

---

**IT-309**  
**EXPERIMENT 5 REPORT**

---

**Indian Institute of Information Technology, Vadodara**

**Group-1**

# Contents

0.1	Experiment -5 : Implementation of Google's PageRank Algorithm . . . .	2
0.1.1	Theory . . . . .	2
0.1.2	Experiment . . . . .	3

## 0.1 EXPERIMENT -5 : IMPLEMENTATION OF GOOGLE'S PAGERANK ALGORITHM

### 0.1.1 Theory

- In this experiment, we run Google's page rank Algorithm
- This involves the computation of the Google matrix Given by -

$$G = \alpha S + (1 - \alpha)(1/n)ee^T \quad (1)$$

where-

- $G$  is the Google Matrix
- $S$  is stochastically adjusted Hyperlink Matrix,  $H$  that prevents Any random surfer stuck at a dangling page go to any other page randomly.
- $\alpha$  is the probability with which a random surfer will follow the hyperlinks and  $(1 - \alpha)$  is the probability with which he will randomly teleport to a new web page.
- $n$  is the number of web pages and  $e$  is a  $1 \times n$  vector with all it's entries 1.
- This makes  $(1/n)ee^T$  a teleportation matrix.
- $G$  is stochastic, aperiodic and irreducible.
- After calculating  $G$  the next step is to solve the following eigen vector problem -

$$\pi^T = \pi^T G \quad (2)$$

and

$$\pi^T e = 1 \quad (3)$$

- This is solved through the power method.
- We have used the Python library to perform the Page Ranking Algorithm on the given data source.
- Since  $G$  is stochastic, and primitive, it takes only about 50 power iterations to converge to a eigen-vector.

- The convergence rate depends on the rate at which  $|\lambda_2/\lambda_1|^k \rightarrow 0$  where  $\lambda_1$  and  $\lambda_2$  are the highest and the second highest eigen values of  $G$ . Since  $G$  is Stochastic,  $\lambda_1 = 1$  and  $\lambda_1 \leq 1$

