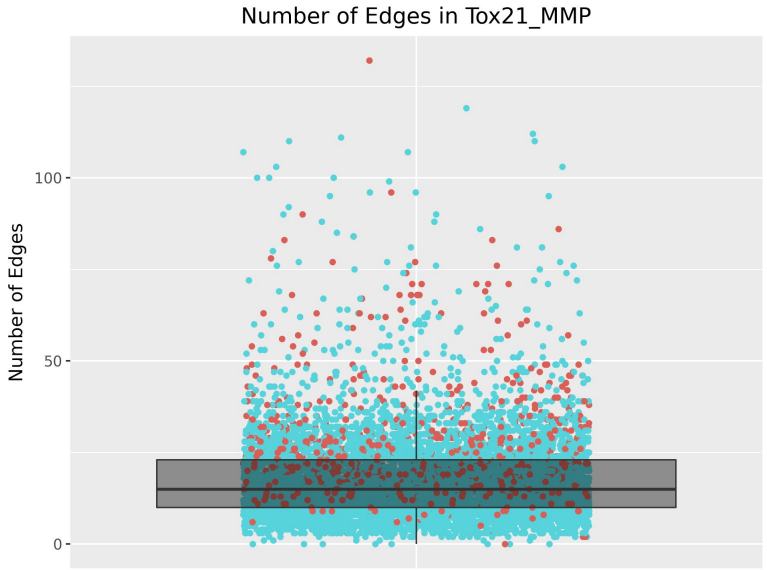

Density Plot of Average Node Degrees in Tox21_HSE
Density
4
2
0
Average Node Degree
Type
Anomaly
Normal

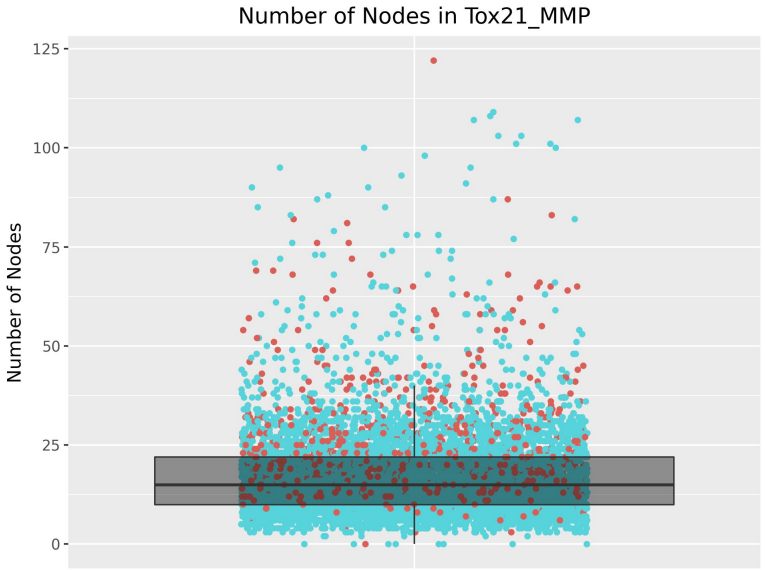
HSE (cont.)


