

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

Step 1: Solr Setup and Indexing the HTML files

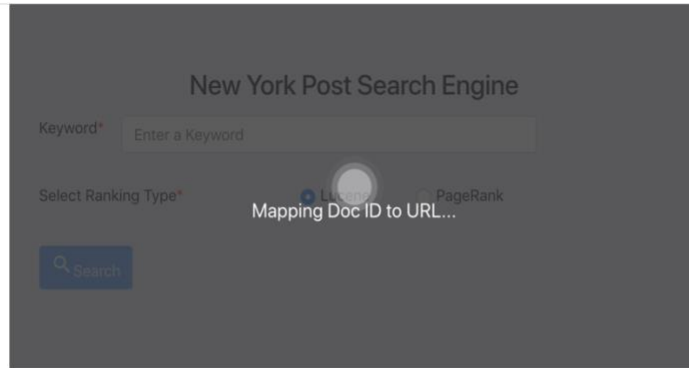
- a) Downloaded files NY Post URL to HTML mapped data and HTML pages from the google drive link. Downloaded and extracted all the crawled files from google drive for nypost.
- b) Downloaded the file solr-7.5.0 and Installed solr-7.5.0.
- c) Created a core by going into the /solr-7.5.0 directory started 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-linked 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
aea27c62-f6dd-4fb3-9c41-c212af5ff6e8.html	f5a772b8-662d-49b6-92f9-dcf15ffd388d.html
c3d89955-6d41-4fc4-879a-824ddee16e91.html	db9f279a-359a-43e2-bcd5-ad75becb6606.html
3ebce63e-fba7-403d-b0b8-50f1db9e67b4.html	40b8760d-53b6-4edb-a9b7-f60c3c13258d.html
1e666eac-ab65-4885-869e-3d057deeb87c.html	6bc5ab6f-4ca4-4588-a3a8-f9a6bb45f6ab.html
e8174513-85d2-4445-a934-aaa4188fbcc9.html	9e1975ef-08c7-45f3-8f77-120bff2a5b32.html
ed3379cc-3531-4936-9234-283f3709e755.html	fc17eac6-ad31-4c0c-b079-3bebbc17fcb9.html
d5d7f5c8-d1a3-450b-9028-30a4b94d846f.html	6dcc0483-03d1-4767-b2dd-d791046f94c1.html
abf7955b-595d-4889-9024-515c73918d28.html	9304c93b-bc36-494c-bd58-4eb8a6425a44.html
50a1285c-9953-42e0-9cd7-06fbf18c131e.html	88f9f9d1-491d-4df3-981d-863a8b1b485a.html
3d620c6c-2630-46b7-8128-3ebd1ff55525.html	e8940ca0-ff7c-4925-aaab-47e9f260d0c1.html

2. LA Lakers:

Lucene	External Page rank
7856e10c-3cee-4e00-93e1-a8538d2c1be9.html	8790e35b-13e7-4277-abec-631c5ca93fd3.html
dbae01ed-fc65-4a63-9fb8-e709fb2229c9.html	14965946-fc00-4d63-a25e-ea27f2efacfd.html
4c86b1ae-2c2d-474f-aa94-87700240e99f.html	a2308fa2-5653-4684-b4cc-c479d3a296b3.html
27f1295b-466b-4636-b9b5-bc7c8df9e0c9.html	c60c3f9d-f05a-4bc1-8d60-c257165e94c2.html
121d01f1-64cb-4550-a6c0-605a11004d6b.html	19f42538-046e-4566-a46f-86dbed6ef25d.html
d90cee60-7435-4cdb-94d8-f6ac1cfbdadc.html	98ceb784-aae1-4a09-add5-d36fdc2c8866.html
636c3495-c1d9-4fc8-95e8-7a64938da7be.html	a0fbd273-4856-450e-bec1-542a8a7ad8ec.html
1123e8f3-8a32-4fab-9f27-75827be2fb70.html	bf89ebee-3c4f-40cb-9cb7-19f7d2fb3514.html
235df70f-bdc7-4645-93bb-99407c30784d.html	b6fcb411-eac2-41f5-9d75-ff04f6f7f777.html
32e0bc21-1473-4d59-8938-7f0f84a4d05a.html	32eaff2d-95a4-4e88-bc6d-10d35ee9fbf5.html

