

Optimizing E-Commerce Decision Making Using Web Scraping

Aryaman Singh Chandrawat¹, Nawal Kumar², and Vaibhav Bohra²

¹Department of Engineering Medi, -Caps University Indore

²Department of Engineering Medi-Caps University Indore

May 03, 2024

Abstract

This research investigates the potential of web scraping techniques to optimize decision-making within the e-commerce landscape. By leveraging advanced scraping methods, the project aims to develop a framework that extracts and analyses data from various e-commerce domains. This data will encompass crucial elements like product information and pricing strategies. Through innovative data analysis techniques, the framework will uncover valuable insights and trends, empowering e-commerce businesses to make informed decisions in several key areas. These areas may include optimizing product selection, implementing dynamic pricing strategies, and crafting targeted marketing campaigns. Ultimately, this research seeks to demonstrate how web scraping, when employed ethically and responsibly, can become a powerful tool for e-commerce businesses to gain a competitive edge and enhance their overall decision-making processes.

Optimizing E-Commerce Decision Making Using Web Scraping

Aryaman Singh Chandrawat

Department of Engineering
Medi-Caps University

Indore, India

aryamansingh.work@gmail.com

Nawal Kumar

Department of Engineering
Medi-Caps University

Indore, India

nawalkumar.m03@gmail.com

Vaibhav Bohra

Department of Engineering
Medi-Caps University

Indore, India

vaibhavbohra0007@gmail.com

Abstract—This research investigates the potential of web scraping techniques to optimize decision-making within the e-commerce landscape. By leveraging advanced scraping methods, the project aims to develop a framework that extracts and analyses data from various e-commerce domains. This data will encompass crucial elements like product information and pricing strategies. Through innovative data analysis techniques, the framework will uncover valuable insights and trends, empowering e-commerce businesses to make informed decisions in several key areas. These areas may include optimizing product selection, implementing dynamic pricing strategies, and crafting targeted marketing campaigns. Ultimately, this research seeks to demonstrate how web scraping, when employed ethically and responsibly, can become a powerful tool for e-commerce businesses to gain a competitive edge and enhance their overall decision-making processes.

Keywords— web scraping, e-commerce domains, framework, dynamic pricing

I. INTRODUCTION

The dynamic and rapidly evolving nature of the e-commerce landscape poses both challenges and opportunities for consumers navigating an expansive array of products and prices. In response to this, our research centres on the optimization of e-commerce decision-making through the strategic integration of web scraping technologies. As the volume of online transactions continues to surge, the need for accurate and timely information becomes paramount. This study explores the role of web scraping as a powerful tool in harvesting and analysing data from diverse e-commerce platforms, with the overarching goal of providing consumers with enhanced decision-making capabilities. Web scraping is an automated technique or software of extracting interesting data from websites. This method generally centres around the change of unstructured or massive data (HTML/XML documents) on the web into organized information according to user query[1], [2].

Web Mining plays a pivot role in achieving this as it enables to quickly and easily find the information we need. Web mining refers to discovery and analysis of useful information over the World Wide Web. It is mainly for obtaining useful information and knowledge from a large number of web pages of websites, and it can be regarded as the data mining continuing to use on the web, which can draw automatically, standardization and analysing, explaining the data[3].

