A project report on

E-COMMERCE SCRAPER

submitted in partial fulfillment

for the

Degree of Bachelor of Science in Computer Science

By

AKASH RAKESH YADAV

ID:21334008

Under the supervision of

Mrs. DHANUSHYA MAXWELL

Assistant Professor

Department of Computer Science

SPICER ADVENTIST UNIVERSITY

Aundh Road, Ganeshkhind,

Pune, Maharashtra - 411067

A project report on

E-COMMERCE SCRAPER

submitted in partial fulfillment

for the

Degree of Bachelor Of Science in Computer Science

By

AKASH RAKESH YADAV

ID:21334008

Under the supervision of

" **Mrs. DHANUSHYA MAXWELL** "

Assistant Professor

Department of Computer Science

SPICER ADVENTIST UNIVERSITY

Aundh Road, Ganeshkhind,

Pune, Maharashtra -411067

ABSTRACT

Title of the project: E-COMMERCE SCRAPER

Price comparison sites are designed to compare the price of goods and services from a range of providers, which will help consumers in making decision to choose products that will save their money through online. Considering the customers' busy lifestyle especially those who are living in the city area, most of the consumers prefer to buy their needs through the internet because it saves their time. Besides, consumers always go for the cheaper price in purchasing products therefore by using price comparison website, customers do not have to travel from shop to shop only to survey the price offered by different shops for the same product. They can just check it from the price comparison website itself and decide where they should buy the products they need. This project, named as Price comparison website using web scrapping is the place where shoppers could find the great deals on the products. The best deals will be clearly highlighted. To obtain best deals from Price comparison websites web scrapping techniques are used to fetch detailed information. This way, paper aims to provide solution for online customers to buy products at good deal and save their valuable time, effort, and money.

ACKNOWLEDGMENTS

I would like to express my gratitude to our Dean, **Dr. SUSAN THOMAS** for giving me the opportunity to do this project on the topic. I also express my sincere thanks to my guide,

Mrs. Dhanushya Maxwell for guiding and helping me throughout this project.

This project has helped me to do a lot of research work and enlightened me with a lot of technical knowledge related to my project.

I would also like to thank all the other teachers, for their kindness and support towards me and my work that helped me to complete my project on time.I would like to thank all my friends for their help and support.

DECLARATION

I, AKASH RAKESH YADAV, certify that this project is my own work, based on my personal study and/or research and that I have acknowledged all material and sources used in its preparation, whether they be books, articles, reports, lecture notes, and any other kind of document, electronic or personal communication. I also certify that this project has not previously been submitted for assessment in any academic capacity, and that I have not copied in part or whole or otherwise plagiarised the work of other persons. I confirm that I have identified and declared all possible conflicts that I may have.

ID: 21334008

Date:

_____

CERTIFICATE

This is to certify that AKASH RAKESH YADAV (21334008) has worked in the, " E-COMMERCE SCRAPER" has successfully completed the project, in partial Fulfillment for the award of the degree of Bachelor of Science in Computer Science under my supervision

Date:

Mrs. DHANUSHYA MAXWELL

Project Guide

CERTIFICATE

This is to certify that AKASH RAKESH YADAV (21334008) has worked in the, " E-COMMERCE SCRAPER" has successfully completed the project, in partial Fulfillment for the award of the degree of Bachelor of Science in Computer Science under my supervision

Date:

Mrs. SUSAN THOMAS

Dean of Science Department

Table of Figures

Table of contents

INTRODUCTION

1.1 INTRODUCTION

Introduction A price comparison website acts as a platform or medium between consumers and sellers. It allows the consumers to view different price lists for the product chosen by the user and helps the consumers to make an informed decision on which one to choose to save money. It also serves as a tool to help consumers increase their price awareness, so they don't feel misled by advertisements from retailers who claim they offer the lowest price, but the reality is different.

Unlike other comparison sites, E- commerce price comparison website (the name of this project) will focus on providing a price list of products which we want to search online and purchase at a cheaper price [1] . Due to the huge increase in online users, it will be of great help to those who have busy office work and don't have much time to check the current prices of products which they want to purchase. According to Social, Digital and Mobile in India research by We Are Social, internet penetration in India is 59% and the average number of hours Indian internet users spend using the internet each week is 19.8 hours. Meanwhile, 21% of Indian internet users access the internet via mobile devices, which means that they can access the internet anywhere with their smartphones. 77% of Indian internet users shared their thoughts about a brand via social media during this study conducted in 2019. The study shows how connected people in India are to the Internet.

As mentioned above, a price comparison website acts as a medium between customers and retailers to allow customers to shop online at specific retailers to shop that offer such services [3].In other words, the price comparison website also has the function of promoting the retailer/store/hypermarket/supermarket to customers. Time and money pressures, especially in today's economy where the cost of living is rising and there is little time for home shopping, a price comparison site is sure to be of great help to consumers. In addition, today's users are very comfortable with the Internet having developed a greater variety of applications. of networking and now provide users with various references [8] .

Compared to other countries, only a few price comparison sites are accessible in India. Most of them compare the price of hotel tariff, holiday package, mobile phone and others. It is important that a comparison website provides results with low prices that match what customers want, but accurate results are also important so that customers get what they really want. It also depends on how often the database is updated; Otherwise, customers will get confused when they compare it with another website. Most workers do not have time to shop for products and other things [16]. As a consumer, you have the right to choose which store offers the best price for specific products that interest you. To the price offered by any business, however, requires a lot of time and due to a limited time, you cannot compare prices and exit the purchase of certain products at a higher price. With a catalog that published online, sellers can save costs . Increase the price awareness among consumers. Most of the running human beings do not have time to purchase products offline because of busy schedule [13]

1.2 PURPOSE OF THE PROJECT

To help online users to grab best deal for their product from multiple ecommerce websites on single web interface.

1.3 OBJECTIVES

The objective of this project is to develop an e-commerce scraping system that automates the process of gathering product data from multiple e-commerce websites. The system should be capable of extracting relevant information such as product details, prices, availability, and customer reviews in a structured and standardized format.