Node Degree in Random Anomaly Graph Tox21_HSE
Node Degrees
4
3
2
1

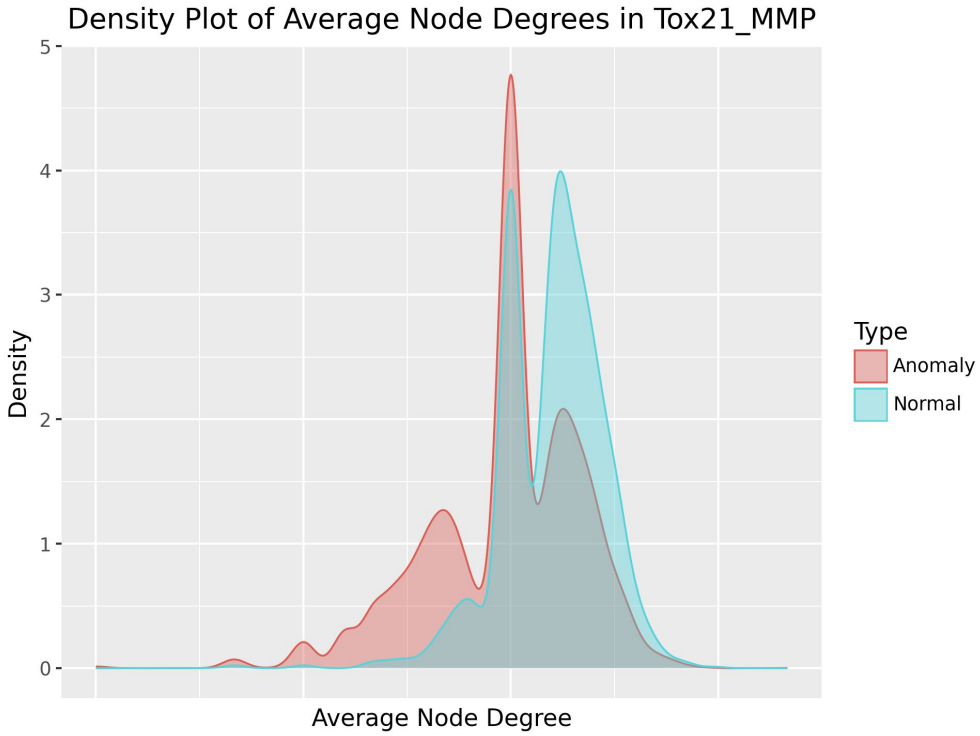

Node Degrees in Random Normal Graph: Tox21_HSE
Node Degrees
3.0
2.5
2.0
1.5
1.0

MMP

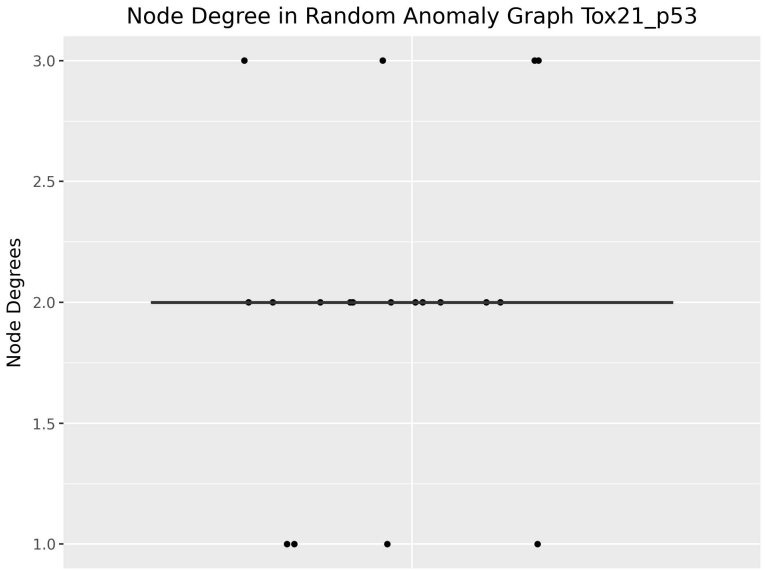

Number of Edges in Tox21_MMP
Number of Edges
100
50
0

