Question 1

An LSTM unit is design for sequential data. LSTM is therefore not a permutation invariant model and is poorly adapted for data structure like set or graph.

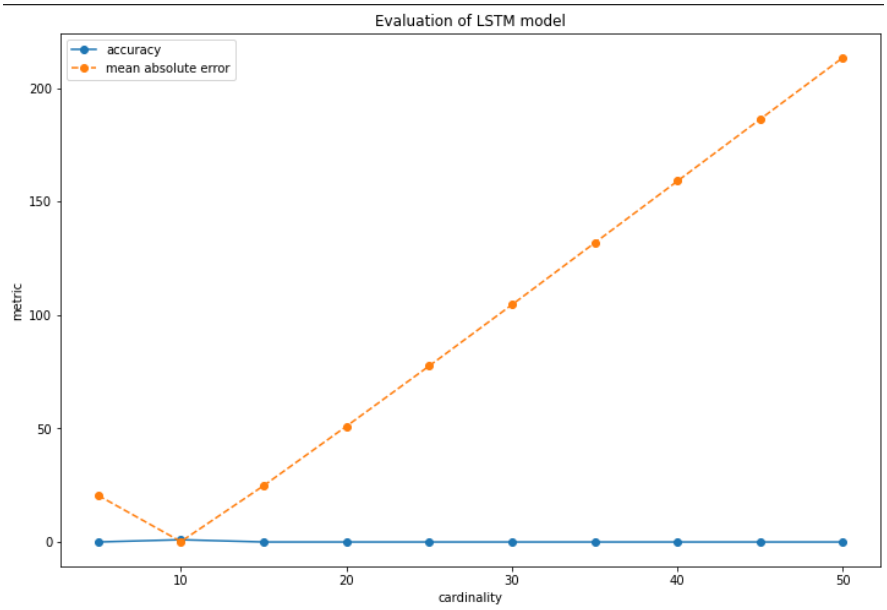

Figure 1: Evaluation of LSTM model

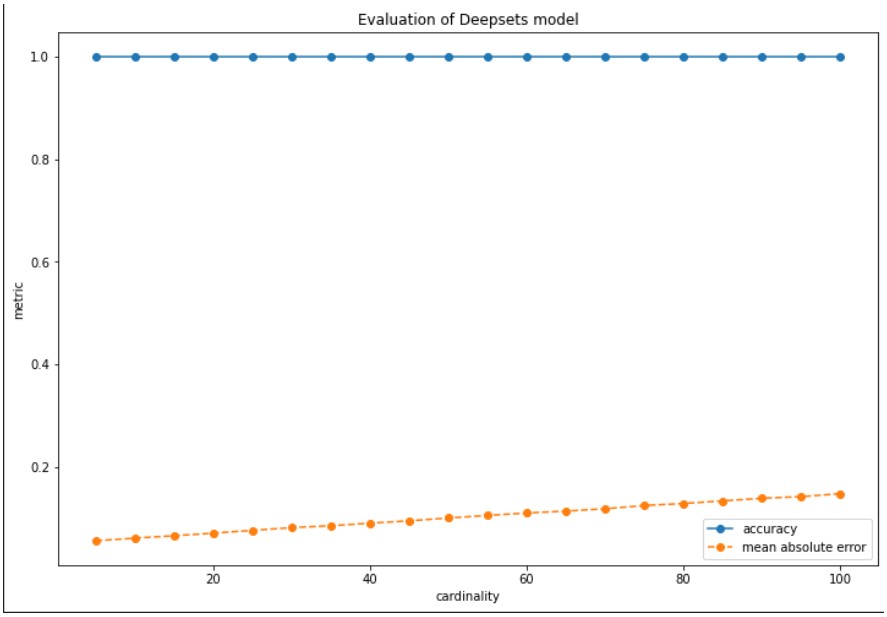

Figure 2: Evaluation od Deepsets model

As expected, LSTM is a poor architecture to deal with data structures like sets.

Question 2

If we consider a set as a graph without edges, then its adjacency matrix is $A = 0$ and we feed $\widetilde{A} = I_n$ and X the vector of the labels of the nodes in a GNN. Then the GNN just learn an embedding for the set like in Deepsets.

Question 3

1. An edge probability matrice P which induces an homophilic cluster structure would be $P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

   An edge probability matrice P which induces an heterophilic cluster structure would be $P = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & /2 \end{pmatrix}$.

2. Inside the block B_i, there are at most $\binom{|B_i|}{2} = \binom{5}{2} = 10$ edges. It gives us the expected number of edges inside this block : $10 \times P_{i,i} = 10 \times 0.8 = 8$ edges.
   Between two blocks B_i and B_j, there are at most $|B_i| \times |B_j| = 5 \times 5 = 25$ edges. It gives us the expected number of edges inside this block : $25 \times 0.05 = 1.25$ edges.

Question 4

If the graph has weights, we can use the generalization of binary cross entropy called cross entropy loss.

References