

Mukesh Dangi(USC ID : 4297380684)
New York Post – Comparing Search Engine Ranking Algorithms

Step 1: Solr Setup and Indexing the HTML files

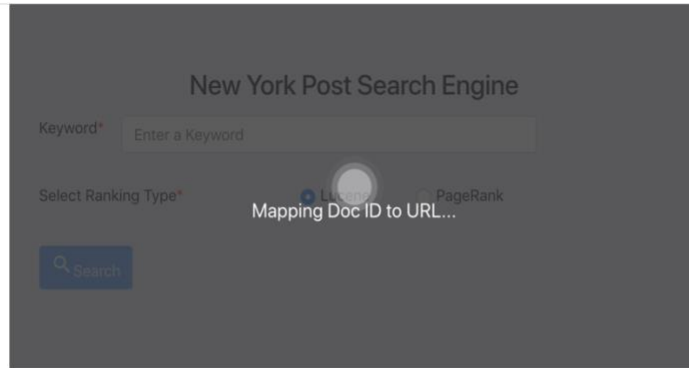
- a) Downloaded files NY Post URL to HTML mapped data and HTML pages from the google drive link. Downloaded and extracted all the crawled files from google drive for nypost.
- b) Downloaded the file solr-7.5.0 and Installed solr-7.5.0.
- c) Created a core by going into the /solr-7.5.0 directory started solr using **bin/solr start** and then created solr core by running **bin/solr -c create nypost**, where nypost is core name.
- d) Edited the managed-schema, solrconfig file in the config folder so that all the text content can be mapped properly. This command will index all the documents on Solr and might take some time depending upon dataset.
- e) Posted all crawled files on solr server using the **bin/post -c nypost -filetypes html path to nypost** folder
- f) Reloaded the nypost core from Solr UI admin to verify changes made were correct

Step 2: Angular 7, NodeJS installation and Front End:

- a) Installed Angular 7, Angular CLI and NodeJS using NPM (Node package manager)
- b) Created new NG project using **ng new nypost**. Added all the dependencies in package.json and ran **npm install**.
- c) Used MVC architecture in the project and created search box, ajax calls and proxy server.
- d) Created solrsearch.js, a node js file which uses express js, csvtojson and many other relevant libraries.
- e)

```
let SolrNode = require('solr-node');  
let client = new SolrNode({  
  host: 'localhost', port: '8983',  
  core: 'nypost', protocol: 'http'  
});
```
- f) Initiated Solr search with

```
let strQuery = client.query().q('La Lakers');  
client.search(strQuery, function (err, result) {results})
```
- g) Made all the API calls to from localhost:4200 which is angular port to localhost:3000 node port using proxy config. Proxy config is being used in the ng serve command.
- h) On search page load, I called **/api/getcsvjson api** to node js server to get all the HTML doc Id to HTML URL mapping.
- i) On page load (load csv, map doc id to URL)



- j) First 10 results are displayed. The URL for the web page file name was obtained by parsing the mappings csv file.
- k) Passing '**&sort = pageRankFile desc**' if user opt for external page rank Algorithms otherwise no sort argument passed.
- l) On Front End, extracted only Title, URL, description and ID from the response object.

Step 3: Generating Edge List:

- a) Wrote a Java program using JSoup library.
- b) Using JSoup API, I extracted all **a[href]** elements and then all the URLs attributes **abs:href** from those Elements. I extracted the outgoing links of each web page by traversing the directory to create the edgeList.txt file.
- c) Used URLtoHTML_nypost.csv to check which of the links are connected to other links. Here I have used doc id for the comparison.
- d) Two webpages will make an entry into Edge list only if the they are inter-liked to each other in some way. Either an edge from page A to page B or vice -versa.

Step 4: Page Rank Calculation:

- a) Used NetworkX library in python to computed page rank and create a web graph of those Edge list.
- b) Created web graph into NetwokX using read_edgelist() function with edgelist.txt and digraph as parameters.
- c) Used pagerank() function to calculate Page Rank and output the result to external_pageRankFile.txt
- d) Wrote python program to compute the page rank values using the NetworkX library.
- e) python code iterates over each line from edgelist.txt and computes the page rank values using the following configuration
- f) ***alpha=0.85, tol=1e-06 personalization=None, nstart=None, max_iter=30, , weight='weight', dangling=None.***
- g) Finally, the python program using NetworkX library created an external_pageRankFile.txt file having format <document_id>=<page_rank_score>.

