

Homework 3 - Machine Learning

Professor: Daniel S. Menasché

Given the dataset on the connection between the webpages, answer the following questions.

1 Centrality

- compute the steady state of the Markov chain using the following three methods
 - power method (P^n)
 - $\pi = \pi P$
 - eigenvalues/eigenvectors
- sort the webpages based on their
 - Pagerank
 - in degree

2 Recommendations

- which page is more more related to Globo: Times or Fox?
 - compute the mean time to reach Globo from Times and from Fox. Which one is smaller?
 - consider Times and Fox as two absorbing states. Starting from Globo, what is the probability of being absorbed by Times? And by Fox?

- consider now the mean time spent by a typical user in each of the webpages given in the following table

Globo	1
CNN	2
Fox	3
ESPN	1
Times	5

- build the generator matrix Q
- from the generator matrix, compute the uniformized matrix P
- starting from Globo, what is the probability that a random surfer will be at each of the other pages at time 10?
 - * answer using uniformization
 - * answer using the generator matrix directly