## 0.1.2 Experiment

- The given data set is an adjacency list of the hyperlink graph. consisting of 2,18,888 nodes.
- The program gives output of the dictionary in which keys are the page numbers (or ids) and values are the fraction of usage.

```

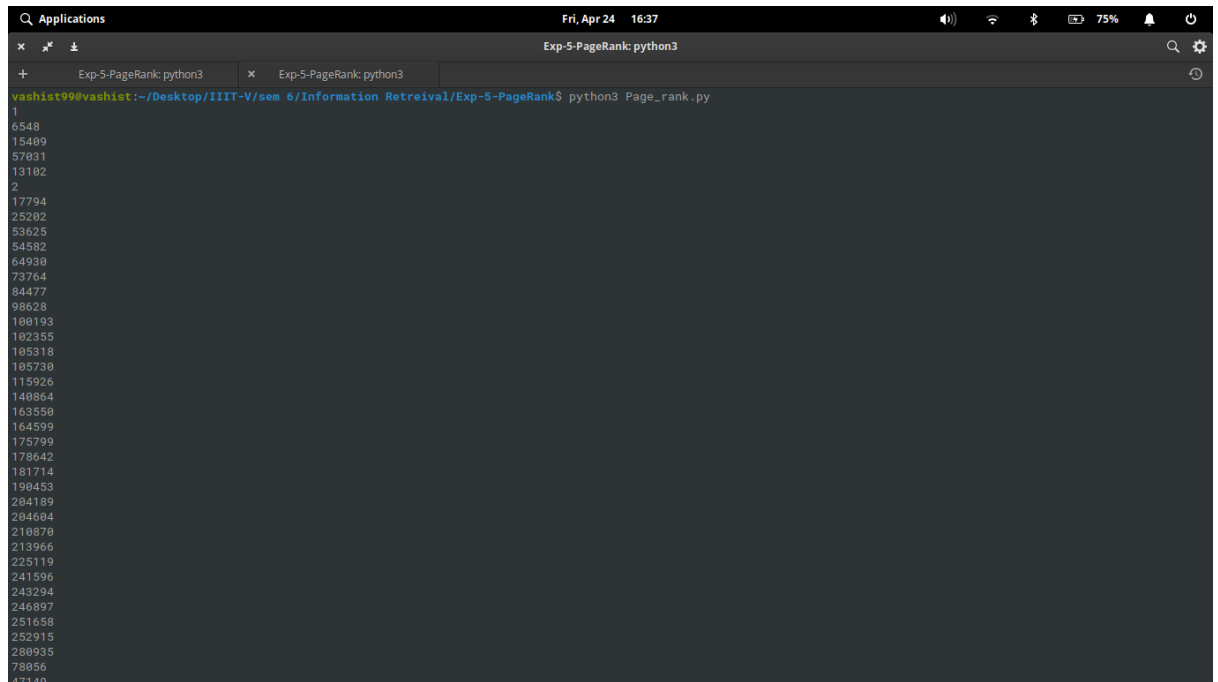
Applications
Fri, Apr 24 16:17
Exp-5-PageRank: python3

+ * Exp-5-PageRank: python3

vashist99@vashist:~/Desktop/IIIT-V/sem 6/Information Retrieval/Exp-5-PageRank$ python3 Page_rank.py
{'1': 3.021237469742247e-06, '6548': 3.3660734540115034e-06, '15409': 3.3660734540115034e-06, '57031': 3.197627640675147e-06, '13102': 3.197627640675147e-06, '2': 2.878258527264566e-05, '17794': 2.655104530532357e-06, '25202': 2.655104530532357e-06, '53625': 2.655104530532357e-06, '54582': 2.655104530532357e-06, '64930': 2.655104530532357e-06, '73764': 2.655104530532357e-06, '84477': 2.655104530532357e-06, '98628': 2.655104530532357e-06, '100193': 2.655104530532357e-06, '102355': 2.655104530532357e-06, '105318': 2.655104530532357e-06, '105730': 2.7955872676759882e-06, '115926': 2.655104530532357e-06, '140864': 2.655104530532357e-06, '163550': 2.655104530532357e-06, '164599': 2.655104530532357e-06, '175799': 2.655104530532357e-06, '178642': 2.655104530532357e-06, '181714': 2.655104530532357e-06, '190453': 2.655104530532357e-06, '204189': 2.655104530532357e-06, '204604': 2.655104530532357e-06, '210870': 2.655104530532357e-06, '213966': 2.732734700383145e-06, '225119': 2.655104530532357e-06, '241596': 2.655104530532357e-06, '243294': 2.655104530532357e-06, '246897': 2.77481979815799e-06, '251658': 2.655104530532357e-06, '252915': 2.655104530532357e-06, '280935': 2.655104530532357e-06, '78056': 9.79107907965568e-06, '47149': 5.5776151429489495e-06, '18816': 4.551452689256762e-06, '182456': 1.1367772407859671e-05, '3': 2.3697250845843806e-06, '74361': 1.3359267486993235e-05, '7213': 3.652833740076763e-06, '10365': 5.263901119471635e-06, '18412': 5.252793048329329e-05, '21315': 5.145572039880443e-06, '25126': 5.51840020084267e-06, '20140': 1.913223544848489e-05, '40745': 3.650210478562326e-06, '47352': 7.224719397026221e-06, '58718': 4.80591658009581e-06, '68036': 9.178769961375991e-06, '83203': 3.9412897652729546e-06, '100770': 5.059052466673771e-06, '120861': 6.163597457273179e-06, '123955': 4.104485494370741e-06, '133747': 9.568972674285411e-06, '146089': 4.426022488913044e-06, '152019': 5.949201138680539e-06, '153990': 9.07363622165814e-06, '176541': 4.570570756936974e-06, '192120': 8.746652618011156e-05, '198821': 6.813805975063184e-06, '215680': 4.436386519028411e-06, '220870': 4.27970388549047e-06, '224643': 6.9998726890861645e-06, '226411': 0.005685657526547046, '230874': 7.294521232221681e-06, '255694': 7.463748030618035e-06, '264216': 5.108742549800079e-06, '266083': 5.222144374994516e-06, '268261': 5.717382827274471e-06, '4': 2.7476780611398565e-06, '35716': 3.958054192814706e-06, '96512': 4.05991097335948e-06, '186750': 0.0014393635870865142, '225872': 0.00143940177228858, '829': 2.860552963534606e-06, '3776': 3.379974486817255e-06, '3869': 2.860552963534606e-06, '4514': 2.816386020764199e-06, '5030': 2.852735220701844e-05, '5682': 2.5635620128565995e-06, '9722': 2.7575973337112727e-06, '10901': 2.6468528512620718e-06, '11020': 2.655568574198629e-06, '12034': 3.82559277926985e-06, '13719': 2.8541938022638252e-06, '13855': 3.368409132421444e-06, '14525': 3.147996243940074e-06, '15781': 3.0403119699867146e-06, '15815': 6.080843253096359e-06, '17253': 3.5708332189853545e-06, '17567': 3.472074781857898e-05, '17745': 2.655568574198629e-06, '18687': 3.81469366108058e-06, '18811': 2.655568574198629e-06, '20206': 3.1443721204882946e-06, '21976': 2.655568574198629e-06, '22429': 2.860552963534606e-06, '23805': 