Assignment 4
Indexing the web using Solr
Shriraj L Salian
USC ID:8705640634

I)   Installation:

1) Virtual Box and Ubuntu Installation
    a. Installed virtual box and Ubuntu 16.04
    b. Downloaded files from the google drive link:
       https://drive.google.com/drive/folders/1ZPWjQZMB7xeegjaeZfzqeJKq9oKuukio
    c. Installed solr-7.2.1
         i. Used to index the data
        ii. Downloaded the file and unzipped it using unzip -q solr-7.2.1
       iii. Created a core by going into the /solr-7.2.1 directory and then running
            bin/solr -c create ircore
        iv. Made the foll changes in the managed-schema file in conf directory:
              1. Uncommented <copyfield source = "*" dest = "_text _" />
         v. Posted the files on solr server using the bin/post -c ircore -filetypes
            html path to crawl folder
        vi. Uncommented the line <str name ="df">_text_</str>
       vii. Reloaded the core to verify changes made were correct

2) Computing the Edge List:
    a. Wrote a program in Java to extract the link structure from downloaded web
       pages to compute EdgeList file, using JSoup library in Java.
    b. Web page files are the vertices and edge between two indicates a link between
       them.

3) Calculating Page Rank
    a. Used NetworkX library in python to compute the pagerank of webpages, from
       Edgelist(obtained from Java program above)
    b. Loaded web graph into NetwokX using read_edgelist() function with edgelist
       and digraph as parameters.
    c. Used pagerank() function to calculate Page Rank and output the result to
       external_pageRankFile.txt
    d. Following parameters were used in pagerank(): alpha=0.85,
       personalization=None, max_iter=30, tol=1e-06, nstart=None,
       weight='weight',dangling=None
    e. The external_pageRankFile.txt has format
       <document_id>=<page_rank_score>

4) Changes to managed-schema and solrconfig.xml for external page rank file
    a. Made the changes in managed-schema for using the external PageRank file for
       ranking
       <fieldType name="external" keyField="id" defVal="0" class="solr.ExternalFileField"/>
       <field name="pageRankFile" type="external" stored="false" indexed="false"/>

