

# Design and Implement Search Engine Optimization Prediction System Using Machine Learning

<sup>1</sup>Jui More, <sup>2</sup>Vaishnavi Gaikwad, <sup>3</sup>Jyoti Lahange, <sup>4</sup>Sayli Shinde,  
<sup>5</sup>Prof. J.Y. Kapadnis

<sup>1,2,3,4</sup>Dept. of Computer Engineering  
PVG'S College of Engineering And SSDIOM  
Nashik -422004.

**Abstract-** Site improvement (Web optimization) alludes to streamlining individual sites and pages to accomplish higher page positions in query items. Sites are generally streamlined for backlinks be that as it may, individual site pages are streamlined for explicit catchphrases. This paper proposes a structure in light of a bunch of rules for watchword examination and backlink age. The proposed structure depends on dataset prepared which recommends that page content ought to be watchword based and webpage traffic ought to be observed by referrals. Achieve better query items assuming your site content and title contain significant and applicable catchphrases and have the right number of backlinks to assist with checking your site traffic You can rank your site to for the system additionally stresses that engineers and fashioners ought to consider legitimate watchword choice and external link establishment while dealing with programming advancement projects.

**Key Words:** Search engine optimization, machine learning, SVM



Published in IJIRMP (E-ISSN: 2349-7300), Volume 11, Issue 3, May-June 2023

License: [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/)



## INTRODUCTION

In the modern era, information technology has facilitated transmission of multi-facet information across the world. Internet community has emerged as a broad society comprising billions of users who use Internet as a communication tool on daily basis. Internet has transformed into a business network where a huge number of traditional business transactions are performed online. A large number of people, known as e-community, use Internet for marketing and business purposes. There are millions of sites on the web and we heavily rely on the search engines for locating specific information and the targeted webpages. The discipline of Search engine optimization (SEO) started in the mid-1990s. Initially, SEO tactics were restricted to tagging the keywords and meta-keywords. The field of SEO started getting matured after 2003 when the strategies of anchoring text and interlinking pages of websites were introduced. Later in 2011, vertical search inclusion was introduced. SEO is analogous to software that gets user queries and returns the required information to the user. SEO is used as a tool to attain higher search rank during online searches for the targeted websites. According to comScore survey, the biggest Internet community is in Turkey where 80% of the population use Internet for marketing and business. In short, SEO mechanism is required to route the Internet search queries to the specialized websites having exclusive content. Hence, search engines play a critical role on steering users to the pertinent websites since 93% of Internet traffic is managed by search engines. SEO is meant for those organizations and companies which want to stay competitive on the web. Millions of websites compete

for placement in the search engines, therefore, only the websites improved and structured through SEO tactics receive increased visibility on the web.

## **EXISTING SYSTEM**

If we have any desire to find something on Google, we just have to enter our desired watchwords to look for, then Google will organize our query items as indicated by the most suitable catchphrases from the top. All the site proprietors maintain that their site should show up on the principal page of the google web crawler. Since, supposing that our site shows up on the principal page of the web search tool then there is a high chance that our site will be perused by many individuals. There are many reasons and benefits in the event that a site is on the web search tool page. They have no clue about that how to make arrangements for ideal Web optimization for their site, Client involves any sum for itself and sit tight for the outcomes, which is off-base.

## **OBJECTIVE OF SYSTEM**

1. Provide an Easy to use system for user to make a quick decision on SEO of website
2. Machine learning Algorithm SVM (support Vector Machine) for improving the accuracy.
3. Search engine optimization technique improvement for user
4. Allow user to rank its website on Top

## **LITERATURE SURVEY:**

“New Technique to Rank Without Off Page Search Engine Optimization” a paper of Asad Nadeem. A paper state that Over the years search engines have become more efficient in delivering relevant pin pointed results. For digital marketers that means more hard work because the actual metrics used to measure the content relevancy and quality are not yet disclosed by the search engines. The paper discusses a new technique named All in Title Query Result Analysis for identifying keywords with low competition. The posts written for such keywords would rank without any off-page search engine optimization. The research spans to almost 36 months of data taken from Google Analytics. The results show the supremacy of the new technique that can help new digital marketers or startups to launch digital platforms with reduction in the off-Site search engine optimization cost.

