Question 3

Use PageRank to find the most influenced vertex of the directed graph defined by the adjacency matrix $\begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix}$. The probabilities are equal. Note that the adjacency matrix is not the matrix that was considered in the videos.

1) First we need to transform our adjacency matrix into a Markov transition matrix P, where each column sums up to 1.

$$P = \begin{bmatrix} 0 & 0 & 1/2 & 1/3 \\ 0 & 1 & 0 & 1/3 \\ 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 1/3 \end{bmatrix} \tag{1}$$

According to Frobenius theorem for any Markov transition matrix P has the following properties:
1) $\lambda = 1$ is an eigenvalue of P,
2) any eigenvalue λ of P satisfies $|\lambda| \leq 1$,
3) there exists an eigenvector g of the eigenvalue 1 all the coordinates of which are greater or equal to 0.

Let us find eigenvalues and corresponding eigenvectors of P :

$$\lambda_1 = 1 \qquad \lambda_2 = 5/6 \qquad \lambda_3 = -1/2 \qquad \lambda_4 = 0$$

$$v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 5/8 \\ -2 \\ 3/8 \\ 1 \end{pmatrix} \qquad v_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad v_1 = \begin{pmatrix} 0 \\ -1/3 \\ -2/3 \\ 1 \end{pmatrix}$$

As it can be clearly seen, our matrix P has all the properties of a Markov transition matrix.
Moreover we can call P regular as its eigenvalue of $\lambda = 1$ has an algebraic multiplicity equal to 1 and all other eigenvalues λ of P have an absolute value less than 1.

2) At the next step we introduce a matrix Q of size $n \times n$ with all entries equal to $\dfrac{1}{n}$. In our case $n = 4$, and the matrix Q looks like:

$$Q = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{bmatrix} \tag{2}$$

3) Now we can replace the matrix P by the following matrix which shall also be a regular one:

$$\begin{gathered} P_\alpha = (1-\alpha)P + \alpha Q \\ \alpha \in [0, 1] \end{gathered} \tag{3}$$

Taking $\alpha = 0.15$ we get the matrix P_α:

$$P_\alpha = (1-0.15)\begin{bmatrix} 0 & 0 & 1/2 & 1/3 \\ 0 & 1 & 0 & 1/3 \\ 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 1/3 \end{bmatrix} + 0.15\begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{bmatrix} =$$

$$= \begin{pmatrix} 0 & 0 & \frac{17}{40} & \frac{17}{60} \\ 0 & \frac{17}{20} & 0 & \frac{17}{60} \\ \frac{17}{40} & 0 & 0 & 0 \\ \frac{17}{40} & 0 & \frac{17}{40} & \frac{17}{60} \end{pmatrix} + \begin{pmatrix} \frac{3}{80} & \frac{3}{80} & \frac{3}{80} & \frac{3}{80} \\ \frac{3}{80} & \frac{3}{80} & \frac{3}{80} & \frac{3}{80} \\ \frac{3}{80} & \frac{3}{80} & \frac{3}{80} & \frac{3}{80} \\ \frac{3}{80} & \frac{3}{80} & \frac{3}{80} & \frac{3}{80} \end{pmatrix} = \begin{pmatrix} \frac{3}{80} & \frac{3}{80} & \frac{37}{80} & \frac{77}{240} \\ \frac{3}{80} & \frac{71}{80} & \frac{3}{80} & \frac{77}{240} \\ \frac{37}{80} & \frac{3}{80} & \frac{3}{80} & \frac{3}{80} \\ \frac{37}{80} & \frac{3}{80} & \frac{37}{80} & \frac{77}{240} \end{pmatrix} \tag{4}$$

4) Knowing that P_α is a regular Markov transition matrix we expect that its eigenvalue of $\lambda = 1$ has an algebraic multiplicity equal to 1 and all other eigenvalues λ of P_α have an absolute value less than 1. Eigenvalues λ and corresponding eigenvectors of P_α confirm that this statement is true:

$$\lambda_1 = 1 \qquad \lambda_2 = 17/24 \qquad \lambda_3 = -17/40 \qquad \lambda_4 = 0$$

$$v_1 = \begin{pmatrix} 40/57 \\ 1669/513 \\ 86/171 \\ 1 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 5/8 \\ -2 \\ 3/8 \\ 1 \end{pmatrix} \qquad v_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad v_1 = \begin{pmatrix} 0 \\ -1/3 \\ -2/3 \\ 1 \end{pmatrix}$$

5) By normalizing obtained vector v_1 corresponding to eigenvalue $\lambda = 1$ we can get vector g which represents the upper bound of vertex influence probability. We call it page rank:

$$\lim_{k \to \infty} (P_\alpha)^k X = g \tag{5}$$

where X is an arbitrary vector such that $X^T = (p_1, \ldots, p_n)$ with $p_i \in [0,1]$ and $\sum_{i=1}^{n} p_i = 1$.

CONCLUSION

For the above calculated matrix $P_{0.15}$ the normalized vector $g = \begin{pmatrix} 0.129 \\ 0.596 \\ 0.092 \\ 0.183 \end{pmatrix}$.

The highest rank has vertex 2, which represents a page with one incoming reference from itself. It can be called a spider trap. Such pages in a loop-refence may get excessive ranking especially in simple models like ours. Similar property has vertex 4, but at least it has 2 incoming refences from other pages in addition to a loop self-refence.
The lowest rank has vertex 3, which represents a page with inly one incoming reference.

For matrix $P_{0.5}$ the normalized vector $g = \begin{pmatrix} 0.214 \\ 0.339 \\ 0.179 \\ 0.268 \end{pmatrix}$

For matrix $P_{0.85}$ the normalized vector $g = \begin{pmatrix} 0.243 \\ 0.265 \\ 0.231 \\ 0.261 \end{pmatrix}$

ADDITIONALLY
We may consider an adapted adjacency matrix where vertex 2 still has one reference but now to vertex, say, 3, rather than to itself. As for the vertex 4 the self-reference is just cleaned up.
The Markov transition matrix P has now the following look:

$$P = \begin{bmatrix} 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1/2 \\ 1/2 & 1 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \end{bmatrix}$$

In this case for $P_{0.15}$ the normalized vector $g = \begin{pmatrix} 0.278 \\ 0.156 \\ 0.288 \\ 0.278 \end{pmatrix}$.

Now we can clearly see the actual rank of the page represented by vertex 2.