



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 Issue: XI Month of publication: November 2023

DOI: <https://doi.org/10.22214/ijraset.2023.56913>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Slickdeals Aggregator Web-Application Using Web Scraping

Aviraj Jagtap¹, Raunak Jaiswal², Vaibhav Kedari³, Nikhil Khajure⁴

^{1, 2, 3, 4}Students, Department of Computer Engineering, Zeal College of Engineering and Research, Pune, Maharashtra

Abstract: Due to the e-commerce industry's explosive growth, there are a large number of online retail platforms that sell a wide variety of goods. Because there are more products and shops available, it can be difficult for customers to compare prices and discounts among different online retailers. It describes how to create a price comparison website that gathers and compares product prices from various e-commerce websites using web scraping techniques. The system uses web scraping to collect product information such as price, reviews, offers, descriptions and availability from certain e-commerce websites. For this purpose, the website is designed with a web scraping module that automates data extraction using Python modules such as BeautifulSoup and Selenium. The web interface primarily helps users make informed purchasing decisions by allowing them to search for products and compare prices. The first issue this article addresses is the content on the website, followed by the oversight of accurate information and the provision of content related to the website's access process. Web scraping's ethical concerns are also covered, with an emphasis on responsible data use and adherence to e-commerce platforms' intellectual property and privacy rights. It demonstrates the value of a pricecomparison website in giving users access to current pricing data from several online retailers. The service makes use of web scraping technologies to make online buying easier and give users the knowledge they need to make wise financial decisions.

Keywords: Web Scraping, E- Commerce, BeautifulSoup, Selenium,

I. INTRODUCTION

Online shopping has become the preferred choice for consumers, but finding the best deals can be overwhelming. Slickdeals Aggregator offers a solution using web scraping technology to compare deals across multiple platforms. The goal is to empower customers with the necessary information to make informed decisions, leveling the playing field between consumers and online sellers. Python is used for web scraping, and web development frameworks like Flask or Django are used for user-friendly interfaces. The scraped data is regularly updated to ensure accurate results, and error handling techniques are employed to handle potential issues. The success of web scraping is heavily reliant on the design of the websites being scraped. Even minor changes to the website structure may require adjustments to the scraping code. Additionally, many websites use JavaScript to load content dynamically after the initial page load, which can result in incomplete data if not accounted for in the scraping approach. Some websites may also detect and reject repeated requests from the same IP address, making it difficult to collect data. Furthermore, the data obtained through scraping is often unstructured or semi-structured and requires further processing and cleaning before it can be used. To ensure accurate and up-to-date data, scraping scripts must be regularly updated to accommodate any changes to the website being scraped.

To obtain pricing data, it is important to explore alternative websites and online merchants. This will provide a comprehensive overview of prices for various products, encompassing both popular and niche businesses. To ensure the accuracy and timeliness of the data, it is advisable to establish a consistent scraping schedule. Additionally, allowing users to comment on any incorrect price information can help maintain data integrity. Employing a rotating IP proxy service can be beneficial as it periodically changes your IP address, mitigating the risk of websites blocking your scraping activities. The primary objective of the team involved in this report is to develop and refine a dependable algorithm for comparing prices across different websites.

II. MOTIVATION

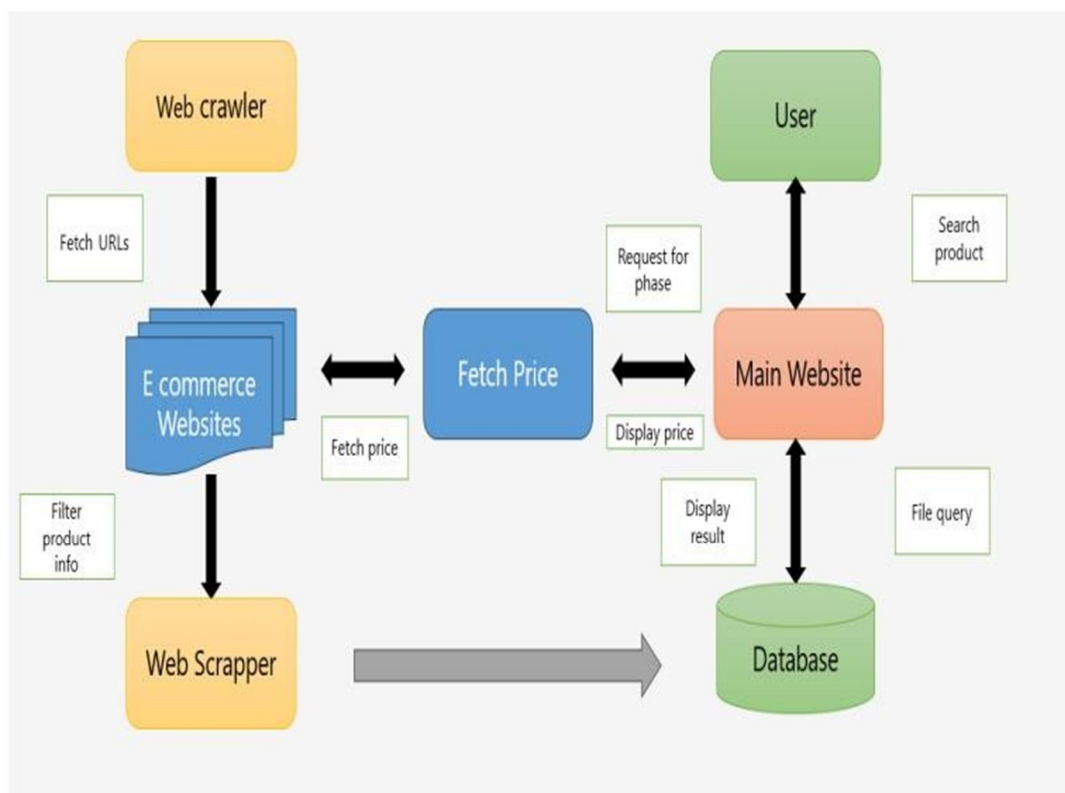
Our web applications provide a comprehensive answer for customers looking for clean get right of entry toto fees, deals, and offers from diverse assets. This platform not best helps informed selection-making but also permits customers to discover fee for money options. Recognizing that clients regularly face time constraints and prioritize convenience, our internet software gets rid of the trouble of trying to find fees across different shops, in the end saving each time and money by way of figuring out the bottom rate to be had.