The objective of this project is to develop a price comparison website that will have the following functions:

1. To provide customers with a list of price comparison and highlight the cheapest price specifically in home groceries products.

2. To increase price consciousness among consumers.

3. To ensure that the price database is updated regularly so that customers will be able to get accurate results.

4. To provide service for users to find the product's price.

5. To provide platform for retailers to promote their products and promotion for free.

1.4 PROBLEM STATEMENT

Problem Statement

In the current era of online business, ecommerce has become a huge market for the people to buy goods online. Increasing use of smart devices and other mediums has paved the way for users to buy products almost from anywhere. This has increased involvement of online buyers evolving e-commerce business. These large numbers of ecommerce websites put users in turmoil to search and choose to buy a single product from multiple ecommerce websites.

Most of the working people do not have time to do shopping for their home groceries. As consumers, they have the right to choose which shop is offering the best price for a certain product that they are interested in. However, to check on price offered by each shop is time consuming and due to limited time that they have, they are not able to compare the prices and end up buying certain product with higher price. Typical mindset of customers nowadays, they see Tesco as the providers for the cheapest product, but the reality is, not every products in Tesco offered at the cheapest price. Sometimes, the smaller shop in the neighborhood offers cheaper price.

1.5 PROPOSED SOLUTION

The proposed solution helps online users to grab best deal for their product from multiple ecommerce websites on single web interface. This will in turn save users time, money, and efforts to find the same product prices on different ecommerce websites. Proposed system uses web scraping technique to extract data from ecommerce web pages and web crawler to links for products.

1.6 SCOPE

The scope of study for this project will be all internet users and it is narrowed down to those who are using internet for business-related purpose. For a price comparison website, it only provides with the useful information to the users and it is up to the users to decide which supermarket or shop they should go.

With internet, a price comparison website is accessible anytime and anywhere. The observation made by the author has resulted that people usually go online when they are on the way back from work especially those who are using public transport. Therefore, visiting price comparison website can be one of their choices to fill up their free time.

Meanwhile for the sellers/retailers, the target will be focusing more on the business that operating at the shop lot. Based on the interview done by the author with some of the owners of the shops, they admit that it will be useful for them if there is service for them to advertise their products on the web for free because they don't have much time and insufficient skills to maintain website/blog on their own.

FEATURES OF PROPOSED SYSTEM

2.1 FUNCTIONAL REQUIREMENTS

The functional requirements for an e-commerce scraping system involve describing the specific features and capabilities that the system must possess to fulfill its intended purpose effectively. These requirements typically outline the functionalities that users expect from the system. Here's a paragraph-format description of functional requirements for an e-commerce scraping system:

The e-commerce scraping system should allow users to input search queries for specific products or categories they want to monitor. Upon receiving a search query, the system should be able to scrape data from various e-commerce websites, including but not limited to Amazon, eBay, and Walmart, to retrieve relevant product information such as prices, descriptions, and availability. The system should provide users with the option to filter and sort the scraped data based on criteria such as price range, brand, and seller ratings. Additionally, users should be able to view historical pricing data trends for products over time. The system should have the capability to automatically update the scraped data at regular intervals to ensure accuracy and timeliness. Furthermore, users should be able to export the scraped data in various formats such as CSV or Excel for further analysis. Finally, the system should have authentication and authorization mechanisms to ensure that only authorized users can access and utilize its functionalities.

2.2 NON-FUNCTIONAL REQUIREMENTS AND CONSTRAINTS

Non-functional requirements and constraints for an e-commerce scraping project define the quality attributes, performance characteristics, and operational limitations of the system. Here are some non-functional requirements and constraints:

- Performance

- Reliability

- Compliance

- Resource Utilization

- User Experience

LITERATURE REVIEW

3.1 OVERVIEW

In order to make this project more credible, previous research and article that related to price comparison website has been gathered and analyzed. Hence, this chapter covers the literature review which is the analytical, critical and objective review of written materials. There are three published research articles and journals that are being used as the main references for the literature review of this project. Each is be further describes in part **2.2**, **2.3** and **2.4.**

3.2 COMPARISON SITES

This research paper was written by Moraga-Gonzalez J.L and Wildenbeest M.R and it was published in July, 2011. The research was focus on the price comparison sites and its connotation towards market efficiency and price competition. The price comparison sites attract all the involved parties no matter suppliers or the consumers to its platform as it has become the aggregator of product information.

Nowadays people usually conduct first research on the internet regarding a particular product or service that they want which is very convenient as compared to what consumers have to do previously when there is no internet. Back then, consumers have to go through books, magazines, newspaper, encyclopedia and other medium to find needed information meanwhile yellow pages, directories, advertisements and others are useful to locate business and their products. Everything has to be done manually and requires a lot of time, energy and resources to find information without the help of internet. Comparison sites are known as the second type of search technologies as it helps users find products and services according to users' have chosen.

It helps consumers to easily compare between the alternatives available in the market and reduce consumer search costs at the same time.

Compared to the price comparison site from its first being established and act as online classifieds, nowadays comparison sites able to do a lot of things such as draw the products' cost, able to trawl sites and give the sites much greater control over the users with the help of current technology. It gathers and aggregate price, products and other relevant information from third-party sellers and presents it to the consumers with its own way.

There are several business models usually used by the price comparison website and the most widely implemented is the charge will be on the sellers if they want to be listed and users can access the sites for free. The fee usually based on a cost-per-click and is paid every time a consumer is referred to the seller's website from the comparison site. Sellers also have the alternative to pay fee if a consumer buys the product which is being implemented by Pricefight.com. Other than that is providing free services for both sellers and buyers and obtain revenue from the advertisers like what being practiced by Google Product Search. The less chosen business model is the sellers are listed for free meanwhile membership fee is charged towards the consumers such as AngeList.com. Whatever model is being used; the most important thing is the agreement from the both sides on the chosen model.

Comparison sites have the potential in providing abundant of data that might be useful for further analysis for example, analyze how exactly consumers search for products and provide indications about the best way to improve the current system.

3.3 THE USE OF PRICE COMPARISON SITES

The research entitled 'The Use of Price Comparison Sites in the UK General Insurance Market' which written by Emily Knight, a strategist for Consumer Intelligence has reported the current performance, media coverage, usage and marketing activity of price comparison sites in the UK General Insurance sector.