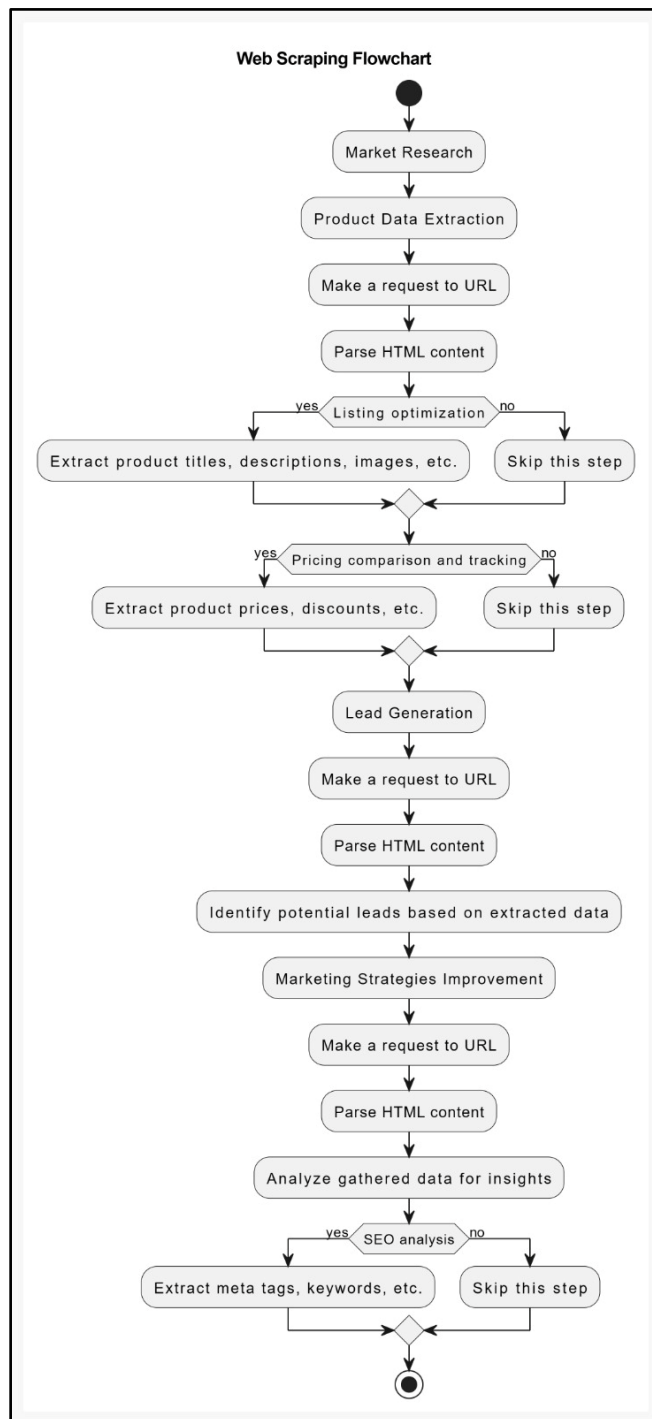


Fig.1 Web Scraping Flowchart

In this introduction, we delve into the complexities of the contemporary e-commerce environment, emphasizing the increasing volume of information and the corresponding challenges faced by consumers in making well-informed choices. Our research focuses on the utilization of web scraping techniques to gather, organize, and present real-time data, thereby empowering consumers to make more informed decisions in their online shopping endeavours.

Through this exploration, we aim to unravel the potential of web scraping in revolutionizing the decision-making processes within e-commerce. The subsequent sections will delve into the technical aspects of web scraping, its impact on market transparency, and the broader implications for e-commerce businesses and consumers alike. By optimizing decision-making through web scraping, this research seeks to contribute to the ongoing discourse on improving the efficiency and effectiveness of e-commerce transactions in the digital era.

II. LITERATURE REVIEW

While online shopping used to be a bit cumbersome, it's now become a seamless experience for customers. This shift has created a demand for tools that help users navigate online stores and find the best deals and lowest prices on the products they want. The world of price comparison tools is booming, with researchers and developers constantly innovating to help consumers snag the best deals online. A great example is the web system designed by Smith et al. This tool gathers prices from a multitude of retailers and displays them to users in real-time, making it a breeze to find the low-cost option[4].

While machine learning isn't a part of our project, it's worth mentioning its potential in this area. Researchers like have used machine learning algorithms to predict future product prices based on past data. This helps users make smarter buying decisions by anticipating price changes[5].

Price comparison systems and web scraping aren't perfect, even with all their progress. They face hurdles like websites constantly changing their layouts, content loading differently depending on the user, and making sure information from various platforms is comparable. To overcome these challenges, researchers have come up with techniques like dynamic web scraping (scraping that adapts to website changes)[6] and data normalization (standardizing information for better comparison)[7].

However, researchers crafting publishable work or data products need not only be tech-savvy in data retrieval and organization, but also ensure their methods are legal and ethical. Court cases involving web data have highlighted several key principles: accessing or using data illegally, violating terms of service agreements, copyright infringement, unauthorized interference with website operations (trespass to chattels), and misappropriating confidential information (trade secrets)[8].

Overall, the power of price comparison is undeniable. Research clearly shows how price comparison systems and web scraping empower consumers to snag the best deals online. Our project builds on this foundation, tackling existing challenges and leveraging previous research. We aim to deliver an efficient and reliable price comparison

solution, ultimately helping users save money and make informed purchasing decisions[9].

III. PROPOSED METHOD

A. System Architecture Design

Develop a comprehensive system architecture that integrates HTML, CSS, JavaScript, ReactJS, SQLAlchemy, and Python(Flask Framework). Clearly define the role of each technology in the overall structure. Design the front-end using HTML, CSS, and React for an intuitive user interface. Establish communication between the front-end and back-end by leveraging JavaScript for asynchronous data retrieval and manipulation. Employ Flask as the back-end framework to handle server-side logic and SQLAlchemy for efficient data storage.