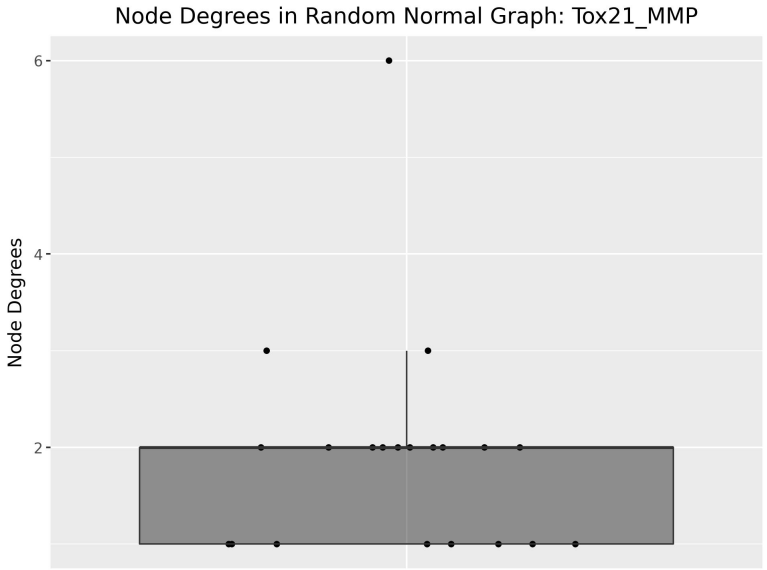

Number of Nodes in Tox21_MMP
Number of Nodes
125
100
75
50
25
0

MMP (cont.)

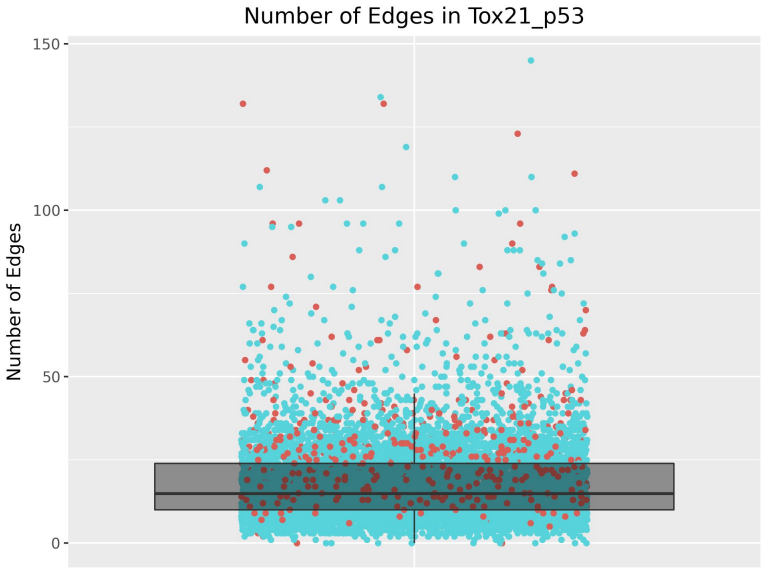

Density Plot of Average Node Degrees in Tox21_MMP
Density
Average Node Degree
0
1
2
3
4
5
Type
Anomaly
Normal

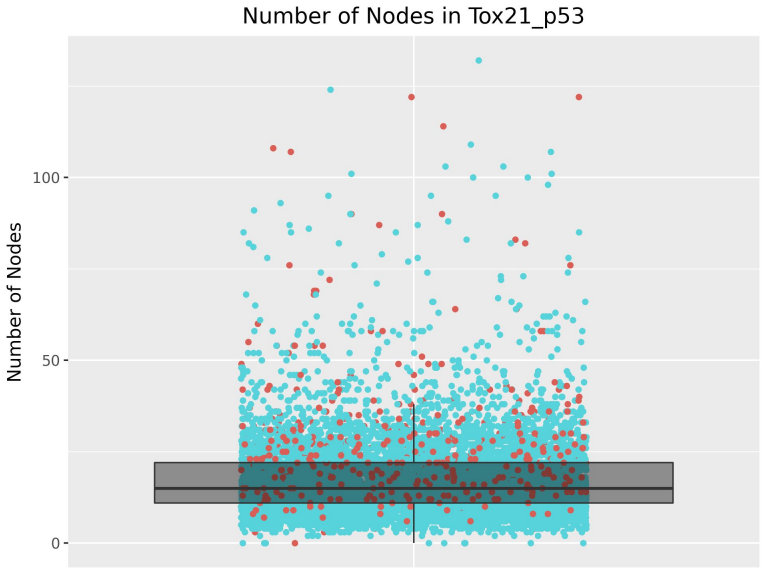
MMP (cont.)