Step 5: Solr Core update with new changes of external PageRank file:

- a) Added the external_pageRankFile.txt in the data folder of the core which we created in step 1. Added necessary changes to managed-schema by adding field, and fieldType params.
- b) Managed Schema changes:
 - a. `<fieldType name="external" keyField="id" defVal="0" class="solr.ExternalFileField"/>`
 - b. `<field name="pageRankFile" type="external" stored="false" indexed="false"/>`
- c) Solr Config Changes (Added external new searcher listeners)
 - a. `<listener event="newSearcher" class="org.apache.schema.ExternalFileFieldRealoder"/>`
 - b. `<listener event="firstSearcher" class="org.apache.schema.ExternalFileFieldRealoder"/>`
- d) Modified the solrconfig.xml file to include listeners as per the instructions.
- e) Reloaded the Solr **nypost** core on the Solr admin page. Next, queried from Solr Admin Page by placing "pageRankFile desc" in sort parameter of the query to get the results based on pageRankFile in descending order.

Explanation for higher page rank values:

We used JSoup to check if two pages are linked to each other. Then we used NetworkX library to generate web graph and External Page. Pages with higher incoming links are ranked higher than other pages. However, In Lucene, Page ranking is taken care by two models, Boolean and vector space model. So we are expecting to get more relevant results from Lucene in most of the cases. 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. PageRank is crucial in showing relevant results to user.

If a web page is pointed by other web pages then that page will receive a higher page rank score, its page rank score tends to increase. Some pages like home page of any website has more incoming links than others and has high reference in network so home page tends to receive higher page rank. However, PageRank algorithm should determine relevant pages than refereeing to home page. Thus, some pages may have a higher page rank if they are pointed by other pages with high scores and relatively fewer outgoing links.

Ten results with higher page rank using Lucene and External Page Rank:

1. Donald Trump:

Lucene	External Page rank
https://nypost.com/2018/10/15/donald-and-melania-trump-begin-tour-of-towns-destroyed-by-hurricane-michael/	https://nypost.com/tech/
https://nypost.com/2018/10/15/donald-and-melania-trump-begin-tour-of-towns-destroyed-by-hurricane-michael/	https://nypost.com/fashion/
https://nypost.com/2016/02/15/donald-trump-has-invented-a-new-way-to-win/	https://nypost.com/columnists/
https://nypost.com/2016/11/09/	https://nypost.com/horoscopes/
https://nypost.com/2017/11/06/how-donald-trump-manipulated-bon-jovi-out-of-bills-purchase/	https://nypost.com/news/
https://nypost.com/2018/10/01/donald-trump-pays-tribute-to-friend-david-wright/	https://nypost.com/living/
https://nypost.com/2016/05/03/donald-trump-is-the-nominee-cruz-and-kasich-should-drop-out/	https://nypost.com/opinion/
https://nypost.com/2016/11/10/	https://nypost.com/media/
https://nypost.com/2017/01/20/	https://nypost.com/entertainment/
https://nypost.com/2016/11/11/	https://nypost.com/

2. LA Lakers:

Lucene	External Page rank
https://nypost.com/2018/09/25/lebron-james-knows-it-wont-come-easy-with-lakers/	https://nypost.com/sports/
https://nypost.com/2018/09/07/warriors-star-emerges-as-lakers-flashy-plan-b-in-free-agency/	https://nypost.com/2018/10/15/tom-bradys-postgame-words-to-gronk-are-an-nfl-nightmare/
https://nypost.com/2018/10/10/klay-thompson-gushes-over-lebron-as-lakers-pursuit-looms/	https://nypost.com/2018/10/15/nfl-flameout-admits-to-having-sex-with-12-year-old-girl-police/
https://nypost.com/tag/nba/	https://nypost.com/2018/10/15/eli-manning-trade-could-save-him-and-the-giants-espn-qb/
https://nypost.com/tag/los-angeles-lakers/	https://nypost.com/2018/10/14/aaron-hernandez-had-sexual-relationship-with-hs-quarterback/
https://nypost.com/2001/02/14/	https://nypost.com/2018/10/15/mlb-youngsters-bring-the-trash-talk-and-much-needed-fun/

