

Comparative Analysis of E-commerce Products and Online Assessment with Hiring Platform

Singh Neha¹, Taneja Sakshi², Shivade Roshani³, Pawar Pratiksha⁴, Vipul Punjabi⁵

¹Student, Department of Information Technology, RCPIT, Shirpur, India

²Professor, Department of Information Technology, RCPIT, Shirpur, India

Abstract: Web mining is the application of data mining techniques to ascertain knowledge from WWW. Web mining is applied in E-Commerce to know the browsing behavior of customers as well as in Online Assessment & Hiring Platform to know behavior of fresher. The user has to open the site, input the product details and the price and all feature comparison lists get displayed. Likewise user has to input the job details & exams And all the details are get displayed with comparison and description. The proposed web browser extension presents a price and feature comparison of all the online products to show us all the best possible prices and different related terms for the products and As well as show for the Online Assessment & Hiring Platform. We will extract the product information from URL using algorithm. The technology used for the system are python and Mongo dB (No SQL database) Web scraping and Web Crawling. The web scraping and crawling are used to extract the raw data from e-commerce websites pages. This web browser extension will allow a simple and good experience of shopping products online. In this way, our aim to provide solution for online customers to buy products at good deal and to search best job & assessment exams and save their valuable time, effort and money.

Keywords: shopping products, customer and Ecommerce

1. Introduction

The continuous growth in the size and use of the World Wide Web imposes new methods of design and development of online information services. Whenever the customer want to buy any product through online from any e-commerce websites he/she does visit the many E-commerce websites for getting the desired product like this customer surf lots of time in visiting of E-commerce websites for getting desired product. The customer only surfs lots of time in visiting of E-commerce sites, and quality of products but also he/she suffers from limited option to choose the product. Likewise, same problem face by user who visits the many websites for the placement and exams. The proposed system, by making use of web mining, offers a solution to these problems. That is for customer and students there is no need to visit many Ecommerce or exam sites. These system acts as an agent on the behalf of many sites. Real time product analysis using web mining provide the stage to customer where the user can get several of variety of a particular product in reasonable price or where can get particular product which fulfil the requirements. This system uses the following technologies:

Web Crawler: The system deals with price and job comparison engine .The first thing required are to gather large amount of data from different websites. It is not possible to manually collect the data from websites. Hence the best way is to create a web crawler that will navigate to these websites. The fetched URL's are send to scrapper for scrapping process.

Web Scrapper: Web Scrapping is used to extract HTML data from URL's and use it for personal purpose. As this is price comparison website, data is scrapped from multiple websites. In this system, Scrapping is done using python libraries like requests and beautifulsoup4. Beautifulsoup4 is a python library which is used for parsing html pages. Using these, product information from different sites is scrapped and stored in database.

MongoDB: MongoDB is a free and open-source cross-platform document-oriented database program. Classified as a NoSQL database program, MongoDB uses JSON-like documents with schemata. MongoDB is developed by MongoDB Inc., and is published under a combination of the Server Side Public License and the Apache License. MongoDB supports field, range query, and regular expression searches.

Django: Django is a free and open-source web framework, written in Python, which follows the model-view-template (MVT) architectural pattern. Django's primary goal is to ease the creation of complex, database-driven websites. Django emphasizes reusability and "pluggability" of components, less code, low coupling, rapid development, and the principle of don't repeat yourself. Comparison of E-commerce Products and Online Assessment & Hiring Platform using web mining is product and price ,job and exams comparison website which is created using Django framework. Products that are been requested by user are queried in mongodb database using an object relational mapper mongo engine.

AdminLTE Template: Best open source admin dashboard & control panel theme. Built on top of Bootstrap 3, AdminLTE provides a range of responsive, reusable, and commonly used components. We will use this to make user attractive website, thus user will efficiently user our website.

2. Literature survey

There are various researchers who have contributed their work for Comparative Analysis of Ecommerce Products &

Online Assessment and Hiring Platform .The researchers have proposed novel techniques with suitable implementations. Now we concentrate on few researchers works as given below. Rajesh R. Gawali and Dr.Shivaji D. Mundhe discussed techniques, process of web mining, applications and opportunities of Data Mining in Ecommerce. Web Personalization is the need of the time due to huge and ever growing volume of web content that is open for access to any user. It will be make the web useful in all walks of life.Web personalization involves prediction of the web content required by a surfer at any time and delivers them simultaneously. Web mining is the application of data mining techniques to extract knowledge from Web. Using Web mining different techniques have been proposed for a variety of applications includes Web Search, Classification and Personalization etc. Most research on Web mining has been from a 'data-centric' point of view. Good web personalization can be achieved by implementing personalization activities at various levels in the web. Web structure mining and web content mining can contribute to complete web personalization [1]. Ahmad Tasnim Siddiqui and Sultan Aljahdali discussed some important mining techniques which are used in data mining. After that explained the proposed architecture which contains mainly four components business data, data obtained from consumers interaction, data warehouse and data analysis. After finishing the task by data analysis module it'll produce report which can be utilized by the consumers as well as the e-commerce application owners. In future this model can be improved more users interactive and applicable in peer to peer applications [2].

