Paper worksheet (highlights)

DropGNN: Random Dropouts Increase the Expressiveness of Graph Neural Networks

DropGNN: Papp PA, Martinkus K, Faber L, Wattenhofer R. DropGNN: Random dropouts increase the expressiveness of graph neural networks. Advances in Neural Information Processing Systems. 2021

(DropGNN, NeurIPS 2021)

DropGNN: Random Dropouts Increase the Expressiveness of Graph Neural Networks

Pál András Papp, ETH Zurich, apapp@ethz.ch
Karolis Martinkus, ETH Zurich, martinkus@ethz.ch
Lukas Faber, ETH Zurich, lfaber@ethz.ch
Roger Wattenhofer, ETH Zurich, wattenhofer@ethz.ch

Abstract

This paper studies Dropout Graph Neural Networks (DropGNNs), a new approach that aims to overcome the limitations of standard GNN frameworks. In DropGNNs, we execute multiple runs of a GNN on the input graph, with some of the nodes randomly and independently dropped in each of these runs. Then, we combine the results of these runs to obtain the final result. We prove that DropGNNs can distinguish various graph neighborhoods that cannot be separated by message passing GNNs. We derive theoretical bounds for the number of runs required to ensure a reliable distribution of dropouts, and we prove several properties regarding the expressive capabilities and limits of DropGNNs. We experimentally validate our theoretical findings on expressiveness. Furthermore, we show that DropGNNs perform competitively on established GNN benchmarks.

6 Conclusion

We have introduced a theoretically motivated DropGNN framework, which allows us to easily increase the expressive power of existing message passing GNNs, both in theory and practice. DropGNNs are also competitive with more complex GNN architectures which are specially designed to have high expressive power but have high computational complexity. In contrast, our framework allows for an arbitrary trade-off between expressiveness and computational complexity by choosing the number of rounds r accordingly.

Societal Impact. In summary, we proposed a model-agnostic architecture improvement for GNNs. We do not strive to solve a particular problem but to enhance the GNN toolbox. Therefore, we do not see an immediate impact on society. We found in our experiments that DropGNN works best on graphs with smaller degrees, such as molecular graphs. Therefore, we imagine that using DropGNN in these scenarios is interesting to explore further.

Figures

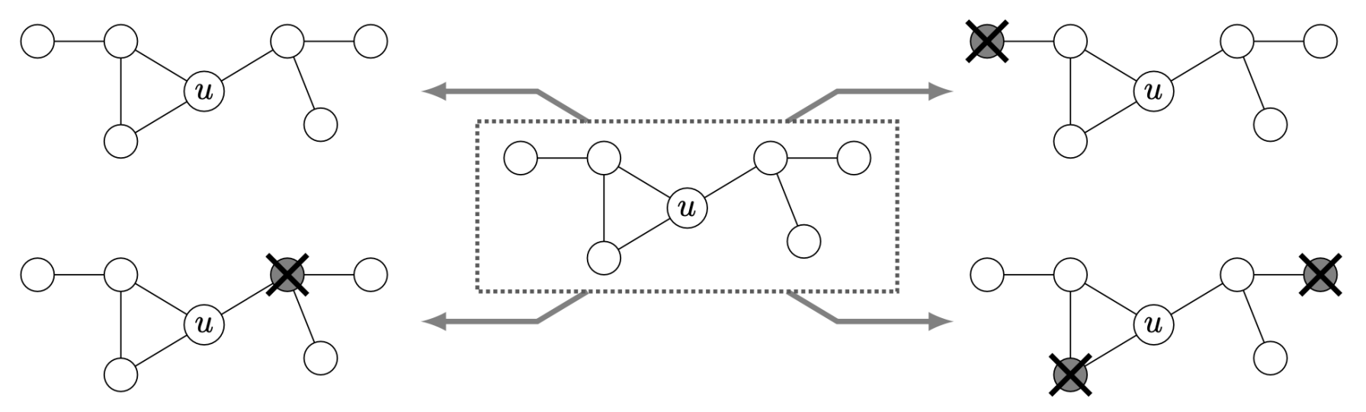


Figure 1: Illustration of 4 possible dropout combinations from an example 2-hop neighborhood around u: a 0-dropout, two different 1-dropouts and a 2-dropout.