b. Added these lines in solrconfig.xml to reload the external file when a file is either reloaded or a new file is loaded
<listener event = "newSearcher" class = "org.apache.solr.schema.ExternalFileFieldReloader"/>
<listener event = "firstSearcher" class = "org.apache.solr.schema.ExternalFileFieldReloader"/>
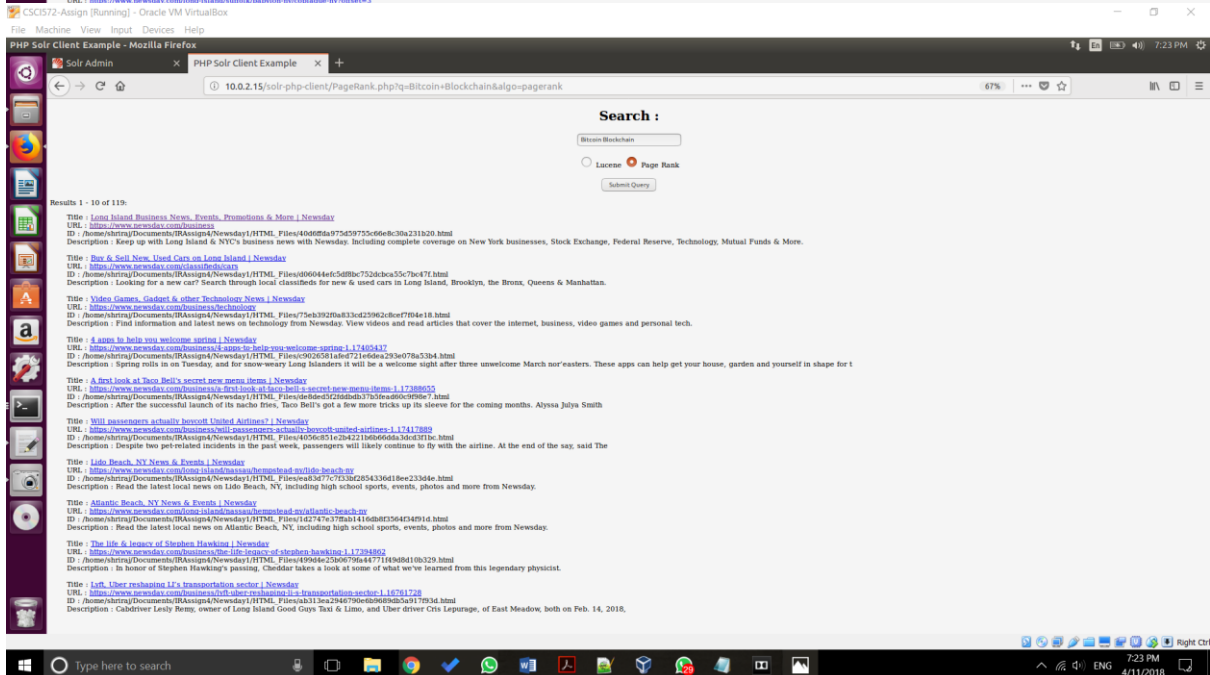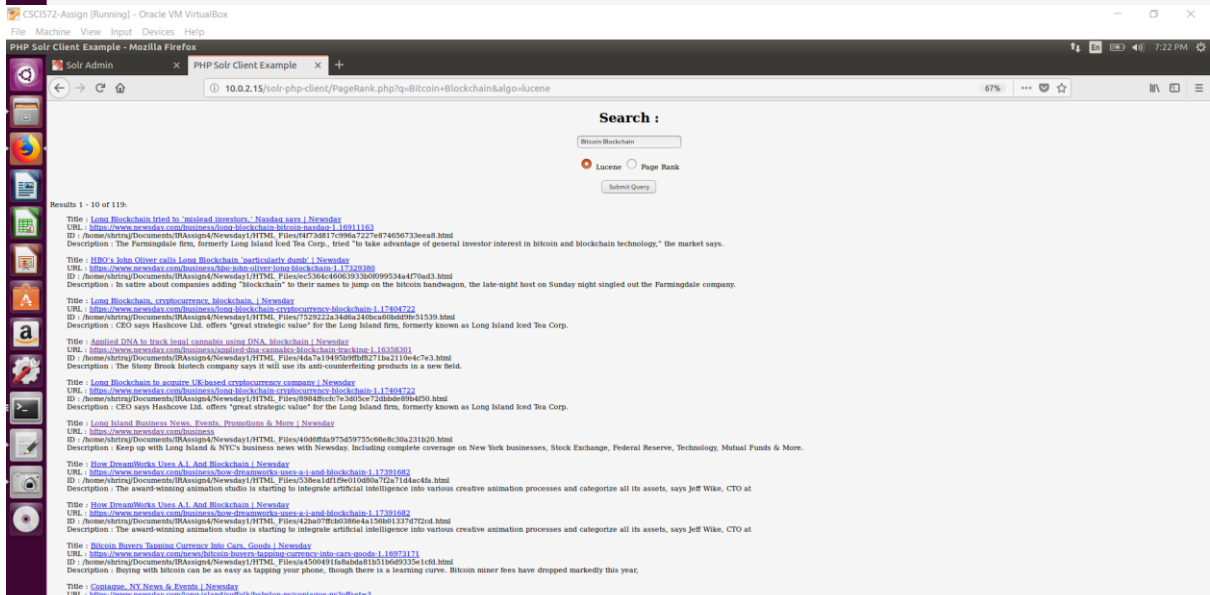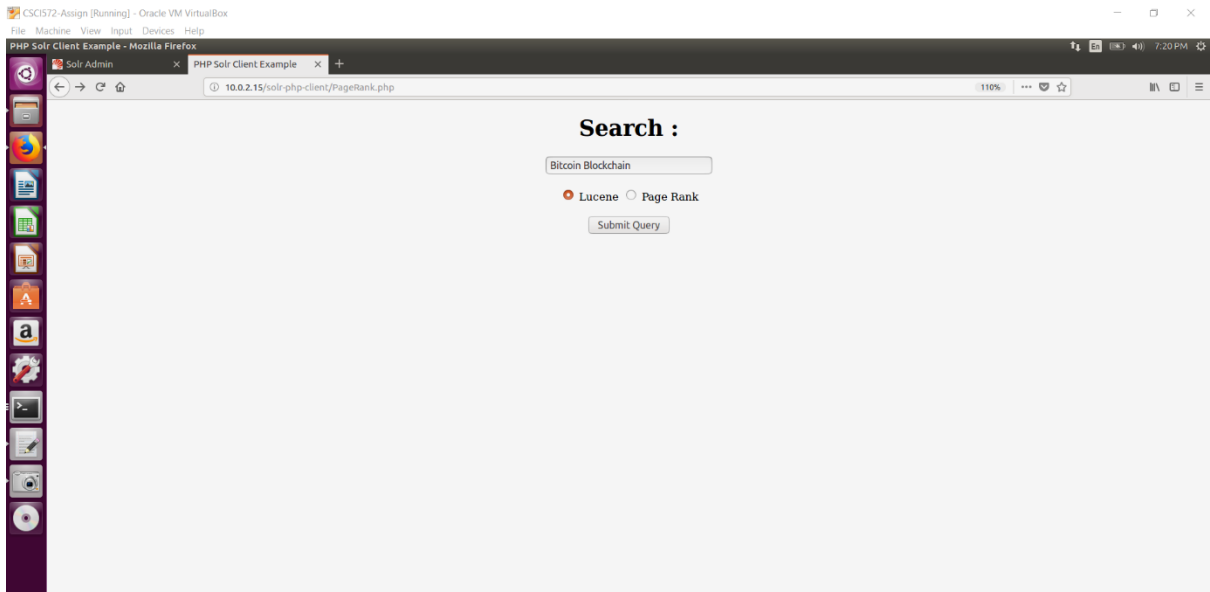
5) Apache Server, Php Installation
