IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, April 2025
Impact Factor: 7.67

Automated Website Suggestion for E-Commerce through Web Scraping

Naveena A[1], Nithya Sri R[2], Subashini S[3], Varsha V[4], Dr. L. R. Sudha[5]
Students, Department of Computer Science and Engineering[1,2,3,4]
Associate Professor, Department of Computer Science and Engineering[5]
Annamalai University, Annamalai Nagar, Tamil Nadu, India

Abstract: *Capturing real time product information from e-commerce platforms is the core task of the project. In this project by using python web crawling technologies like beautiful soup and selenium we crawl the details of requested product from different e-commerce platforms such as Amazon, Beautiful Soup, which is a python library is used to get requests from users. After getting the requests, Requests library is used to fetch the HTML content of the target e-commerce pages. From the fetched HTML content, specific products details like name, price, description, ratings, discounts, delivery charges, and reviews are extracted and saved as a csv file. Then the details of the product with minimum price is identified after conducting a comparative analysis and displayed. This automation enhance consumer decision making by aggregating and analyzing product data from online retailers in real time. The performance of the proposed project will be evaluated by using error rate and by measuring speed retrival.*

Keywords: Web Scraper, E-commerce, Price Comparison

I. INTRODUCTION

The need for on-line shopping over the traditional Shopping is increasing day by day. To perform this people are utilizing electronic gadgets such as tablets, cell phones and PCs to get into an E-Commerce sites through the Internet. Web scrapping is a new technique developed for gathering and downloading the web information for data analysis in on-line shopping .

The Product Comparison Program is designed to automate the process of comparing products across different e-commerce platforms, allowing users to make informed purchasing decisions. The program begins by initializing the necessary configurations and libraries required for web scraping, user input handling, and data presentation. The user enters the names of the products they wish to compare, which are then used to search for relevant products on popular e-commerce sites like Amazon and Flipkart. The program extracts key details such as product titles, prices, and ratings from the product pages using web scraping techniques powered by libraries like Beautiful Soup and Selenium.

Once the data is fetched, the system compares the products based on the extracted attributes and present the comparison results in a user-friendly display dashboard. This dashboard presents the product information in a clear, structured format, helping users quickly identify the product that offers the best value. After displaying the results, the comparison data is saved to a text file for future reference. The program ensures a seamless and efficient experience by automating what would otherwise be a tedious and time-consuming manual process. By combining web scraping, data analysis, and intuitive presentation, this system empowers users to make better purchasing decisions without the hassle of browsing through multiple online stores.

II. LITERATURE SURVEY

Web scraping, the process of extracting information from websites, has evolved significantly over the years, and has become a critical tool for data collection in various domains. This survey explores the techniques and tools associated with web scraping to provide a comprehensive understanding of its current state and future directions.

DOI: 10.48175/IJARSCT-24801

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, April 2025
Impact Factor: 7.67

In[1]**, Dhruv Agarwal and Falguni Joshi** Developed a Product comparison website utilizing webscraping techniques to extract product information &Techniques like Beautiful Soup, Flask, Selenium. Data extracted from multiple ecommerce Platforms including product prices, product features, user ratings.

In[2], **Fan Chen,** Explored Python libraries Beautiful Soup, Selenium for Price comparison Highlights challenges in dynamic content extraction and antiscraping. Taobao, JD.com, and Amazon collected using web scraping tools including product prices, product names, product details.

In[3], **Dr.K.V.Nagendra, et al.,** Developed a price Comparison system leveraging Web scraping and node.js, for Enhancing data accuracy and Scalabilty. Amazon, Flipkart, and ebay including product prices& product details.

In[4], **Aryaman Singh Chandrawat, et al.,** Proposed a framework using web Scraping to optimize e- commerce Decisions by analyzing product data and pricing strategies, focusing on ethical compliance. Data Scraped from e-commerce websites including product names, prices, images, and competitor data.

In[5], **Arman Shaikh, et al.,** Developed a Price comparison website using web scraping to help consumers to find the best deals through online.Price data from Amazon, flipkart, Snapdeal and Croma for including product title, price, URL, historical prize data.

In[6], **Arman Shaikh, et al.,** Developed a Price Comparison website using web scraping techniques to help consumers to find the best deals through online, saving time and Money. E-commerce platform for price comparison including product title, price, URLs and historical price data.

In[7], **Aswad Shaikh, et al.,** Developed a Product Comparison website that utilizes webscraping to collect data from various e- commerce platforms.The System Employs a customized algorithm considering factors like prices.Various Product websites including product prices, product features, user ratings.

In[8], **Pavan Sai, et al.,** Developed a System utilizing web crawling and Scraping techniques to fetch product details from various e-commerce websites. The Extracted data is stored in MongoDB. Data extracted from Multiple E-commerce platforms including product details, product prices, product urls**.**

In[9], **k.Varun, et al.,** Proposes a Selenium based system for Price comparison with real-time data extraction and user friendly features Amazon and Flipkart including product prices, product details, product URLs.

