Homework 5: Message Passing Neural Networks

CS 1470/2470

Due November 9, 2020 at 11:59pm AoE

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. How is the lifting function similar to an embedding matrix and why does improve performance compared to using raw node features? (3-5 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 multiple graphs with varying numbers of nodes/edges simultaneously. Even though frameworks like DGL can only process a single graph at a time, they are able to support batching of arbitrary numbers of graphs. How is this possible? (Hint: Think about the graphs in your batch as adjacency matrices). (3-5 sentences)

2   Ethical Implications

One practice of using graph neural networks for social good is fake news detection on Twitter. As delineated in a July 2020 publication here, without relying on text information, GNNs have the capability to achieve comparable or superior outcomes to state-of-the-art methods when differentiating real and fake news. Furthermore, despite GNNs' tendency to perform poorly on unseen data, there exists a solution to attain balanced performance on both new and existing datasets.

1. Although intentional misinformation via fake news may be reduced by the network described above, what are some potential unintended adverse outcomes? (3-7 sentences)

2. What are some possible biases that these GNNs could inherit from their data sources? (3-7 sentences)

3. Do you believe that a (mostly) automated system can be implemented to fairly identify fake news and solve an issue present in society today? Please specify. (4 sentences minimum)

4. Similar automated algorithms have been used by Youtube to identify and label videos that may contain conspiracy theories, which prompts an informative banner below the video to link to a reliable source. Would a similar implementation of an imperfect fake news detection algorithm be justified given the damaging effects of unregulated spread?

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 how 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 classifcation as a whole? Justify your answers.