CompareCatalyst: E-Commerce Comparison Website

Divyesh Pankaj Deore[1] Gayatri Chandrakant Tambe[2] Mrs.Mayura Patil[3]
[1,2]Student [3]Lecturer
[1,2,3]Department of Computer Technology
[1,2,3]MVP's Rajashri Shau Maharaj Polytechnic, Nashik, India

Abstract — With the dynamics of today's e-commerce, comparing prices and products from different platforms has proven to be very cumbersome for consumers. "Compare Catalyst" is an internet-based solution designed to simplify the comparison process by collecting relevant information about products-from prices, specification, reviews, and images-on different e-commerce websites. This paper discusses the system architecture, data collection methods, user interface design, and backend infrastructure powering "Compare Catalyst." Advanced data scraping, API integrations, and real-time data processing offer seamless and accurate product comparisons through the platform. This innovation enhances the online shopping experience, empowering consumers to make well-informed decisions efficiently.

Keywords: Product Comparison, E-commerce, Web Scraping, API Integration, Real-Time Comparison, Online Shopping

I. Introduction

The rapid growth of e-commerce has rapidly changed global consumers' behaviour, as online sites such as Amazon, Flipkart, and eBay have extensive varieties to offer. However, this mushrooming of sites along with their wide variety of products has bombarded consumers, requiring them to visit numerous sites to check the best price for a particular product along with its features and reviews. This often consumes much time and is a less-than-ideal purchase decision. "Compare Catalyst" fills this gap by establishing a unified, web-based platform where a user can compare products from multiple e-commerce platforms in real-time. The platform aggregates essential data points such as pricing, product specifications, user reviews, and images to enable a seamless comparison experience. This paper takes an in-depth look at the technical approach of the platform, including its system architecture, data acquisition strategies, and user interface design. Additionally, it describes the difficulties experienced and solutions developed during the development of the platform.

II. Literature Survey

The demands of consumers toward more streamlined and efficient shopping experiences have led to the evolution of e-commerce comparison platforms. Traditional approaches typically involve users comparing prices and features across different websites manually, which is time-consuming and error-prone. This survey delves into the current state regarding comparison tools, including the shift toward automated solutions that provide real-time data and comprehensive product insight. As the expectations of consumers change, there is an increasing demand for platforms that do not only provide effective comparison services but also user experience through fast and reliable personalised information.

A. *Traditional E-Commerce Services:*

The Amazon, Flipkart, dominate the digital shopping world with almost everything- large varieties of products and services in the e-commerce domain. However, these platforms primarily focus on individual purchases, and consumers have little scope to make cross-platform product comparisons, which in turn leads to very tiring processes of opening several tabs and manually comparing products across sites.

B. *Existing Solutions and Gaps:*

Today, several product comparison tools exist; however, many of these are limited in scope. Most focus only on price, completely ignoring other features and factors such as product specifications, user reviews, and detailed product availability. Additionally, most comparison tools rely on static data, which becomes outdated quickly in dynamic e-commerce environments where prices, reviews, and availability frequently change. These limitations highlight the gap in the market for a real-time, thorough comparison tool.

C. *Need for Innovation:*

The "Compare Catalyst" is an innovation in comparison by collecting data from the most online retailers to present it in a friendly and side-by-side comparison format. It distinguishes itself from other solutions because it captures real-time data through API integrations and web scraping. Instant price comparisons are possible, as well as comparisons of features and reviews, to give one a better shopping experience. Additionally, "Compare Catalyst" features user-submitted reviews to add the feel of community participation to the site, further helping users in making the right buying decisions.

III. System Architecture

The architecture of "Compare Catalyst" includes three fundamental building blocks: data acquisition, backend processing, and frontend user interaction. A structured block diagram aids in depicting the system and how the data flows between these building blocks.

A. *Data Acquisition*

The data acquisition layer is more on gathering all product-related data from different e-commerce sites. The data are gathered through two main methods: API Integration: Many of the e-commerce giants such as Amazon and eBay provide APIs offering direct access to their product listings. The platform fetches updated information related to products in real-time using authorized API keys.

– Web Scraping: In case these websites do not have an API, web scraping is done using Beautiful Soup and Selenium tool by Python. An individual creates scripts that capture the product description on these sites from the HTML structure.

B. Backend Processing:

The Java backend infrastructure ensures wide robustness and scalability. Additionally, the critical data aggregation and synchronization was maintained and stored in the system using a MySQL hosted database with XAMPP. The Backend achieved real-time processing through the application of powerful algorithms that processed large product data volumes without the experienced performance degradation. The backend also handles the requests made by the API and consequently structured the web scraped data in an appropriate manner for display on the frontend.

