Price Comparison Website Using Web - Scraping Algorithm

Arush Sirotiya
Department of Computer Science and Engineering with specialisation in Information and Technology
S.R.M. Institute of Science and Technology
Chennai, India
as2733@srmist.edu.in

Sanyog Dani
Department of Computer Science and Engineering with specialisation in Information and Technology
S.R.M. Institute of Science and Technology
Chennai, India
sd6513@srmist.edu.in

Abstract: Price comparison platforms are specifically designed to assess the cost of commodities and services from a variety of providers, aiding customers in their selection of products that offer the most savings when shopping online. Given the fast-paced lives of urban residents, a significant portion of consumers favor online purchases as a means to conserve time.
They can conveniently peruse these price comparison websites and determine the best source for the items they require. Optimal deals are prominently showcased. While not all consumers exclusively shop online, it remains an effective method for enhancing consumer price awareness.
Consequently, this benefits consumers who are consistently informed about the prevailing prices of particular items, safeguarding them from potentially misleading claims made by physical stores. In addition, this platform provides a valuable opportunity for grocers and retailers to promote their merchandise. This project seeks to empower consumers with the knowledge and tools needed to make informed purchasing decisions, ultimately improving their financial well-being and shopping experiences.

Keywords: Web Scraping, Price Comparison, Ecommerce, Beautifulsoup,

I. INTRODUCTION

In the contemporary digital landscape, it's become a common practice for every internet user to seek out the most advantageous bargains when contemplating a purchase. Among the paramount determinants that sway the decision to acquire a product is its price or cost. Prospective buyers habitually engage in a comprehensive price evaluation prior to making a purchase. Nevertheless, the process of scouring multiple price comparison websites for a specific item can be quite laborious. Hence, a compelling solution to streamline and automate this entire procedure is imperative.

Enter the necessity for an automated system that expedites this intricate task, saving users valuable time and effort in their pursuit of the best deals. This innovative solution is poised to revolutionize the online shopping experience, enabling consumers to effortlessly access and compare prices across a multitude of platforms, rendering their purchasing decisions more informed and economically advantageous. By harnessing the power of automation and cutting-edge technology, this project aims to redefine the way consumers engage with ecommerce and empower them to make savvy financial choices with ease and efficiency.

a) *Objective:* This platform serves as a valuable tool for comparing prices across diverse Ecommerce websites. It caters specifically to the needs of frequent online shoppers, offering the convenience of centralized price comparisons from a multitude of online retailers. This system efficiently aggregates pricing data from various sellers, enabling users to pinpoint the most cost-effective source for their desired products. Once the data from these websites is collected, it is presented on the platform in the form of an easily navigable price comparison.

   In the realm of E-commerce, applications consist of several essential components, including a database server, a web application server, and the Payment Gateway Interface (PGI) for facilitating online transactions. The pervasive influence of the internet has fundamentally transformed the way individuals and businesses approach their operations and decision-making processes. This project embodies the fusion of technology and consumer empowerment in the ever-evolving world of online shopping.

b) *Problem Statement:* An individual is seeking to purchase a Titan watch, but he's encountering varying price listings on different websites. Additionally, he's frustrated by intrusive ads and the risk of stumbling upon fraudulent websites, resulting in a significant waste of 15 to 30 minutes.

   Develop an open-source software tool that helps users avoid intrusive ads, identify counterfeit and fraudulent websites, and obtain up-to-date, accurate prices for authentic and high-quality products from online retailers.

c) *challenges:* Price comparison website developers encounter several challenges while creating and maintaining their platforms:

1. Data Aggregation: Gathering and updating links of different ecommerce and their classes can be challenging due to differences in data formats.
2. Data Accuracy: Ensuring the accuracy of price and product information is crucial to gain user trust, and this requires continuous data validation and cleansing.
3. Scalability: As the number of products and retailers increases, the platform needs to scale efficiently to handle the growing data volume and user traffic.
4. Website Performance: Maintaining a responsive and fast loading website, especially when dealing with large datasets, can be a technical challenge.
5. User Experience: Designing an intuitive and user-friendly interface is crucial for retaining and attracting users. This involves addressing navigation, search functionality, and overall usability.
6. Monetization: Finding a sustainable revenue model, such as affiliate marketing, sponsored listings, or premium subscriptions, is important to support the platform's operation.
7. Data Privacy and Security: Handling sensitive user data and adhering to privacy regulations is essential. Developers must implement strong security measures to protect both user data and the platform itself from cyber threats.
8. Competition: Staying competitive in the crowded price comparison website market and differentiating from other platforms can be challenging.
9. SEO and Marketing: Attracting users to the platform requires effective search engine optimization (SEO) and marketing strategies to increase visibility and drive organic traffic.
10. Legal Compliance: Adhering to copyright, trademark, and data usage laws when scraping data from third-party websites is essential to avoid legal issues.
11. Continuous Maintenance: Keeping the platform up-to-date with changing retailer websites, adding new features, and fixing bugs is an ongoing challenge.
12. Mobile Responsiveness: Ensuring the platform is mobile-friendly is increasingly important as more user access price comparison websites through smartphones and tablets.
13. Price Tracking: Implementing price tracking features to alert users when the price of a product drops or when a deal becomes available can be complex.
14. Handling User Reviews and Feedback: Incorporating user reviews and feedback effectively while maintaining their credibility and authenticity can be challenging.
15. Price Discrepancies: Dealing with situations where retailers provide different prices for the same product or have outdated information requires careful handling.
16. Web Scraping Challenges: Overcoming website structure changes and implementing efficient scraping techniques.
17. Database Management: Designing an optimized schema and maintaining data consistency.
18. User Experience: Ensuring the website is user-friendly and responsive.