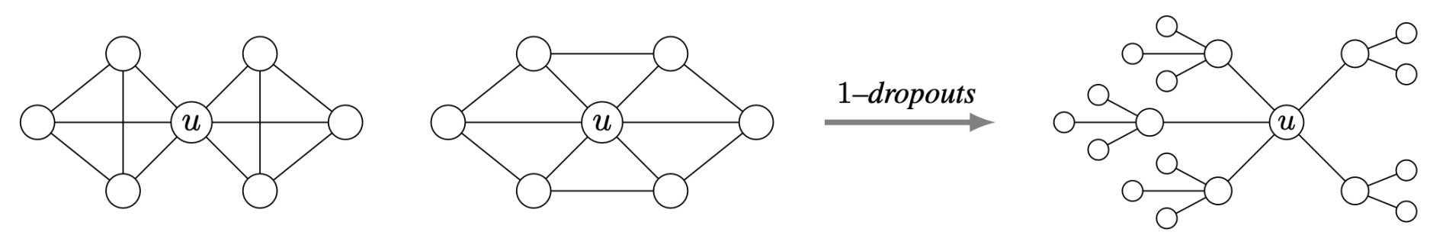


Figure 3: Example of two graphs not separable by 1-dropouts (left side). In both of the graphs, for any of the 1-dropouts, u observes the same tree structure for $d = 2$, shown on the right side.

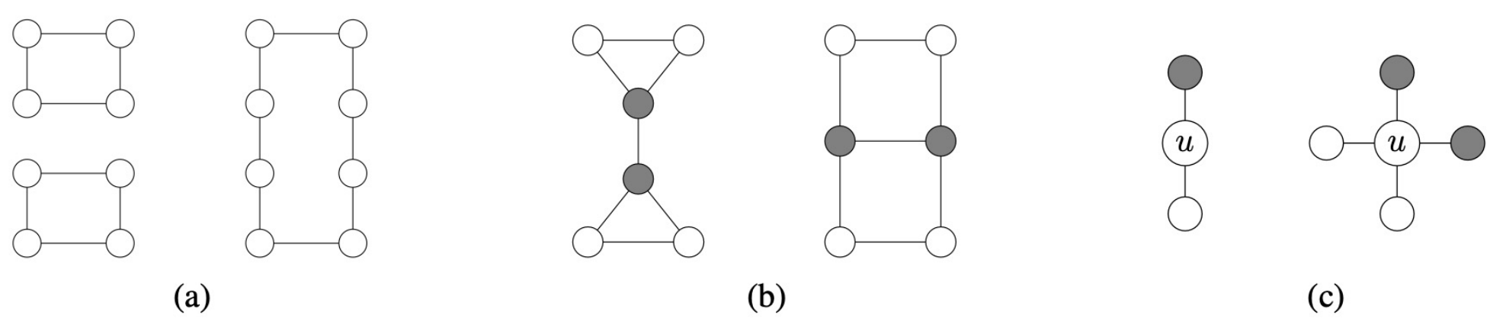


Figure 2: Several example graphs which show that DropGNNs are more expressive than standard GNNs in various cases. Different node colors correspond to different node features.

Theory

Theorem 3 *There exists a pair of neighborhoods that cannot be distinguished by* 1- *and* 2-*dropouts.*

Theorem 4 *In the setting of Theorem 1, a DropGNN with port numbers can distinguish any two non-isomorphic d-hop neighborhoods.*

Aggregation with `mean`. We have seen in Section 3.4 that even with `mean` aggregation, DropGNNs can sometimes distinguish 1-hop neighborhoods (that is, multisets S_1 and S_2 of features) which look identical to a standard GNN. One can also prove in general that a similar separation is possible in various cases, e.g. whenever the two multisets have the same size.