By aggregating charges from multiple retailers, our platform fosters transparency within the commercial enterprise panorama. In addition to comparing charges, customers also can evaluate product capabilities, specs, and reviews, making our internet site a one-stop platform for comprehensive product studies. Moreover, we increase our offerings to outlets by using providing them with the great equipment to show off their products and provide certain insights into consumer behavior. To further enhance consumer level in, we're considering integrating functions that permit customers to customize their choices and acquire notifications while costs drop or unique objects pass on sale.

III. OBJECTIVE

The challenge's overarching objective is to set up and perform a platform with specific desires dedicated to enhancing the net buying level in for clients. The number one awareness is on developing a user-friendly charge reviewing utility, fostering a continuing and fun surfing enjoy for customers. This platform will empower clients to effortlessly examine costs and product opinions throughout diverse online retailers, emphasizing a visually appealing and smooth-to-navigate interface to ensure person delight and engagement. Integral to the venture's assignment is the commitment to offering users with accurate and updated pricing information. This aligns with the second objective, aimed toward permitting customers to make well- knowledgeable shopping decisions and in the long run store money. By delivering reliable statistics on product prices, the platform seeks to contribute to a greater obvious and customer-friendly on line market. To accomplish these objectives, the assignment will put in force a sturdy seek and filter gadget, constituting the third goal. This gadget will empower customers to correctly discover the pleasant deals on a numerous variety of merchandise, streamlining the consumer enjoy and facilitating a more green decision-making manner. In precis, the undertaking's multifaceted goals converge on growing a person-friendly platform that no longer simplest simplifies charge evaluation however also prioritizes accuracy, timeliness, and efficiency, thereby helping customers navigate the expansive panorama of on line retail applications. These algorithms have the ability to able to handle dynamic processes and multiple data types, each with its own structure and format.. The data collection process begins with identifying sources, which are typically e-commerce websites, inventory, and other online businesses. These algorithms are responsible for collecting information from various sources and integrating this information into a unified and user-friendly interface.

IV. SYSTEM ARCHITECTURE



V. METHODOLOGY

- 1) *Web-Scrapping Algorithm:* Implement error handling to manage situations where the website's structure changes or network issues arise. Websites may change their structure, requiring you to adapt your scraping code. Implement rate limiting to avoid overwhelming the website's servers with too many requests. Use the library's functions to navigate the HTML structure and locate the elements containing the data you want to scrape. If you're scraping data from multiple pages (pagination), implement a loop to repeat the scraping process for each page. Test your scraper on a small scale and ensure it's extracting the data accurately. Choose a web scraping library such as BeautifulSoup or Scrapy (Python) to assist in parsing HTML and navigating the website's structure. Then we have determine the website which we want to scrape.
- 2) *Data Matching Algorithm:* It begins with statistics extraction, where in product information is gathered from one of kind resources and normalized to make sure consistency. Statistics facts matching in evaluation web packages is a vital technique that involves aligning and identifying the equal or comparable products from diverse online stores. Weighted matching can prioritize sure attributes, and in some instances, gadget studying models are used to learn styles and similarities. String matching algorithms, together with specific matching and fuzzy matching, are hired to locate similarities, accounting for variations and typos. Tokenization breaks down textual facts into smaller units for comparison, while function extraction focuses on applicable attributes like product names, logo names, version numbers, and UPC codes.
- 3) *Data Aggregation and API Integration Algorithms:* Data collection and API integration algorithms play an important role in the performance of price comparison web Figure represents about the gadget architecture and the workflow of the proposed machine. The front a part of the The machine provides a graphical user interface (GUI) that's a internet site where the client can have interaction with the gadget and net crawling and scraping techniques to extract contextual records within the heritage From e-commerce web sites. Contains statistics about removals from e-trade sites stored in MySQL database. Customers request products from the main internet site and the product is produced Searches are performed from nearby databases. Product facts is displayed on the main net web page of the website on the way to. The customer or consumer can discover the charge of the specified product at one area inside the available e-trade web sites.