https://nypost.com/2014/05/02/	https://nypost.com/2018/10/15/empty-vows-alderon-back-stab-inside-mets-gm-search-mess/
https://nypost.com/2012/11/12/	https://nypost.com/2018/10/15/angry-about-the-subways-send-cuomo-a-message/
https://nypost.com/2015/06/26/	https://nypost.com/2018/10/15/a-giants-trade-for-ex-patriots-quarterback-would-make-sense/
https://nypost.com/2017/02/21/	https://nypost.com/2018/10/15/how-one-veterans-vietnam-horrors-inspired-this-is-us/

3. Star Wars:

Lucene	External Page rank
https://nypost.com/2015/12/18/	https://nypost.com/real-estate/
https://nypost.com/2015/12/15/	https://nypost.com/tech/
https://nypost.com/2015/12/17/	https://nypost.com/fashion/
https://nypost.com/2015/12/16/	https://nypost.com/columnists/
https://nypost.com/2015/12/22/	https://nypost.com/photos/
https://nypost.com/2015/12/14/	https://nypost.com/horoscopes/
https://nypost.com/2015/11/24/	https://nypost.com/metro/
https://nypost.com/2015/04/20/	https://nypost.com/sports/
https://nypost.com/2015/10/20/	https://nypost.com/news/
https://nypost.com/2015/04/16/	https://nypost.com/living/

4. LeBron James:

Lucene	External Page rank
https://nypost.com/2018/09/25/lebron-james-knows-it-wont-come-easy-with-lakers/	https://nypost.com/metro/
https://nypost.com/2018/04/04/lebron-saban-feud-over-shop-talk-show-in-copyright-dispute/	https://nypost.com/sports/
https://nypost.com/tag/los-angeles-lakers/	https://nypost.com/2018/10/13/100m-cybersecurity-center-slated-to-open-in-soho-next-year/
https://nypost.com/2010/07/09/	https://nypost.com/2018/10/15/judge-tosses-stormy-daniels-defamation-suit-against-trump/
https://nypost.com/2018/07/02/	https://nypost.com/2018/10/15/beach-house-still-standing-after-hurricane-michael-devastates-nearby-homes/

https://nypost.com/2018/10/10/klay-thompson-gushes-over-lebron-as-lakers-pursuit-looms/	https://nypost.com/2018/10/15/ex-nightlife-king-plans-to-open-a-private-club-in-noho/
https://nypost.com/2010/07/08/	https://nypost.com/2018/10/15/nyc-goes-entire-weekend-without-a-shooting-for-first-time-in-14-years/
https://nypost.com/2014/07/11/	https://nypost.com/2018/10/15/new-york-city-restaurant-scene-teeming-with-new-entries/
https://nypost.com/2010/07/01/	https://nypost.com/2018/10/15/knicks-season-preview-what-theyre-really-playing-for/
https://nypost.com/2010/07/07/	https://nypost.com/2018/10/12/little-kids-are-wearing-grown-up-makeup-is-it-ok/

5. 2018 World Cup:

Lucene	External Page rank
https://nypost.com/2018/07/12/theres-a-knicks-battle-and-bet-over-the-world-cup-final/	https://nypost.com/mobile-apps/
https://nypost.com/video/this-psychic-pig-just-predicted-the-world-cup/	https://nypost.com/real-estate/
https://nypost.com/2018/06/05/nycfc-cant-wait-to-get-monkey-off-their-back-in-u-s-open-cup/	https://nypost.com/tech/
https://nypost.com/2018/10/04/stricken-ryder-cup-fan-i-thought-id-be-trampled-by-tiger-woods-fans/	https://nypost.com/fashion/
https://nypost.com/2018/07/13/son-of-stanley-cup-winning-devil-chasing-nhl-dream/	https://nypost.com/video/
https://nypost.com/2014/06/12/	https://nypost.com/columnists/
https://nypost.com/2014/06/11/	https://nypost.com/business/
https://nypost.com/2018/10/04/capitals-stanley-cup-tribute-was-a-little-bit-petty/	https://nypost.com/contact/
https://nypost.com/2018/10/02/nhl-preview-these-stanley-cup-contenders-are-racing-the-clock/	https://nypost.com/photos/
https://nypost.com/tag/tiger-woods/	https://nypost.com/horoscopes/