In[10], **M.Sowmiya, et al.,** Proposed a system using web Scraping and crawling for price comparison, Featuring AES for Security and Discount notifications for users. Flipkart, Shop Clues, and Snapdeal including product prices ,discounts and user reviews.

III. METHODOLOGY

The Automated Website Suggestion for E-commerce through Web Scraping system extracts and compares product details from online platforms like Amazon to help users make informed purchasing decisions. The process begins with the user selecting a product, followed by automated searches on e-commerce website like Amazon using web scraping tools like Beautiful Soup and Selenium. The system fetches relevant product details, including product name, prices, ratings, and reviews using sentimental Analysis and saves them in a CSV file for further analysis. A comparison algorithm then evaluates the extracted data to identify the best deals, which are displayed on a dashboard with filters, sorting options, and graphical representations. The system can also incorporate automated periodic updates to track price fluctuations over time. By leveraging Python, and visualization tools like Matplotlib and Plotly, this approach enhances the online shopping experience by providing real-time, data-driven product recommendations.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, April 2025
Impact Factor: 7.67

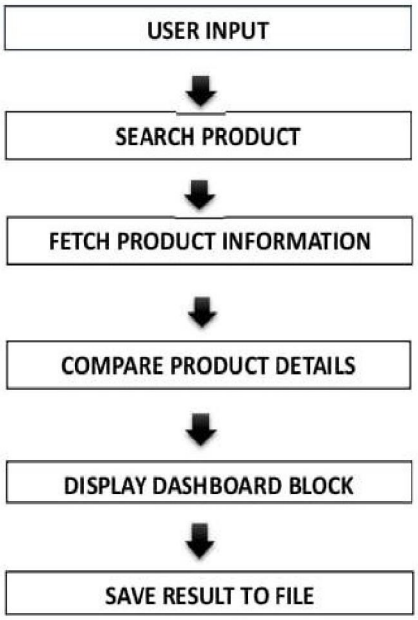

Figure 1. Block Diagram

1. SEARCH PRODUCTS

The Search Product Module automates retrieving product details from Amazon based on user input. It sends search queries, navigates results, and extracts data like name, price, ratings, and product URL using web scraping. Pagination handling ensures comprehensive data collection. Error-handling techniques manage network issues, CAPTCHAs, and dynamic content loading. The system ensures accurate and efficient data extraction for analysis. This module forms the foundation for product comparison and recommendations..

Figure 2. Select Product

2. SELECT PRODUCT

The Select Product Module in the Automated Website Suggestion for E-commerce allows users to enter a product name to search exclusively on Amazon. It ensures accurate input validation to prevent errors and refine search accuracy. The system processes the user's request and retrieves relevant product details using web scraping techniques. An auto complete or suggestion feature can enhance user experience by providing popular search terms. Once the product is selected, the system fetches and stores the extracted data for comparison

This module serves as the starting point for generating meaningful product insights.Requests library is used to send HTTP requests to web servers and retrieve HTML content or other data from web pages.Beautiful Soup is used to parse HTML and XML documents, making it easy to navigate and extract specific data from web content.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, April 2025
Impact Factor: 7.67

3. FETCH PRODUCT

This module extracts key product details from the fetched URLs and scraps the following information from the product pages.

- **Title:** the product's full name or description.
- **Price:** the current price of the product.
- **Rating:** the customer rating or review score.
- **Review:** the customer reviews of product.
- **Delivery cost:** Delivery cost of product

These details are essential for meaningful product comparison.

The fetched results are stored in a csv file for future reference by this module. Product names, prices, ratings, reviews, delivery cost and comparison summary allows users to keep a record of the comparison for decision-making.

4. COMPARE PRODUCT DETAILS

This module is implemented to compare the scraped details of the two products.

Title Comparison: Ensures the right products are being compared.

Price Comparison: Identifies the cheaper product or highlights price differences.

Rating Comparison: Helps users choose based on customer reviews.

Review Comparison: Analyzes customer feedback to highlight the product with the most positive sentiments.

Delivery cost comparison: identifies the product with the lowest delivery fee or compares shipping challenges.

Displays the comparison results in a user-friendly format.

5. DISPLAY DASHBOARD

Product comparison results are displayed in a clear, user-friendly format from the comparison block as shown in Table 1. This enhances user experience through structured text, tables, or graphical dashboards using libraries like Tkinter .

