ABBOUDI Mohammed Amine
mohammed-
amine.abboudi@polytechnique.edu

1 Question 1

It is possible to modify the random walk model so that it becomes impossible to walk immediately back into the same node by adding a memory variable that keeps track of the current node right after sample the next node. The variable, denoted by c, is initialized with $None$. Then the sampling space at each position in the walk becomes $\Omega = \{neighbors(v_i^{(t)}) \setminus \{c\}\}$. After sampling uniformly from the Ω, c is set to the chosen node.

2 Question 2

Much like the case of word/document embedding, it is possible to extend the implemented architecture to embed and classify graphs. The point is to add another matrix to train, dubbed the **Graph Matrix** which lives in $\mathcal{R}^{\mathcal{N}\times\mathcal{D}}$ where N is the number of documents and D is the graph feature vector dimension. Each row in the matrix is only used when looking at the corresponding graph, while node vectors (rows of **X** and W_0) are shared across all graphs. The task is therefore to predict the class of the graph given its context and the corresponding graph vector. The training of both sets of vectors is done by stochastic gradient descent and the gradient is obtained via backpropagation.
This would allow to train feature vector representations for both graphs and nodes.

3 Question 3

The GNN did not outperform the DeepWalk + logistic regression approach in the node classification task. Accuracy of 85.7% for DW + LR vs. 28.57% for GNN. The issue seems to come from the initialization of the feature vectors, where all nodes are given the same initial feature vector.

4 Question 4

The performance of the GNN after changing the initial feature vectors matrix $\mathbf{X}$ has maximally improved, reaching 100% accuracy after a few epochs of training. Giving each node a unique starting feature vector representation has allowed the matrix to update differently over each SGD backpropagation and has therefore caused it to converge towards a more representative ”embedding”.

References