6. North Korea

Lucene	External Page rank
https://nypost.com/2018/08/30/a-glimpse-at-life-inside-north-korea/slide-1/	https://nypost.com/real-estate/
https://nypost.com/2018/08/24/state-department-reportedly-blindsided-by-trump-canceling-north-korea-trip/	https://nypost.com/sports/
https://nypost.com/2018/06/07/former-pow-calls-on-trump-to-reclaim-us-ship-captured-by-north-korea/	https://nypost.com/2018/10/15/tom-bradys-postgame-words-to-gronk-are-an-nfl-nightmare/
https://nypost.com/2018/10/07/pompeo-progress-made-with-kim-jong-un-on-north-korea-trip/	https://nypost.com/2018/10/15/nfl-flameout-admits-to-having-sex-with-12-year-old-girl-police/
https://nypost.com/2018/09/10/north-korea-marks-the-hermit-kingdoms-70th-anniversary/slide-1/	https://nypost.com/2018/10/15/eli-manning-trade-could-save-him-and-the-giants-espn-qb/
https://nypost.com/2017/09/09/how-democrats-left-america-naked-before-north-koreas-nukes/	https://nypost.com/2018/10/14/aaron-hernandez-had-sexual-relationship-with-hs-quarterback/
https://nypost.com/2017/09/03/	https://nypost.com/2018/10/15/mlb-youngsters-bring-the-trash-talk-and-much-needed-fun/
https://nypost.com/2018/10/02/trumps-love-for-kim-is-undercutting-his-own-teams-necessary-toughness/	https://nypost.com/2018/10/15/empty-vows-alderon-back-stab-inside-mets-gm-search-mess/
https://nypost.com/2017/09/04/	https://nypost.com/2018/10/15/how-eddie-lampert-led-sears-into-bankruptcy-disaster/
https://nypost.com/2018/10/04/north-korean-hackers-linked-to-global-banking-heist/	https://nypost.com/2018/10/15/the-south-bronx-is-in-the-midst-of-a-development-boom/

7. Hurricane Florence

Lucene	External Page rank
https://nypost.com/2018/09/14/hurricane-florence-storms-north-carolina/slide-1/	https://nypost.com/real-estate/
https://nypost.com/2018/09/11/the-carolinas-brace-for-hurricane-florence/slide-1/	https://nypost.com/photos/
https://nypost.com/2018/09/13/hurricane-florence-closes-in-on-carolina-coast/	https://nypost.com/news/
https://nypost.com/2018/09/10/hurricane-florence-looks-wild-from-the-international-space-station/	https://nypost.com/2018/10/15/surveillance-video-shows-moment-before-model-shot-in-neck-crashes-car/