Product Comparison Dashboard

Feature	Product 1: Apple iPhone 15	Product 2: Samsung Galaxy S23
Title	Apple iPhone 15	Samsung Galaxy S23
Price	₹79,999	₹74,999
Rating	4.5 Stars	4.3 Stars
Reviews (Sentiment)	Mostly Positive	Mixed
Delivery Cost	Free	₹199
Key Insights	Better Performance & Design	Affordable with Good Features

Table 1. Product Comparison

IV. EXPERIMENTAL RESULTS AND ANALYSIS

Figure 3, 4 and 5 shows the price, ratings and sentiment comparison of top 4 products which will help the user to take decisions about buying products. Data accuracy and scrapping speed calculated by using the equations 1 and 2 are the metrics used to analyse the proposed system

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, April 2025
Impact Factor: 7.67

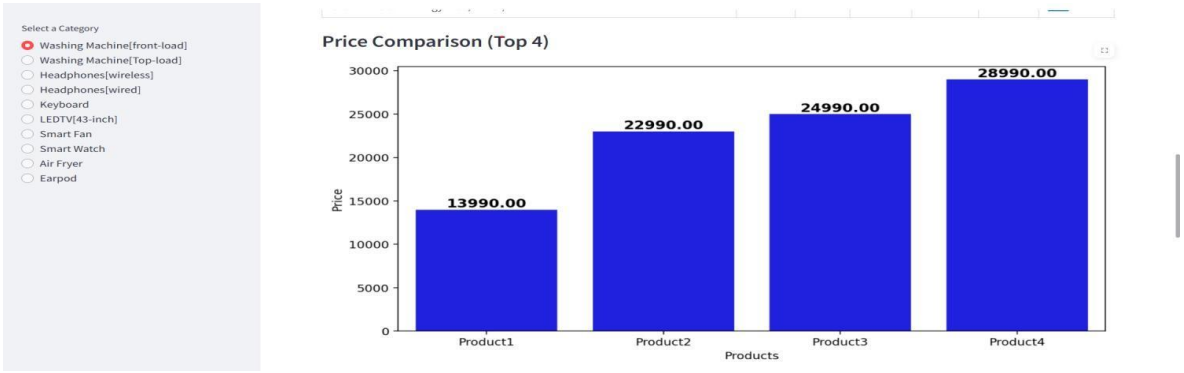

Figure 3. Price Comparison

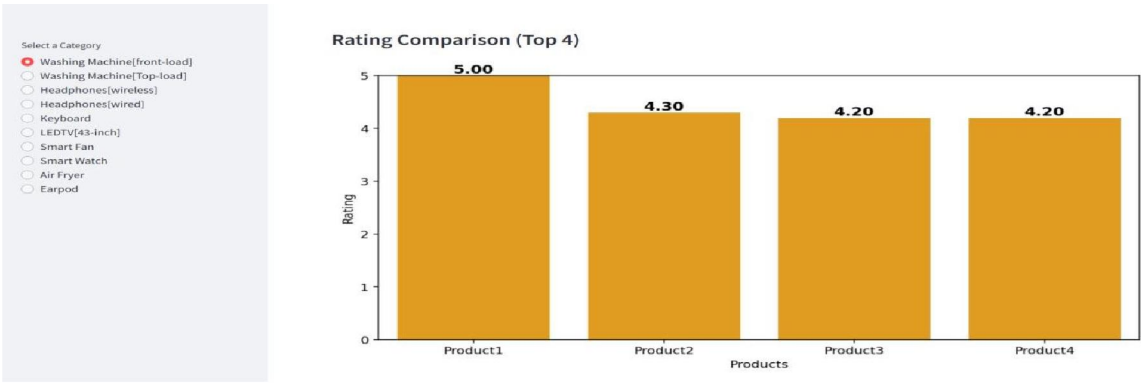

Figure 4. Rating Comparison

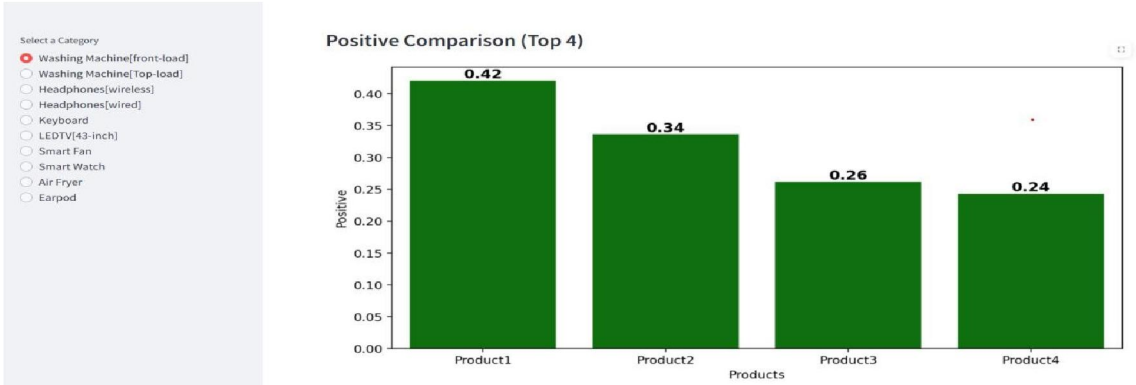

Figure 5.Sentiment Comparison

Data Accuracy

Data accuracy ensures reliable product comparisons by maintaining precise and up-to-date information. Errors may arise due to website changes, missing product details, or price fluctuations. Implementing data validation, duplicate removal, and periodic updates helps improve accuracy. Standardizing price formats and verifying extracted information against multiple sources reduce inconsistencies. Handling missing or incorrect data through error detection techniques ensures better recommendations. Maintaining high data accuracy enhances user trust and improves decision-making.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, April 2025
Impact Factor: 7.67

$$\text{Data Accuracy} = \left(\frac{\text{Correct Data Scraped}}{\text{Total Data Scraped}}\right) \times 100 \quad (1)$$

Figure 6. Data Accuracy

Scraping Speed

The Scraping Speed focuses on the efficiency of data extraction from Ecommerce platform like Amazon. This evaluates the time taken to retrieve product details. Scraping speed depends on factors like website response time. Using Asynchronous requests, headless browsing, and multi-threading can significantly improve performance. Dynamic content loading may slow down scraping, requiring Selenium or Scrapy with wait times. Optimizing request intervals and caching previously fetched data helps reduce delays. Efficient scraping ensures faster product comparisons and a smoother user experience. We obtained a speed of 2.5–3.5 seconds for Single product and ~3 sec per product for multiple products.

$$\text{Scraping Speed} = \frac{\text{Total Time}}{\text{Number of Products Scraped}} \quad (2)$$

Figure 7. Scraping Speed

Error Handling

