CS224W-Fall 2021 Homework 1 Solutions

Tianjian Li

February 2022

1 Link Analysis

1.1 Personalized PageRank I

We first calculate the PageRank vector for each page $i \in \{1,2,3,4,5\}$.

$$v_1 = \frac{1}{3}v_a + \frac{1}{3}v_c + v_d$$

$$v_2 = \frac{1}{3}v_a$$

$$v_3 = \frac{1}{3}v_a + \frac{1}{3}v_b$$

$$v_4 = \frac{1}{3}v_b + \frac{1}{3}v_c$$

$$v_5 = \frac{1}{3}v_b + \frac{1}{3}v_c$$

Therefore the Personalized PageRank of Eloise, whose interests are represented by the teleport set $\{2\}$, is $v_e = v_2 = \frac{1}{3}v_a$

1.2

$v_f = v_5 = \frac{1}{3}v_b + \frac{1}{3}v_c$

1.3

$v_g = 0.1v_1 + 0.2v_2 + 0.3v_3 + 0.2v_4 + 0.2v_5 = \frac{1}{30}v_a + \frac{1}{30}v_c + \frac{1}{10}v_d + \frac{1}{15}v_a + \frac{1}{10}v_a + \frac{1}{10}v_b + \frac{1}{15}v_b + \frac{1}{15}v_c + \frac{1}{15}v_b + \frac{1}{15}v_c = \frac{1}{5}v_a + \frac{7}{30}v_b + \frac{1}{6}v_c + \frac{1}{10}v_d$

1.4

As we have shown above, we can compute the personalized PageRank of the items that belongs to the user's teleport set.

1.5

Since $\sum_{i=1}^{N} r_i = 1$ we have $\mathbb{1}^\top r = 1$ Therefore:

$$r = Ar$$

$$r = (\beta M + \frac{1-\beta}{N}\mathbb{1}\mathbb{1}^\top)r$$

$$r = \beta Mr + \frac{1-\beta}{N}\mathbb{1}(\mathbb{1}^\top r)$$

$$r = \beta Mr + \frac{1-\beta}{N}\mathbb{1}$$

2 Relational Classification I

After first iteration: $v_1 = [\frac{1}{2}, \frac{1}{2}], v_2 = [\frac{3}{8}, \frac{5}{8}], v_3 = [\frac{5}{8}, \frac{3}{8}], v_4 = [\frac{15}{32}, \frac{17}{32}], v_8 = [\frac{31}{64}, \frac{33}{64}], v_9 = [\frac{31}{128}, \frac{97}{128}]$

After second iteration: $v_1 = [\frac{1}{2}, \frac{1}{2}], v_2 = [\frac{51}{128}, \frac{77}{128}], v_3 = [\frac{81}{128}, \frac{47}{128}], v_4 = [\frac{241}{512}, \frac{271}{512}], v_8 = [\frac{365}{1024}, \frac{659}{1024}], v_9 = [\frac{365}{2048}, \frac{1683}{2048}]$

2.1

$$P(Y_3 = +) = \frac{81}{128}$$

2.2

$$P(Y_4 = +) = \frac{241}{512}$$

2.3

$$P(Y_8 = +) = \frac{365}{1024}$$

2.4

Vertices 1, 2, 4, 8, 9 would be classified to ”-” after the second iteration.

3 Relational Classification II

3.1

The initial Node - Label pairs are Node1 - 0, Node2 - 1, Node3 - 1, Node4 - 1, Node5 - 1, Node6 - 0, Node7 - 0.

3.2

The link vectors for each node are LinkV1 = (0, 1), LinkV2 = (1, 0), LinkV3 = (1, 1), LinkV4 = (1, 0), LinkV5 = (1, 1), LinkV6 = (1, 1), LinkV7 = (1, 1)

After the first iteration, the predicted labels are Node1 - 0, Node2 - 0, Node3 - 0, Node4 - 0, Node5 - 0, Node6 - 0, Node7 - 0.

3.3

After the second iteration, the link vector for each node are the same (1, 0)
The predicted labels are all 0.

3.4

The labels assigned by the iterative classification algorithm have converged.

4 GNN Expressiveness

4.1

The two nodes would have different representations after 3 iterations(layers) of message passing. The red node in the first graph would have representation 18, and the red node in the second graph would have representation 17.

4.2

The random walk matrix M is defined as follow

$$\text{M} = \begin{bmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{bmatrix}$$

Where the ith row corresponds with node i in the graph.
The eigenvector has an eigenvalue of 1 is $[\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}]$

4.3

The transition matrix $M = D^{-1}A$

4.4

The new transition matrix $\tilde{M} = I + D^{-1}A$ where I is the identity matrix with same dimensions as A.

4.5

A non rigorous proof:
Since the graph is connected and has no bipartite components, the transtion matrix $M = D^{-1}A$ has eigenvalue 1 with eigenvector r. After a sufficient amount of iterations, the node features would converge to r and applying the transition matrix M to it does not change since $r = Mr$.

4.6

The aggregation function for Breadth-First-Search is defined as follow

$$h_i^{(l+1)} = \begin{cases} 0 & \sum_{j \in \mathcal{N}_i} h_j^{(l)} \\ 1 & \text{else} \end{cases}$$

5 Node Embedding and its relation to Matrix Factorization

5.1

Here the decoder is the inner product of the encoded vectors $z_u^T z_v$

5.2

The minimization objective becomes $\min_z ||A - Z^T W Z||_2$

5.3

(I am not sure of my answer right here.) If we want the matrix factorization of $\min_z ||A - Z^T W Z||_2$ equivalent to the learning of the eigendecomposition of adjacent matrix A, we need the weight matrix W to be an diagonal matrix with A's eigenvalues on its diagonal entries.

5.4

Let A' be the matrix where $A'_{ij} = 1$ if i and j are connected or i is j's two hop neighbor. Then the matrix factorization problem would be

$$\min_z ||A - Z^T Z||_2$$

5.5

The nodes in the left clique would have a representation, whereas the nodes in the right clique would have an entirely different representation, even through they are structurally the same.

5.6

In node2vec, you can arrive at any node within that clique with same probability.

In struct2vec, you can also arrive at any node within the second clique. But the probability for the green node is not the same as the probability for the blue node.

5.7

Because different g_k capture different types of structrual information.

5.8

All of the nodes with same degrees are represented with the same color(shown in the picture in blue), whereas the node with an additional degree have a different representation(shown in green).