

Review Paper- Google Web Scraping in Marketing Strategy

Mr. Abdul Karim Khatib¹

Prof. Venkateswara Gupta Alamuri²

¹Student of Master of Management Studies, Alamuri Ratnamala Institute of Engineering and Technology

Mumbai University,
Karimsayyed888@gmail.com

Abstract –

Web scraping, also known as data mining, is the process of collecting large amounts of data from the web and then placing it in databases for future analysis and later use. Web scraping offers insight into price data, market dynamics, prevailing trends, practices employed by your competitors and the challenges they face.

Big data analytics enables firms to analyze large data volumes and gain new knowledge. It addresses fundamental questions concerning corporate operations and performance. In large datasets or combinations of them, it also aids in the discovery of unknown patterns. Big data approaches are becoming more and more important to apply and analyze for organizational success in today's data-driven world. Web scraping represents a fundamental approach in this regard. Web scraping can be a time-consuming process.

Keywords: web scraping, organization, big data, data driven

I. INTRODUCTION-

Web scraping is the technique of autonomously extracting structured data from the internet. Many people refer to it as web data extraction or scraping. The most popular applications of web scraping include price tracking, price information retrieval, news tracking, lead generation, market research, and so on.

Web scraping has become a significant business technique because your target audience is on the internet and the websites they visit collect and keep data. Web scraping provides access to that data, and this article will teach you how to use it to outperform your competitors.

Web scraping is the automatic extraction of data from specific web pages using bots in accordance with predetermined rules. A bot is a software written in any programming language whose logic is designed to retrieve information from websites.

Web scraping can be used to collect any type of information that is relevant to organizations, such as data about users, products, competitors, and so on. Scraping allows us to gather everything we discussed in data-driven marketing and will discuss further below.

Web scraping is the process of extracting data from the internet using software such as scraper API and storing it for further analysis. This

technique can occur on a smaller scale, such as copying and pasting text from the web into a spreadsheet. However, the primary principle of web scraping occurs on a big scale, utilizing automation tools. The objective of gathering such large amounts of data is to inform corporate choices, market research, competitive analysis, and other activities.

When a web scraping tool receives a target URL, it initiates a GET request. The website server that receives this request then returns website content. The web scraping software then examines the website's HTML structure to retrieve pertinent data. The information it collects is then saved in a readable format.

The idea of web scraping is broad, and the kind of information gathered depends on the demands of the collector. In addition, various programming languages are employed in the development of the necessary tools.

Web scrapers can take on diverse shapes due to differences in development methods, programming languages, and purposes, among other factors. These are; scraper API, cloud-based software, browser extensions, web crawler and self-made software.

The process of web scraping can take a lot of time. For any market researcher, however, having data on your end is an invaluable tool.

With the help of cutting-edge technology, web scraping can extract hundreds, millions, or even billions of data points from the infinite internet. This contrasts with the laborious manual data collection approach. Social media scraping, for instance, is the practice of gathering data by scraping social media networks like Facebook, Instagram, LinkedIn, and Twitter, among others. Social media sites' interfaces, algorithms, and policies are regularly updated, making it challenging to scrape data from these platforms. This is how social media scraping varies from other forms of web scraping.

II. APPLICATION OF WEB SCRAPING-

A wide range of uses for web scraping exist, such as price comparison on the internet, tracking weather patterns, detecting website changes, research, merging data from various sources, obtaining deals and discounts, obtaining job posting details from job portals, brand monitoring, and market analysis. Additionally, it is a fast and effective way to gather data. There are many uses for web scraping across different industries. Prior to big data analytics, it serves as a requirement. Below are some of the many domains where web scraping is employed.

In healthcare

The field of healthcare no longer exclusively depends on direct physical interaction. Rather, it has transitioned to digital in a distinctive way. Web scraping in healthcare can save many lives in this data-driven environment by enabling logical decisions to be made. Healthcare professionals often view gathering data from a large number of patients as a laborious and time-consuming procedure. Despite the fact that clinical data is more important than ever, it is practically hard to collect given the present patient load.

In order to achieve this, the author suggests putting in place a system that gathers clinical data from SARS-CoV2 patients who

visit the hospital on their own, automatically, in preparation for upcoming studies [3].

Another use of web data extraction techniques in the healthcare field is found in Dascalu et al.'s research [4], where drug leaflets are extracted by crawlers.

