

SEO Tool

¹Yash Jeurkar, ²Dnyaneshwar Chavan ³Mohan Tikande, ⁴Faizan Sayyed
¹Student, ²Student, ³Student, ⁴Student
¹Department of computer Engineering
¹Mathoshri college of Engineering and Research centre, Eklahare, Nashik, India

Abstract— Now a day's to increase the business website is must. Website takes business at global level. Website should get popularity and high rank in search result given by search engine. Search engines like Google, Bing, Ask, Yahoo etc. Use the procedure to decide the rank of websites. If Website comes at high priority in search results then website will get more visitors that help us to expand the business. It is very difficult task to increase website rank in SERP (Search engine result page). SEO is well defined complicated process which is used for improving website ranking in search engines and also helps to increase traffic to website using search engine. The SEO tool addresses or helps in resolving issues related to SERP by generating standard SEO report of website. Report will suggest exactly where improvement is needed to rank high in SERP. For more improvement schema creator and rich snippets are provided in tool. Our proposed system will take input URL of website which has to optimize and keyword related to area of website which are used by web user. For that purpose we used certain technique like text Filtering, Stemming and algorithm relevancy and keyword frequency based approach.

Index Terms— Google page Ranking, SERP, Schema, and Rich Snippets.

I. INTRODUCTION

The main aim of the project is improve the visibility of webpage in SERP (Search Engine Result Page). SEO is well define complicated process used to improve webpage ranking in Search Engine like yahoo, Bing, Google etc. SEO Tool is application which will help to web user to improve the ranking of website in SERP. SEO is technique which helps to search engine to improve rank of websites which comes in higher than millions of other sites. SEO helps to recover the traffic of search engine. SEO is all about to optimizing the website. With this SEO reduce the human efforts and it takes less time. SEO works on some factors like keywords, back links, visitors, Meta tags etc. The main purpose of SEO is to study the content of website. Those contents are improved by SEO. It is possible for only visible websites. SEO is well defined complicated process which is used for improving website ranking in search engine and also helps to increase traffic to website using search engine. Search engine resolved the issues related to SERP by generating standard SEO report of website. Report will suggest exactly where improvement is needed to rank high in SERP. For more improvement schema creator and rich snippets are provide.

II. LITRATURE SURVEY

A literature review is a critical and an evaluative summary of the themes, issues and arguments of a specific clearly defined research topic obtained from the published (and unpublished) literature.

Search Engine Optimization task has been attempted through many different approaches like Page Ranking, word frequency count, text pre-processing, web based stemming, rich snippets, schema creator etc. Page ranking techniques give brief idea about how websites are ranked in SERP. But these page ranking techniques are not sufficient for improving website rank. For obtaining high rank in SERP you should not only focus on Google's page ranking but also you should develop or modify website such a that it will pass through all page ranking criteria and rank high in SERP. Sagar Pandit, Toshi Jain in '*Study of search engines that travel World Wide Web*' clarified the concept of Raking Technique, Search Engine Optimization technique [3]. This Paper gives knowledge of various page ranking technique but this paper have limitation over knowledge of Search Engine Optimization. Later, Prof. Punit Patel, in '*Research of Page ranking algorithm on Search engine using Damping factor*' studied concept of Page rank, web mining which are helpful for Improving Page Rank but gain there are limitation the concept of Search Engine Optimization factors, how to do analysis of website [1]. In next research of Vignesh. J, Deepa. V. in '*Search Engine Optimization to Increase Website Visibility*' the concept of Ranking Metrics and the Search Guidance for Learning Object Repositories, web content analysis are cleared. The paper also gives knowledge of fetching web content from sever side and analyse the content but this paper also have limitation this system does not give idea about schema creator, rich snippets [4]. These concepts got very much cleared in Prof Nilima V. Pardakhe's '*Enhancement of Web search engine Results Using keyword Frequency Based Ranking*'. Techniques like MCM WUM WSM are focused on in this paper. The paper also helps to find Keyword and help to add schema definition and rich snippets [2].

III. EXISTING SYSTEM

While studying existing system, we come to know that there are various tools available to improve web page ranging. Existing systems are only focus on structured SEO, but there are many things that effect on SEO that dropped down website rank. There are various systems for improving web page rank but they are not covered all aspect that need for improving ranking of webpage in SERP. For this we need SEO experts. Cost and time of SEO expert is Equal to cost and time of website development or sometimes more than that and it is no guarantee of SEO expert will implement this process perfectly. It effects on SERP ranking of website. More over these Tools do not work on the content of the website such as tags, frequency count of keywords in website. Comparison of content, tags with other top websites are also not provided in these tools. The cost of these tools is very high.