Arjun Sidana and Dr.Himanshu Aggarwal reviewed various web mining algorithms and techniques that have been used by the previous researchers. In-depth analysis of the data mining and web mining help to identify the benefits and limitations of these techniques. Along with this, provided a proper process of web usage mining with three phases including the preprocessing, pattern discovery and pattern analysis [3]. Shweta Rewatkaret. al. determined Product Analysis and Comparison using web mining is web based system which will help users in decision making while buying products online. This website will facilitate users to analyze prices that are present on different e-commerce shopping websites so that they get to know the cheapest price of product with best deal. The website will also have the facility of comparing products with all its specifications that belong to same category. This will surely save buyers efforts and valuable time. Ultimately, this will bring together strategies, best offers and deals from all leading online stores and will help buyers to shop online [4]. Hitendra A. Chavanet. al. proposed web browser extension presents a price comparison of all the online products to show us all the best possible prices for the products. The extract the product information from URL using Pattern matching algorithm. The technology used for the system are Node.js, Angular.js and Mongo dB (No SQL database) Web scraping and Web Crawler. The web scraping and crawler are used to

extract the raw data from e-commerce websites pages. This web browser extension will allow a simple and good experience of shopping products online[5]. Y. Thushara and V. Ramesh analyzed and also identified whether Google Analytics can be considered as a state-of-the art alternative to collect data for web usage mining. The principle is to cluster customer segments by using automatic discovery and analysis of patterns in E-Commerce website which input data comes from web log of various e-commerce websites. The authors analyzed one of the leading E-Commerce Software (OPENCART) to track the information of the web users. The authors implement into the Google Analytics Tool for the report of traffic information of the users [6]. Jawahire Nakash Karim Zohraet. al. determined the Real Time product Analysis using Data mining is a price comparison engine that aims to facilitate the buyers to compare products from different E-commerce websites and purchase the product at the cheapest rate with best price. This way, the buyer has more power in his/her hands and can take better decision on different products at different price. Thus this project saves buyers efforts, time and money and also avoids user to physically visit each and everywhere Ecommerce website. The beauty of this project is that it can be customize for a specific business segment and also can be used. It also helps the different E-commerce applications to boost their business by providing them a platform to compete and do business in a more reasonable manner. By involvement of data intelligence and also comment and predict about the services and quality provided by the different E-Commerce. So that user can be help to choose a better deal [7].

Ron Kohavi et. al. determined the architecture of Blue Martini Softwares ecommerce suite has supported data collection, data transformation, and data mining since its inception. With clickstreams being collected at the application-server layer, high-level events being logged, and data automatically transformed into a data warehouse using metadata, common problems plaguing data mining using weblogs (e.g., sessionization and conflating multi-sourced data) were obviated, thus allowing us to concentrate on actual data mining goals. The paper briefly reviews the architecture and discusses many lessons learned over the last four years and the challenges that still need to be addressed. The lessons and challenges are presented across two dimensions: business-level vs. technical, and throughout the data mining lifecycle stages of data collection, data warehouse construction, business intelligence, and deployment. The lessons and challenges are also widely applicable to data mining domains outside retail e-commerce [8]. Nasrin Jokar et. al. proposed methods and techniques which are required for the use of this information and extract new information from them by creating and deploying Web and a significant increase in the volume of data. Web mining as new knowledge and the practical tools has been emerged to help users and webmasters. Web mining is divided into three methods: 1. web usage mining, 2. web structure mining, 3. web content mining which web usage mining

techniques include association rules, sequential pattern, clustering, that can be used to develop the site. There are several functional areas in web mining, e-commerce and customer relationship management are the most ones and numerous articles have been written in these fields. Now, many researches are conducting in web mining to destroy these challenges. Although the explaining all methods and applications in this area is not possible, this article can give the reader an overview of web mining and resources and guide him to resources which are his interests. Web mining is faced with challenges such as incorrect and inaccurate data, lack of tools, custom tools, lack of the required resources, management and so on [9]. Riya Shah et. al. determined comparison of E-commerce products using web mining is web based system which will help users in decision making while buying products online. This website will facilitate users to analyze prices that are present on different e-commerce shopping websites so that they get to know the cheapest price of product with best deal. The website will also have the facility of comparing products with all its specifications that belong to same category. This will surely save buyers efforts and valuable time. Ultimately, this will bring together strategies, best offers and deals from all leading online stores and will help buyers to shop online [10]. Hsinchun Chen et. al. designing a class project for students in a graduate course to use open Web APIs for developing Web mining applications. Overall, the results are encouraging. Through these Web mining projects, we observe that most students were able to build innovative Web applications within a short period of time; which would be impossible if the students had to develop the systems from scratch. The students developed the abilities to create interesting business models and integrate necessary system components to implement them. In addition to the three suggested APIs from Google, eBay, and Amazon, students also successfully identified and incorporated other Web APIs and data mining/visualization tools in the projects [11].

