

Web Site Ranking Feedback System

Shaik Muzammil*, Sai Kiran Yerramaneni

Department of Computer Science and Engineering, Vasireddy venkatadri Institute of Technology, Guntur,
Andhra Pradesh, India

ABSTRACT

We all search on google for something and get the results in the form of different websites with some description. We generally click on first or second website links if results are not found we go down on google page. The website ranking is given by search engines by different criteria. Last website won't be seen by none of the people in most cases and first website will be having a great market compared to last website. So we need to help the last website to be moved up in results and help in generating revenue and have good rank on searching by giving feedback. This system will provide the difference between the first website and last website on the google results and will provide the feedback to the last website like content, links, images used by first website which helps the last website to be used in his webpage. This system is user friendly which is built on HTML as front end and Python flask as back end and used python package beautiful soup to parse HTML data and to automate browser behaviour with python. This system is done on web mining which has three categories firstly web content mining in which we scan the web pages and get to know the links, text, images used. Secondly web usage mining in which reports are generated after analysis which contain the details of text, images, links. Finally the web structure mining states that structural summary of website.

Keywords : Beautiful Soup Package, Web Crawling, Feedback, Python- Flask, Searching Results, Web Mining, Website

I. INTRODUCTION

World Wide Web is a major source to retrieve information. It has very large amount of data. For the user, it is a difficult task to find a particular page or a set of pages from this massive amount of data. It is highly impossible to remember the URLs of all these web pages. To solve this difficulty, search engine acts as an intermediary between the users and WWW.

The search engines are able to extract the information from the WWW for the users' given query. They are used to collect the web pages and display them to the user. The order of the URLs depends upon the rank given to the pages. To

perform this task, search engines use different page rank algorithms. Some ranking algorithms use content mining and some use web structure mining. But all the existing algorithms have some limitations. The system gives feedback to the results which was clicked by user here the first the user gives query and it was searched on google automatically and user have to choose the web links to show the difference between it and provide the feedback to last website to improve his ranking. In typical Search engine environment Web pages having matched keywords are fetched from the predefined database and ranked. Wide range of relevant Web pages having unmatched keywords are omitted in this procedure. Neglected Web pages having relevant information are covered

by the proposed method. Relevant Web pages are fetched and displayed irrespective of keyword matching.

Web Content Mining (WCM) concerns with retrieval of information from WWW in a structured form and thereby indexes the information for quick retrieval. It is the process to extract useful information from web documents contents. The documents consist of images, audio, video, text files, structured records etc. WCM technique is used on web documents and the results page obtained from a search engine.

II. METHODS AND MATERIAL

Web site ranking feedback system provides the user to increase their rank in the search results and user have to search the query as follows.

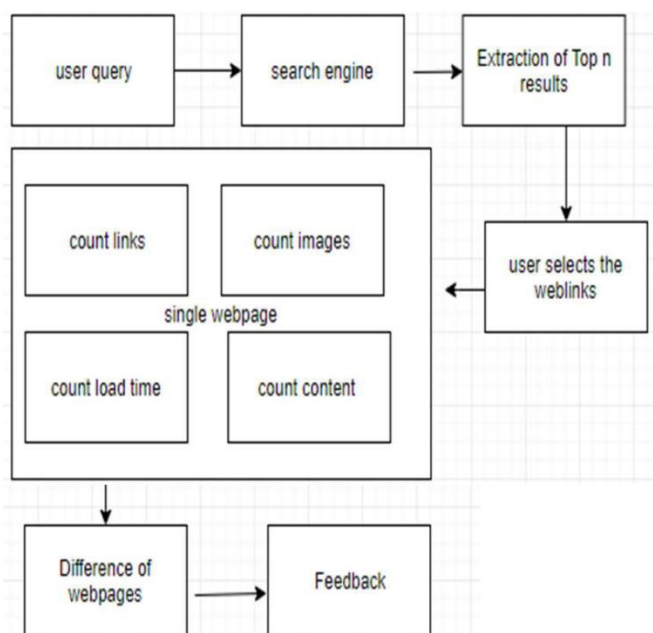


Figure 1 : workflows of system

User query is searched using google search engine and extract the top n results from search query and user selects the web links by checking the weblinks and there system will count the following factor like links, images, load time, content etc ,. Then comparison will take place here the feedback is

provided to webpage to increase its rank with respect to first page.

