

Web and Android Application for Comparison of E-Commerce Products

Aditya Ambre¹, Praful Gaikwad², Kaustubh Pawar³, Vijaykumar Patil⁴

^{1,2,3} Student, Department of Information Technology, Bharati Vidyapeeth College of Engineering, Navi Mumbai.

⁴ Assistant Professor, Department of Information Technology, Bharati Vidyapeeth College of Engineering, Navi Mumbai.

Abstract— E-Commerce websites are currently the most popular sources for shopping all kinds of products online. Currently, users have many different websites to access and search for the desired products available in the market. Nowadays, different kinds of strategies are being implemented to analyze and understand the customer's behavior to increase growth of the business. As, variety of websites are currently available it becomes difficult for the users to purchase the desired product for an affordable price. The paper presents, Comparison of E-Commerce products available on these different websites in order to help users to grab their desired products for the best affordable price. Techniques like Web Crawling and Web Scraping are adopted to collect detailed product information from the websites and MongoDB (NoSql Database) is used to store the scraped details of the products. Libraries like Requests and BeautifulSoup4 were implemented for crawling and scraping techniques using Python and Indexing method is used in MongoDB to acquire best possible results thereby savings customers time, efforts and money.

Keywords— Indexing, MongoDB, Python, Web Crawling, Web Scraping.

I. INTRODUCTION

Nowadays, due to rapid growth and advancement in the upcoming technologies internet has becoming the vital and useful in numerous fields like E-Commerce, Finance, Business, Social Networks, etc. Currently, E-Commerce has benefited many consumers all over the world to buy, sell their products on different available websites on the online platform thus making shopping easier than the traditional way, wherein the consumer needed to manually visit every local store and search for the desired product and buy if for the least affordable price. Due to the recent advancement and demand in E-Commerce, many shopping websites are available with hundred thousands categories of different products to choose from and order on the go. Thus, it becomes a tedious process for the consumers to manually visit and search the same product on different websites, to buy it at an affordable price. Therefore, it was necessary to develop price

comparison systems to help consumers to buy the products with the best deal.

Many Price Comparison systems are now available in the market. Price comparison can be done in multiple ways. Hence, these price comparison sites have made the shopping experience far easier and more convenient for customers in all aspects whether it may be payment, return of the purchased product or and in case of any further queries. Even the consumers are also satisfied with the prices and the deals they get online. The online retailers too, maintain a good relationship with the customers. It has become a common marketing gig now a days that, some of the big electronic firms launch their products directly on the E-commerce websites, because of the large number of consumers shopping/buying products online and trusting the brand.

Moreover, there are systems, extensions available they have shopping assistance which helps you suggests the best products but are not likely to compare the prices from all other E-commerce websites.

The proposed system compares the product details from different websites and provides users with an overview of the complete specifications about the product and their prices on the particular websites. It also displays about the ongoing deals and allows the users to add any desired product to the wish list in order to get notified when price drop occurs. The brand wise filter allows users to view the available products according to the brand category on the website.

II. LITERATURE SURVEY

The Comparison of E-Commerce products proposed by Riya Shah, describes about e-commerce products comparison using web mining. They created a price comparison website which was built using Django which is a Python's web framework. The data was scraped from different websites using Web Crawler and Scraper and stored in MongoDB. Another feature included was user could add same category of products to compare and analyze its details and specifications [1].

Jianxia Chen designed a Price Comparison System Based on Lucene. The system provided consumers with the

price comparison of similar products available on different online shopping malls. They adopted Lucene which is a full-text search library, to search different products based on indexing and rank or query relevance and MySQL at the backend to store the data [2].

Lucene and Deep Learning based commodity information analysis system was proposed by Jiangzhong Cao. The system adopted web crawler technology to capture the details of the commodities and used it to build a resource library along with patent information. Based on its resource library, deep learning techniques and Lucene is used to analyze information of commodities from the respective images and text [3].

Leo Rizky Julian used web scraping for comparing prices in computer parts and assembly. The paper describes about the application which allows to buy computer parts available at the cheapest price and good in quality at the online stores. Pentaho Software was used as a tool for web scraping and the application was build using PHP and javascript with MySQL as database [4].

An Evaluation of Lucene for Keywords Search in Large-scale short text storage was proposed by QIAN Liping which describes about mining huge data of the short text generated from blogs, google buzz. It focuses about Lucene indexing and searching the short-text and gives a comparison between Lucene and Oracle Text [5].

Tobias Bruggemann proposed Mobile Price Comparison Services which describes about the importance of price comparison on the electronic commerce back in year 2005. The paper focuses on importance and benefits of price comparison of the products available on the online market [6].

III. METHODOLOGY

The Comparison of products from different e-commerce websites requires the product details to be fetched from those particular websites. Thus, Web Crawling and Web Scraping techniques are adopted to fetch product's details available on different e-commerce websites. The Crawler crawls the products URL's and feed it to the Scraper, further the scraped details of the products are filtered and HTML data is extracted and saved to the local MongoDB using PyMongo. The frontend user interface is designed using PHP with login and signup options to maintain users wish listed product's data. The backend uses MongoDB for storing the user's data and products information.

Thereafter, CRON files are deployed in order to periodically update the product details and price variations on the different websites to the database. The Fig. 3.1, shows the Architecture of the proposed system.