3. System architecture

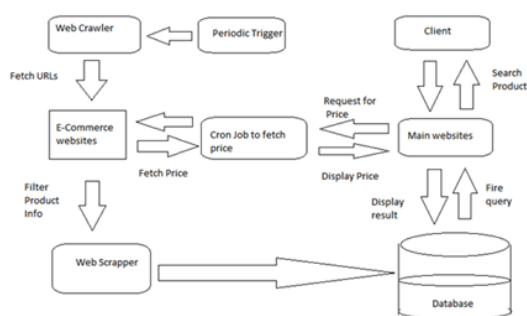


Fig. 1. System architecture

It describes Fig. 1, system architecture and its detailed working procedure. The front end system provides a graphical user interface (GUI) in the form of website where clients

interact with the system whereas the backend consists of web crawling and scrapping techniques in order to extract product information from different websites. The extracted information of products is stored in Mongo DB database. Client requests for desired product from main website and query is fired in local database. Product Information is displayed on main web page. Client can see prices, job, and exam of required product at one place present on different firms. Another feature is provided on the website that compares products. User can add products of same the category to compare. User may also analyze the product for its details and specifications [3].

4. Implementation

Working of the proposed system is as follows: The backend system consists of two important techniques web crawling and web scrapping. Web scrapping is a technique that is used to extract information in the human readable format and display it on destination terminal. But before scrapping the output, Web Crawlers are responsible to navigate to the destination once the crawler reaches the correct page and matches up with the products, scrapping process starts. Crawler periodically fetches information from e-commerce websites so as to check for updates. If updates are available crawlers carries those updates and makes necessary changes in the database. Web scrapping essentially consists of two tasks: first is to load the desired web page and second is to parse HTML information of the page to locate intended information. In this system Scrapping is done using python as it provides rich set of libraries to address these tasks. “Requests” is used to load the urls and “Beautiful soup” library is used to parse the web page. After scrapping the products information from different e-commerce websites the data is stored in Mongo DB database. Using pymongo connectivity data is scrapped and stored in database. The frond end consists of Main website. The client searches for the required product in search bar and query is fired in local database i.e. Mongo DB. The website is designed using Django web framework which is written in python. Communication is done between python web framework and Mongo DB using Mongo engine which is a python object –document-mapper working with Mongo DB. Required results are retrieved and displayed on Main website. The client can then compare prices of products that are available on e-commerce websites. A soon as client selects on best deal according to them, they will be redirected to the original e-commerce website. Another feature provided is, Clients can compare products that belong to same category so as to differentiate specifications and choose accordingly.

5. Conclusion

Web Mining is a developing domain Web mining is increasingly reaching in mostly aspects of today’s environment. In depth analysis of the data mining and web mining help to identify the benefits and limitations of the techniques. Along with this, it could be provided with a proper process of web

usage mining with three phases including the preprocessing, pattern discovery and pattern analysis. The Existing systems provide good functionalities, but they have certain drawbacks too. The current existing systems generate more traffic because the technique used by them is web crawling. The immense popularity of the price comparison systems means that these sites are here for long and continue providing amazing cost savings to consumers on the e-commerce ecosystem and Likewise, for the online assessment and hiring platforms. Our price and job comparison extension will be very useful to consumers/users which will enhance their shopping experience which is done online

References

- [1] Hsinchun Chen, Xin Li, Michael Chau, Yi-Jen Ho, Chunju Tseng. Using Open Web APIs in Teaching Web Mining.
- [2] Rajesh R. Gawali and Shivaji D. Mundhe(2013). Web Mining techniques, process and applications in Ecommerce.
- [3] Ahmad Tasnim Siddiqui and Sultan Aljahdali Volume 69 No.8, May 2013. Web Mining Techniques in E-Commerce Applications.
- [4] [https://en.wikipedia.org/wiki/Python_\(programming_language\)](https://en.wikipedia.org/wiki/Python_(programming_language)).
- [5] <https://en.wikipedia.org/wiki/HTML>.
- [6] Hitendra A. Chavan, Harshad Bhojane, Rohan Naik, Osho Priyadarshi "Commodity Search Using Web Browser Extension".
- [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] Ansari Mohd Azam Mohd Salim Reshma Bano(11CO53), Jawahire Nakash Karim Zohra(11CO56), Shaikh Anas Munawwar Aisha(11CO38), Siddiqi Muzammil Ahmad Mansoor Ahmad Razia Begum(11CO55). "Real Time Product Analysis using Data Mining".
- [9] Ron Kohavi, Llew Mason, Rajesh Parekh "Lessons and Challenges from Mining Retail E-Commerce Data".
- [10] Nasrin Jokar, Ali Reza Honarvar, Shima Aghamirzadeh, Khadijeh ESFANDIARI, Bulletin de la Socit des Sciences de Lige, Vol. 85, 2016, p. 321-328, "Web mining and Web usage mining techniques"
- [11] Han, J., Kamber, M., Pei, J. (2011). Data mining: concepts and techniques. Elsevier.