A. Implementation

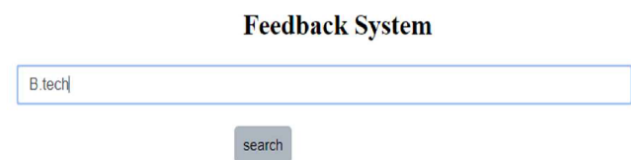


Figure 2 : Search bar to search queries

Here we can search for anything we get the weblinks what the google has described in the google search.

1	☒	http://home.bt.com/tech-gadgets/computing/how-to-perform-a-clean-installation-of-windows-10-11364000633913
2	☐	https://www.windowcentral.com/how-to-clean-installation-windows-10
3	☐	https://www.youtube.com/watch?v=89-OT_hsyEk
4	☐	https://www.youtube.com/watch?v=SKbR6XT7fcA
5	☐	https://www.wikihow.com/Install-Windows-10
6	☐	https://www.microsoft.com/software-download/windows10startfresh
7	☐	https://www.howtogeek.com/197559/how-to-install-windows-10-on-your-pc/
8	☐	https://www.expertreviews.co.uk/software/operating-systems/1401617/how-to-clean-install-windows-10-and-create-boot-media/page/0/1
9	☐	https://www.itpro.co.uk/operating-systems/25067/how-to-download-and-install-windows-10
10	☒	https://www.digitalcitizen.life/how-install-windows-10

Figure 3. List of weblinks

Here we are get the list of weblinks which user searched and here user can select the weblinks why their link was least when compare to others.

The weblinks will be crawled and we will find the no of links, images, page load time and found the content of web page and keywords and find the social networks present in the web page and some other factors which help the web page to improve his ranking and if the new one want to design a web page he can use this system how the content should be and how the factors to be taken in his web page to have a better ranking in search results.

https://www.sharda.ac.in/programmes/btech-mba-dual-degree	[34, 0.066697344779960382, 255]
https://www.sharda.ac.in/course/b-tech-mba-integrated-14	[34, 0.03018808364860344, 255]
https://www.ipu.in/programmes/engineering/integrated-b-tech-mba-in-computer-science	[19, 0.14443375680114746, 353]
https://careernobindia121.com/career-institutes/list_colleges_institutes.asp?group=3&educat=B_Tech_M_Tech_Dual_Degree_Colleges	[47, 0.47324109077450513, 262]
https://www.quora.com/Why-do-people-do-MBA-after-B-Tech	[3, 0.7266077993500295, 129]
https://www.quora.com/Does-a-B-Tech-graduate-qualify-for-a-teaching-job-after-doing-B-Ed	[0, 0.587323472003174, 80]
https://www.quora.com/Can-I-pursue-B-Ed-after-B-Tech	[0, 0.5837490558624268, 119]
https://www.quora.com/What-is-the-scope-of-an-MBA-after-getting-a-B-Tech-in-mechanical-engineering	[0, 0.7091835618621826, 129]
https://www.quora.com/What-is-the-difference-between-integrated-M-Tech-and-dual-degree-B-Tech-M-Tech-courses-in-ITs	[0, 0.66799406551758, 91]

Figure 4 : Factors of different webpages

Here the factors are calculated for every user checked weblinks which helps the user why the least webpage are having the rank with first page.

Algorithm

1. Input: user query
2. Display: Top n google results
3. user: choose the weblinks
4. calculate the rank factors
5. Display: Difference of webpages
6. Display: Feedback to least factor of webpage

Here same as google search the user enters query in search bar and it display the top n results and user choose the weblinks to find why the rank is difference and later it find the rank factors and difference of webpages are made using the results of rank factor and feedback is given to webpage whose rank factors are low.