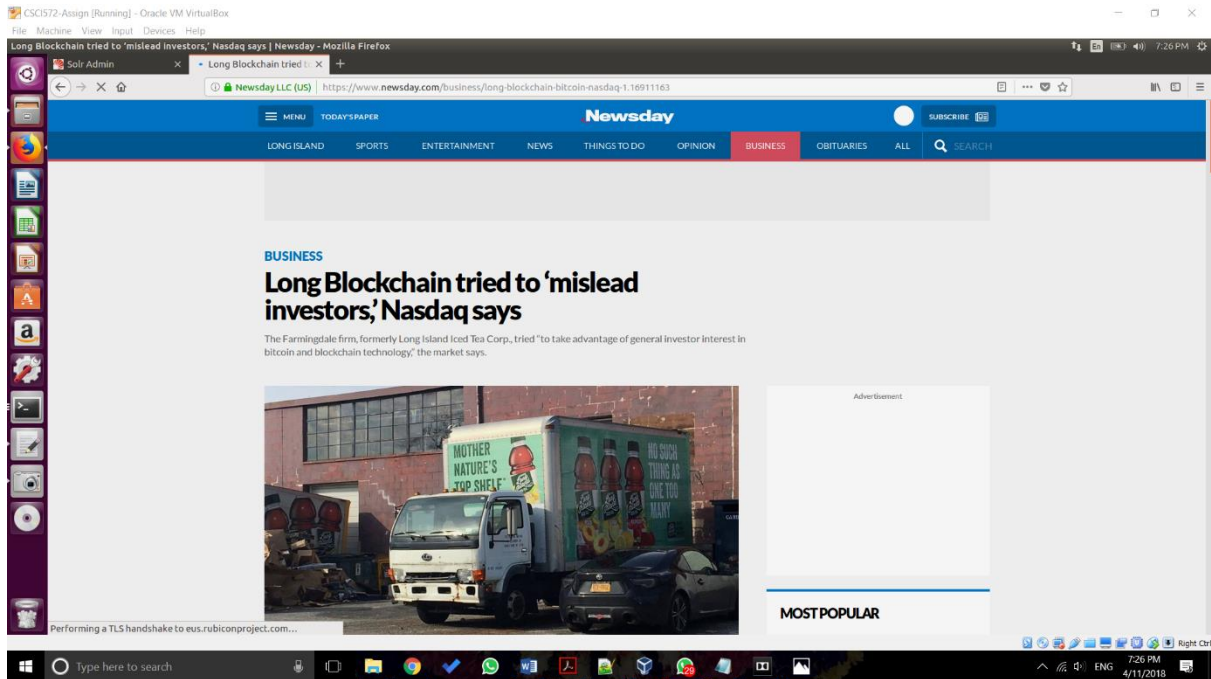   a. Installed the apache server and php on Ubuntu.
   b. Cloned the git repository in appropriate folder
   c. Started with the sample code then added radio buttons, aligned the elements and added php script to check and handle if url, id, title and description were null. The core name and path carefully written
   d. Added additional parameters 'sort' => 'pageRankFile.txt desc' for sorting the webpage results based on the external_ pageRankFile.txt file

