

Web Scraping for E-Commerce Websites

Atharva Bankar¹, Prof. Sameer Kakade²

¹*Dept of MCA-Trinity Academy of Engineering, Pune, India*

²*Assitant Professor, Trinity Academy of Engineering, Pune, India*

Abstract

The internet is a rapidly expanding environment where a vast amount of info is available, yet obtaining it might be challenging. Web scraping is therefore a crucial technique for obtaining data. Data can be extracted from different websites using a process called web scraping. Depending on the technology used, end users can obtain the data in spreadsheet, csv, json, xml, and database forms.

Numerous fields, including e-commerce, market research, brand monitoring, and others, use web scraping. Our solution suggests a way to compare product details that are retrieved from e-commerce websites. Various tools are used for data extraction, including BeautifulSoup, Selenium, Scrapy, and others. BeautifulSoup is used by our system to extract data. Data is stored in a MySQL database following extraction.

Our webapp then presents this data in a format that is similar. It takes time to visit each website individually and compare product details; therefore, in order to help the user compare products, our system will show all of the product details from different websites.

I. INTRODUCTION

The system's objective is to gather product information from several e-commerce websites and present it on a single web interface so that it may be compared. One way to gather information from multiple sources is through data extraction. The dynamic nature of dynamic websites makes data extraction extremely difficult. The product name entered into this system will be sent to the different websites. Following a search on each website for pertinent input, pertinent product data is scraped and entered into a database. Through an API, the stored data is subsequently shown in a different format on a website. The database used by our system is MySQL. MySQL-connector facilitates the connection between MySQL and Python.

II. LITERATURE SURVEY/BACKGROUND

Web Scraping: Current Developments and Potential Applications

The primary goal of data scraping is to extract the information and process it appropriately in the appropriate format, such as a spreadsheet, database, or CSV file. This essay explains several web scraping methods. It examines several approaches, categories, and instruments that are now in use. The paper employs many methodologies, including imitation, weight measurement, differential, and machine learning approaches. Spider, data miner, dexi.io, and many more are examples of built-in tools that can be utilized. covers the areas in which web scraping is used as well. This suggests that the journalism industry needs web scraping more.

A Review of Web Scrapping and Its Applications

The internet is a vast data source, and gathering it is very difficult. To overcome this, web scraping is used. This essay explores the many software and tool options for web scraping while concentrating on different facets of the practice. It also goes into detail on the different kinds, benefits, drawbacks, and uses of web scraping. There are three methods for doing web scraping here. first with the aid of programming language libraries, then with the use of frameworks, and finally with desktop environments. Web scraping also makes use of a number of technologies, including regular expressions and rvest, which are included in the paper. In addition to the previously described details, the advantages and disadvantages of web scraping are also discussed.

Comparison of Web-Mining Products for E-Commerce:

The comparison of products from different e-commerce websites is presented in this article. Because there are so many websites available, users may become somewhat confused. Users can choose the best product at the lowest cost by comparing the pricing of these products across multiple websites at the same time. Techniques like web scraping and crawlers are employed to obtain product details. The purpose of this paper is to help internet shoppers save time, effort, and money by helping them purchase things at the greatest price.

Data Analysis by Web Scrapping using Python

The technique for obtaining problem-solving solutions is data analysis. Data must be divided into steps such as specification assembly, organization, cleaning, reanalysis, application of models and algorithms, and final result in order for analysis to be completed. In order to analyze this data, it must first be extracted. Web scraping is utilized for this purpose. Information obtained from the web host is used by the web scraper program. It could be because the majority of website pages are designed with human users in mind rather than the minimum amount of robot use. That's how the toolset for web data scraping was created. The primary obstacle stemming from the web's autonomous and heterogeneous character was uncovering the buried data. This project's primary outputs included an easy-to-use search interface, indexing, query processing, and an efficient method for extracting data from web forms using web structure, form submission analysis, and a new submission strategy. This thesis proposes a completely automated, domain-dependent proto-type system to integrate and retrieve the data hidden behind the search form.

A Framework for Retrieving Petroleum Information Based on Web Scrapping Using Python: This paper provides a framework for extracting petroleum information that will be utilized by businesses and researchers studying petroleum development. The system is composed of two parts: first, information is scraped using the open-source Scrapy framework, and then, only relevant data is stored in the system with the aid of the k-means algorithm. As a result, we have an automated and flexible system that increases work efficiency.