In social media

Businesses find that social media data extraction really aids in enhancing their marketing campaigns. Businesses can immediately assess consumer sentiment regarding their products in this fast-paced world and enhance audience engagement and public relations. Using web scraping technologies, they developed an application for downloading Instagram account data that is web-based and can be used by many parties. The researchers selected the site scraping

technique in order to avoid utilizing Instagram's Application Programming Interface (API), which

has a number of limitations on how users can access and retrieve data from the platform. The Instagram account data grabber application was successfully constructed using the web scraping technique. In this study, application testing was done on 15 accounts totaling between 100 and 11,000 publications. Based on the analysis's findings, the web scraping solution was able to successfully obtain.

In finance

In order to address some of the shortcomings of the current indicators, the author suggested a preliminary method for creating web-based innovation indicators. They specifically developed a method for quickly and cheaply discovering product innovator enterprises. Next, using conventional company-level information from a survey-based

German Community Innovation Survey online texts categorized as "product innovator" or "no product innovator" were used to train an ANN classification model. After that, by examining the internet texts of hundreds of thousands of German enterprises, they projected, using their categorization model, whether or not these organizations are product innovators. Afterwards, they contrasted their forecasts with firm-level patent information and benchmark data calculated using regional innovation indicators and survey data. The results demonstrate that this approach produces reliable estimates due to its scope and spatial specificity, and it may prove to be a beneficial and affordable supplement to the current collection of innovation indicators [5].

In marketing

Boegershausen et al. [7], in the paper, write about the large amount of consumer data in the

form of a digital footprint available to evaluate customer behavior and to answer customer research questions. In their study, Saranya et al. [8] suggest employing machine learning models

to predict customers' purchasing intentions when they make online transactions. The

Since the content on the Web is unstructured, web scraping is used to obtain data. To forecast the intention to purchase, the data is further examined.

III. BENEFITS OF WEB SCRAPING FOR MARKET RESEARCH-

A common method for obtaining data from websites is web scraping. Scraping can be used to build a database of leads, gather information for marketing research, and much more!

Although it is possible to perform web scraping manually, using a web scraping application is typically more efficient. There are a lot of web scraping programs out there, but not all of them are made equally.

Easy to Use: A software program's usability and ease of use set it apart from its competitors. The more friendly the tool's interface is, the easier it will be to use.

Fast: A good tool will have a short response time. They will be able to extract data fast rather than waiting for the website to load.

Accuracy: Another crucial feature of a scraping tool is accuracy. With this tool, data extraction accuracy will be maximized.

Reliable: The greatest web scraping tools will be dependable. They will be able to extract data reliably without fear of the website falling down.

IV. OBJECTIVE OF STUDY –

To study the Google web scraping in marketing strategy.

The main objective is to determine the various application of web scraping and benefits of web scraping.

V. REVIEW OF LITERATURE-

Web scraping technologies have been designed to provide deep crawling and extraction capabilities. For example, the Google search engine employs a web crawling tool (also known as a web scraper) to index billions of pages on the internet.

Business departments primarily gather photographs, reviews, product information, videos, text, and customer feedback. On the other side, marketing departments collect data such as bounce rate, clicks, keywords, and so on. So, the type of data you collect is determined by your business requirements.

In today's competitive landscape, businesses strive to outperform themselves. This is likely a common concern among businesses. In an industry.

In a customer-centric environment, competitor analysis is necessary rather than optional. Having access to extensive data can provide you a competitive advantage in your field (Samanta, 2021).

Web scraping technologies have been designed to provide deep crawling and extraction capabilities. For example, the Google search engine employs a web crawling tool (also known as a web scraper) to index billions of pages on the internet.

Whether you use a scraper API or create your own online scraping solution, you can get a competitive advantage in terms of market share and domination. Here are usage instances when web scraping gives you an advantage:

- Product and price comparison
- Brand management
- Market research

Web scrapers can extract product information for comparison. For example, if you own an e-commerce business, you can use your competitors' data for price analysis, copywriting, reviews, and so on. You can also find out when your competitors run out of a certain stock, giving you an opportunity to refill and advertise.

Online reviews are also excellent metrics to track. Monitoring metrics allows you to learn how your customers utilize your product. You can use these reviews into your marketing plan, as modern buyers value user-generated content.

To achieve a significant market share, it is critical to broaden your company's reach. To obtain new clients, you must frequently adopt new business methods or launch new enterprises. Web scraping can assist you learn about the next frontier. Market research can also provide you with facts about your present firm that you were previously unaware of, potentially leading to higher revenue.