II)   Why PageRank of some pages is more than Page Rank of others?

PageRank is an algorithm used by search engines to rank websites in their search engine results and is a way of measuring the importance of website pages. The page algorithm assigns a score to the web pages based on the number in-links and outlinks (Though recently Google uses many other factors to assign a score) If a web page is pointed by other web pages with a higher page rank score, its page rank score tends to increase. The EdgeList was used to construct a graph using the network library in Python. We computed the page rank based on this graph using the pagerank function. Thus, some pages may have a higher page rank if they are pointed by other pages with high scores and relatively fewer outlinks. A damping factor

III)   ScreenShots depicting flow of process: (On next Page)

PHP Solr Client Example - Mozilla Firefox

Solr Admin | PHP Solr Client Example +

① 10.0.2.15/solr-php-client/PageRank.php

Search :

Bitcoin Blockchain

● Lucene ○ Page Rank

Submit Query

PHP Solr Client Example - Mozilla Firefox

Solr Admin | PHP Solr Client Example +

① 10.0.2.15/solr-php-client/PageRank.php?q=Bitcoin+Blockchain&algo=lucene

Search :

Bitcoin Blockchain

● Lucene ○ Page Rank

Submit Query

Results 1 - 10 of 119:

Title : Long Blockchain tried to 'mislead investors,' Nasdaq says | Newsday
URL : https://www.newsday.com/business/long-blockchain-bitcoin-nasdaq-1.16911163
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/f4f73d817c996a7227e874656733eea8.html
Description : The Farmingdale firm, formerly Long Island Iced Tea Corp., tried "to take advantage of general investor interest in bitcoin and blockchain technology," the market says.

Title : HBO's John Oliver calls Long Blockchain 'particularly dumb' | Newsday
URL : https://www.newsday.com/business/hbo-john-oliver-long-blockchain-1.17329380
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/ec5364c46063933b0099534a4f70ad3.html
Description : In satire about companies adding "blockchain" to their names to jump on the bitcoin bandwagon, the late-night host on Sunday night singled out the Farmingdale company.

Title : Long Blockchain, cryptocurrency, blockchain, | Newsday
URL : https://www.newsday.com/business/long-blockchain-cryptocurrency-blockchain-1.17404722
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/7529222a34d6a240bca60bdd9fe51539.html
Description : CEO says Hashcove Ltd. offers "great strategic value" for the Long Island firm, formerly known as Long Island Iced Tea Corp.

Title : Applied DNA to track legal cannabis using DNA, blockchain | Newsday
URL : https://www.newsday.com/business/applied-dna-cannabis-blockchain-tracking-1.16358301
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/4da7a19495b9fb8271ba21100e4c7e3.html
Description : The Stony Brook biotech company says it will use its anti-counterfeiting products in a new field.

Title : Long Blockchain to acquire UK-based cryptocurrency company | Newsday
URL : https://www.newsday.com/business/long-blockchain-cryptocurrency-blockchain-1.17404722
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/8984ffccfc7e3d05ce72dbbde89b4f50.html
Description : CEO says Hashcove Ltd. offers "great strategic value" for the Long Island firm, formerly known as Long Island Iced Tea Corp.

Title : Long Island Business News, Events, Promotions & More | Newsday
URL : https://www.newsday.com/business
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/40d6f8da975d59755c66e8c30a231b20.html
Description : Keep up with Long Island & NYC's business news with Newsday. Including complete coverage on New York businesses, Stock Exchange, Federal Reserve, Technology, Mutual Funds & More.

Title : How DreamWorks Uses A.I. And Blockchain | Newsday
URL : https://www.newsday.com/business/how-dreamworks-uses-a-i-and-blockchain-1.17391682
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/538ea1df1f9e010d80a7f2a71d4ac4fa.html
Description : The award-winning animation studio is starting to integrate artificial intelligence into various creative animation processes and categorize all its assets, says Jeff Wike, CTO at

Title : How DreamWorks Uses A.I. And Blockchain | Newsday
URL : https://www.newsday.com/business/how-dreamworks-uses-a-i-and-blockchain-1.17391682
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/42ba07ffcb0386e4a156b01337d7f2cd.html
Description : The award-winning animation studio is starting to integrate artificial intelligence into various creative animation processes and categorize all its assets, says Jeff Wike, CTO at

Title : Bitcoin Buyers Tapping Currency Into Cars, Goods | Newsday
URL : https://www.newsday.com/news/bitcoin-buyers-tapping-currency-into-cars-goods-1.16973171
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/a45004911a8abda81b51b6d93335e1cfd.html
Description : Buying with bitcoin can be as easy as tapping your phone, though there is a learning curve. Bitcoin miner fees have dropped markedly this year;

Title : Copiague, NY News & Events | Newsday
URL : https://www.newsday.com/long-island/suffolk/babylon-n/copiague-ny?offset=3

PHP Solr Client Example - Mozilla Firefox

Solr Admin | PHP Solr Client Example +

① 10.0.2.15/solr-php-client/PageRank.php?q=Bitcoin+Blockchain&algo=pagerank

Search :

Bitcoin Blockchain

○ Lucene ● Page Rank

Submit Query

Results 1 - 10 of 119:

Title : Long Island Business News, Events, Promotions & More | Newsday
URL : https://www.newsday.com/business
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/40d6f8da975d59755c66e8c30a231b20.html
Description : Keep up with Long Island & NYC's business news with Newsday. Including complete coverage on New York businesses, Stock Exchange, Federal Reserve, Technology, Mutual Funds & More.

