Website Design Ranker

Using Machine Learning

Adhyaksh Guhan - 7 , Anet Eliza Johny - 23 , Dharwish Raj - 47 ,
Joel J Padayattil - 60

Department of Computer Science and Engineering
FISAT

16 SEPTEMBER 2019

Introduction

- Our project is aimed at ranking websites in terms of its design which is evaluated baised on certain parameters.
- Since a perfect model for wedsite ranking is not in practice this follows ranking according to submissions by critics.
- Thus we compare ranking implementation done by Google that analyses a webpage's content.
- We will be comparing against:
- A Website that ranks other design according to submissions by critics
- A ranking implementation done by Google that analyses a web-page's contents

Problem Statement

- Our main problem is to evaluate website designs using an algorithm that uses machine learning
- It must take into account various parts of the website to use as parameters.
- For example: pepole won't be interested to visit a site if they are berated with ads.Thus ads above a degree will be considered as the parameter for ranking.
- There are no existing algorithms or methods that manage to analyze and evaluate websites in this way.

Why "Website Design Ranker" ?

- Website Design Ranker will rank set of input websites based on certain parameters.
- It will be helpful to find best website among list of websites which have same content.
- We can compare our website design with other competing websites.
- We can see how a website's design may improve in an area.

Related Works : Google Page Layout

- Google introduced Page Layout Algorithm to analyze website readability.[1]
- Looks for the layout of the web page and the amount of content we see in the page once we click on a result.[1]
- Focuses to reduce the difficulty of users to find the actual content.[1]
- The websites which does not have a lot of visible content above-the-fold and dedicates a large fraction (above a normal degree) to ads will be affected.[1]

Related Works : Google Page Layout

Figure: One of the criteria of GPL Algorithm[Fig:1]

Proposed System

- Website Design Ranker will rank set of input websites based on certain parameters.
- The parameters we are focussing on are color and grid .
- Then we move on for public review.
- This will be helpful to finding the best website among list of websites.
- We can compare our website design with other competing websites.
- We can see how a website's design may improve in an area.

Explanation

- It is an objective analysis of website designs by ranking them based on a parameter.
- The website's CSS file is scrapped via a web scrapper and then read through.
- The file is then parsed to find hex codes based on a regular expression.
- Once the codes are found, they are counted and printed using a variable.

Input Data

Algorithm

1. Start
2. Using a website scraper to accept the various website addresses
3. Scraping through the source code of each website via CSS files
4. Find the hex codes of all elements of the website and count them with a count variable
5. If count $==$ 0 then give mark as 0
6. else if count $>$ 5 then also give mark as 0
7. else if count $<=$ 5 then give mark as 1
8. Stop

Output

Testing

Testing was done in several stages of this project. It was helpful to determine suitable language and library to collect data from different websites to find out ranking

- A website template's 'style.css' file was downloaded and then used as subject to testing the colour counting code.
- A website was created to have public ranking of various websites with a possible ranking of 1-5 (1 being the lowest, 5 the highest). This could then be used as ranking data for the program.

Experimental Results

- Number of colours counted:

Compiled languages are languages typically processed by compilers, though theoretically any language can be compiled or interpreted. The important ones are:

- Ada
- C
- C++
- Fortran
- Java

Data Flow Diagram

Future Work

Conclusion

- Here we can see the logical differences in the approaches that our algorithm takes versus any existing methods.
- Our method relies on an objective and automated method that is consistent in nature as opposed to the subjective methods of the existing methods.

References

1 - Google Page Layout Algorithm: Everything You Need to Know "https://www.searchenginejournal.com/google-algorithm-history/page-layout/close"

Fig:1 - https://cdn.searchenginejournal.com/wp-content/uploads/2017/10/google-algorithm-above-the-fold-380x238.png