

Search Engine Using R



Team Members

Omkar Nandkumar Dixit

ond170030

Dhwani Raval

dsr170230

Karan Yashwant Kanani

kyk170030

Introduction

A search engine is a software, typically accessed on the Internet, that searches a database of information according to the user's query. The engine provides a list of results that best match what the user is trying to find.

TF-IDF

TF-IDF, short for term frequency - inverse document frequency, is a numerical statistic that is intended to reflect how important a word is a document in a collection or corpus. It is often used as a weighting factor in searches of information retrieval, text mining, and user modelling. The tf-idf value increases proportionally to the number of times a word appears in the document and is offset by the number of documents in the corpus that contain the word, which helps to adjust for the fact that some words appear more frequently in general. TF-IDF is one of the most popular term-weighting schemes today, 83% of text-based recommender systems in digital libraries use tf-idf.

Files

searchEngine.R

Implementation Details

Libraries: [™], dplyr, tidytext, wordcloud, SnowballC

Scripting Language: R

Screenshot

The screenshot shows a web application interface with a search bar at the top right. Below the search bar, there is a table of search results. The table has three columns: 'doc', 'score', and 'text'. The results are sorted by score in descending order. A tooltip is visible over the first result, showing the full text of the document.

doc	score	text
758	0.3320578	The film stars Vincent D'Onofrio as "Bob", a cab drivin...
860	0.2827504	Homicide inspector Lindsay Boxer , teams up with
712	0.2819559	The Frozen Ground is based on the true story of
741	0.2406872	Actors: China Doll;Carmel Rose The film tells thre
1258	0.2011886	In New York City during the middle of a hot summ
710	0.1920929	My Apocalypse is the surrealistic story of Stewart
318	0.1814335	The director of an asylum offers to the serial kille
1223	0.1770895	Prospector Juan Valdez is murdered by The Cactu
525	0.1553519	The life of a professional killer becomes complicat
656	0.1153786	Big was a Buddhist Monk, but he gave up this occupat...

Showing 1 to 10 of 10 entries

Console Terminal x

```
~/   
Press 'Q/q' to exit. Enter your Query :: serial killer  
Press 'Q/q' to exit. Enter your Query :: |
```

MovieSearch.R

results[1:10,]

Filter

	doc	score	text
841	24155746	0.1670865	A series of murders in the West End of London baffle t...
117	17477023	0.1474779	Albert Dehousse has grown up on heroic novels, unfor...
429	248613	0.1426928	During both World War I and World War II, the Women's ...
609	23449178	0.1210858	Two swindlers meet and fall in love in Gangwon Provinc...
1079	26187979	0.1206325	A small-town policeman works to uncover the truth be...
1159	20067217	0.1122264	In the times of austerity after the Second World War a ...
215	36388452	0.1119345	After opening with newspaper articles, footage
1391	10068547	0.1095739	In World War II, Japanese troops occupied Sout
275	32817844	0.1060122	On Armistice Day, some years after World War
1229	32987372	0.1011271	Two friends who live near the Murrumbidgee River enco...

Showing 1 to 10 of 10 entries

Console

Terminal

```

~/
Press 'Q/q' to exit. Enter your Query :: war crimes
Press 'Q/q' to exit. Enter your Query :: |

```

MovieSearch.R

results[1:10,]

Filter

	doc	score	text
1076	14481527	0.2267376	The Invisible Children of Love is a love story set agains...
356	10387338	0.1795775	It tells the story of a secretary who married with her b...
1278	19631398	0.1672057	The story of a mysterious woman named Laura, who e...
672	31713107	0.1608415	Inspired by the "I Love You" computer virus, the film tel...
496	6224371	0.1526267	The story takes place in ol
312	7256290	0.1177350	The story is of an English r
1024	5984222	0.1140465	The story follows a detecti
993	33449153	0.1093411	The film tells story of a Kor
798	10855054	0.1055996	The film tells the story of t
411	25230351	0.1054340	Jyothi , a teenage girl who is in love with Ravi , suddenl...

Showing 1 to 10 of 10 entries

Console

Terminal

```

~/
Press 'Q/q' to exit. Enter your Query :: love stories
Press 'Q/q' to exit. Enter your Query ::

```