Title : Buy & Sell New, Used Cars on Long Island | Newsday
URL : https://www.newsday.com/classifieds/cars
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/d06044efc5d8fbc752dcbca55c7bc47f.html
Description : Looking for a new car? Search through local classifieds for new & used cars in Long Island, Brooklyn, the Bronx, Queens & Manhattan.

Title : Video Games, Gadget & other Technology News | Newsday
URL : https://www.newsday.com/business/technology
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/75eb392f0a833cd25962c8cef7f04e18.html
Description : Find information and latest news on technology from Newsday. View videos and read articles that cover the internet, business, video games and personal tech.

Title : 4 apps to help you welcome spring | Newsday
URL : https://www.newsday.com/business/4-apps-to-help-you-welcome-spring-1.17405437
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/9026581afed721e6dea293e078a53b4.html
Description : Spring rolls in on Tuesday, and for snow-weary Long Islanders it will be a welcome sight after three unwelcome March nor'easters. These apps can help your house, garden and yourself in shape for t

Title : A first look at Taco Bell's secret new menu items | Newsday
URL : https://www.newsday.com/business/a-first-look-at-taco-bell-s-secret-new-menu-items-1.17388655
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/8ded5f28ddbdb37b5fead60c9f98e7.html
Description : After the successful launch of its nacho fries, Taco Bell's got a few more tricks up its sleeve for the coming months. Alyssa Julya Smith

Title : Will passengers actually boycott United Airlines? | Newsday
URL : https://www.newsday.com/business/will-passengers-actually-boycott-united-airlines-1.17417889
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/4056c851e2b4221b6b66dda3dcd3f1bc.html
Description : Despite two pet-related incidents in the past week, passengers will likely continue to fly with the airline. At the end of the say, said The

Title : Lido Beach, NY News & Events | Newsday
URL : https://www.newsday.com/long-island/nassau/hempstead-ny/lido-beach-ny
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/ea83d77c7f33bf2854336d18ee233d4e.html
Description : Read the latest local news on Lido Beach, NY, including high school sports, events, photos and more from Newsday.

Title : Atlantic Beach, NY News & Events | Newsday
URL : https://www.newsday.com/long-island/nassau/hempstead-ny/atlantic-beach-ny
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/1d2747e37f8ab1416d8f3564f34f91d.html
Description : Read the latest local news on Atlantic Beach, NY, including high school sports, events, photos and more from Newsday.

Title : The life & legacy of Stephen Hawking | Newsday
URL : https://www.newsday.com/business/the-life-legacy-of-stephen-hawking-1.17394802
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/499d4e25b0679fa44771f49d8d10b329.html
Description : In honor of Stephen Hawking's passing, Cheddar takes a look at some of what we've learned from this legendary physicist.

Title : Lyft, Uber reshaping LI's transportation sector | Newsday
URL : https://www.newsday.com/business/lyft-uber-reshaping-li-s-transportation-sector-1.16761728
ID : /home/shriraj/Documents/IRAssign4/Newsday1/HTML_Files/ab313ea294679e4b9689db5a917f93d.html
Description : Cabdriver Lesly Remy, owner of Long Island Good Guys Taxi & Limo, and Uber driver Cris Lepurage, of East Meadow, both on Feb. 14, 2018,

