

Homework 5: Message Passing Neural Networks

CS 1470/2470

Due November 13, 2019 at 11:59pm

1 Conceptual Questions

1. What is the benefit of using a Message Passing Neural Network over a vector representation of the graphs (e.g stacking all the features of the nodes/edges together indiscriminately) and a fully connected network? (3-4 sentences)
2. Describe how a message passing update is a generalization of a convolution. (2-3 sentences)
3. The purpose of the lifting function (known as `f_init` in the slides) is to transform the vertex/edge data into a new space. Why is operating in this space better than feeding raw features directly into the message passing network? (4-6 sentences)
4. As mentioned in lecture, batching with an MPNN can be quite difficult. This is primarily due to the fact that you would need to trigger message passing across a number of graphs with varying numbers of nodes and edges. Still, even when we pad our graphs to all have the same number of nodes, a framework like DGL can only process a single graph at a time. Nevertheless, they still support batching. Give an explanation of how this could be done, and why it works (Hint: Think about the graphs in your batch as adjacency matrices). (3-5 sentences)

2 Ethical Implications

While this assignment focuses on molecular structure, it turns out that graphs are great representations for social networks, which makes social networks a very relevant application of deep learning on graphs (e.g. [predicting influence](#) or [recommendations](#)). (If you're curious, feel free to read about a variety of other cool applications [here!](#))

Suppose you're working on such a system: given a social network represented as a graph, your task is to implement a neural net such that you can answer questions like: Who is likely to buy a product? Who are "influential" people in this network? What kind of messaging are parts of this network receptive to?

1. Why might an algorithm like this be useful? (1-2 sentences)
2. In what ways might this algorithm be dangerous? What kinds of questions might be asked—and by whom—that may be harmful, and what kind of harm could be caused? (2-3 sentences)
3. Are there any ways to mitigate some of these harms, at least partially? (2-3 sentences)
4. For you, personally, are there any boundaries you might draw in terms of what kinds of questions you're willing to ask (or answer)?

3 CS2470-only Questions

1. Our readout layer is essentially taking every node vector, feeding it through a linear layer, summing the result across all nodes, and then returning it as the "representation" of the graph. However, some researchers have criticized global operations like these (e.g taking the max/min/mean across the graph) because your network just spent time learning local relationships through message passing, only for this to be "thrown away" at the global level. In lecture, we have explored some ideas about this ought to be done.

Consider the following algorithm for a more iterative readout function: After computing the state at each node through rounds of message passing, we cluster these nodes using a clustering algorithm like K-means. Then, we turn every resulting cluster into a single output node by averaging over the state of each node in the cluster.

We posit that this scheme will likely not work. Why is that? And how could you fix this?

(Hint: You may find [this paper](#) helpful in formulating your answer:)

2. Imagine that you have a graph G with two nodes H_1 and H_2 as follows. If an MPNN is run on this graph and only classifies based on the structure of the graph (not the node features), will H_1 and H_2 be classified similarly or differently? Is this a disadvantage or an advantage for graph/node classification as a whole? Justify your answers.