The results shows that there is increased on the advertising spend and competition and it gives adverse effect on the financial performance towards the price comparison sites. Meanwhile the number of consumers using price comparison sites for quotes has remains high and its average number of sites used are increasing over time. Based on the research, 8 out of 10 people are likely to get quotes from the price comparison sites in the future which show good sign of price comparison sites to stay in business in future.

To compete with other brands through the internet, big brands usually have to face big advertising spends. Based on Nielsen advertising data, back in 2006 it requires £35m has increased to £85m in 2009. The amount of spend needed to compete with the big player with the big name has make it tough for other competitors unless they have a good marketing budget. Some of the comparison sites have launched various campaigns to promote their sites and get more traffic for example '*Compare the Market*' that creates a personality designed to appeal to consumers and improve their impression towards the brand at the same time increase its web traffic and boost conversion rates.

3.4 CONSUMERS PLAN MORE WEB RESEARCH BEFORE BUYING

This article was written by Reuther T. – Senior Editor of internetretailer.com, a portal related to e commerce. It is based on the findings of a survey made by Deloitte LLP, concludes that a fifth of online consumers plan to conduct more web research this year (2012) as compared to year 2011 before buying. Delaoitte LLP, a consulting, auditing, and financial services firm, has conducted the online survey on 5-12 July 2012 towards 1, 314 parents of children in kindergarten through the 12th grade. Around 20% of the survey respondents plan to shop online this year but the web is playing the important role in giving influence for the purchases either it is done online or offline (going to the shop to purchase). The result shows that a third of the parents who responded plan to visit e-commerce sites, retail blogs and other web locations to learn before buying the items.

Not only that, but the survey also resulted that 57% of the respondents owned smart phones and use it as shopping tools for seeking price information with 63%, getting coupons and discounts with 45%, and finding store locations with 38% were among the most popular mobile shopping activities. From this result, it shows that there is big number of people who are using smart phone to check on the price information of a certain product. Therefore, based on this article, it can be concluded that most of the customers, as well as potential customers of a product will make some research online before making a purchase.

METHODOLOGY

4.1 OVERVIEW

This chapter includes the research methodology, software requirements specifications for developing the system, key milestones, and the Gantt chart for the project.

4.2 RESEARCH METHODOLOGY

4.2.1 Stage 1: Project Planning

This is the stage where the problem related to the project is identified and the significance of the study is determined. The objective and also scope of study are outlined and the feasibility of the project work is ensured to be developed within the time frame given. The solution to the problem statement is studied and the types of system to be develop and tools used for developing the system are also identified through literature reviews.

4.2.2 Stage 2: Data Gathering and Analysis

Series of studies has been performed to gain further knowledge on the energy and electricity consumption. Also, readings was done to get better understanding on what comparison site is all about, how it helps people to solve problem before buying home groceries product, and example of existing comparison sites to check on the competitors. The data regarding previous researches are gained from research papers written by scholars and have been explained in details in the literature review section earlier. Meanwhile the data regarding consumers were obtained through survey and interview that made online as well as meeting the respondents (shop owners in Perak's area) face to face.

4.2.3 Stage 3: Research on any Existing Similar Systems

Next is the study performed to check if any similar system exists. The main objective of doing research on similar existing system is to know how it works, what concept is being applied in the system, what is being computed by the system and how the system helps solving the problem.

4.2.4 Stage 4: Drafting the Main Components of the System

After the research on similar existing systems, the next step needed is to identify what will be the main component that made up the system to be developed. In order for the visitors of this website to use it services, they have to register with their basic information such as name, email, etc. The registered users will be subscribed automatically to this website's newsletter. The users shall be able to choose product and related information will be displayed.. Therefore user will be able to get directly the information that they are interested in. The main components of this system are:

- Database to store products and users' information

- User able to search the product that they are interested in

4.2.5 Stage 5: Develop System Architecture

The next phase to develop the architecture on how the system will works. This will give the clear picture and understanding on how the system will operate and to avoid developing a system that does not solving the problem it intended to solve.

Registered Users

The website requires visitors to sign up before they can use the functions in the website. The registered users will have personal account to this website and they may login using their username and password. The actions that registered users can do with this websites are;

- Login to the website

- Able to edit profile

- Search products

- View promotion and sales

- Compare prices

<u>Admin</u>

The admin of this website will be responsible to maintain both website and the database. The roles of admin for this website are;

- Receive price lists from sellers

- Update database regularly

- Maintenance of website content and functionality

- Maintain web servers and website

- Manage advertisements regarding sales and promotion from the sellers

- Write report and analyze website user statistics

Sellers/Retailers

Any parties that would like to join and become partner with this project. Their roles are to
submit the price lists to the admin. They are also responsible to inform the admin if they
want to advertise the new products and also if there are sales and promotion being held at the
shop.

Figure 1: System Architecture Design

4.3 SOFTWARE REQUIREMENT SPECIFICATIONS

4.3.1 HARDWARE

Processor (CPU):

Minimum: Dual-core processor.

Recommended: Quad-core processor.

Memory (RAM):

Minimum: 4GB RAM.

Recommended: 8GB RAM.

Storage (HDD/SSD):

Minimum: 100GB of available storage space.

Recommended: 250GB SSD.

Network Interface:

Wired Ethernet or Wi-Fi connectivity.

Monitor:

Recommended: Full HD (1920x1080).

4.3.2 SOFTWARE

Programming languages:

- Python.

- Html.

- Js.

- Css.

Python Libraries:

- Requests.

- BeautifulSoup or Scrapy.

- Pandas.

- Selenium (optional).

Other relevant libraries based on project requirements.

Integrated Development Environment (IDE):

Examples: PyCharm, Visual Studio Code, Jupyter Notebook.

Web Browser:

Examples: Google Chrome, Mozilla Firefox, Safari.

Database Management System (DBMS):

Examples: SQLite, MySQL, PostgreSQL.

Dependency Management:

A package manager such as pip (Python) or npm (Node.js) is necessary for installing and managing project dependencies.