“Exploring the Effectiveness of Search Engine Optimization Tactics for Dynamic Websites in Sri Lanka” a paper of Ushadi Niranjika. A paper present the Visibility of a business, both in online and offline is important to grow businesses. In this modernizing digital era, businesses need to understand the significance of web presence in search engines through Search Engine Optimization (SEO) strategies where it can make better ROI compared to traditional marketing. Therefore, than the international methodical SEO strategies, locally tactics have to optimize singly. The purpose of this study is to explore SEO tactics for dynamic websites in Sri Lankan context & evaluating its effectiveness over international SEO practices in terms of Google.lk. Further, as the methodology survey questionnaire was distributed to 20 SEO experts & 100 internet users who supported to gather the experimental data for [www.officestationery.lk](http://www.officestationery.lk) web site. The research proceeded with simple random sampling technique and SPSS software used to analyze the data. Based on the SEO checklist it was implemented over the experimental website and proved that content optimization from on-page factors and backlinks from off-page factors are the underlying factors to rank in Google.lk. This also revealed that there is no sense of keep relying on outdated SEO tactics as Meta tags or deep technical factors to rank in search engines.

“Search Engine Optimization Using Unsupervised Learning” is a paper of Asad Nadeem. It state that, Nowadays, web has emerged as the most demanding tool for retrieving information over a large repository. As the amount of information on the world wide web grows, it becomes increasingly difficult to accurately find what we want. The existing search engines mostly display the content based on many factors and not just the quality of the content. These include sponsored links, advertisements, paid appreciation, etc. Our project aims at developing a tool to rank the search results solely on the basis of the content, and not by keeping in consideration that which article would the user be most likely to click. Hence there will be no problem of Clickbait. Therefore, the aim is to create a tool that could scan the web on a specific topic and create a synthesis of the content found. We do this by gathering search results from various search engines using a web crawler and processing the results obtained from them by using our custom made ranking algorithm which clusters

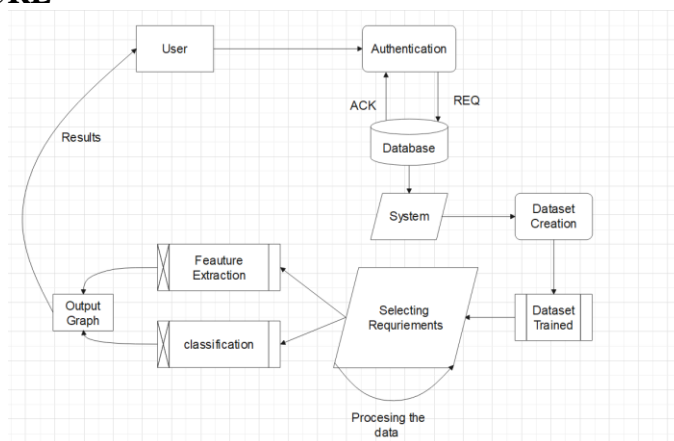
the results and ranks them on the basis of the content quality. After crawling through the web and retrieving the information, we will be using ‘term frequency – inverse document frequency’ as our weighting algorithm, followed by ‘singular value decomposition’, for decomposition of the weighted matrix. Lastly, we will be using ‘spherical K-means’ and custom ranking algorithm to display rich content. In order to give more efficient results, our project presents a new algorithm to rank web pages in accordance to the relevance of the user’s query.

Sales forecasting is an important aspect of different companies engaged in retailing, logistics, manufacturing, marketing and wholesaling. It allows companies to efficiently allocate resources, to estimate achievable sales revenue and to plan a better strategy for future growth of the company. In this paper, prediction of sales of a product from a particular outlet is performed via a two-level approach that produces better predictive performance compared to any of the popular single model predictive learning algorithms. The approach is performed on Big Mart Sales data of the year 2013. Data exploration, data transformation and feature engineering play a vital role in predicting accurate results. The result demonstrated that the two-level statistical approach performed better than a single model approach as the former provided more information that leads to better prediction.

### PROPOSED SYSTEM

- The dataset for framework and prepared the framework and make model for dissecting the prerequisite of client.
- Then we are taking the prerequisites of client to investigation the client keywords.
- Framework will then, at that point, give a shrewd Web optimization plan to client in form of report

### SYSTEM ARCHITECTURE



**Fig -1:** System Architecture Diagram

### IMPLEMENTATION DETAILS (Modules)

1. **Register and Login:** Here we are allowing user to register first to our system which will be a security protocol used by us.
2. **Dataset Creation and Trained:** This module is based on machine learning where we create the dataset for system and trained the system and create model for analyzing the requirement of user.
3. **Keyword Search:** Here we are taking the requirements of user input of keywords
4. **Processing:** Here we will do the Requirement extraction and matching it with the dataset trained model.

### ADVANTAGES

1. To keep improving in accuracy and efficiency. This lets them make better decisions.
2. As the amount of data you have keeps growing, your algorithms learn to make more accurate predictions faster.