VI. CHALLENGES OF WEB SCRAPING-

- It can be difficult to determine which sites to scrape. Consider scraping either the data owner or an aggregator.
- The internet is dynamic. Websites have unique structures that can alter at any time.
- Legal concerns and informing website owners.

VII. WEB SCRAPING TECHNIQUES-

There are many techniques use for scraping data on internet.

- Traditional copy and paste
- Text regular expression
- Document Object Model (DOM)Parsing
- Web Scraping Software
- Vertical aggregation platforms
- Semantic annotation recognizing
- Computer vision web-page analyzers

There are various strategies available, but the goal is to select one that yields accurate results within an acceptable timeframe.

Web scraping software

You can choose between using pre-existing software or creating your own script. There are various software tools available for use.

scrape data from the internet. There are two popular approaches for scraping data from the web: using generic web scraping software or writing code.

Use generic web scraping software

Web scraping software is classified into two groups.

There are two options: locally installed on your PC or cloud-based. It's browser-based.

Writing code

You can either hire a developer or create your own unique data extraction program. The most well-known languages used in web scraping are:

- R programing
- Python (Selenium)
- Ruby
- Java

VIII. FEATURES AND SCOPE OF THE STUDY-

The research study contains current trends being embraced by major market players in the Animation and Web Scraping Services sectors, such as the usage of cutting-edge technology, government financing for R&D, and an increased emphasis on sustainability.

Market Research Intellect's Global Web Scraping Services Market research gives a complete review of the present market and the fundamental forces driving it.

It provides deep insights into the global market size, market dynamics, and market segments, along with a detailed assessment of the competition landscape, key players, and regional trends.

The research is intended to give critical market knowledge that will assist organizations in developing strategies to capitalize on development possibilities in the Global Web Scraping Services Market.

The global Web Scraping Services Market research provides in-depth analysis of the competitive environment, market dynamics, value chain, and major segments to help readers understand the current condition of the global market.

It gives light on how firms can benefit on trends and untapped potential in the global Web Scraping Services market.

In the near future, web scraping will be a critical element in the lead generation process. The web scraping technology may conduct market research on specific products/services and provides huge benefits in the marketing area.

IX. CONCLUSION-

This study, which builds on earlier thematic work [41, 42], examines recent research on the uses of web scraping in several fields, web scraping methods, and programs that use web scraping methods. In the process of using this study to enhance our online scraping procedure, we found that the majority of web scrapers are generally pretty similar and general in nature, made to perform simple and straightforward tasks. After evaluating the functionality and performance of several tools and frameworks, we discovered that Scrapy outperforms them all since it is quick, flexible, and strong. Scrapy processes requests asynchronously, allowing for quick scraping of the results. Moreover, Scrapy design is built around a web crawler, making data extraction simple. You can use Scrapy selectors, such as XPath and CSS, to extract the necessary data. Because of its flexible and expandable characteristics, which facilitate simpler integration with VPNs and proxies, Scrapy is an ideal solution for large projects. Moreover, Scraper API worked with proxies, browsers, and CAPTCHAs, so you could use a single API call to obtain raw HTML from any website.

"Web scraping is wizardry if programming is magic" (Mitchell, 2015) Ryan Mitchell said.

The Internet's existence resulted in an increase in the amount of information available for access, making information seeking one of the most popular and time-consuming hobbies (Adam, 2012). We will always remember the Internet as the first platform that allowed us to gather vast amounts of data at low cost or low energy costs.

The utilization of web scraping techniques will be crucial for success in both e-commerce and e-marketing, as they offer valuable insights into the target market and aid in decision-making.

REFERENCE

- Saranya, G., Gopinath, N., Geetha, G., Meenakshi, K., and Nithya, M. (2020, December). Prediction of Customer Purchase Intention Using Linear Support Vector Machine in Digital Marketing. In Journal of Physics: Conference Series (Vol. 1712, No. 1). IOP Publishing:012024.
- <https://venturebeat.com/datadecisionmakers/how-to-use-web-scraping-for-marketing-and-product-analytics/>
- Phan, H. (2019). Building Application Powered by Web Scraping. Doctoral thesis
- <https://attentioninsight.com/3-ways-web-scraping-can-help-you/>
- <https://www.scrapingdog.com/blog/web-scraping-for-market-research/>
- <https://inkbotdesign.com/web-scraping/>