4.4 PROJECT DEVELOPMENT METHODOLOGY

Considering the time constrains to complete this project on time, the most suitable methodology is the Rapid Application Development (RAD) method. There is a set of management techniques that are optimized for speed in RAD which are;

- Prototyping - which is an approach based on creating a demonstrable result as early as possible and do refinements based on how the prototype works, whether it returns the desired result or not.

- Iteration – which is the incremental development based on refinement.

This project requires a rapid prototyping which will involves methods like iterative development and software prototyping. This methodology is also chosen due to the possibility if there might be functionality and performance compromising so the process of fixing the problems will be done promptly. The benefit of using this methodology is it allows any changes to be made during the development phase if there is needs to review and recheck at any other phase of project development. This is important as it provides flexibility throughout completing the project such as debugging process.

There are four main phases in RAD which are analysis and quick design, prototyping cycles, testing and implementation phase.

4.4.1 Design analysis

This phase has been covered up during the FYPI whereby the process of defining goals into defined functions and operation of the intended web system were being done. The

designs of the website including its layout, process diagrams and other documentations also included in this phase.

Basic layout for the website has Three main menus which are "About Us"- this page will basically have the information regarding the website and its functionality. Next is "price-check" page that will bring users to the website's main function which is users will be able to search the product that the users wanted to know. Then the website will check on the database and retrieve the information needed and it will give return results of the shops offering the products. However, this function requires member's login therefore users need to create an account first and login to this website before they can use the website's functions.The "Contact Us" page will provide contact information of the website's admin in order for users who wants to give further enquiries regarding the system as well as for the sellers that want to join in advertising with website.

4.4.2 Prototyping development

This phase consists of 3 main categories namely; build, demonstrate and refine. These 3 steps is performed in cycle order, where its starts with building the system. The development of the system began and is performed part by part. The website was developed starting from the HTML according to the designated layout for its main interface. Then, the HTML will be connected with Python language and the MySQL as the database to store all the information especially the products' information or we can store the data in json file than import it on the website. Each completed part is then demonstrated to check for its functionality. The results should be matched with the expected results whereby the information will be filtered according to the location chosen by users and it needs to be sorted according to its price with the cheapest will be at the

top. If there are any requirements or new functionality being identified, the system will be refined, where the building process will start again from where it has initially completed. These 3 steps cycles after one another until its functionality achieve satisfaction.

4.4.3   Testing Phase

This phase checks for errors, bugs and functionality of the website. There are five types of testing which are; regression test, internal testing, unit testing, application testing and stress testing. Regression test will focus on the following;

Focus	Purpose
Internal	Make sure all internal's components work well
Unit	Make sure all customers' components work well
Application	Make sure the system can complete all scenarios

Table 1: Regression Test

Internal testing get the every function or component being tested which also called as white-box testing because all details are visible to the test. Meanwhile unit testing tests the interaction of many functions but impound

the test within one unit which also known as black-box testing because it focused on the details of the interface that visible to the test. Next is the application testing that deals with the entire system. The project is ready for implementation after pass all of these tests.

Test No	Type of System Testing	Purpose	Status
1	Black Box testing	To test for system requirement and functionality without considering the internal architecture of the system.	Passed
2	White Box Testing	To test for internal functionality of the system. This testing included the coding of the system.	Passed
3	Unit Testing	To test for the functionality of each separated system page.	Passed
4	Acceptance Testing	To ensure the system is completed and performed as requested by user.	Passed
5	Functional Testing	To ensure the system able to performed all its intended functionality.	Passed
6	Usability Testing	To ensure the system is understandable and can be easily used by the user	Passed

7	Integration Testing	To verify that all different part of the system pages can integrate with each other and function with no error	Passed

Table 2: Type of System Testing

4.4.4   Implementation Phase

After all the testing performed is completed and passed, the system is ready for the implementation. The website has received a positive response and will be considered to be used soon. This is the final phase of the system development and hence, the system is expected to be fully functioning as it intended for.
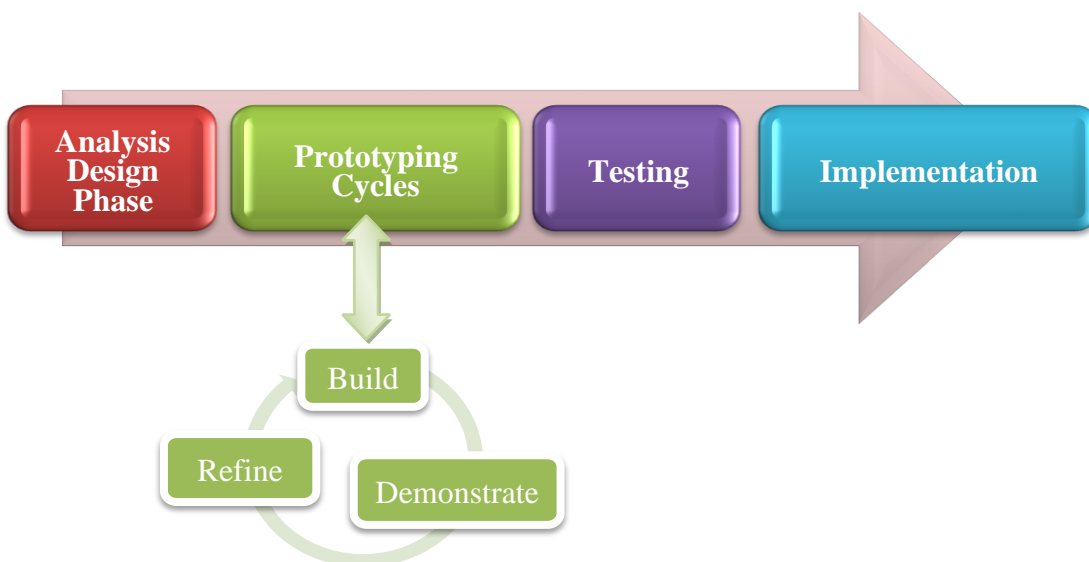
Figure 2: Rapid Application Development Cycle