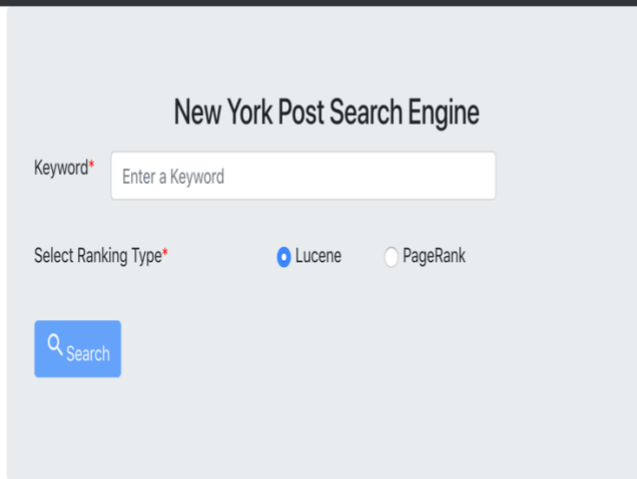
https://nypost.com/2018/09/20/unc-football-players-lend-hand-to-hurricane-florence-victims/	https://nypost.com/2018/10/15/beach-house-still-standing-after-hurricane-michael-devastates-nearby-homes/
https://nypost.com/2018/09/14/astronaut-snaps-picture-of-hurricane-florence-as-it-makes-landfall/	https://nypost.com/2018/10/15/ex-nightlife-king-plans-to-open-a-private-club-in-noho/
https://nypost.com/2018/09/15/kitten-named-survivor-rescued-with-owner-from-hurricane-florence/	https://nypost.com/2018/10/15/sears-death-of-an-american-icon/slide-1/
https://nypost.com/2018/09/24/hurricane-florence-isnt-done-with-the-carolinas-just-yet/	https://nypost.com/2018/10/15/heres-what-elizabeth-warrens-dna-test-shows/
https://nypost.com/2018/09/19/north-carolina-hog-lagoons-are-exploding-poo-after-hurricane-florence/	https://nypost.com/2018/10/11/michael-leaves-florida-panhandle-in-shambles/slide-1/
https://nypost.com/tag/hurricanes/page/2/	https://nypost.com/2018/10/10/hurricane-michael-starts-pounding-the-florida-panhandle/slide-1/

8. Paul Allen

Lucene	External Page rank
https://nypost.com/2018/10/15/microsoft-co-founder-paul-allen-dead-at-65/	https://nypost.com/tech/
https://nypost.com/2018/10/15/microsoft-co-founder-paul-allen-dead-at-65/	https://nypost.com/columnists/
https://nypost.com/2018/10/15/microsoft-co-founder-paul-allen-dead-at-65/	https://nypost.com/sports/
https://nypost.com/tag/paul-allen/	https://nypost.com/news/
https://nypost.com/2018/03/06/microsoft-co-founder-finds-long-lost-wwii-aircraft-carrier/	https://nypost.com/
https://nypost.com/author/carleton-english/	https://nypost.com/2018/10/15/microsoft-co-founder-paul-allen-dead-at-65/
https://nypost.com/tech/page/2/	https://nypost.com/2018/10/15/worlds-youngest-male-billionaire-is-25-year-old-model/
https://nypost.com/tech/page/3/	https://nypost.com/2018/10/15/sears-death-of-an-american-icon/slide-1/
https://nypost.com/tech/	https://nypost.com/2018/10/15/state-department-sends-out-email-featuring-cookie-monster-cat/
https://nypost.com/2018/03/06/microsoft-co-founder-finds-long-lost-wwii-aircraft-carrier/	https://nypost.com/2018/10/15/de-blasios-nycha-strategy-was-deceptive-from-the-start/

Screenshots that describes the whole flow:

1. Initial Page



The screenshot shows a web interface for the 'New York Post Search Engine'. It features a title 'New York Post Search Engine' at the top. Below the title is a 'Keyword*' label next to a text input field containing the placeholder 'Enter a Keyword'. Underneath the input field is a 'Select Ranking Type*' label with two radio button options: 'Lucene' (which is selected) and 'PageRank'. At the bottom left of the form is a blue button with a magnifying glass icon and the text 'Search'.

1. Search button will be enabled once system maps all the doc ids to urls from provided csv file and user enter valid key word.
2. Select one of the two Page Ranking algorithms(Lucene and Page Rank by NetworkX)

2. External Page Rank Results:

New York Post Search Engine

Keyword*

LA Lakers

Select Ranking Type*

Lucene

PageRank

Search

Results 1 - 10 of 4555

Query - **LA Lakers** 2 ms

Title : Sports | New York Post

URL : <https://nypost.com/sports/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/8790e35b-13e7-4277-abec-631c5ca93fd3.html

Description : N/A

Title : Patriots plans: Tom Brady tells Rob Gronkowski they'll never retire

URL : <https://nypost.com/2018/10/15/tom-bradys-postgame-words-to-gronk-are-an-nfl-nightmare/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/14965946-fc00-4d63-a25e-ea27f2efacfd.html

Description : It's not what the rest of the NFL wanted to hear. After the Patriots outlasted the Chiefs 43-40 in a shootout Sunday, Tom Brady wasn't ready to stop having big nights with his dominant ...