IV) Comparison of top 10 results for Default and PageRank Algorithms

1. Bitcoin Blockchain

Default Solr Algorithm	External Page Rank using NetworkX
https://www.newsday.com/business/long-blockchain-bitcoin-nasdaq-1.16911163	https://www.newsday.com/business
https://www.newsday.com/business/hbo-john-oliver-long-blockchain-1.17329380	https://www.newsday.com/classifieds/cars
https://www.newsday.com/business/long-blockchain-cryptocurrency-blockchain-1.17404722	https://www.newsday.com/business/technology
https://www.newsday.com/business/applied-dna-cannabis-blockchain-tracking-1.16358301	https://www.newsday.com/business/4-apps-to-help-you-welcome-spring-1.17405437
https://www.newsday.com/business/long-blockchain-cryptocurrency-blockchain-1.17404722	https://www.newsday.com/business/a-first-look-at-taco-bell-s-secret-new-menu-items-1.17388655
https://www.newsday.com/business	https://www.newsday.com/business/will-passengers-actually-boycott-united-airlines-1.17417889
https://www.newsday.com/business/how-dreamworks-uses-a-i-and-blockchain-1.17391682	https://www.newsday.com/long-island/nassau/hempstead-ny/lido-beach-ny
https://www.newsday.com/business/how-dreamworks-uses-a-i-and-blockchain-1.17391682	https://www.newsday.com/long-island/nassau/hempstead-ny/atlantic-beach-ny
https://www.newsday.com/news/bitcoin-buyers-tapping-currency-into-cars-goods-1.16973171	https://www.newsday.com/business/the-life-legacy-of-stephen-hawking-1.17394862

| https://www.newsday.com/long-island/suffolk/babylon-ny/copiague-ny?offset=3" | https://www.newsday.com/business/lyft-uber-reshaping-li-s-transportation-sector-1.16761728 |

2. **Stephen Hawking**

Default Solr Algorithm	External Page Rank using NetworkX
https://www.newsday.com/news/world/physicist-stephen-hawking-dies-at-age-76-1.17381457	https://www.newsday.com/business
https://www.newsday.com/entertainment/tv/stephen-hawking-tv-roles-1.17384428	https://www.newsday.com/opinion
https://www.newsday.com/news/world/theoretical-physicist-stephen-hawking-has-died-at-76-1.17380191	https://www.newsday.com/entertainment
https://www.newsday.com/news/world/theoretical-physicist-stephen-hawking-dies-at-76-1.17380408	https://www.newsday.com/news/nation
https://www.newsday.com/business/the-life-legacy-of-stephen-hawking-1.17394862	https://www.newsday.com/entertainment/music
https://www.newsday.com/opinion/editorial/stephen-hawking-death-1.17387052	https://www.newsday.com/entertainment/celebrities
https://www.newsday.com/news/world/theoretical-physicist-stephen-hawking-has-died-at-76-1.17380191	https://www.newsday.com/entertainment/movies
https://www.newsday.com/news/world/stephen-hawking-dead-1.17380130	https://www.newsday.com/entertainment/tv
https://www.newsday.com/entertainment/tv/stephen-hawking-tv-roles-1.17384428	https://www.newsday.com/business/technology
https://www.newsday.com/news/world/theoretical-physicist-stephen-hawking-dies-at-76-1.17380408	https://www.newsday.com/entertainment/books

3. **Falcon Heavy**

Default Solr Algorithm	External Page Rank using NetworkX
https://www.newsday.com/opinion/columnists/michael-dobie/elon-musk-spacex-1.16670802	https://www.newsday.com/long-island/education/school-by-school-spending-rule-1.17484061
https://www.newsday.com/opinion/columnists/michael-dobie/elon-musk-spacex-1.16670802	https://www.newsday.com/lifestyle/restaurants/perennial-review-garden-city-1.17402910
https://www.newsday.com/long-island/towns/freshmen-win-locust-valley-high-school-float-contest-at-homecoming-parade-1.6329245	https://www.newsday.com/sports/basketball/knicks/knicks-young-players-going-all-out-for-wins-1.17488993
https://www.newsday.com/entertainment/movies/tomb-raider-review-1.17387308	https://www.newsday.com/sports/football/jets/jets-deal-with-colts-to-get-no-3-pick-1.17448060
https://www.newsday.com/business/air-industries-sells-its-amk-subsidiary-for-up-to-6-million-1.13039083	https://www.newsday.com/lifestyle/restaurants/feed-me/global-buffet-levittown-all-you-can-eat-1.17463506
https://www.newsday.com/entertainment/tv/wet-hot-american-summer-prequel-on-netflix-funnier-than-the-movie-1.10689017	https://www.newsday.com/entertainment/music/my-fathers-place-concerts-1.17424391
https://www.newsday.com/entertainment/movies/tomb-raider-review-1.17387308	https://www.newsday.com/sports/columnists/bob-glauber/nfl-free-agency-1.17433781
https://www.newsday.com/entertainment/celebrities/notable-long-islanders-1.2342650	https://www.newsday.com/lifestyle/restaurants/long-island-cheap-eats-restaurant-reviews-1.4445645
https://www.newsday.com/long-island/suffolk/centerport-house-fire-power-line-1.17205951	https://www.newsday.com/news/world/column-another-victim-of-russian-doping-bites-the-dust-1.17498449
https://www.newsday.com/lifestyle/home-and-garden/winter-bedding-1.14848287	https://www.newsday.com/lifestyle/restaurants/2-spring-review-oyster-bay-1.17208606