IV. PROPOSED SYSTEM

The SEO tool is very useful in area of Search Engine Optimization. SEO Tool accepts input as URL and keyword this input is forwarded to web content analyser, web content analyser will fetch the given website from web server and also fetch website that are optimized for given input, then applying HTML parsing HTML source code is converted into JAVA code .After applying HTML parsing machine learning techniques for text pre-processing are applied. Text reprocessing mainly content stemming in which keyword are processed. Tokenizing in which tokens are created for given data. Grammar recognizing in which semantic words are processed. A chart is generated in form of keyword and token for both given URL and also optimized URL for given input. This chart is then forwarded to SEO suggestion. In SEO suggestion appropriate steps are generated like where website is lagging in website where contents are less. According to that SEO suggestion are generated like where to add content or where to remove contents. This is also converted to standard report by referring this report we can do SEO for website. For further enhancement of website features like schema creator, rich snippets are also provided in this SEO tool. Schema is what just a small piece of code which is adds to HTML. It indicates that certain page or paragraph is actually what about. It is a new way which we use for further enhance of Artificial intelligence for search engine. Rich snippets is design to summarized the content of web page and understand what page about the web result. This depth information provided to Webmaster.

A. System Architecture

In the design phase the architecture is established. This phase starts with the requirement document delivered by the requirement phase and maps the requirements into architecture. The architecture defines the components, their interfaces and behaviour. The deliverable design document is the architecture. The design document describes a plan to implement the requirements. This phase represents the "how" phase.

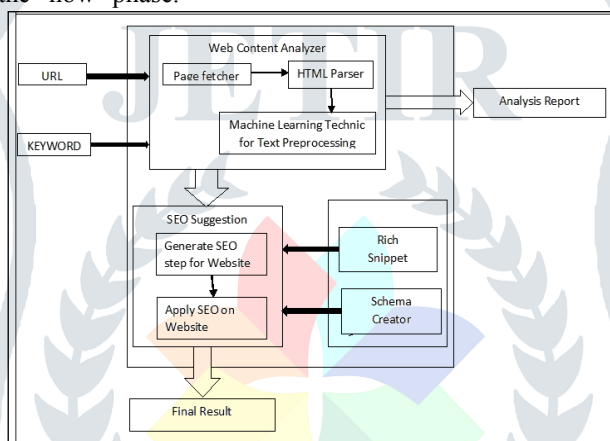


Fig. 1. System Architecture

B. Modules of the System

1) *SEO Input:* In this module URL and keywords for which website is analysed are given as input.

2) *Web Content Analyser:* This Module performs three functions. Those are given below According to given URL web Page fetcher fetches the HTML web page from the server.

- HTML Parser parses the HTML content to JAVA content.
- Machine learning technique for text processing involves semantic matching, removing meaningless keyword, remove suffix, count the keyword frequency, identify exactly where keyword is appearing in tags.

3) *Result Analyser:*

Result Analyser performs the main task of producing result. Generated result is in two manners tag wise and keyword wise.

- Tag wise result show exactly in which tag keyword is appeared and how many time it is appeared in that tag. Result is in tabular form as shown in fig (2).

Tag/Keywords	engineering	college	nashik
Title	1	1	1
Body	2	1	0
Div	1	0	0
Img	0	0	0
Ul	1	1	0
Span	0	0	0
Script	0	0	1

Fig. 2. Tag wise keyword count

- Keyword wise result shows how many time keywords is appeared in page. i.e. keyword frequency count. Result is in tabular form as shown in fig (3).

Keywords	Count
Engineering	5
College	3
Nashik	2

Fig. 3. Keyword wise frequency count

4) Advance Services

Some advance services are also provided in SEO Tool to obtain high rank in Google. Those are as follows

- Schema Creator
- Rich Snippets

The purpose of Rich Snippet is to present user with more information about the content that exist on a page so that they can better decide which result is more relevant for their query. This may result in additional traffic to the website.

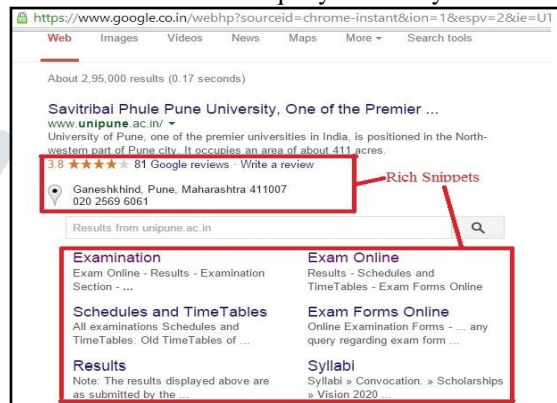


Fig. 4. Rich Snippet and Schema Creator

5) Final Result

Final result of the system is suggestion list, which will suggest where in which Tag keywords should be added or removed, on which keyword work should be done.