Algorithm :

calculate the rank factors

1. Take the first link
2. scrap the webpage
3. count the links
4. count the images
5. count the load time
6. repeat step 1 to second link
7. Difference of webpages using factors

Rank factors plays a major role in ranking of webpage here most of the users don't know what is it? First of all we need to scrap the webpage and then find page load time and count the images and count the web links for short urls and find the content it should be minimum of 1800 words in a webpage and there must be social connect in webpage which helps in increase the page rank.

III. RESULTS AND DISCUSSION

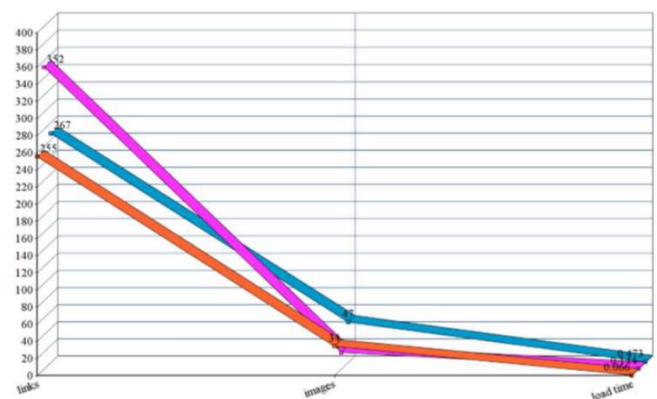


Figure 5. plotting the factors with webpages

Here the graph is drawn with factors with the different webpages which makes the user can identify the drawbacks of the least webpage.

We can easily identify the difference by using this graph why the web page having least rank and how to improve his ranking by using the first web page factors and he can have the good rating over the webpage and have the better income.

IV. CONCLUSION

Web search engines are generally designed to provide services to all users without taking consideration of the user interests. Personalized web search includes an individual user's interests when determining relevant results to return. The main focus of project is to increase the rank of last webpage of the google result with the help of feedback given to the user

checked query. Which helps the others who are designing the web page what to put in the webpage and how the content resides on webpage. For the future work, consider the SEO tools and machine learning techniques for ranking the web page ranking not by factors.

V. REFERENCES

- [1]. Gyanendra Kumar, Neelam Duhan, A. K. Sharma "Page Ranking Based on Number of Visits of Links of Web Page" 2011 2nd International Conference on Computer and Communication Technology (ICCCT-2011)
- [2]. S. Geetharani, Dr M. Soranamageswari "Locationbased Ranking Method (LBRM) for Ranking Search Results in Search Engines" at 2016 10th International Conference on Intelligent Systems and Control (ISCO)
- [3]. Nagappan V.K, Dr. P. Elango "Agent Based Weighted Page Ranking Algorithm for Web Content Information Retrieval"
- [4]. Nandnee jain, Upendra Dwivedi "Ranking Web Pages Based on User Interaction Time" at 2016 10th International Conference on Intelligent Systems and Control (ISCO)
- [5]. Tuhena Sen, Dev Kumar Chaudhary, Tanupriya Choudhury "Modified Page Rank Algorithm: Efficient Version of Simple Page Rank with Time, Navigation and Synonym Factor" at 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)
- [6]. Ricardo Baeza-Yates and Emilio Davis ,"Web page ranking using link attributes" , In proceedings of 13 international World Wide Web conference on Alternate track papers & posters, PP.328-329, 2004.
- [7]. Harmunish Taneja and Richa Gupta "Web Information Retrieval using Query Independent Page Rank Algorithm",

International Conference on Advances in Computer Engineering, IEEE, 2010

- [8]. Frank McSherry "A Uniform Approach to Accelerated PageRank Computation", ACM, 2005

Cite this article as :

Shaik Muzammil, Sai Kiran Yerramaneni, "Web Site Ranking Feedback System", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 5 Issue 1, pp. 443-447, January-February 2019.

Available at doi :

<https://doi.org/10.32628/CSEIT195135>

Journal URL : <http://ijsrcseit.com/CSEIT195135>