## APPLICATION

- **Personal:** Our system will be used by person who need to have his business on top of Google web pages, which will help him to get more client and increase it revenue.
- **Organization:** Digital marketing organization can use our system for managing their clients and to generate leads for them.
- **Company:** Our system can use by single company which will generate leads for them and increase their revenue, and also flash there company website on the top of existing business.

## ALGORITHM/TECHNOLOGY

- **SVM**

Support Vector Machine (SVM) is a supervised machine learning algorithm used for both classification and regression. Though we say regression problems as well its best suited for classification. The objective of SVM algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points.

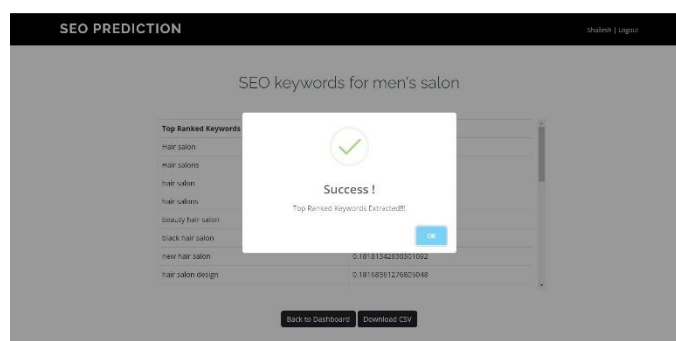
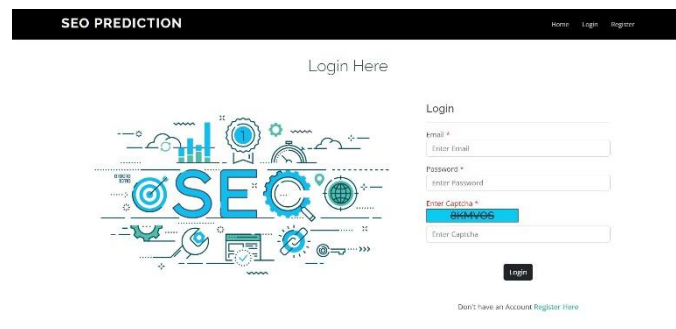
## RESULT



Search Keywords

SEO Keywords for ^  
SEO keywords for mens salon

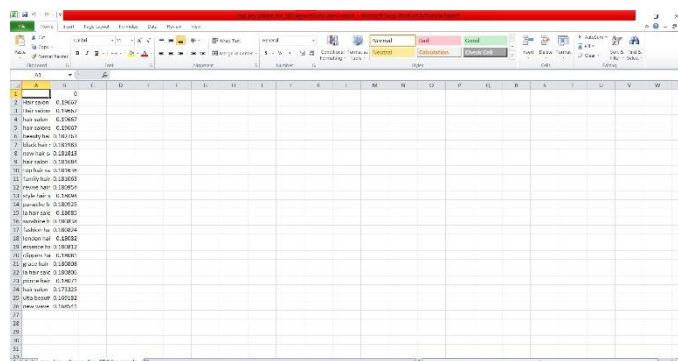
Search Keywords



### SEO keywords for men's salon

Top Ranked Keywords	Scores
hair salon	0.19657010829197142
hair salons	0.19657010829197142
hair salons	0.19657010829197142
hair salons	0.19657010829197142
beauty hair salon	0.18276270552154063
black hair salon	0.1815832104243108
near hair salon	0.1818134283031052
hair salon design	0.18158321278803048

[Back to Dashboard](#) [Download CSV](#)



## CONCLUSION

After doing research through qualitative methods by taking a few quotes from journals and international conferences about the effects of Search Engine Optimization (SEO). The conclusion is that there are many techniques that can be done to do SEO and the most important techniques are On-site Optimization techniques such as making headlines accurately, and Off-site optimization techniques such as backlinking. After implementing SEO, the effect that we will get such as increasing traffic on the website and make the website more popular

Hence, we are overcoming the drawback of exiting system and provide better solution in low cost.

## REFERENCES:

1. Mathematical Model For Prediction the Dynamis of Organic Traffic at E-commerce Web-site in the Process of its Search Engine Optimization ,2020
2. The Effect and Technique in Search Engine Optimization,2020
3. Using Machine Learning for Wed Page Classification in Search Engine Optimization,2021
4. Search Engine Optimization Using Unsupervised Learning, 2019
5. New Technique to Rank Without Off Page Search Engine Optimization, 2020
6. Exploring the Effectiveness of Search Engine Optimization Tactics for Dynamic Websites in Sri Lanka, 2019
7. Optimization of Information Retrieval Algorithm for Digital Library Based on Sematic Search Engine, 2020
8. Inspecting Engineering College Websites for Effective Search Engine Optimization, 2019