3. Star Wars:

Lucene	External Page rank
ee74e3bb-ea2c-43c1-bc83-f7b6ace3b460.html	db9f279a-359a-43e2-bcd5-ad75becb6606.html
4491a6de-3d7b-4988-ab2f-94458c8f2046.html	40b8760d-53b6-4edb-a9b7-f60c3c13258d.html
ba259ebd-c57b-4cae-b905-ae1d739737b4.html	6bc5ab6f-4ca4-4588-a3a8-f9a6bb45f6ab.html
688728e6-c354-4dc7-98e7-f2ef4a2ab5d4.html	fc17eac6-ad31-4c0c-b079-3bebbc17fcb9.html
9ac36575-0add-47e9-a872-53e8eb628a32.html	88f9f9d1-491d-4df3-981d-863a8b1b485a.html
772ed371-1a0f-4052-a451-bdf4889fa554.html	e8940ca0-ff7c-4925-aaab-47e9f260d0c1.html
133dc680-99d6-4e55-9110-014b90469839.html	72897ce7-e9b1-489f-810d-bba4f5261645.html
d5278590-fcfd-4723-aadd-34c2e8f43711.html	8790e35b-13e7-4277-abec-631c5ca93fd3.html
9426d89a-91ad-4558-948c-3b1db285f7bb.html	2f65e692-a9f9-4f32-9591-198ecc3a0d9b.html
ba57043a-4b7d-4575-ab0b-0a9500edf8c4.html	08328725-0507-4b37-a111-b1eda05fd1fc.html

4. LeBron James:

Lucene	External Page rank
7856e10c-3cee-4e00-93e1-a8538d2c1be9.html	72897ce7-e9b1-489f-810d-bba4f5261645.html
3fc52765-999e-4a16-889e-6ac2b4a223d1.html	8790e35b-13e7-4277-abec-631c5ca93fd3.html
121d01f1-64cb-4550-a6c0-605a11004d6b.html	0152ac12-1507-4bbd-8471-ad819fb7a010.html
7a97608c-841f-4403-a782-e858a53fa23e.html	c6cf8280-d576-4ac1-9fb4-4291561097d9.html
571472d8-790a-43c4-9ba9-caa3c1eb0693.html	b2e7b626-de2a-4610-a9f8-127c85231751.html
4c86b1ae-2c2d-474f-aa94-87700240e99f.html	63481d75-f371-4ce9-98c6-189db19fd80b.html
4e8b8f21-979d-438d-990d-9a190cba5b70.html	f9594a03-4efc-46a1-9049-91482921def0.html
0bebe75e-be86-41b4-8082-947641dcbccf.html	3bc082a7-0349-469b-97b3-bc242a031422.html
920837b4-deb9-4c2f-af2e-b0ce334df29e.html	95052138-85f4-4972-88e8-5c2688c11783.html
1c1d3046-4401-40ad-ba10-6351044f881b.html	b640d196-95ef-4dcd-ba99-fef56657b0f4.html

5. 2018 World Cup:

Lucene	External Page rank
cbfe590e-bef5-4798-ac49-11e4062906df.html	f5a772b8-662d-49b6-92f9-dcf15ffd388d.html
8a52ca34-45d7-4e5a-836f-c0c1707da00e.html	db9f279a-359a-43e2-bcd5-ad75becb6606.html
40641f08-3f72-4af8-8eee-d802969aa528.html	40b8760d-53b6-4edb-a9b7-f60c3c13258d.html
57fcb62a-129a-4342-bc8a-a689f87239b0.html	6bc5ab6f-4ca4-4588-a3a8-f9a6bb45f6ab.html
8d871909-86a4-4638-a508-b94c9466ece2.html	9e1975ef-08c7-45f3-8f77-120bff2a5b32.html
d8dd177a-435d-40a9-ae34-6edb0d919c1d.html	fc17eac6-ad31-4c0c-b079-3bebbc17fcb9.html
95c15d1d-60b1-4b22-8bde-aec38b5c916b.html	6dcc0483-03d1-4767-b2dd-d791046f94c1.html
2369952c-dbf2-4352-a3ea-14a34b7fb131.html	9304c93b-bc36-494c-bd58-4eb8a6425a44.html
7ad33c91-d65b-4f1e-b48a-00d0e9635727.html	88f9f9d1-491d-4df3-981d-863a8b1b485a.html
5ad7e0c5-34ab-4b12-99bb-7bf7110d8c7a.html	e8940ca0-ff7c-4925-aaab-47e9f260d0c1.html