4.5 SYSTEM DEFINITION/ANALYSIS (HIGH LEVEL DESIGN)

SYSTEM ARCHITECTURE
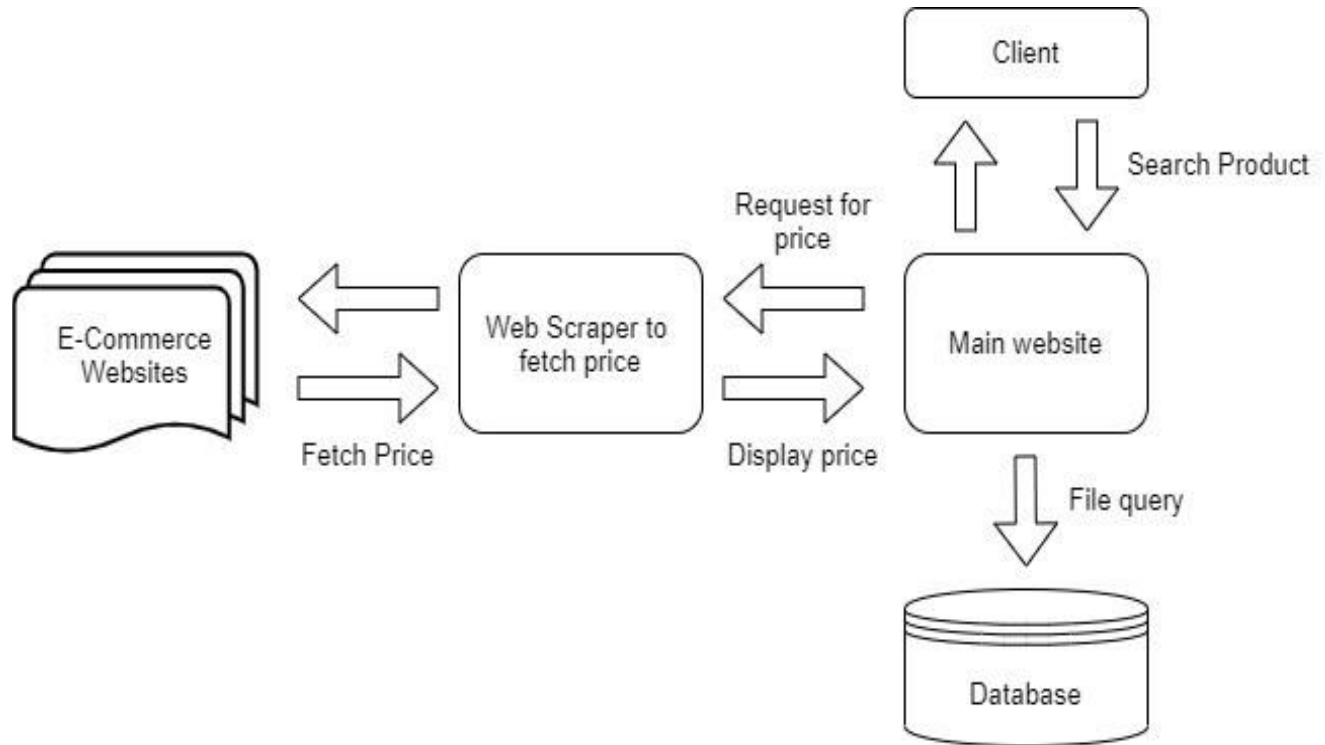
Fig-3 System architecture

Figure 1 describes 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 e-commerce websites. The extracted information of ecommerce products is then displayed on website. 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 of required product at one place present on different E-commerce firms. Another feature is provided on the website is price alert, which user can set, to get notified by the website whenever the suitable price comes up.

UML Diagrams

Use-Case diagram.

A use case diagram illustrates the interactions between users (actors) and a system to achieve specific goals or functionalities. Here's a simplified use case diagram for an e-commerce scraping system:

Actors:

User: Represents any individual or entity interacting with the e-commerce scraping system.

Use Cases:

Configure Scraping Parameters:

Description: Users can configure parameters such as e-commerce websites, product categories, filters, and scheduling options for scraping tasks.

Actors: User

Preconditions: User is logged in to the system.

Trigger: User initiates the configuration process.

Basic Flow:

User selects e-commerce websites to scrape.

User specifies product categories, brands, and other filters.

User sets scheduling options for automated scraping tasks.

Postconditions: Scraping parameters are saved for future use.

Run Manual Scraping Task:

Description: Users can manually initiate a scraping task to retrieve data from e-commerce websites based on configured parameters.

Actors: User

Preconditions: User has configured scraping parameters.

Trigger: User initiates a manual scraping task.

Basic Flow:

User selects the option to run a manual scraping task.

System retrieves data from e-commerce websites based on configured parameters.

System stores scraped data for further processing.

Postconditions: Scraped data is available for analysis.

View Scraped Data:

Description: Users can view scraped data, including product information, prices, availability, and other relevant details.

Actors: User

Preconditions: Scraped data is available in the system.

Trigger: User requests to view scraped data.

Basic Flow:

User accesses the system's interface for viewing scraped data.

System presents scraped data in a structured format.

User can browse, search, and filter scraped data as needed.

Postconditions: Users have access to scraped data for analysis and decision-making.

Manage Scraping Tasks:

Description: Users can manage existing scraping tasks, including editing, pausing, or deleting scheduled tasks.

Actors: User

Preconditions: User has configured scraping parameters and scheduled scraping tasks.

Trigger: User accesses the management interface for scraping tasks.

Basic Flow:

User views a list of existing scraping tasks.

User can edit, pause, resume, or delete scheduled tasks as needed.

Postconditions: Scraping tasks are updated or removed based on user actions.

This use case diagram provides a high-level overview of the interactions between users and the e-commerce scraping system, outlining the key functionalities and user roles involved in the process.