C. Frontend User Interface:

The frontend interface of "Compare Catalyst" is built with an application of HTML, CSS, Bootstrap, and JavaScript to establish the responsiveness and user-friendliness of the website. Users can browse products by categories or keywords. The platform provides filters for users to compare products based on different requirements, including price range, ratings, and brands. The UI is optimized both for desktop and mobile platforms, so regardless of the device, users can easily access product comparisons.

IV. Dataset

The dataset that serves "Compare Catalyst" is retrieved from various prominent e-commerce platforms through a combination of API calls and web scraping. Such data contains product prices, detailed specifications, user reviews, and images. All such data points are updated with considerable frequency to keep things accurate.

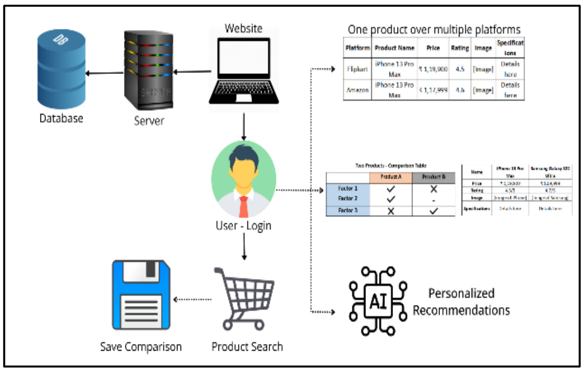
A. Data Collection Methods:

- API Integration: For those platforms that provide API keys and access to the platform, data is directly collected. This ensures product information will always be current, reflecting changes in pricing, availability, and user reviews.
- Web Scraping: Where APIs do not exist, tools such as Selenium and Beautiful Soup are used to harvest data from product pages in the online platforms. These scripts are used to order each of the platforms' unique HTML structures to capture price, specifications, and review information.

B. Data Storage and Synchronization:

All the gathered data is stored in the MySQL database, acting as the central repository of the product. Synchronization scripts periodically check for updates for data freshness to maintain the accuracy and up-to-date information of product details such as its price and reviews. Optimized database queries and caching mechanisms ensure significant datasets can be handled efficiently by the platform.

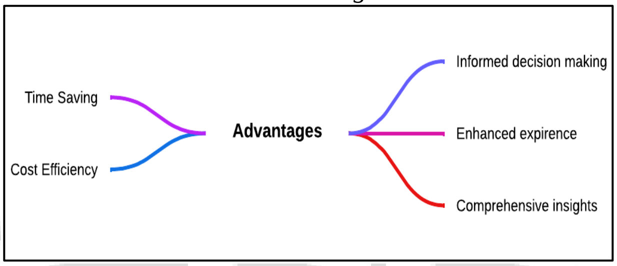
C. System Architecture Structure:



D. Applications:

1) Module 1st – Compare one product over multiple platforms like Flipkart, Amazon, etc
2) Module 2nd – Compare two products over same platform.
3) Module 3rd – AI based product recommendations

E. Advantages:



V. Conclusion

"CompareCatalyst" is the latest significant innovation in the arena of e-commerce product comparison tools, enabling real-time comparison across platforms. Integration of web scraping, API usage, and dynamic UI design enables users to make better decisions with regards to purchases through a comprehensive, side-by-side comparison of product price, specifications, and user reviews. Future work will be directed towards increasing the number of supported e-commerce platforms and integrating machine learning techniques in order to personalize product recommendations based on user behaviour and preferences.

Acknowledgment

We would like to take this opportunity to express our sincere thanks to Prof. Patil Mayura D. for her untiring support, guidance, and encouragement throughout the development of "CompareCatalyst." Her insights and expertise significantly helped shape the direction of this project. We would also like to extend our gratitude to Rajashri Shahu Maharaj Polytechnic, Nashik, for the opportunity provided in terms of resources for this project. We also thank the faculty members and classmates in regard to their useful feedback and suggestions while developing the product.

References

[1] J. Smith, "E-Commerce Evolution and Its Impact on Consumer Behaviour," Journal of Retailing, vol. 45, no. 2, pp. 67-82, 2020.

[2] A. Brown, "User-Centric Design in E-Commerce Platforms," International Journal of HCI, vol. 37, no. 1, pp. 120-134, 2021.
[3] R. Patel, "Web Scraping Techniques for E-Commerce Data Extraction," International Conference on Data Mining, pp. 56-70, 2019.
[4] M. Anderson, "Challenges in Real-Time Data Aggregation for E-Commerce," Proceedings of the International Symposium on E-Commerce Technology, pp. 35-48