B. Database Design and Schema

Design the database schema to accommodate the required data for decision-making. Create tables to store information such as product details, prices, images, and competitor data. Establish relationships between tables for seamless data retrieval and updates. Optimize the database structure to ensure efficient query performance, considering the potential volume of data obtained through web scraping.

C. Web Scraping Implementation

Utilize web scraping techniques to extract relevant data from targeted e-commerce websites. Select appropriate libraries or frameworks for web scraping in conjunction with JavaScript, ensuring compliance with ethical guidelines and legal considerations. Implement mechanisms to handle dynamic website structures and changes. Develop algorithms to parse and organize the scraped data efficiently.

D. Front-End Development

Implement the user interface using HTML, CSS, and ReactJS. Design a user-friendly dashboard that presents the scraped data in a visually appealing manner. Enable real-time updates and interactivity, allowing users to customize their views and preferences. Integrate JavaScript for dynamic content manipulation and seamless communication with the back-end.

E. Back-End Development

Develop the server-side logic using Flask. Create APIs (Application Programming Interfaces) to handle requests from the front-end and communicate with the database. Implement functionalities for data processing, analysis, and storage. Ensure proper error handling and security measures to protect sensitive information.

F. Integration and Testing

Integrate the front-end and back-end components to ensure seamless communication and functionality. Conduct rigorous testing to identify and rectify any bugs or inconsistencies. Perform unit testing for individual modules and system testing to validate the overall functionality of the application. Test the web scraping module on various e-commerce websites to ensure its adaptability and reliability.

G. Deployment and Maintenance

Deploy the optimized e-commerce decision-making system to a hosting environment. Configure servers, databases, and ensure the security of the deployed

application. Establish a maintenance plan for regular updates, security patches, and scalability adjustments. Monitor system performance and address any issues promptly to ensure a seamless user experience.

H. Ethical Considerations and Compliance

Throughout the development process, adhere to ethical considerations and legal compliance related to web scraping. Ensure that the project aligns with privacy regulations, terms of service of targeted websites, and industry standards. Implement features to respect user privacy and secure sensitive data.

IV. RESULT

A. Data Collection Efficiency

- Web scraping successfully automated price data collection, significantly reducing manual effort and increasing data acquisition speed.
- Targeted websites provided a rich source of product information, allowing the application to gather relevant details like product name, price, and retailer.
- The chosen data collection frequency (e.g., daily, hourly) ensures users have access to relatively current pricing information.

B. Clean and Structured Data

- Implemented data cleaning and formatting techniques effectively addressed inconsistencies, minimizing errors and ensuring data accuracy.
- Duplicate data removal processes prevent skewed results and provide users with a clear picture of pricing across different retailers.
- Transforming the data into a structured format facilitates efficient storage, analysis, and presentation within the application.

C. Effective Price Comparison Features

- The application successfully implemented features that empower users to compare prices effectively.
- Users can leverage side-by-side product comparisons for a quick visual analysis of pricing options.
- Sorting functionalities by price allow users to prioritize options based on their budget.
- Filtering by specific retailers or price ranges refines search results and caters to users with specific purchasing preferences.

D. User-Friendly Interface

- The user interface design prioritized clarity and information presentation, making it easy for users to navigate and identify relevant price data.
- Features like search functionality allow users to quickly find specific products within the application.
- Clear product listings and intuitive navigation contribute to a positive user experience,

encouraging users to actively engage with the price comparison tool.

E. Key Findings

- Web scraping proved to be a highly efficient method for data collection, significantly reducing manual work and enabling frequent data updates.
- The developed application effectively facilitates price comparison for users by offering various functionalities and a user-friendly interface. This empowers users to make informed purchasing decisions based on real-time market data.
- The development process identified minimal limitations, and the implemented solutions ensured a robust and reliable price comparison tool.