V. ALGORITHM

An Algorithm is a step-by-step procedure for calculations. Algorithms are used for calculation, data processing, and automated reasoning. More precisely, an algorithm is an effective method expressed as a finite list of well-defined instructions for calculating a function. Starting from an initial state and initial input (perhaps empty), the instructions describe a computation that, when executed, will proceed through a finite number of well-defined successive states, eventually producing "output" and terminating at a final ending state. The transition from one state to the next is not necessarily deterministic; some algorithms, known as randomized algorithms, incorporate random input.

The algorithms used in the system are as follows:

1. Keyword Searching

Fetch the page on web server

1. Start
2. The input of the web server is:
 - a. URL name: Name_of_URL="URL_name"
 - b. Keywords: char Keyword={ K1,k2,..... Kn };
 - c. Optimized URL from the web server that knows the content of that websites.
3. In next steps the web server are fetch the all content of the websites.
4. Fetched content of the website are optimized URL for given keyword.
5. Store the content in tag of websites in data structure.

Tag (t1, t2, t3, t4 tn)
6. Using this tag content or URL page content pre-processing can done. In pre-processing the ungrammatical word (in URL) can be removed used by next step.
7. Generate the tag wise frequency count of given URL in the form of table.
8. Generate the tag wise frequency count of given Keyword in the form of table.
9. Compare frequency count table of given URL and frequency count table of optimized URL.

If given URL frequency count is less than optimized URL then improve the content of given URL.
10. Generate the SEO suggestion for improvement of given website.
11. Stop

2) Stemming Keyword modification algorithm

Stemming is the process of textual data preceding, text mining and natural language. This algorithm changes the different grammatical word to natural word for example CEED -> Used for carried out on same steps e.g. succeeds replaced by success, ING -> to remove the ing for the steps e.g. scoring replaced by score many more grammatical word are removing by the stemming. Let us discuss the algorithm of stemming.

1. Start
2. Enter Grammatical Keywords (K1, K2, K3.... Kn).
3. Select ending characters of word.
4. Examine the final letter.
5. Check applicability of rule of stemming word.
6. The ending characters are not match by the stemming word then words are not replaced by the natural language.
7. Apply rule.
8. Delete the end character of keyword and replaces by satisfied word you want.
9. The expected output of keyword is given then terminates the process of stemming.
10. Stop

VI. CONCLUSION

Thus system SEO Tool will be design for improve the webpage rank in SERP, which take website URL and Search keyword are inputs and give the report contain SEO suggestion step as output that help to user improve website ranking.

ACKNOWLEDGMENT

We take immense pleasure in thanking our college management for having permitted us to carry out this project work. Words are inadequate in offering our thanks to our respected co-operation in carrying out project work in the starting phase of the project like requirement analysis and literature survey. We also express our at most gratitude to **Dr. V. H. Patil** Head of department of Computer Engineering, for this valuable co-operation in selecting the project topic and guiding us stepwise ahead. Inspiration and guidance are invaluable in every aspect of life, especially in field of education, which we have received from our respected project guide **Mr. G. D. Nikam** who has helped a lot In first two phase of project like topic selection then information gathering and guide us to throughout the project work and gave earnest co-operation whenever required. We would like to express sincere gratitude towards him. At last, we would like to take this opportunity to convey thanks to all my staff members, who directly or indirectly encouraged and helped us project work. Finally yet importantly, we would like to express heartfelt thanks to our beloved parents for their blessings, our friends/classmates for their and helps and wishes.

REFERENCES

- [1] Prof Punit Patel, in “*Research of Page ranking algorithm on Search engine using Damping factor*” International journal of Advance Engineering and Research Development (IJAERD) Volume 1 Issue 1, February 2014, ISSN: 2348 – 4470.
- [2] Prof Nilima V. Pardakhe’s, “*Enhancement of Web search engine Results Using keyword Frequency Based Ranking*”, IJCSMC, Vol. 3, Issue. 5, May 2014, pg.395 – 403.
- [3] Sagar Pandit, Toshi Jain, Rimpal chugga, Pargya Bagti in “*Study of search engines that travel World Wide Web*”, International Journal Of Computer Architecture And Mobility (0012-0187), Volume 1-Issue 1, November 2012
- [4] Vignesh. J, Deepa. V. in “*Search Engine Optimization to Increase Website Visibility*”, International Journal of Science and Research (IJSR), Volume 3 Issue 2, February 2014.
- [5] <http://files.pxlpartner.ch.s3-eu-west-1.amazonaws.com/78746?CFID=70d45639-45bb-4b25-bafc-1ed30a946fb0&CFTOKEN=0>