4. **Syrian War**

Default Solr Algorithm	External Page Rank using NetworkX
https://www.newsday.com/news/world/syria-united-nations-1.17436072	https://www.newsday.com/news/nation
https://www.newsday.com/news/world/syria-united-nations-1.17436072	https://www.newsday.com/sports/olympics
https://www.newsday.com/news/world/turkey-syria-afrin-1.17463820	https://www.newsday.com/long-island/history
https://www.newsday.com/news/nation/president-donald-trump-we-re-not-going-into-syria-1.13444720	https://www.newsday.com/opinion/submit-your-letter-1.2516352
https://www.newsday.com/news/world/turkey-syria-afrin-1.17463820	https://www.newsday.com/sports/high-school/girls-basketball/baldwin-girls-make-

	li-history-by-repeating-as-state-class-aa-champs-1.17483760
https://www.newsday.com/news/world/russia-uk-diplomats-spy-poisoning-1.17448346	https://www.newsday.com/entertainment/celebrities/recent-notable-deaths-1.1236203
https://www.newsday.com/news/world/north-korea-photos-from-a-reclusive-country-1.13637395	https://www.newsday.com/long-island/nassau/passover-baskets-1.17486468
https://www.newsday.com/news/nation/sean-spicer-claims-even-hitler-didn-t-use-chemical-weapons-1.13431614	https://www.newsday.com/news/nation/president-trump-first-year-moments-1.16257304
https://www.newsday.com/news/world/the-latest-russia-rejects-ultimatums-over-spy-poisoning-1.17381984	https://www.newsday.com/travel/weekend-getaways-near-new-york-city-1.5299723
https://www.newsday.com/long-island/suffolk/iraq-helicopter-crash-westhampton-1.17456031	https://www.newsday.com/news/world

5. **LA Rams**

Default Solr Algorithm	External Page Rank using NetworkX
https://www.newsday.com/sports/football/super-bowl/super-bowl-history-1.170249	https://www.newsday.com/classifieds/cars
https://www.newsday.com/sports/football/giants/alec-ogletree-traded-giants-rams-1.17191446	https://www.newsday.com/sports/football/jets
https://www.newsday.com/sports/football/giants/alec-ogletree-traded-giants-rams-1.17191446	https://www.newsday.com/sports/football/giants
https://www.newsday.com/sports/football/jets/jets-isaiah-crowell-trumaine-johnson-1.17379143	https://www.newsday.com/sports/soccer
https://www.newsday.com/sports/football/giants/alec-ogletree-odell-beckham-teammates-1.17406118	https://www.newsday.com/opinion/submit-your-letter-1.2516352
https://www.newsday.com/sports/football/jets/jets-isaiah-crowell-trumaine-johnson-1.17379143	https://www.newsday.com/entertainment/celebrities/recent-notable-deaths-1.1236203
https://www.newsday.com/sports/football/giants/giants-overhaul-to-offensive-line-has-been-a-smooth-ride-1.17412710	https://www.newsday.com/lifestyle/restaurants/feed-me/beginnings-atlantic-beach-goodfellas-1.17419673
https://www.newsday.com/sports/football/giants/alec-ogletree-odell-beckham-teammates-1.17406118	https://www.newsday.com/sports/football/jets/jets-deal-with-colts-to-get-no-3-pick-1.17448060

https://www.newsday.com/sports/football/nfl-free-agency-cousins-brees-1.17225240	https://www.newsday.com/lifestyle/restaurants/barbecue-restaurants-on-long-island-our-critics-favorites-1.8916757
https://www.newsday.com/sports/football/super-bowl/super-bowl-game-winning-drives-1.6477576	https://www.newsday.com/travel/show-us-your-travel-photos-1.12098434

6. **Black Panther**

Default Solr Algorithm	External Page Rank using NetworkX
https://www.newsday.com/entertainment/movies/black-panther-box-office-1.17471969	https://www.newsday.com/long-island
https://www.newsday.com/entertainment/movies/hempstead-students-see-black-panther-1.16901584	https://www.newsday.com/entertainment
https://www.newsday.com/entertainment/movies/black-panther-review-1.16582314	https://www.newsday.com/lifestyle
https://www.newsday.com/entertainment/movies/black-panther-1.16552223	https://www.newsday.com/long-island/nassau
https://www.newsday.com/entertainment/movies/black-panther-box-office-1.17471969	https://www.newsday.com/news/health
https://www.newsday.com/entertainment/movies/black-panther-1.16552223	https://www.newsday.com/classifieds/real-estate
https://www.newsday.com/entertainment/movies/hempstead-students-see-black-panther-1.16901584	https://www.newsday.com/entertainment/movies
https://www.newsday.com/entertainment/movies/black-panther-review-1.16582314	https://www.newsday.com/entertainment/tv
https://www.newsday.com/long-island/nassau/hempstead-students-treated-to-black-panther-showing-1.16901559	https://www.newsday.com/sports/boxing
https://www.newsday.com/entertainment/movies/latest-movie-reviews-1.2333132?offset=2	https://www.newsday.com/lifestyle/home-and-garden