3.5599168754920646e-06, '24206': 3.0489427648575665e-06, '24838': 3.627458417914987e-06, '26090': 3.377458739863081e-06, '29597': 2.561830500797892e-06, '31926': 2.655568574198629e-06, '34500': 2.556093971871244e-06, '35616': 3.367397347754529e-06, '39657': 2.860552963534606e-06, '42593': 3.2074582849968892e-06, '45686': 3.0453787994372617e-06, '47795': 2.860552963534606e-06, '49182': 3.5773502400753416e-05, '49561': 2.860552963534606e-06, '50574': 3.6752633861822554e-06, '53654': 3.196833212345464e-06, '54456': 2.655568574198629e-06, '55457': 3.2774316701686176e-06, '57328': 3.367397347754529e-06, '58846': 3.4461020982681847e-06, '59590': 2.858223931349329e-05, '60562': 2.655568574198629e-06, '61586': 3.0403119699867146e-06, '62941': 3.0378999843259214e-06, '63835': 3.5228318778543355e-06, '64599': 3.288056906583252e-06, '65243': 2.655568574198629e-06, '66757': 2.655568574198629e-06, '70290': 2.860552963534606e-06, '70671': 2.561830500797892e-06, '70903': 2.9591620197080160e-05, '71430': 2.946762312900903e-06, '71716': 3.0403119699867146e-06, '73205': 3.386502677442047e-06, '75702': 3.501467326440020e-06, '76836': 2.655568574198629e-06, '77816': 2.860552963534606e-06, '83071': 3.284966926886813e-06, '84820': 2.655568574198629e-06, '85879': 2.655568574198629e-06, '86333': 2.860552963534606e-06, '90918': 3.8378999843259214e-06, '92226': 2.574909787097778e-06, '92827': 2.573076595325495e-06, '94452': 3.1109094465758375e-06, '95869': 2.860552963534606e-06, '96811': 3.8378999843259214e-06, '96899': 2.860552963534606e-06, '99059': 3.8378999843259214e-06, '99371': 3.2074582849968892e-06, '99515': 2.655568574198629e-06, '99622': 3.1850318344946076e-06, '104318': 2.860552963534606e-06, '106804': 2.860552963534606e-06, '107669': 2.655568574198629e-06, '113763': 3.0403119699867146e-06, '115739': 2.655568574198629e-06, '118374': 3.5228318778543355e-06, '119687': 3.0453787994372617e-06, '121298': 3.423464129658465e-06, '122938': 2.840361734222764e-05, '124603': 3.107827842930056e-06, '124927': 2.655568574198629e-06, '125782': 3.2546971040627474e-06, '126731': 2.655568574198629e-06, '127058': 3.322270447634523e-06, '130889': 2.860552963534606e-06, '134079': 2.655568574198629e-06, '134231': 2.655568574198629e-06, '136090': 2.655568574198629e-06, '136387': 3.1348438062874045e-06, '138909': 3.6752633861822554e-06, '139074': 2.655568574198629e-06, '141160': 3.6752633861822554e-06, '145296': 2.8481255362158408e-05, '145736': 2.701556610711461e-06, '145790': 3.451492718727649e-06, '146785': 2.690268046822346e-06, '149922': 2.860552963534606e-06, '150649': 3.837899984325921e-06, '152190': 3.448948286948828e-06, '152357': 2.655568574198629e-06, '153026': 2.973363983411923e-05, '153865': 4.3445153744398944e-06, '153889': 3.408753628434772e-06, '158429': 2.655568574198629e-06, '160056': 4.9693423486507795e-06, '161514': 2.880245020872997e-05, '164644': 3.0403119699867146e-06, '165022': 2.860552963534606e-06, '167080': 3.367397347754529e-06, '167818': 2.923285168426531e-05, '172912': 3.3198839259236273e-06, '172998': 2.655568574198629e-06, '174328': 2.860552963534606e-06, '176422': 2.5660426436535574e-06, '178388':

```

- Sorting the keys according to the values in descending order gives us the rank



```
Applications
Fri, Apr 24 16:37
Exp-5-PageRank: python3
Exp-5-PageRank: python3
vashist99@vashist:~/Desktop/IIIT-V/sem 6/Information Retrieval/Exp-5-PageRank$ python3 Page_rank.py
1
6548
15489
57831
13182
2
17794
25282
53625
54582
64938
73764
84477
98628
100193
102355
105318
105730
115926
140864
163558
164599
175799
178642
181714
190453
204189
204684
210878
213966
225119
241596
243294
246897
251658
252915
280935
78056
37148
```

- Run Time of the Algorithm for this data set: 43.85716223716736 seconds