Node Degree in Random Anomaly Graph Tox21_p53
Node Degrees
3.0
2.5
2.0
1.5
1.0

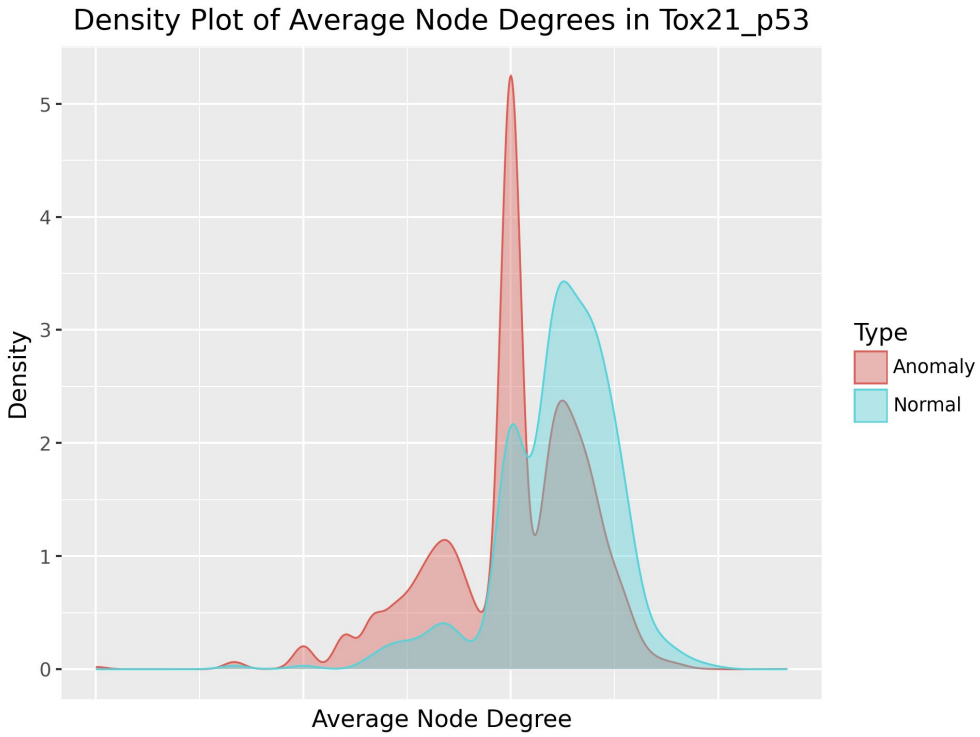

Node Degrees in Random Normal Graph: Tox21_MMP
Node Degrees
6
4
2

p53

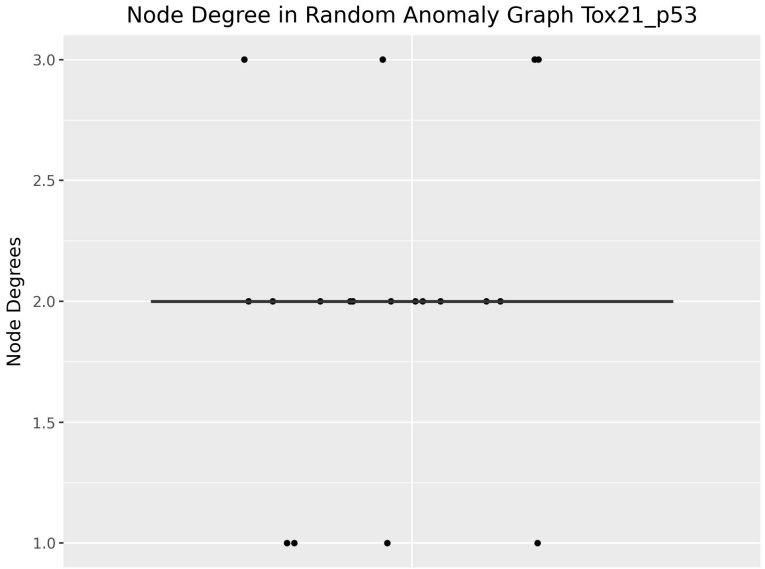

Number of Edges in Tox21_p53
Number of Edges