Title : Falcons flameout Justin Crawford admits to having sex with 12-year-old girl: police

URL : <https://nypost.com/2018/10/15/nfl-flameout-admits-to-having-sex-with-12-year-old-girl-police/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/a2308fa2-5653-4684-b4cc-c479d3a296b3.html

Description : A former standout college running back admitted to having sex with a 12-year-old girl, telling investigators that the sexual encounter was her idea, police said. Justin Crawford, 23, was spotted by...

Title : Giants should trade Eli Manning to Jaguars: Matt Hasselbeck

URL : <https://nypost.com/2018/10/15/eli-manning-trade-could-save-him-and-the-giants-espn-qb/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/c60c3f9d-f05a-4bc1-8d60-c257165e94c2.html

Description : Just what Eli Manning needs ... to be reunited with Ereck Flowers. ESPN analyst Matt Hasselbeck suggested Sunday that the Giants trade Manning to the Jaguars to give the 37-year-old another sho...

Title : Aaron Hernandez 'had sexual relationship with high school quarterback'

URL : <https://nypost.com/2018/10/14/aaron-hernandez-had-sexual-relationship-with-hs-quarterback/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/19f42538-046e-4566-a46f-86dbed6ef25d.html

Description : NFL star-turned-killer Aaron Hernandez hid a longtime sexual relationship with his high school quarterback, according to a new report that also said the troubled late footballer was physically and ...

Title : MLB youngsters bring the trash talk — and much-needed fun

URL : <https://nypost.com/2018/10/15/mlb-youngsters-bring-the-trash-talk-and-much-needed-fun/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/98ceb784-aae1-4a09-add5-d36fdc2c8866.html

Description : HOUSTON — Remember how much the Yankees struggled to fill the void left by Aaron Judge's injury? While it won't make Yankees fans feel any better, Alex Bregman has found a way to replicate one of J...

Title : Empty vows, Alderson back-stab: Inside Mets' GM search mess

URL : <https://nypost.com/2018/10/15/empty-vows-alderson-back-stab-inside-mets-gm-search-mess/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/a0fbd273-4856-450e-bec1-542a8a7ad8ec.html

Description : The most important questions the Mets should be asking in their interview process for a new general manager are to themselves. Are they sure they know what they are offering? Are they sure they kno...

Title : Angry about the subways? Send Andrew Cuomo a message

URL : <https://nypost.com/2018/10/15/angry-about-the-subways-send-cuomo-a-message/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/bf89ebee-3c4f-40cb-9cb7-19f7d2fb3514.html

Description : Are you happy with Gov. Andrew Cuomo's stewardship of the subways? More than a year after the state-run Metropolitan Transportation Authority launched its emergency Subway Action Plan, and nearly s...

Title : A Giants trade for ex-Patriots quarterback would make sense

URL : <https://nypost.com/2018/10/15/a-giants-trade-for-ex-patriots-quarterback-would-make-sense/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/b6fcb411-eac2-41f5-9d75-ff04f6f7f777.html

Description : At some point, the Giants will not be able to sit back and allow their quarterback situation to settle like all those leaves gently falling from the trees. They will need to shake that darn tree. T...

Title : How one veteran's Vietnam horrors inspired 'This Is Us'

URL : <https://nypost.com/2018/10/15/how-one-veterans-vietnam-horrors-inspired-this-is-us/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/32eaff2d-95a4-4e88-bc6d-10d35ee9fbf5.html

Description : Last season's "This Is Us" featured a few scenes of Jack Pearson's (Milo Ventimiglia) 1971 tour of duty in Vietnam. Tuesday night's episode will focus entirely on Jack's time in Southeast Asia. Ent...