Use case diagram
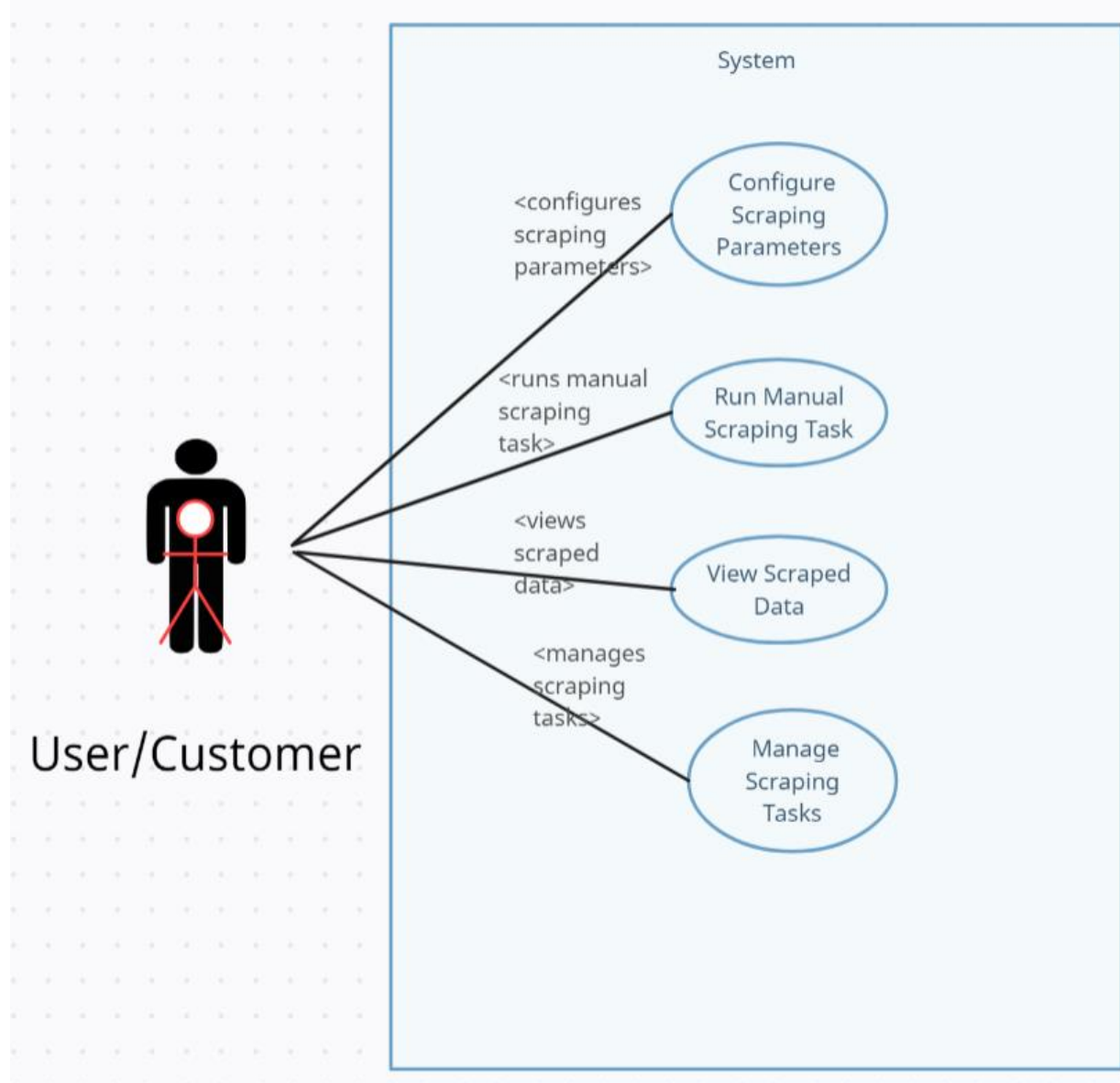
Fig-4 UML-use case diagram

Activity Diagram
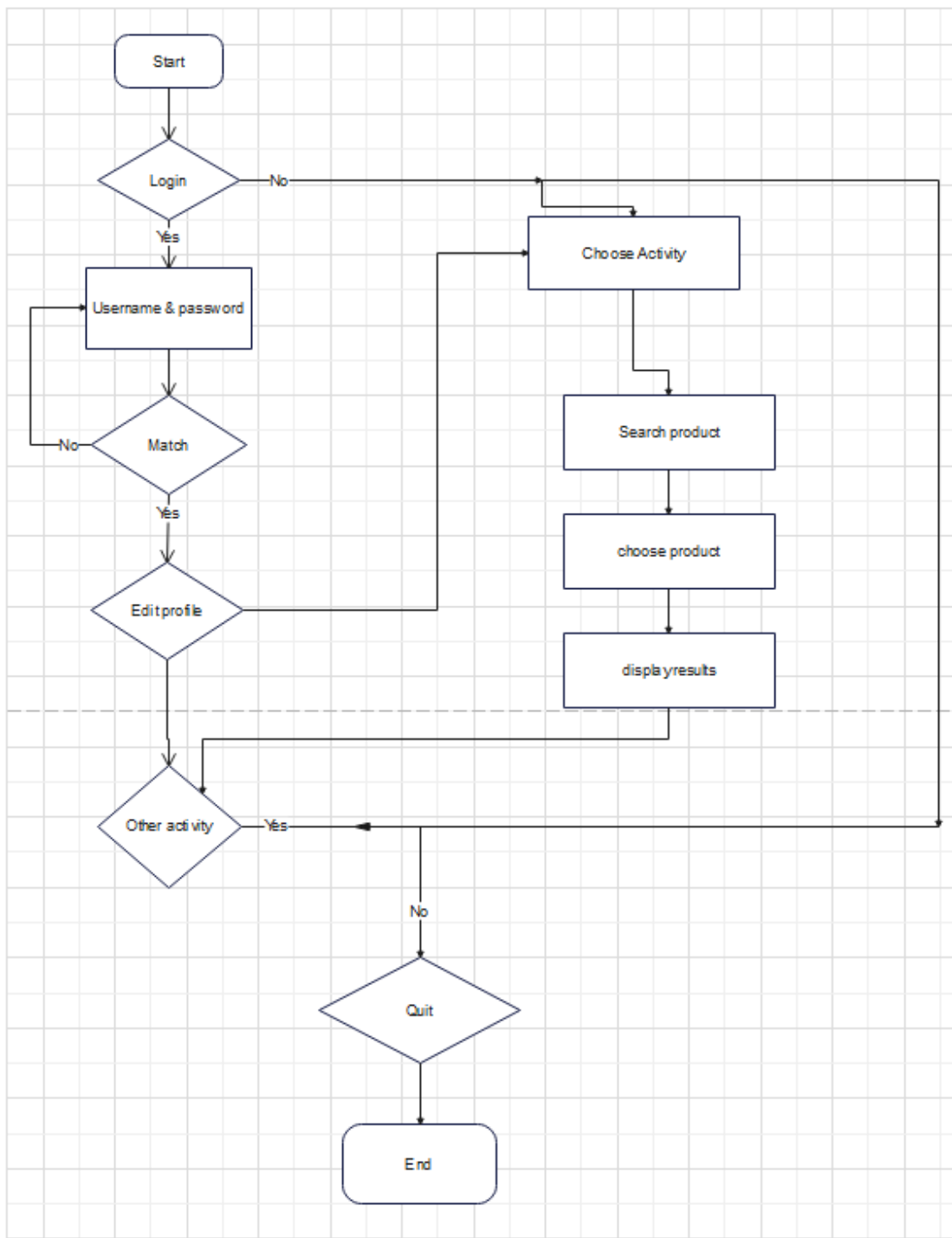
Fig-5 process flow diagram