Overcoming these challenges requires a combination of technical expertise, effective data management, user-centric design, and a commitment to maintaining data accuracy and user trust.

II. MATIRALS AND METHODS

It seems like you are working on a script that compares prices of different products from websites. The script retrieves information from a MySQL database, fetches data from the specified URLs, and then extracts and compares the prices using BeautifulSoup.

Here are some methods and materials used in your script:

Methods:
1. Database Connection:
- `mysql.connector.connect`: Establishes a connection to the MySQL database.

2. Data Retrieval from Database:
- `cursor.execute`: Executes SQL queries to retrieve data from the database.

3. Web Scraping:
- `requests.get`: Sends HTTP requests to the specified URLs.
- `BeautifulSoup`: Parses HTML content and extracts information from the websites.

4. Data Extraction:
- `find`: Searches for HTML elements with the specified class type and class name.

5. Printing and Display:
- `print`: Displays information and messages on the console.

Materials:
1. MySQL Database:
- Host: localhost
- User: root
- Password: your_password
- Database: project_database

2. Products and URLs:
- Titan Watch: URLs retrieved from the database using the product_id 'watch1' and 'watch2'.
- Victus Laptop: URLs retrieved from the database using the product_id 'laptop1' and 'laptop2'.
- Boat Airdopes: URLs retrieved from the database using the product_id 'boatairdropes1' and 'boatairdropes2'.

3. Web Scraping Libraries:
- `requests`: Used for making HTTP requests to fetch the HTML content of the websites.
- `BeautifulSoup`: Used for parsing HTML and navigating the parse tree.

Output Display:
1. Product Information:
- Displayed product details such as brand, model, specifications for each product.

2. Data Fetching Confirmation:
- Confirmed whether data was fetched successfully from the websites.

3. Price Comparison:
- Extracted and compared prices from the two websites.

4. Final Display:
- Displayed prices and URLs for both websites.

Additional Notes:
- The script is structured using the new Python 3.10 `match` statement for better readability.
- The script exits when the user selects option "4."

Make sure that the URLs and the HTML structure of the websites remain consistent for the script to work correctly. Additionally, handle exceptions and errors for robustness in a production environment.

III. Conclusion

our price comparison website project is a significant achievement, demonstrating our skills in web scraping, database management, and web development. It provides a valuable service to users by simplifying their product research and saving them time and money. As we continue to work on this project and make improvements, it has the potential to become a popular destination for online shoppers seeking the best deals. It showcases our ability to create practical, data-driven solutions and opens the door to various opportunities for expansion and monetization. Keep refining and enhancing our project to stay competitive in the dynamic e-commerce landscape.

Acknowledgments

We express our heartfelt thanks to our honorable Vice Chancellor Dr. C. MUTHAMIZHCHELVAN, for being the beacon in all our endeavors.

We would like to express my warmth of gratitude to our Registrar Dr. S. Ponnusamy, for his encouragement.

We express our profound gratitude to our Dean (College of Engineering and Technology) Dr. T. V.Gopal, for bringing out novelty in all executions.

We would like to express my heartfelt thanks to Chairperson, School of Computing Dr. Revathi Venkataraman, for imparting confidence to complete my course project

We wish to express my sincere thanks to Course Audit Professors Dr. Vadivu. G , Professor, Department of Data Science and Business Systems and Dr. Sasikala. E Professor, Department of Data Science and Business Systems and Course Coordinators for their constant encouragement and support.

We are highly thankful to our Course project Faculty Dr. Manickam. M, Assistant Professor, Department of Networking and Communications, School of Computing School of Computing, for his/her assistance, timely suggestion and guidance throughout the duration of this course project.

We extend my gratitude to our HoD Dr. Annapurani Panaiyappan. K (Networking And Communication) and my Departmental colleagues for their Support.

Finally, we thank our parents and friends near and dear ones who directly and indirectly contributed to the successful completion of our project. Above all, I thank the almighty for showering his blessings on me to complete my Course project.

References

[1] Priya Matta, Sonal Sharma, Nitin Uniyal, "Comparative Study Of Various Scraping Tools: Pros And Cons," in 2023 IEEE Conference on Internet of Things, Green Computing and Communications (IoTGC), 2023, pp. 1-5, doi: 10.1109/IoTGC57244.2023.2915787.

[2] Harikirshnan.K, Nagavigneshwar.R, Vignesh.R, Dr.R.Santhosh, R.Reshma, "Intelligent Online Shopping using ML-based Product Comparison Engine," in 2022 IEEE International Conference on Artificial Intelligence in Information and Communication (ICAIIC), 2022, pp. 585-589, doi: 10.1109/ICAIIC55531.2022.10104789.

[3] Ayush Asawa, Swapnil Dabre, Shravani Rahise, Manisha Bansode, Kiran T. Talele, Priya Chimurkar, "Co-Mart - A Daily Necessity Price Comparison Application," in 2021 IEEE International Conference on Computer, Communication and Electrical Technologies (ICCCET), 2021, pp. 1-4, doi: 10.1109/ICCCET52102.2021.9614129.

[4] Harsh Khatter, Dravid, Akshat Sharma, Ajay Kumar Kushwaha "Web Scraping based Product Comparison Model for E-Commerce Websites," in 2020 IEEE 18th International Conference on Advanced Communication Technology (ICACT), 2020, pp. 1-5, doi: 10.1109/ICACT50364.2020.9272524.