III. PROPOSED WORK/SYSTEM

Real-time data is essential for staying competitive and making well-informed decisions in the ever-changing world of e-commerce. One effective way to retrieve useful data from a variety of online sources is through web scraping. This proposal describes how web scraping will be used in our e-commerce website project to improve decision-making, increase functionality, and improve user experience.

Applications Of Web Scraping For E-Commerce Websites:

Web scraping is revolutionizing e-commerce by improving user experience, insights, and competitiveness.

Our e-commerce website project's planned integration of web scraping is a revolutionary step forward, utilizing real-time data retrieved from rival websites to enhance our product catalog, maximize pricing tactics, and obtain priceless industry insights. Our objective is to improve the user experience by continuously adjusting our prices, maintaining competitiveness, and giving customers accurate and current product listings. To achieve this, we plan to automate the extraction of product information, pricing data, and customer feedback. In the constantly changing world of e-commerce, we hope to open up new possibilities for growth, efficiency, and consumer happiness through data-driven decision-making made possible by web scraping.

The potential for web scraping to completely transform e-commerce operations is enormous. Businesses can obtain important insights into consumer behavior, pricing tactics, and market trends by methodically pulling data from competing websites. With the use of this data, dynamic pricing modifications can be made to maximize profitability while preserving competition. Additionally, web scraping makes it easier to add thorough and current information to product catalogs, which improves user experience by offering precise specs, reviews, and descriptions. Enhanced search capabilities, fueled by data collected from the internet, make it easier for customers to locate the things they want. Businesses can also improve resource allocation and strategic planning by using scraped data analysis to drive decisions about product development, marketing tactics, and inventory management. All things considered, web scraping used in e-commerce helps companies remain ahead of the competition, provide great customer experiences, and spur growth through data-driven decision-making.

IV. RESULT AND DISCUSSIONS

Using web scraping for e-commerce websites has revolutionary effects. Businesses can obtain significant insights into consumer preferences, pricing tactics, and market dynamics by methodically pulling data from competing websites. Dynamic pricing adjustments are made possible by this information in order to maximize profitability and maintain competitiveness. Web scraping also improves the user experience by adding current and accurate details, such as reviews, specs, and descriptions, to product catalogs. Enhancements to search capability facilitated by data scraping guarantee that clients may locate desired products with ease. Web scraping is a topic of discussion in e-commerce because of its critical role in helping to guide strategic decisions, maximize marketing efforts, and promote business expansion. To ensure the ethical and lawful use of scraped data, however, ethical factors, such as adherence to data protection rules and respect for terms of service, must be properly addressed.

V. CONCLUSION

Our system scrapes the data from different websites and compares the product detail on a website. The application of web scraping in the context of e-commerce in our project shows its significant influence on user experience, competitiveness, and strategic decision-making. We have improved our product portfolio, refined our pricing tactics, and obtained priceless market insights by utilizing real-time data from rival websites. In addition to improving consumer satisfaction, this revolutionary strategy puts us in a successful and long-term growth position in the ever-changing e-commerce industry. To guarantee the responsible use of scraped data in the future, ethical and legal adherence to data privacy standards will be crucial.

REFERENCES

1. Rabiyatou DIOUF, Edouard Ngor SARR, Ousmane SALL, Babiga BIRREGAH, Mamadou BOUSSO, Sény Ndiaye MBAYE, "Web Scraping: State-of-the-Art and Areas of Application" -2019 IEEE International Conference on Big Data (BigData).
2. David Mathew Thomas, Sandeep Mathur "Data Analysis by Web Scraping using Python"-Third International Conference on Electronics Communication and Aerospace Technology [ICECA 2019].
3. Riya Shah, Karishma Pathan, Anand Masurkar, Shweta Rewatkar, Prof. (Ms.) P.N.Vengurlekar "Comparison of E-commerce Products using web mining"-International Journal of Scientific and Research Publications, Volume 6, Issue 5, May 2016.
4. https://en.wikipedia.org/wiki/Web_scraping
5. **Fig1:** <https://images.app.goo.gl/WeSEFBWg6zeXn1tq9>