This measure evaluates how well the scraper manages missing or incorrect data. Errors may occur due to website structure changes, anti-scraping mechanisms, missing data, or price fluctuations. These issues can lead to incorrect product comparisons and outdated recommendations. To minimize errors, techniques like dynamic scraping, data validation, and periodic updates are used. The error rate is measured by tracking failed data retrieval attempts against total requests. Continuous monitoring and optimizations ensure a more accurate and efficient system.

$$\text{Error Handling Success Rate} = \left(\frac{\text{Handled Errors}}{\text{Total Errors}}\right) \times 100 \quad (3)$$

Figure 8. Error handling

V. CONCLUSION

In conclusion, web scraping serves as more than just a tool to locate the most optimal bargains; it functions as a revolutionary force actively reshaping the commerce landscape. By equipping consumers with actionable insights and revealing the concealed mechanisms of online marketplaces, it establishes the groundwork for a shopping experience that is not only more knowledgeable but also more transparent. Our project stands as a testament to the potential for transformation that web scraping possesses, as it effectively demonstrates how this technique can provide individuals with the means to make more intelligent purchasing decisions and enable them to navigate the digital marketplace with astuteness.

REFERENCES

[1] Dhruv Agarwal, Falguni Joshi, “web Scraping based Product Comparison Model for E-Commerce Website “, Journal of Emerging Technologies and Innovative Research (JETIR), [2024], Volume 11, Issue 4.

[2] Fan Chen, “Research on Real-time E-commerce Price Comparison System Using python Web Scraping Technology “, Warwick Evans Publishing, [2024] ,Volume 4 ,Issue 2.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, April 2025
Impact Factor: 7.67

[3] Dr.K.V. Nagendra, Dr.M. Ussenaiah, "Research on Real-time E-commerce Price Comparison System Using Python Web Scraping Technology ", Warwick Evans Publishing ,[2024], Volume 4, Issue 2.
[4] Nawal kumar, Vaibhav Bohra, "Optimizing E-Commerce Decision Making Using Web Scraping", Authorea [2024].
[5] Arman Shaikh, Vaibhav Sonaje, " E-commerce Price Comparison website using web Scraping ", International Journal of Innovative Research in Engineering and Multidisciplinary Physical Devices,[2023], Volume 11 ,Issue 3 .
[6] Arman Shaikh, Raihan khan, "E-commerce Price Comparison Website using Web Scraping ", International Journal of Innovative Research in Engineering and Multidisciplinary Physical Devices , [2023] ,Volume 11, Issue 6.
[7] Aswad Shaikh,Aniket Sonmali, " Product Comparison Website using Web scraping and Machine learning" , International Research Journal of Engineering and Technology (IRJET) ,[2023] ,Volume 10, Issue 11.
[8] Sneha Sanju, Dr. S.Vijay kumar , " PRODUCT PRICE COMPARISON ON MULTIPLE E-COMMERCE WEBSITES ", International Research Journal of Modernization in Engineering Technology and Science, [2023], Volume 5 ,Issue 5.