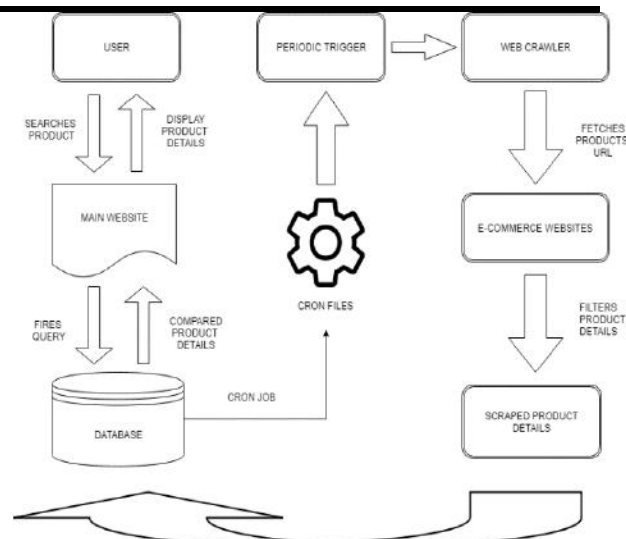


Fig. 3.1: Architecture Diagram.

A. Web Crawling

To compare products data needs to be fetched from different e-commerce websites. The amount of data is very large and cannot be collected manually by visiting the websites. Therefore, building a web crawler is beneficial, as it will automatically fetch URL's data and feed it to the scraper for scraping process. Multi-threaded crawler can be implemented to grab and fetch many URL's simultaneously to the Scraper. Requests library from Python can be used to fetch and load the data from the respective URL's.

B. Web Scraping

The Scraper extracts the text from the HTML data, collected by the crawler from the website's URL. The product's URL fed by the crawler are parsed and all the required data is collected from the websites. The collected data contains HTML tags therefore, Python's BeautifulSoup4 library is used to parse and filter out only the required data which is in the form of text and the extracted data is directly saved into MongoDB.

C. MongoDB

MongoDB is a NoSql database which stores data in document oriented format with keys and values. The JSON format used by it, is beneficial for storing large amount of unstructured data collected during the scraping process. Data extracted from different websites by the scraper is saved into MongoDB.

D. Comparison Logic

The Comparison of E-Commerce products are done on the basis of the different products attributes viz. name, price, specifications, etc. The user searches for the desired product and the query is fired to the local MongoDB database. Separate databases are allotted in MongoDB for storing the product details from different E-Commerce websites according to their categories. Therefore, the query is fired to the different databases simultaneously

and products are parsed according to their above mentioned attributes, categories and the resultant comparison between them is displayed. To make the process more efficient and faster during searching, Indexing Method is implemented in MongoDB. Due to the large amount of data stored in the database, it becomes difficult to search and handle the data during the retrieval. Therefore, indexing is applied on particular attributes of the products to efficiently filter and speed up the search process.

Pseudo Code:

Step 1: Set the URL to the desired E-Commerce website.

Step 2: Crawl and fetch all the data from the website.

Step 3: Scrape the required product details from the fetched data.

Step 4: Create new database according to the names of the E-Commerce website.

Step 5: Save the Scraped data into respective databases in MongoDB.

Step 6: Repeat the process for different E-Commerce Website.

Step 7: User searched query is fired to MongoDB.

Step 8: Product will be searched with name & category wise in the available different databases.

Step 9: If product is available in either of the database compare and display the results else display NA message.

Step 10: Periodic triggers to the CRON files to update the MongoDB with the latest available data of the products.

IV. CONCLUSION

The paper proposes, Comparison of E-Commerce Products that benefits users by allowing them to compare products available on the different e-commerce websites. Furthermore, users can filter the products according to their categories or brands available thus, allowing them to easily find and compare amongst variety of products available in the market. The wish list option provided, helps users to keep a track on the product prices and get instantly notified in case of price drops on any of the e-commerce websites. This will help to save the customers time, efforts and money. In future, the scope can be extended by including more e-commerce websites thereby providing the finest results with the best affordable deal available in the market.

REFERENCES

- [1] Riya Shah, Karishma Pathan, Anand Masurkar, Shweta Rewatkar, P.N. Vengurlekar, "Comparison of E-Commerce Products using Web Mining", International Journal of Scientific and Research Publications, Volume 6, Issue 5, May 2016 640 ISSN 2250-3153.
- [2] Jianxia Chen, Ri Huang, "Price Comparison System based on Lucene", The 8th International Conference on Computer Science & Education (ICCSE 2013) April 26-28, 2013 Colombo, Sri Lanka.
- [3] Jiangzhong Cao, Jinjian Lin, Suxue Wu, Mingxiang Gaun, Qingyun Dai, Wenxian Feng, "Lucene and Deep Learning based Commodity Information Analysis System", 2016 IEEE International Conference on Consumer Electronics-China (ICCE-China).
- [4] Leo Rizky, Friska Natalia, "The use of Web Scraping in Computer Parts and Assembly Price Comparison", Tangerang, Banten 15810, Indonesia.
- [5] Qian Liping, Wang Lidong, "An Evaluation of Lucene for Keyword Search in Large Scale Short Text Storage", 2010 International Conference on Computer Design and Applications (ICCD 2010).
- [6] Tobias Bruggemann, Michael Breitner, "Mobile Price Comparison Service", Second IEEE International Workshop on Mobile Commerce and Services (WMCS'05).
- [7] Y. Thushara and V. Ramesh, Volume 149 No.11, September 2016. A Study of Web Mining Application on E-Commerce using Google Analytics Tool.
- [8] Jos'e Ignacio Fern'andez-Villamor, Jacobo Blasco-Garc'ia, Carlos 'A. Iglesias, Mercedes Garijo "A Semantic Scrapping Model for Web Resources" Spain.
- [9] Li Mei, Feng Cheng, "Overview of WEB Mining Technology and Its Application in E-commerce", Proceedings of the 2010 IEEE 2nd International Conference on Computer Engineering and Technology, Vol. 7, 2010.