150
100
50
0

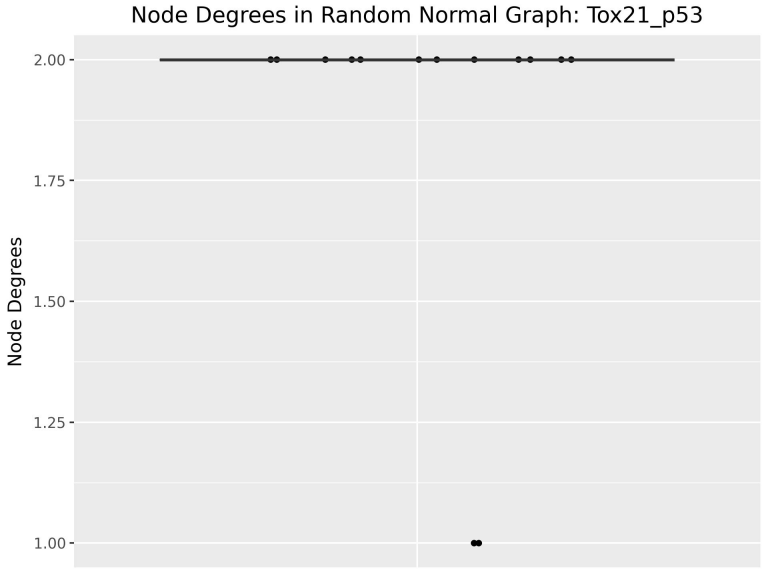

Number of Nodes in Tox21_p53
Number of Nodes
100
50
0

p53 (cont.)

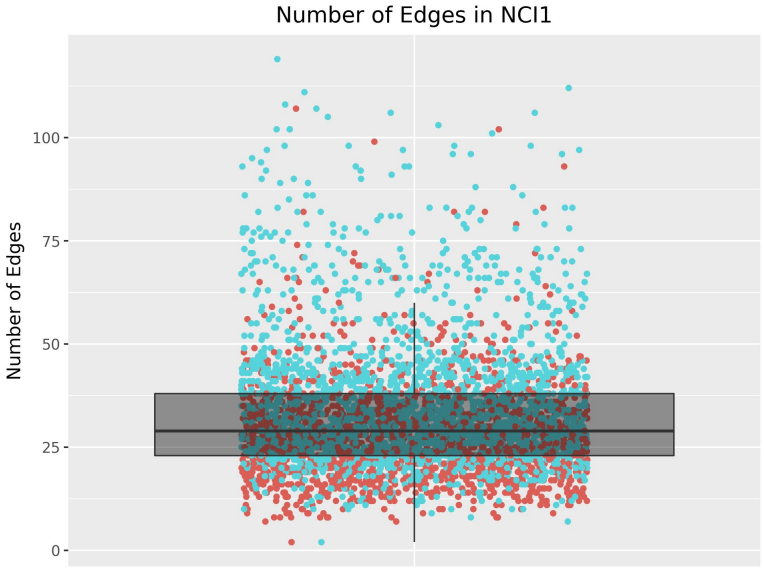

Density Plot of Average Node Degrees in Tox21_p53
Density
0
1
2
3
4
5
Average Node Degree
Type
Anomaly
Normal