Lemma 1 *Let $S_1 \neq S_2$ be two multisets of feature vectors with $|S_1| = |S_2|$. Then S_1 and S_2 can be distinguished by a DropGNN with `mean` neighborhood aggregation.*

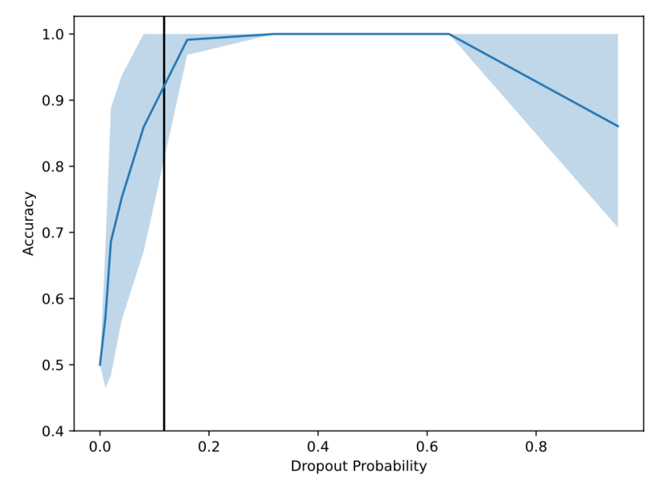
However, in the general case, `mean` aggregation does not allow us to separate any two multisets based on 1-dropouts. In particular, in Appendix C, we also describe an example of multisets $S_1 \cap S_2 = \emptyset$ where the distribution of means obtained from 0- and 1-dropouts is essentially identical in S_1 and S_2. This implies that if we want to distinguish these multisets S_1 and S_2, then the best we can hope for is a more complex approach based on multiple-node dropouts.

Results

5.3 Graph classification

Model	Complexity	MUTAG	PTC	PROTEINS	IMDB-B	IMDB-M
WL subtree [37; 31]	$O(n)$	90.4 ±5.7	59.9 ±4.3	75.0 ±3.1	73.8 ±3.9	50.9 ±3.8
DCNN [2]	$O(n)$	-	-	61.3 ±1.6	49.1 ±1.4	33.5 ±1.4
PatchySan [23]	$O(n)$	89.0 ±4.4	62.3 ±5.7	75.0 ±2.5	71.0 ±2.3	45.2 ±2.8
DGCNN [39]	$O(n)$	85.8 ±1.7	58.6 ±2.5	75.5 ±0.9	70.0 ±0.9	47.8 ±0.9
GIN [37]	$O(n)$	89.4 ±5.6	64.6 ±7.0	76.2 ±2.8	75.1 ±5.1	**52.3 ±2.8**
DropGIN (ours)	$O(rn), r \approx 20$	**90.4 ±7.0**	**66.3 ±8.6**	**76.3 ±6.1**	**75.7 ±4.2**	51.4 ±2.8
1-2-3 GNN [22]	$O(n^4)$	86.1	60.9	75.5	**74.2**	49.5
PPGN [21]*	$O(n^3)$	**90.6 ±8.7**	**66.2 ±6.5**	**77.2 ±4.7**	73 ±5.8	**50.5 ±3.6**

Table 2: Graph classification accuracy (%). The best performing model in each complexity class is highlighted in bold. *We report the best result achieved by either of the three versions of their model.



(b) 4—CYCLES

Write a summary of the paper using your own words

List the new discoveries/conclusions

1.

a) Why the authors included these 3 figures?
b) Can you think of another scenario in Fig 3 where 1-dropout leads to the same computational graph?
c) Can you think of other figures that could be included?

Paper outline (backbone)

a) What you do you like about the outline and why?
b) Which sections you find most interesting and why?

	1	2	3	4
C				
I				
S				
L				

C: clarity (level of clarity)
I: technical/theoretical innovation (level of innovation)
S: simplicity (simple idea)
L: logic (well presented)

1. What is the problem to solve?

2. What are the novel contributions of the paper?
3. What is the logic/rationale behind the new ideas?

4. Experimental design

5. What are the limitations of the proposed method?

Illustrate the theorems where possible and write your reflections

Analyze the plot figure

https://basira-lab.com/

YouTube

Search "BASIRA Lab"

https://github.com/basiralab/DGL