4.6 FUTURE ENHANCEMENTS

It is recommended to have more research on this project and determine what can be improved from the existing comparison site. One of the functions that could be included into this project is by allowing users to set price trigger alert where by customer will have the ability to set a certain price for a particular item, and once there is seller offers the price match with the parameter set by the user, E-Commerce scraper will send a notification alert through email saying that the product's current price has triggered the alert. This type of alert is widely being used by the sites that provide tools and guidance for stock exchange business. When the price of a certain stock has reached a certain value set by the user, it will trigger the alert and notify the user. Therefore user does not need to stay and watch the price movement all the time.

Other recommendation for the price comparison website is to be developed as mobile application for android or IOS phone. Extended function to be added is for the users can make order and purchase through this application by using Google wallet – a new product by Google that allows users to shop online, and there will be delivery services to the customer's house like what have been done by Tesco Homeplus in South Korea where they are using the concept of 'Let the Store Come to People'. Customers only have to scan the QR code of the products with their smart phone then the products will be automatically lands in the customer's online cart. When the purchase is done, the item will be delivered to the customer's house. This system helps the customers to shop easier just by using their smart phone.

RESULTS AND DISCUSSION

5.1 OVERVIEW

The results and discussion regarding this project will be explained and discussed in details in this section. This section will cover the IPO (Input, Process, Output) Diagram, the Process Flow Diagram, the Interface Design of PriceWar.com, the Sample of Search's Results, and Survey's Results analysis.

5.2 IPO (Input, Process, Output) Diagram

IPO diagram represent a process by demonstrating the relationships between input and output elements which also known as 'factors' and 'responses' respectively. It is very useful to define a process as an activity that transforms inputs to generate corresponding outputs. Refer the figure below for Projects IPO diagram.
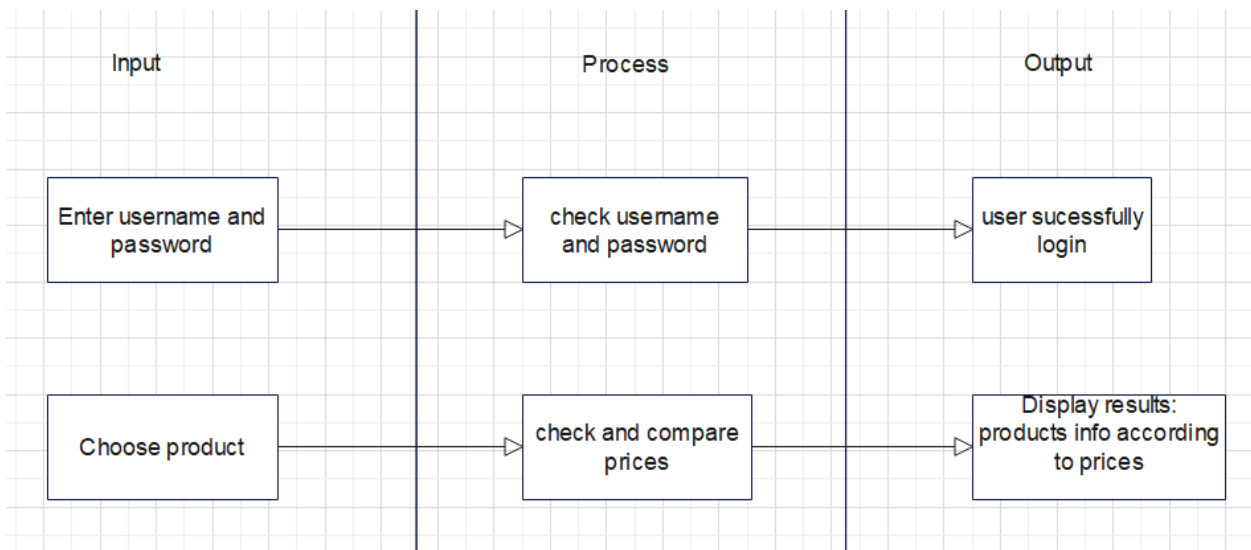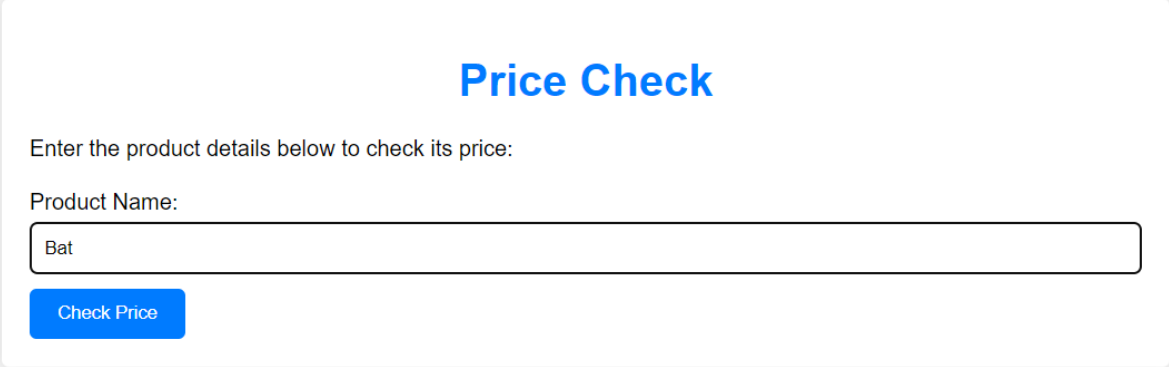
Figure 6: IPO Diagram