Show More

3. Solr Lucene Results(Default):

New York Post Search Engine

Keyword*

LA Lakers

Select Ranking Type*

Lucene

PageRank

Search

Results 1 - 10 of 4555

Query - **LA Lakers** 13 ms

Title : [LeBron James knows it won't come easy with Lakers](#)

URL : <https://nypost.com/2018/09/25/lebron-james-knows-it-wont-come-easy-with-lakers/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/7856e10c-3cee-4e00-93e1-a8538d2c1be9.html

Description : EL SEGUNDO, Calif. — LeBron James realizes the size of the challenge he accepted when he joined the Los Angeles Lakers. After eight consecutive NBA Finals appearances, he knows his new team isn't #82...

Title : [Lakers plan to pursue Klay Thompson in free agency](#)

URL : <https://nypost.com/2018/09/07/warriors-star-emerges-as-lakers-flashy-plan-b-in-free-agency/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/dbae01ed-fc65-4a63-9fb8-e709fb2229c9.html

Description : The Lakers could already be plotting their Plan B for Kawhi Leonard. It's presumed the Lakers will have the new Raptors star on top of their wish list and are also high on Warriors sharpshoot...

Title : [Klay Thompson gushes over LeBron James as Lakers' pursuit looms](#)

URL : <https://nypost.com/2018/10/10/klay-thompson-gushes-over-lebron-as-lakers-pursuit-looms/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/4c86b1ae-2c2d-474f-aa94-87700240e99f.html

Description : The Lakers could be the one team that could lure Klay Thompson away from the Warriors. The Golden State sharpshooter, who hits free agency next season, showed his excitement about watching LeBron J...

Title : [Nba | New York Post](#)

URL : <https://nypost.com/tag/nba/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/27f1295b-466b-4636-b9b5-bc7c8df9e0c9.html

Description : N/A

Title : [Los Angeles Lakers | New York Post](#)

URL : <https://nypost.com/tag/los-angeles-lakers/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/121d01f1-64cb-4550-a6c0-605a11004d6b.html

Description : N/A

Title : [February 14, 2001 | New York Post](#)

URL : <https://nypost.com/2001/02/14/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/d90cee60-7435-4cdb-94d8-f6ac1cfbdadc.html

Description : N/A

Title : [May 2, 2014 | New York Post](#)

URL : <https://nypost.com/2014/05/02/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/636c3495-c1d9-4fc8-95e8-7a64938da7be.html

Description : N/A

Title : [November 12, 2012 | New York Post](#)

URL : <https://nypost.com/2012/11/12/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/1123e8f3-8a32-4fab-9f27-75827be2fb70.html

Description : N/A

Title : [June 26, 2015 | New York Post](#)

URL : <https://nypost.com/2015/06/26/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/235df70f-bdc7-4645-93bb-99407c30784d.html

Description : N/A

Title : [February 21, 2017 | New York Post](#)

URL : <https://nypost.com/2017/02/21/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/32e0bc21-1473-4d59-8938-7f0f84a4d05a.html

Description : N/A

Show More

Closer Look at one of the results

Query - **LA Lakers** 9 ms

Title : [LeBron James knows it won't come easy with Lakers](#)

URL : <https://nypost.com/2018/09/25/lebron-james-knows-it-wont-come-easy-with-lakers/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/7856e10c-3cee-4e00-93e1-a8538d2c1be9.html

Description : EL SEGUNDO, Calif. — LeBron James realizes the size of the challenge he accepted when he joined the Los Angeles Lakers. After eight consecutive NBA Finals appearances, he knows his new team isn't easy...

Title : [Lakers plan to pursue Klay Thompson in free agency](#)

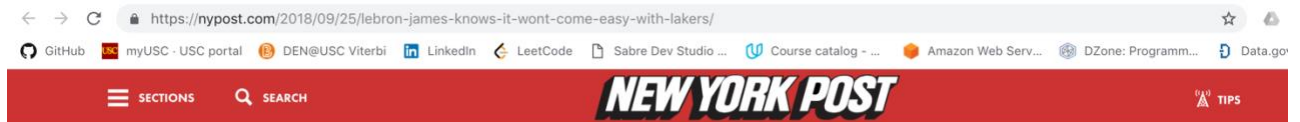
URL : <https://nypost.com/2018/09/07/warriors-star-emerges-as-lakers-flashy-plan-b-in-free-agency/>

ID : /Users/mukesh/Office/Tools/Solr/solr-7.5.0/nypost/nypost/dbae01ed-fc65-4a63-9fb8-e709fb2229c9.html

Description : The Lakers could already be plotting their Plan B for Kawhi Leonard. It's presumed the Lakers will have the new Raptors star on top of their wish list and are also high on Warriors sharpshoot...