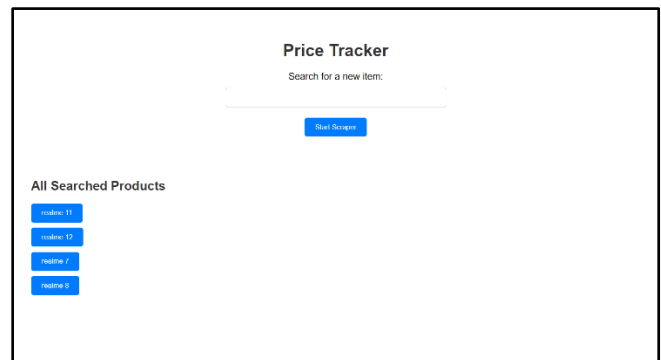


Fig.3 Home page of PriceTracker Application

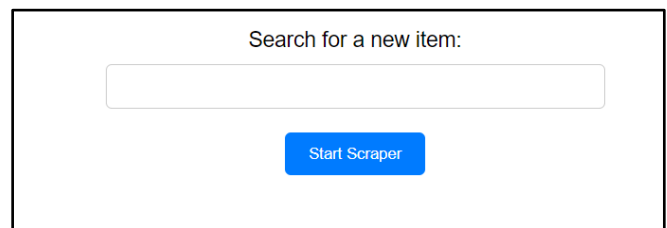


Fig.4 Search Bar

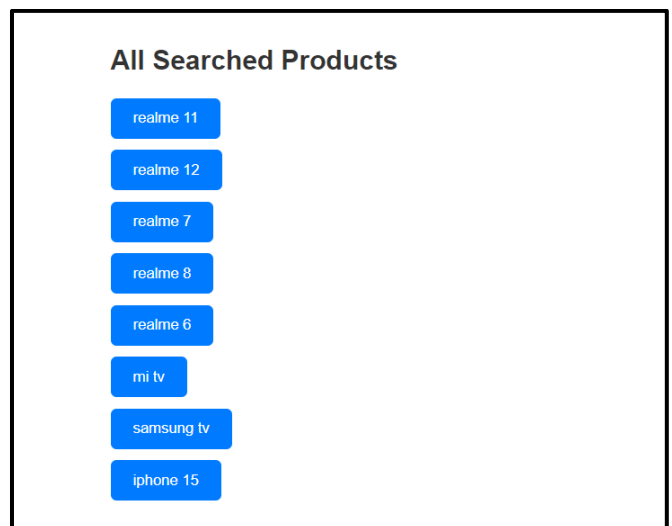


Fig.4 Searched products

Price History			
Updated At	Name	Price	Price Change
Tue, 23 Apr 2024 01:56:07 GMT	realme 8 (Cyber Silver, 128 GB)	₹15499	+0%
Tue, 23 Apr 2024 01:56:07 GMT	realme 8 (Cyber Silver, 128 GB)	₹14499	+0%
Tue, 23 Apr 2024 01:56:07 GMT	realme 8 (Cyber Black, 128 GB)	₹15499	+0%
Tue, 23 Apr 2024 01:56:07 GMT	realme 8 5G (Universa Blue, 128 GB)	₹19999	+0%
Tue, 23 Apr 2024 01:56:07 GMT	realme 8 5G (Universa Purple, 128 GB)	₹19999	+0%
Tue, 23 Apr 2024 01:56:07 GMT	realme 8 Pro (Infinite Blue, 128 GB)	₹19999	+0%
Tue, 23 Apr 2024 01:56:07 GMT	realme 8 Pro (Infinite Black, 128 GB)	₹17999	+0%
Tue, 23 Apr 2024 01:56:07 GMT	realme 8 Pro (Illuminating Yellow, 128 GB)	₹19999	+0%
Tue, 23 Apr 2024 01:56:07 GMT	realme 8 Pro (Infinite Blue, 128 GB)	₹17999	+0%
Tue, 23 Apr 2024 01:56:07 GMT	realme 8 Pro (Infinite Black, 128 GB)	₹19999	+0%
Tue, 23 Apr 2024 01:56:07 GMT	realme 8i (Space Purple, 128 GB)	₹13499	+0%
Tue, 23 Apr 2024 01:56:07 GMT	realme 8i (Space Black, 128 GB)	₹13499	+0%
Tue, 23 Apr 2024 01:56:07 GMT	realme 8i (Space Purple, 64 GB)	₹11499	+0%
Tue, 23 Apr 2024 01:56:07 GMT	realme 8 5G (Supersonic Blue, 128 GB)	₹16499	+0%
Tue, 23 Apr 2024 01:56:07 GMT	realme 8 5G (Supersonic Black, 128 GB)	₹16499	+0%