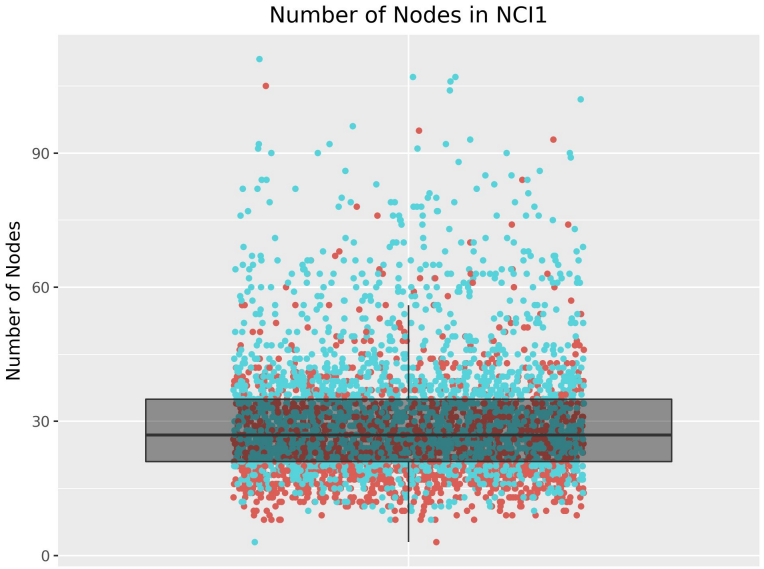
P53 (cont.)


Node Degree in Random Anomaly Graph Tox21_p53
Node Degrees
3.0
2.5
2.0
1.5
1.0

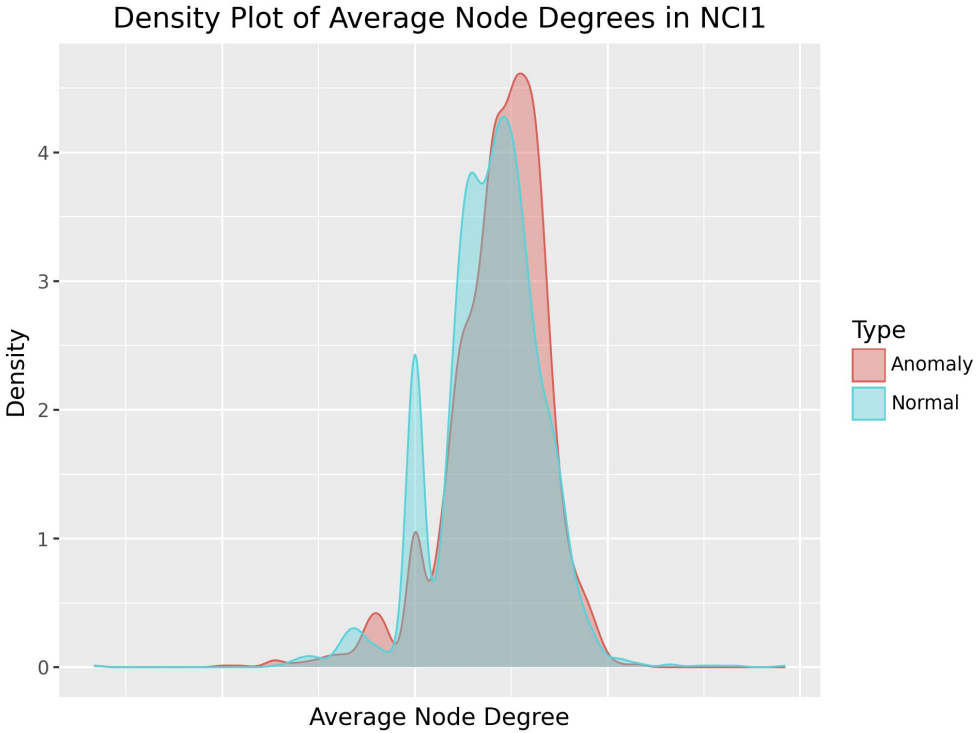

Node Degrees in Random Normal Graph: Tox21_p53
Node Degrees
2.00
1.75
1.50
1.25
1.00