6. North Korea

Lucene	External Page rank
578d17b6-aae3-450f-9d00-e819317735b9.html	db9f279a-359a-43e2-bcd5-ad75becb6606.html
1b9dbf54-30f0-4089-8120-42e883ad3190.html	8790e35b-13e7-4277-abec-631c5ca93fd3.html
7338c1c4-f450-4cdc-b0f8-4785229f58f8.html	14965946-fc00-4d63-a25e-ea27f2efacfd.html
04e9fa64-57f3-46f2-958c-fc72e485180c.html	a2308fa2-5653-4684-b4cc-c479d3a296b3.html
ba10023b-02ef-4c41-bfa7-6c082bb6ccce.html	c60c3f9d-f05a-4bc1-8d60-c257165e94c2.html
c162ea51-5228-4e03-a3ae-94e7b410a1d1.html	19f42538-046e-4566-a46f-86dbed6ef25d.html
518b63b0-82e1-4282-91b4-d07f531fe49a.html	98ceb784-aae1-4a09-add5-d36fdc2c8866.html
280fbb2e-f7c6-4d95-afb8-fc495a6dad44.html	a0fbd273-4856-450e-bec1-542a8a7ad8ec.html
898fd909-cb83-4221-9065-a48cca028b48.html	ce677a5e-7c95-431c-96ac-5f4151e0940f.html
cc803d21-d568-467f-8b49-6516e07f2810.html	125a4ac4-54f1-47d3-a99e-dfdc86c714ea.html

7. Hurricane Florence

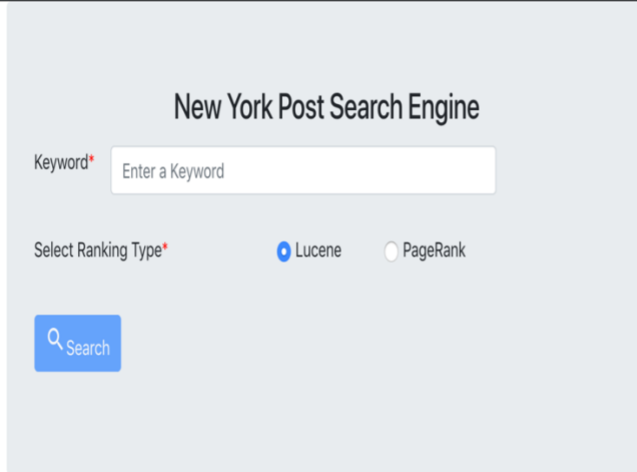
Lucene	External Page rank
c7ceb758-465e-41dc-8ef7-4fdaa80b0fa9.html	db9f279a-359a-43e2-bcd5-ad75becb6606.html
8eceb803-7343-498c-a53c-8bf25ff65b13.html	88f9f9d1-491d-4df3-981d-863a8b1b485a.html
fec5f5e8-2847-453a-a6bf-5b41312610bd.html	2f65e692-a9f9-4f32-9591-198ecc3a0d9b.html
9200d7aa-1f72-4cea-a0be-45c1bd1b5a18.html	9a67d0ce-ea79-476c-a348-551c7eeb7e9b.html
b44adbc3-a830-4c15-9387-b20a5970f2d5.html	b2e7b626-de2a-4610-a9f8-127c85231751.html
e54fe440-bcef-464a-b381-8865f08faec1.html	63481d75-f371-4ce9-98c6-189db19fd80b.html
6d7c7533-5c12-44ac-858c-c5cf3ebfa72e.html	2d83166a-b564-45a4-8641-36d04bb99766.html
5881054c-8610-417c-b1d2-228b07e3314e.html	e8b84f6a-5701-4283-8f52-1d947632760b.html
f68ee1fb-cd8d-4754-9399-b68323af283f.html	790f0b67-eaf1-446c-9452-81e078b0a16a.html
1a2fa94c-8709-4d3d-bb1d-a9f119637222.html	62d6d7c2-bba1-4c4e-9efa-a19d4748bf14.html