7. Austin Bombings

Default Solr Algorithm	External Page Rank using NetworkX
https://www.newsday.com/news/police-search-for-motive-in-austin-bombings-1.17372261	https://www.newsday.com/news/nation
https://www.newsday.com/news/police-search-for-motive-in-austin-bombings-1.17372261	https://www.newsday.com/
https://www.newsday.com/news/nation/austin-texas-explosion-1.17496366	https://www.newsday.com/
https://www.newsday.com/news/nation/austin-texas-explosion-1.17496366	https://www.newsday.com/long-island/education/school-by-school-spending-rule-1.17484061
https://www.newsday.com/news/police-tripwire-may-have-set-off-austin-blast-1.17498369	https://www.newsday.com/long-island/suffolk/andreas-o-keeffe-iraq-chopper-crash-1.17487480
https://www.newsday.com/news/federal-agents-swarm-austin-explosion-area-1.17497925	https://www.newsday.com/entertainment/celebrities/celebrities-at-sxsw-1.17186200
https://www.newsday.com/news/austin-texas-explosion-injures-two-1.17497374	https://www.newsday.com/lifestyle/family/netflix-kids-streaming-now-1.10684524
https://www.newsday.com/news/nation/austin-package-bombs-1.17336169	https://www.newsday.com/news/region-state/cuomo-gun-background-check-1.17489680
https://www.newsday.com/news/friend-hawking-was-full-of-life-to-the-end-1.17393821	https://www.newsday.com/news/nation/president-trump-first-year-moments-1.16257304
https://www.newsday.com/news/showbiz-minute-locklear-american-idol-scientology-1.17361523	https://www.newsday.com/sports/columnists/bob-glauber/nfl-free-agency-1.17433781

8. Paris Climate Deal

Default Solr Algorithm	External Page Rank using NetworkX
https://www.newsday.com/news/nation/trump-withdraws-u-s-from-paris-agreement-andrew-cuomo-barack-obama-scott-pruitt-more-respond-1.13700637	https://www.newsday.com/business
https://www.newsday.com/news/nation/trump-u-s-will-withdraw-from-paris-climate-accord-1.13699039	https://www.newsday.com/sports
https://www.newsday.com/opinion/columnists/michael-dobie/withdrawing-from-climate-deal-wrongheaded-1.13700910	https://www.newsday.com/news/nation
https://www.newsday.com/news/nation/trump-withdraws-u-s-from-paris-agreement-andrew-cuomo-barack-obama-scott-pruitt-more-respond-1.13700637	https://www.newsday.com/classifieds/cars

https://www.newsday.com/news/nation/trump-u-s-will-withdraw-from-paris-climate-accord-1.13699039	https://www.newsday.com/sports/football/jets
https://www.newsday.com/opinion/editorial/after-paris-pullout-americans-can-do-their-part-on-climate-1.13701237	https://www.newsday.com/sports/football/giants
https://www.newsday.com/opinion/eight-ideas-to-deal-with-climate-change-and-save-the-planet-1.11176829	https://www.newsday.com/sports/football
https://www.newsday.com/opinion/eight-ideas-to-deal-with-climate-change-and-save-the-planet-1.11176829	https://www.newsday.com/sports/soccer
https://www.newsday.com/opinion/columnists/michael-dobie/the-terrible-what-ifs-of-climate-change-1.11068353	https://www.newsday.com/entertainment/tv/reality-tv
https://www.newsday.com/long-island/environmentalists-li-is-closest-to-problem-of-climate-change-1.13729504	https://www.newsday.com/sports/golf

V) **Overlaps**

(I assume an overlap occurs when the url and id of the webpage is same)

Query no.	Query	No. of Overlaps
1	Bitcoin Blockchain	1
2	Stephen Hawking	0
3	Falcon Heavy	0
4	Syrian War	0
5	LA Rams	0
6	Black Panther	0
7	Austin Bombings	0
8	Paris Climate Deal	0

Number of overlaps between results