NCI1

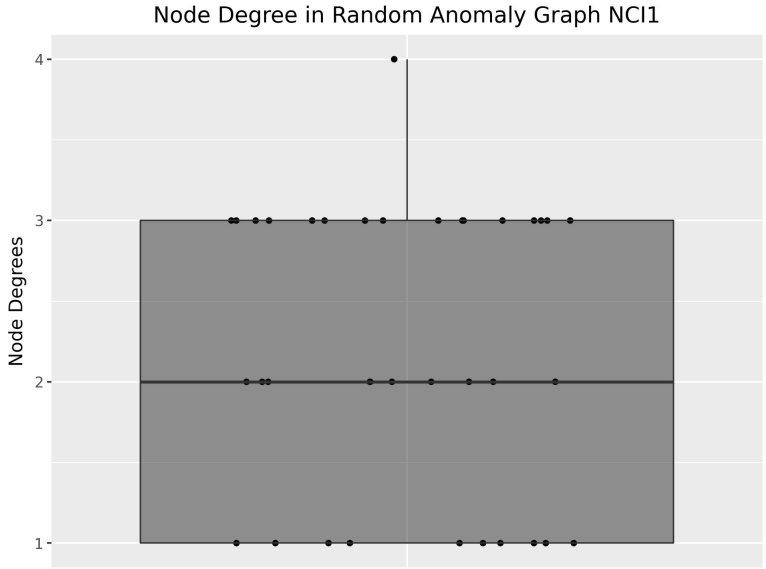

Number of Edges in NCI1
Number of Edges
100
75
50
25
0

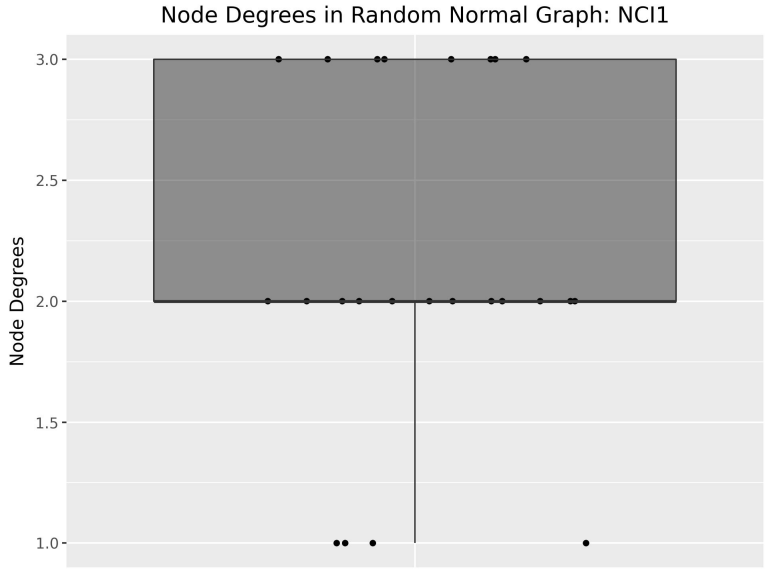

Number of Nodes in NCI1
Number of Nodes
90
60
30
0

p53 (cont.)


Density Plot of Average Node Degrees in NCI1
Density
0
1
2
3
4
Average Node Degree
Type
Anomaly
Normal

NCI1 (cont)


Node Degree in Random Anomaly Graph NCI1
Node Degrees
4
3
2
1


Node Degrees in Random Normal Graph: NCI1
Node Degrees
3.0
2.5
2.0
1.5
1.0

Datasets Analysis

- Edge and node count appear to not vary that much for normal and anomaly graphs.
- Nonetheless…
  - Node degree does appear to play a role.
  - Perhaps this was the key to the naive classifier's 'relative' success (not worse than flipping a coin).

Brainstorming (cont.)

- Not a lot of information about what makes graphs anomalous on each dataset.

Brainstorming (cont.)

- Can we use GANs for the task of generating graphs?
  - Proven to work better than simple encoders and decoders.
- Check:
  - X. Ma et al., "A Comprehensive Survey on Graph Anomaly Detection With Deep Learning," in IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 12, pp. 12012-12038, 1 Dec. 2023, doi: 10.1109/TKDE.2021.3118815.
  - Only source I found of people using GANs (not really related to GAD)
    - Zheng, P., Yuan, S., Wu, X., Li, J., & Lu, A. (2018). One-Class Adversarial Nets for Fraud Detection. https://doi.org/10.48550/arxiv.1803.01798