Martin BRICENO
jbriceno@telecom-paris.fr

Lab session # 6
ALTEGRAD 2020

02/02/21

1 Question 1

In the lab we implemented a version of Graph autoencoder with a loss of type "MSE" and with an activation function in the final layer of type "identity". In this way, we treat the problem as if it were a regression in which the element $\hat{A}_{ij}$ must be approximated to the element of the same position A_{ij}.

Following the same idea we could see this problem as a classification problem where $\hat{A}$ is obtained applying the sigmoid's activation function to each element of the last layer of the GAE model, this is a element wise operation, so:

$$\hat{A} = \sigma(ZZ^\top) \tag{1}$$

And thus we would have n^2 classifications where each element of A is the true label of $\hat{A}$ and hence our loss would now be the total sum of BCE (binary cross entropy loss) between each pair of elements, thus:

$$loss = \frac{1}{n^2} \sum_{i,j} BCE(\hat{A}_{i,j}, A_{i,j}) \tag{2}$$

In fact we could see this problem as a multi label problem since each classification is independent and the matrix A has several ones.

2 Question 2

Based on the Bi-decoder model of article [1], we see that for a convolutional graph autoencoder, with M layers between encoder and decoder, it is possible to reconstruct the feature matrix as well as the adjacency matrix. In our case we do not perform convolutions on the learning layers, but the idea is the same.

So the encoder will remain the same, i.e. the output is Z, while **the decoder will be constituted of two parts: the first will be the decoder that we implemented in the laboratory to reconstruct A, and the second will be a decoder that will be implemented as the inverse (mirror) process of the encoder**, i.e. if the linear layer of the encoder has dimension (N1,N2), then the linear layer of the decoder will have dimension (N2,N1), with the same amount of hidden layers. In the final layer of this decoder we will be able to apply an identity function with which we will reconstruct the features matrix.

$$H = ReLU(\tilde{A}ZW^2) \quad where \quad W^2 = dim(h2, h1) \tag{3}$$

$$\hat{X} = \tilde{A}HW^3 \quad where \quad W^3 = dim(h1, n\_feat) \tag{4}$$

Finally the loss of our model will also include the loss of this second decoder, which can be the MSE since the feature matrix is a real-valued matrix or it can also be the frobenius norm as proposed in the article.

$$loss_X = MSE(X, \hat{X}) \quad or \quad \left\| X - \hat{X} \right\|_F^2 \tag{5}$$

then:

$$total\_loss = loss_A + loss_X \tag{6}$$

3 Question 3

Z_{G_1} **takes the first 3 rows,** Z_{G_2} **takes the next 4 rows and** Z_{G_3} **takes the last 2 rows of** Z.

- Sum

$$Z_{G_1} = sum\left(\begin{bmatrix}0.77 & -1.26 & -0.63\\ 1.15 & -1.90 & -0.94\\ 0.77 & -1.26 & -0.63\end{bmatrix}, axis = 0\right) = \begin{bmatrix}2.69 & -4.42 & -2.2\end{bmatrix} \tag{7}$$

$$Z_{G_2} = sum\left(\begin{bmatrix}1.15 & -1.90 & -0.94\\ 1.15 & -1.90 & -0.94\\ 1.15 & -1.90 & -0.94\\ 1.15 & -1.90 & -0.94\end{bmatrix}, axis = 0\right) = \begin{bmatrix}4.6 & -7.6 & -3.76\end{bmatrix} \tag{8}$$

$$Z_{G_3} = sum\left(\begin{bmatrix}0.77 & -1.26 & -0.63\\ 0.77 & -1.26 & -0.63\end{bmatrix}, axis = 0\right) = \begin{bmatrix}1.54 & -2.52 & -1.26\end{bmatrix} \tag{9}$$

- Mean

$$Z_{G_1} = mean\left(\begin{bmatrix}0.77 & -1.26 & -0.63\\ 1.15 & -1.90 & -0.94\\ 0.77 & -1.26 & -0.63\end{bmatrix}, axis = 0\right) = \begin{bmatrix}0.90 & -1.47 & -0.73\end{bmatrix} \tag{10}$$

$$Z_{G_2} = mean\left(\begin{bmatrix}1.15 & -1.90 & -0.94\\ 1.15 & -1.90 & -0.94\\ 1.15 & -1.90 & -0.94\\ 1.15 & -1.90 & -0.94\end{bmatrix}, axis = 0\right) = \begin{bmatrix}1.15 & -1.90 & -0.94\end{bmatrix} \tag{11}$$

$$Z_{G_3} = mean\left(\begin{bmatrix}0.77 & -1.26 & -0.63\\ 0.77 & -1.26 & -0.63\end{bmatrix}, axis = 0\right) = \begin{bmatrix}0.77 & -1.26 & -0.63\end{bmatrix} \tag{12}$$

- Max

$$Z_{G_1} = max\left(\begin{bmatrix}0.77 & -1.26 & -0.63\\ 1.15 & -1.90 & -0.94\\ 0.77 & -1.26 & -0.63\end{bmatrix}, axis = 0\right) = \begin{bmatrix}1.15 & -1.26 & -0.63\end{bmatrix} \tag{13}$$

$$Z_{G_2} = max\left(\begin{bmatrix}1.15 & -1.90 & -0.94\\ 1.15 & -1.90 & -0.94\\ 1.15 & -1.90 & -0.94\\ 1.15 & -1.90 & -0.94\end{bmatrix}, axis = 0\right) = \begin{bmatrix}1.15 & -1.90 & -0.94\end{bmatrix} \tag{14}$$

$$Z_{G_3} = max\left(\begin{bmatrix}0.77 & -1.26 & -0.63\\ 0.77 & -1.26 & -0.63\end{bmatrix}, axis = 0\right) = \begin{bmatrix}0.77 & -1.26 & -0.63\end{bmatrix} \tag{15}$$

4 Question 4

Based on the Z matrix of the previous problem, we see that the nodes that have the same number of edges in each graph also have the same representation. Thus the node with 2 edges has the same representation z = [1.15 -1.90 -0.94] in each graph.

In our case we see that G1 and G2 are not equal but they are constituted entirely of nodes with 2 edges and with the same amount of them, therefore no matter what READOUT function we use, **the representation of both graphs Z_{G_1} and Z_{G_2} will be the same**.

References

[1] Yunxing Zhang Xuelong Li, Rui Zhang. Graph convolutional auto-encoder with bi-decoder and adaptive-sharing adjacency. In *arXiv:2003.04508v2 [cs.LG]*, 2020.