1 Question 1

To do so, we store v_{i-1} when computing v_{i+1} and make sure v_{i-1} is not selected when choosing a random neighbour of v_i. The remark was implemented in the `random_walk` function.

2 Question 2

Similarly to when handling text data, a network outputting word-level labels can be modified to output class-level labels by adding a pooling layer. If the data is of the form (batch_size $\times n_{\text{edges}} \times n_{\text{features}}$), then the pooling has to be done on the first axis.

3 Question 3

The network reaches an accuracy of 1, which is the same performance as for the DeepWalk approach. My explanation is that the problem is not that hard and therefore any cleverly designed architecture can fully capture the information available. Provided the nature of the problem, message passing is particularly adapted, as it is a good way to model influences over a group of people.

4 Question 4

The GNN does not perform wall anymore, scoring 0.28. The chosen features does not permit the GNN to differentiate between the nodes, therefore the model loses all its interest.