VI. LIMITATIONS

- 1) *Data Integrity and Reliability:* Website downloads depend on the layout of the website, and Any changes to the layout of the Site may and should result in errors and inaccuracies in the creation process data that may misrepresent pricing information to users.
- 2) *Legal and Ethical Issues:* Deleting the site may violate the sites being deleted, including legal ones consequences. It's about making sure legal and ethical restrictions on web scraping are adhered to It is important to avoid legal disputes and maintain a good reputation.
- 3) *Data Volume and Scalability:* Handling the amount of data generated by the network Scraping can be difficult, and so can scaling infrastructures to accommodate increasing volume Data can be complex and resource intensive.
- 4) *Access Restrictions:* Some websites may use anti-scraping mechanisms, e.g. CAPTCHA challenge or IP blocking, in order to prevent data exceptions, makes it difficult to access From that data these areas.
- 5) *Data Quality Control:* It is possible to ensure data quality and reliability. It is difficult, because errors and inconsistencies in the generated data can affect the user as a whole Experience and reliability of a price comparison website
- 6) *Technical Fixes and Updates:* The web scraping script still works, especially with It changes the layout and configuration of source websites and requires constant maintenance and updated to keep pace with data that is accurate about it.

VII. CONCLUSION

A price comparison website can be the most effective tool in the current digital age geography, allowing consumers to make informed purchases with ease and confidence. These platforms can provide a comprehensive and user-friendly experience that benefits both consumers and the entire e-commerce ecosystem.

By certain standards, product Descriptions, and user reviews from multiple stores, these sites allow users to a More information to make informed decisions. The primary value of pricing comparison website is to empower users by giving them more control The purchase decision. They are generated through convenient research materials, tailored to everyone information, and side-by-side product comparisons, all of which simplify decision-making and save customers valuable time and money. and provide what is appropriate Up-to-date pricing information, users can rely on these platforms to find the best deals and. Optimal purchasing options for their preferred products.

REFERENCES

- [1] Jose Alenjandro Cano, Abraham Allec, Landono-Pineda, authored the research paper titled 'Sustainable business model of e-marketplaces: An analysis from the consumer perspective' [2023]
- [2] Eve Sarah Troll, Julius Frankenbach, Malte Friese, David D. Loschelder, authored the research paper titled 'A meta-analysis on the effects of just-below versus round prices' [2023]
- [3] Arman Shaikh, Raihan Khan, Komal Panokher, Mritunjay Kr Ranjan, Vaibhav Sonaje, authored the research paper titled, 'E-commerce Price Comparison Website Using Web Scraping' [2023]
- [4] Mrs. M. Sowmiya, Srinandhan Cs, Mugesh Raja M, Sudheekshan Kumar S, authored the research paper titled 'Price Comparison for Products in Various Ecommerce Website' [2023]
- [5] Pavan Sai Rayalla, Setty Vasu Vivek, Sreeja G, Sneha Sanju, Dr. S. Vijay Kumar, authored the research paper titled 'Product Price Comparison On Multiple E-Commerce Websites' [2023]
- [6] Federico mangio, giandomenico, Di Domenico, authored the research paper titled 'All that glitters not real affiliation : How to handle affiliate marketing programs in the era of falsity' [2022]
- [7] Nagender Singh, Omprakash Sahu, authored the research paper titled 'Feasibility assessment for Ecommerce: A data collection from developing country (Ethiopia)' [2022]
- [8] Gandhe Vineeth Kumar, Hema M S, Aishwarya R, authored the research paper titled 'Web Scrapping for E-commerce Website' [2022]
- [9] S. Rajendar, K. Manikanta, M. Mahendar, Assistant Prof. (Mrs.) K. Madhavi, authored the research paper titled 'Price Comparison Website for Online Shopping' [2021]
- [10] Kasereka Henrys, authored the research paper titled 'Importance of scrapping in ecommerce and marketing' [2021]
- [11] Yidan Wang, authored the research paper titled 'Research on E-commerce Platform of Online Shopping Consumers' [2021]
- [12] Deli wang, Wuwei Li, authored the research paper titled 'Optimization algorithm and simulation of supply chain coordination based on cross border E-commerce network platform' [2021]
- [13] Priya, Matta, Nikita Sharma, Devyani Sharma, Bhasker Pant, authored the research paper titled 'Web Scrapping: Applications and Scrapping Tools' [2021]
- [14] Zhongcheng Lei, Hong Zhou, Shenwang Ye, Wenshan Hu, Guo-Ping Liu, authored the research paper titled 'Cost-Effective Server-side Re- deployment for Web based Online Laboratories Using NGINX Reverse Proxy' [2020]
- [15] Aditya Ambre, Praful Gaikwad, Kaustubh Pawar, Vijaykumar Patil, authored the research paper titled 'Web and Android Application for Comparison of E-Commerce Products' [2019]





10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)