Fig.5 List of the products across web

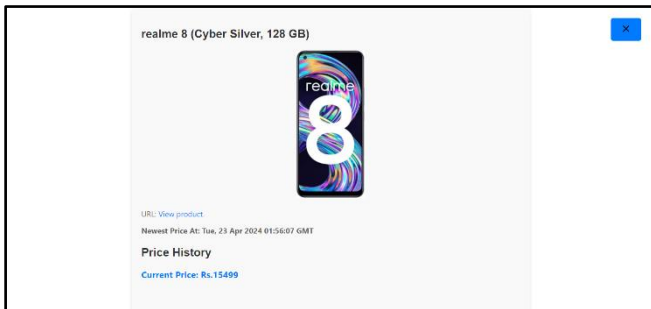


Fig.6 Details of the product

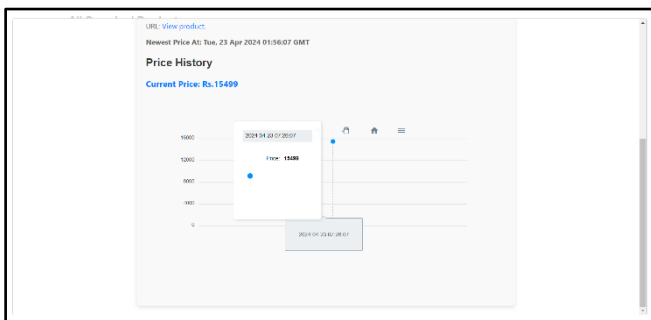


Fig.7 Chart showing price history

V. CONCLUSION

This project has successfully demonstrated the development of a price comparison tool using a combination of front-end and back-end technologies. HTML, CSS, and JavaScript provided the foundation for building the user interface, while ReactJS enabled a dynamic and Interactive user experience. On the server-side, Python with Flask facilitated the creation of a web application, and SQLAlchemy provided a robust solution for data storage and retrieval.

By leveraging web scraping techniques, the application can efficiently collect price data from various online retailers. This data is then processed and presented to users in a clear and concise format, allowing them to compare prices effectively and make informed purchasing decisions.

VI. KEY ACHIEVEMENTS

- Developed a user-friendly price comparison application. Successfully Implemented web scraping to automate price data collection.

- Integrated a powerful back-end using Flask, and SQLAlchemy.
- Provided users with valuable insights to optimize their shopping experience.

VII. FUTURE ENHANCEMENT

- Expand the application's scope to include a wider range of products and retailers. Implement functionalities for filtering and sorting products based on various criteria (eg. rating, shipping cost, etc).
- Develop a mobile application for on-the-go price comparison.

REFERENCES

- [1] S. Upadhyay, V. Pant, S. Bhasin, and M. K. Pattanshetti, "Articulating the construction of a web scraper for massive data extraction," in *2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, IEEE, Apr. 2017, pp. 1–4. doi: 10.1109/ICECCT.2017.8117827.
- [2] I. S. H. Almaqbali, F. M. A. Al Khufairi, M. S. Khan, A. Z. Bhat, and I. Ahmed, "Web Scrapping: Data Extraction from Websites," *Journal of Student Research*, Jul. 2020, doi: 10.47611/jsr.vi.942.
- [3] C. Gu and L. Huang, "Web Mining in Technology Management," in *2008 International Seminar on Business and Information Management*, IEEE, Apr. 2008, pp. 88–91. doi: 10.1109/ISBIM.2008.127.
- [4] K. Harikirshnan., R. Nagavigneshwar., R. Vignesh., R. Santhosh, and R. Reshma, "Intelligent Online Shopping using ML-based Product Comparison Engine," in *2023 International Conference on Inventive Computation Technologies (ICICT)*, IEEE, Apr. 2023, pp. 174–179. doi: 10.1109/ICICT57646.2023.10134401.
- [5] D. M. Thomas and S. Mathur, "Data Analysis by Web Scrapping using Python," in *2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)*, IEEE, Apr. 2019, pp. 450–454. doi: 10.1109/ICECA.2019.8822022.
- [6] E. Uzun, "A Novel Web Scrapping Approach Using the Additional Information Obtained From Web Pages," *IEEE Access*, vol. 8, pp. 61726–61740, 2020, doi: 10.1109/ACCESS.2020.2984503.
- [7] V. Singrodia, A. Mitra, and S. Paul, "A Review on Web Scrapping and its Applications," in *2019 International Conference on Computer Communication and Informatics (ICCCI)*, IEEE, Apr. 2019, pp. 1–6. doi: 10.1109/ICCCI.2019.8821809.
- [8] V. Krotov, L. Johnson, and L. Silva, "Legality and Ethics of Web Scrapping," *Communications of the Association for Information Systems*, vol. 47, pp. 539–563, 2020, doi: 10.17705/ICAIS.04724.

- [9] N. Singh, A. Rana, and A. Chaudhary, "Price Comparison Using Web Scraping and Machine Learning," in *2023 International Conference on Computer Science and Emerging Technologies (CSET)*, IEEE, Apr. 2023, pp. 1–5. doi: 10.1109/CSET58993.2023.10346784.