8. Paul Allen

Lucene	External Page rank
14c83721-0dc6-45ad-959d-c834b1917d8a.html	40b8760d-53b6-4edb-a9b7-f60c3c13258d.html
02423896-fc91-4f04-94dc-4bb9136ae283.html	fc17eac6-ad31-4c0c-b079-3bebbc17fcb9.html
d6304d6f-f459-4993-a533-c223b6d6be02.html	8790e35b-13e7-4277-abec-631c5ca93fd3.html
c5ca3dc2-3b72-4026-9be1-1f0c775e3777.html	2f65e692-a9f9-4f32-9591-198ecc3a0d9b.html
cfe9b62f-c461-42f4-9e49-1fead46936f8.html	dc52408d-235a-4535-a9bd-4b08b1c777b0.html
971ee8c4-34fd-4ff2-a3fa-84fc9a69b49e.html	d6304d6f-f459-4993-a533-c223b6d6be02.html
c0965d22-7ff1-46dd-8995-91f53fb5106f.html	7fbf6e8f-6f2a-49f2-9ae1-d4e195cfd49f.html
8cfd526a-8915-4966-b7de-d30813b0e4ce.html	2d83166a-b564-45a4-8641-36d04bb99766.html
40b8760d-53b6-4edb-a9b7-f60c3c13258d.html	7f94956c-c400-42e8-9641-c945ea9e07c5.html
6a5ec53a-e9eb-43e9-93db-f5f7cd5413d0.html	96510681-810e-4c59-902c-a7ac8d741e6b.html

Screenshots that describes the whole flow:

1. Initial Page



The image 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 followed by 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 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](#)

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

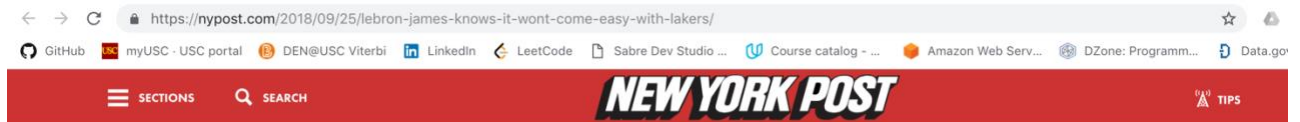
URL : <https://nypost.com/2018/09/07/warriors-star-emerges-as-lakers-flashy-plan-b-in-free-agency/>

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...

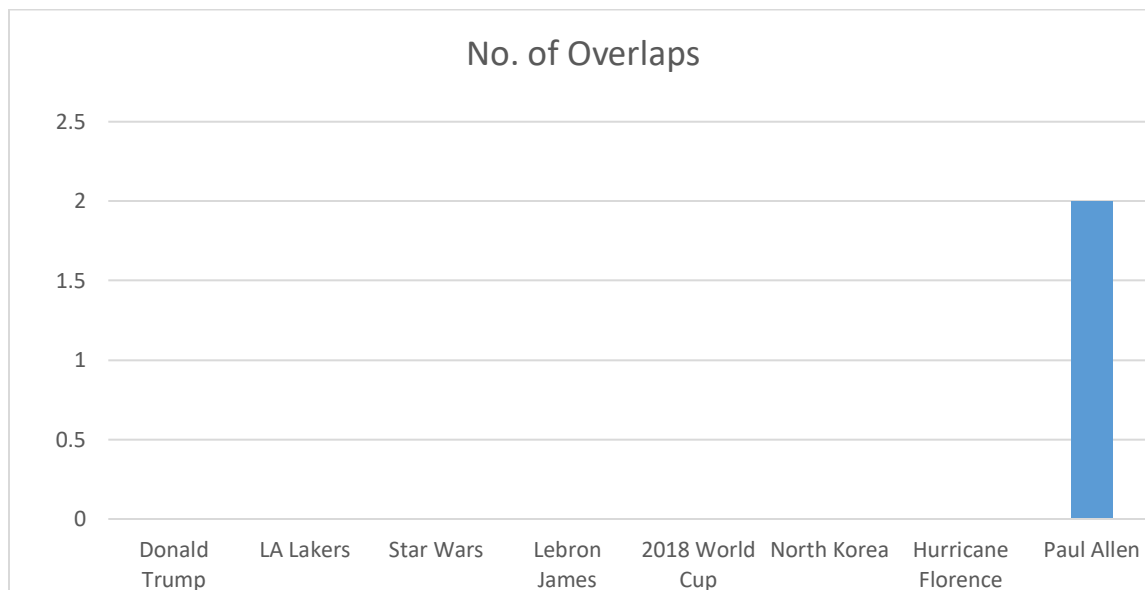
Title : Klay Thompson gushes over LeBron James as Lakers' pursuit looms

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.