**WEB MINING**

**HARSH TRIPATHI**

**16BCE1186**

**PAGERANK ALGORITHM**

**d = 0.85**

**size = outlinks.shape[0]**

**page\_ranks = [1 for i in range(size)]**

**out\_degrees =[ ]**

**for i in range(size):**

**sums = 0**

**for j in range(size):**

**sums += outlinks[i][j]**

**out\_degrees.append(sums)**

**print('Initial page ranks:')**

**print(page\_ranks)**

**for \_ in range(100):**

**for j in range(size):**

**temp = 0**

**for i in range(size):**

**if outlinks[i][j] == 1:**

**temp += page\_ranks[i] / out\_degrees[i]**

**temp \*= d**

**temp += (1-d)**

**page\_ranks[j] = round(temp, 4)**

**return page\_ranks**

**outlinks = [0, 1, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0]**

**outlinks = np.reshape(outlinks, (5, 5))**

**page\_ranks = calculate\_PageRank(outlinks)**

**print()**

**print('The converged page rank is:')**

**print(page\_ranks)**

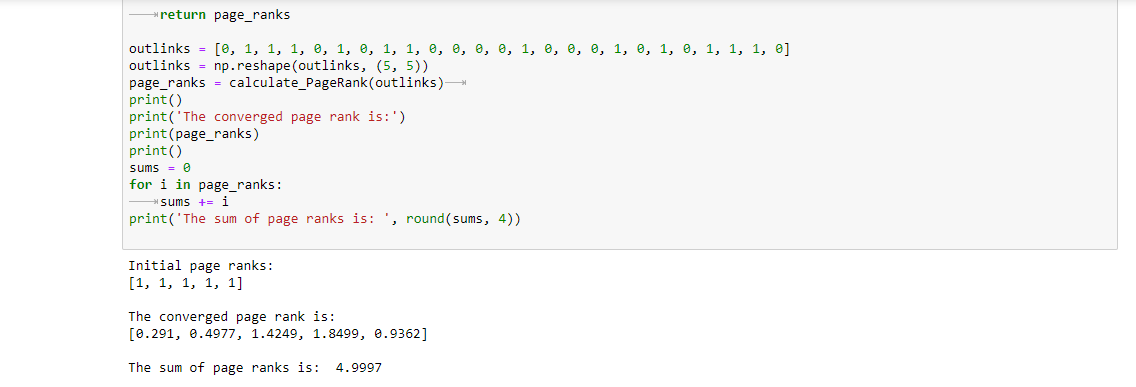
**print()**

**sums = 0**

**for i in page\_ranks:**

**sums += i**

**print('The sum of page ranks is: ', round(sums, 4))**

****

**HITS ALGORITHM**

**import numpy as np**

**size = outlinks.shape[0]**

**hub\_scores = [1.0 for i in range(size)]**

**authority\_scores = [1.0 for i in range(size)]**

**print(hub\_scores)**

**for \_ in range(100):**

**for j in range(size):**

**temp\_auth = 0.0**

**for i in range(size):**

**if outlinks[i][j] == 1:**

**temp\_auth += hub\_scores[i]**

**authority\_scores[j] = temp\_auth**

**auth\_sum = sum(authority\_scores)**

**for i in range(len(authority\_scores)):**

**authority\_scores[i] /= auth\_sum**

**for i in range(size):**

**temp\_hub = 0.0**

**for j in range(size):**

**if outlinks[i][j] == 1:**

**temp\_hub += authority\_scores[j]**

**hub\_scores[i] = temp\_hub**

**return authority\_scores, hub\_scores**

**n = int(input('Enter the size of the matrix:\t'))**

**outlinks = []**

**for i in range(n\*n):**

**temp = int(input('Enter the element:\t'))**

**outlinks.append(temp)**

**outlinks = np.reshape(outlinks, (n, n))**

**authority\_scores, hub\_scores = authority\_hub\_score(outlinks)**

**print("Authority Scores:")**

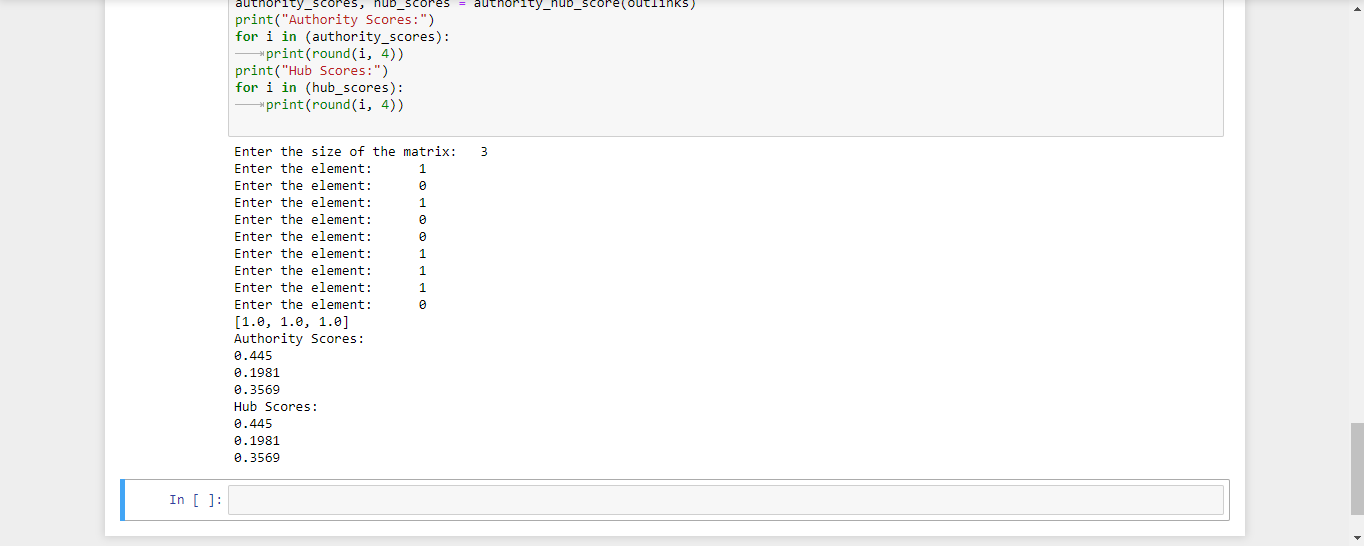
**for i in (authority\_scores):**

**print(round(i, 4))**

**print("Hub Scores:")**

**for i in (hub\_scores):**

**print(round(i, 4))**

****