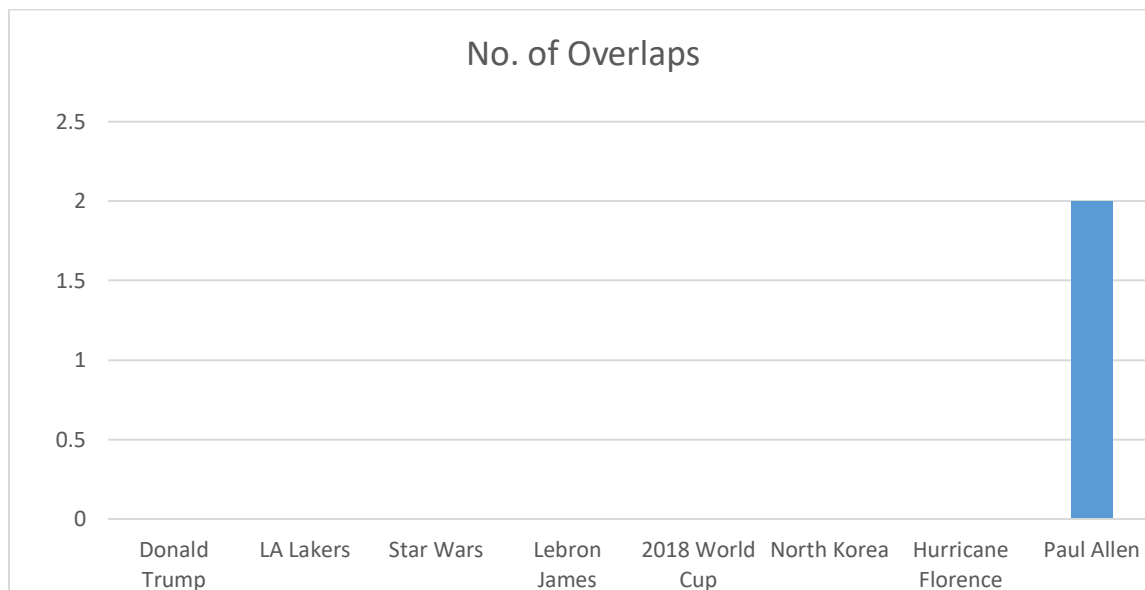
Title : [Klay Thompson gushes over LeBron James as Lakers' pursuit looms](#)

4. Actual Webpage (from La Lakers results, default page rank)



Query Response Overlap

#	Query	No. of Overlaps
1	Donald Trump	0
2	LA Lakers	0
3	Star Wars	0
4	Lebron James	0
5	2018 World Cup	0
6	North Korea	0
7	Hurricane Florence	0
8	Paul Allen	2



Instructions to execute the code:

1. Install NPM, NodeJS and Angular 7, Angular CLI, Git, Visual Studio Code, Nodemon
2. Git Clone : <https://github.com/mukeshkdangi/nypost-searchengine>
3. Run **npm install** to install all the dependencies from package.json
4. Run **npm start** to compile and run the angular code, default port:4200. This will trigger ng serve command with Proxy created: /api -> http://localhost:3000
5. Run **nodemon src/solrsearch.js** to run the NodeJs server, default port: 3000
6. Both Node and Angular are up on respective ports. Before running Node server, make sure to edit csv path in **src/solrsearch.js** from your local machine or any other server
7. Start your Solr using instruction mentioned on Step 1 and 2
8. I have included only **src** folder of the Angular Project for reference purpose. For execution purpose please clone the above mentioned git project.