First International Conference on Computer Engineering
International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
Volume 5 Issue 8, November-December-2020

SEO Report Generator and Optimizer

Anshuman Vats[1], Kshitij Motke[1], Pragati Tamboli[1], Pranav Gholap[1]
[1]Department of Computer Engineering, Dr. D. Y. Patil School of Engineering, Lohegaon, Maharashtra India

ABSTRACT

In this project, we will be working on to create a fully automated SEO report generator based on the guidelines given by the search engines (Google/Bing/ Gropher) and create an indexing chart by going through the source code of the given website and ranking it in aspects of performance, SEO, best practices and availability.
The secondary objective of the project is to recommend keywords based on the given website description (meta description tag from HTML file). To create the report we will need to rank the result that comes up after searching someone's name or their website and categorize them into three categories Good, Bad, and Critical. These can be flagged to the administrator team for content removal. For categorization, we will be building upon the Compromise NLP engine based on the NODE JS environment.
Keywords: SEO, NLP, NODE JS, HTML

I. INTRODUCTION

The current process of SEO optimization and report generation is a manual process. When we search for their online reputation, the search result gets categorized based on textual context and the effect on one's reputation. If they have a personal website or organisation's website we go to that website and based on the search engine guidelines (Google/ Bing/ Gropher) we rank the website in various aspects. After collecting all the data a report is generated that is then delivered to the client.

Pain points:

- Manual process
- Report Generation is a repetitive task
- Data collection from various sources is time consuming

Our model will fix all these issues by automating time consuming and repetitive tasks with the help of web scraping and NLP engine for contextual understanding.
The system starts with this creating selenium web driver instances and then running them simultaneously to measure different aspects of website performance, search engine optimization guidelines, categorizing search results of a person/company name and thus does the scoring for one's reputation.

We use the NODE environment to create the backend that initializes all the required systems and takes in all the factors that we need to work on. We also create a parallel puppeteer instance that will provide us with an headless browser interface which inturn can be automated by the selenium web driver.
These systems will be self reliant and will produce a report in which the website will be scored for given

factors and then that would be re-written back to the document object model of the website for the users to see and have an option to download.

II. SYSTEM

Initializing all the required systems in the NODE environment.

Selenium web driver (NPM)

```
const webdriver = require('selenium-webdriver');
const chrome = require('selenium-webdriver/chrome');
const firefox = require('selenium-webdriver/firefox');

let driver = new webdriver.Builder()
   .forBrowser('firefox')
   .setChromeOptions(/* ... */)
   .setFirefoxOptions(/* ... */)
   .build();
```

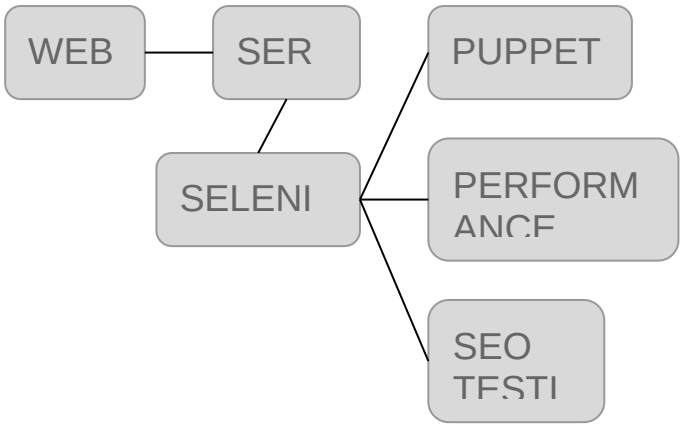
Puppeteer JS (NPM)

```
const puppeteer = require('puppeteer');

(async () => {
  const browser = await puppeteer.launch();
  const page = await browser.newPage();
  await page.goto('https://example.com');
  await page.screenshot({path: 'example.png'});

  await browser.close();
})();
```

Brain JS (NPM)

```
var brain = require("brain.js");
var net = new brain.NeuralNetwork();
```

We use all these to initialize our system and have all the dependencies running.



III. METHOD ANALYSIS

In our method of performance testing, SEO Testing and ORM Testing we start by gathering required data using the web user interface that data is transferred to the server which creates multiple instances of puppeteer driven by selenium.

Automated processing of natural language is a challenging problem. Systems performing various tasks in this domain have existed for several decades, but until relatively recently they were predominantly rule-based. They are time consuming and labour intensive and that is why we will go with a simpler and faster approach.

With our approach of automation of this process we will be able to deliver a solution that is fast and reliable.

Conventional approaches have proven to be tedious and have low accuracy rate in all aspects. Our system will be able to counter these challenges and provide a better solution.

IV. CONCLUSION

In this we have demonstrated to create an automated system that can perform automated performance testing , SEO testing and also use an NLP engine to dissect whether a person or company's reputation is categorically good, bad or critical.

V. REFERENCES

[1]. D. Pratiba, Abhay M.S, Akhil Dua, Giridhar K. Shanbhag, Neel Bhandari, Utkarsh Singh, "SEO TECHNIQUES FOR VARIOUS APPLICATION-A COMPARATIVE ANALYSIS AND EVALUATION," Published in IEEE 20-22 Dec 2018

[2]. Gowtham Aashirwad Kumar ,Dr A Ravikumar, "AN ANALYTICAL STUDY OF SEARCH ENGINE OPTIMIZATION (SEO) TECHNIQUES: TO MAXIMIZE NUMBER OF TRAVELERS ON AN E-CONTENT MATERIAL WEBSITE," Published Volume 11, Issue 1, January 2020

[3]. Peng Qi*, Yuhao Zhang*, Yuhui Zhang, Jason Bolton, Christopher D. Manning, "A PYTHON LANGUAGE PROCESSING TOOLKIT FOR MANY HUMAN LANGUAGES ," Published Stanford University Stanford, CA 94305 23 Apr 2020