5.3 SAMPLE OF SEARCH'S RESULTS

Comparison of product prices from different ecommerce websites and result is displayed on single web interface. This website aims at providing the best possible deal to the users for the required product by comparing the product price and displaying the minimum price from various E-commerce websites such as Amazon, Flipkart and Croma, which are leading and some of the best websites to shop. To achieve this result web mining is done to fetch the required product details and concept of web crawler and web scraper is used to extract information of these products available on different ecommerce websites. System will allow users to redirect to original website of that specific product selected by the user as a best deal. Thus, website serves as a time - saving tool for frequent online buyers as they can compare the prices at one - stop instead of searching for the same product on various websites. Following images show how product analysis and comparison of e-commerce sites is done.

Object/Product input:

Figure 7: sample price check

31

Background process:

```
obj = ps.ECommerceScrapper()
        obj.flipkart_scrapper('Cricket Bat', 2)

print(ECommerceScrapper().flipkart_scrapper('Cricket Bat', 2))

obj = ps.ECommerceScrapper()
        obj.snapdeal_scrapper('Cricket Bat', 2)

print(ECommerceScrapper().snapdeal_scrapper('Cricket Bat', 2))
```
Figure 8: sample background

Results:

Flipkart

	Name	Price	Original Price	Number of Ratings
0	EmmEmm Size 3 Plastic Cricket Bat for Kids upt...	Rs. 299	Rs. 650	(34)
1	Prospo Tennis Ball Cricket Bat And Tennis Ball	Rs. 999	Rs. 1,299	(6)
2	Rmax Light Green Tennis Ball Kashmir Willow Cr...	Rs. 1,399	Rs. 2,399	(2)
3	LYGER Plastic Cricket kit for All Age Groups a...	Rs. 849	Rs. 899	None
4	Trex Bolt Designer Scoop Poplar Willow Free Te...	Rs. 938	Rs. 1,599	(12)
5	NAS Nude Popular Willow Cricket Bat (For Hard...	Rs. 1,025	Rs. 1,399	(319)
6	EmmEmm Premium Quality Black Heavy Cricket Bat...	Rs. 439	Rs. 899	(21)
7	Prospo Cricket Heavy Duty Bat	Rs. 999	Rs. 1,199	(5)
8	Rmax Gully Plastic Bat, Plastic Cricket Bat Ha...	Rs. 419	Rs. 999	(4)
9	LYGER Kashmir Willow Leather Cricket Bat, with...	Rs. 2,160	Rs. 2,799	None
10	TREX BOLT 1000 DESIGNER CRICKET BAT (POPLAR WI...	Rs. 649	Rs. 1,499	(1)

Figure 9: sample flipkart

Snapdeal

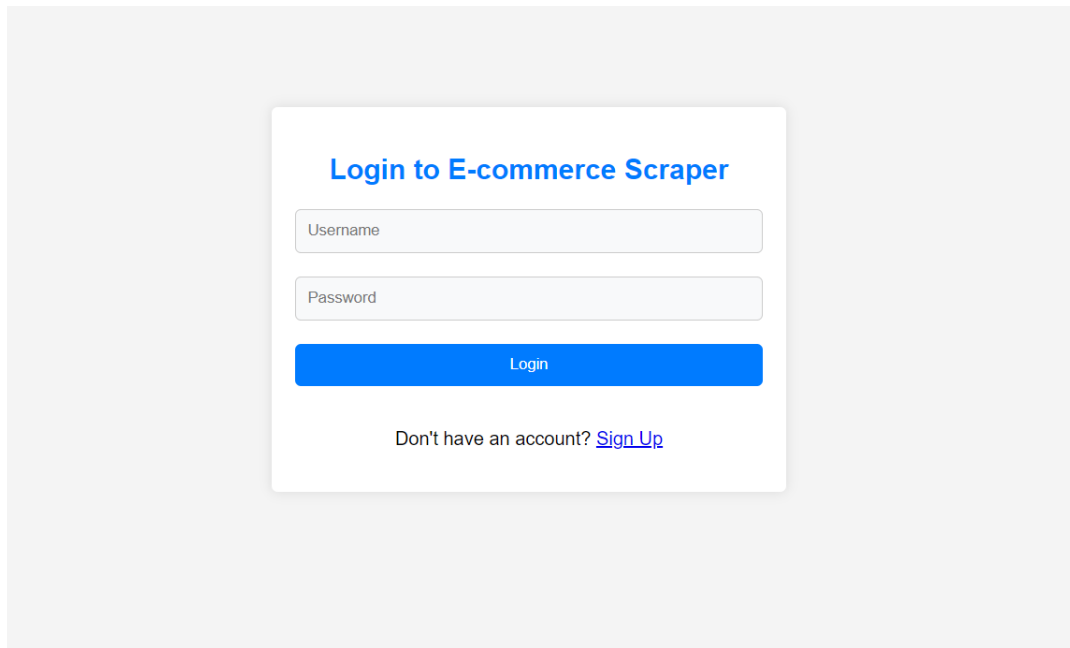|	Name	Price	Original Price	Description	Rating
0	LYCAN Junior Size 3 For Age Group 8 Years PVC/...	₹249	₹269	400 g	3.7
1	HF THE ORIGINAL Slog Full Size (Ideal for 15+)...	₹399	₹2,149	850 g	4
2	MRF Genius Virat Kohli bat|Size-4| For 8-11 Ye...	₹322	₹407	950-1100 g	3.6
3	KNK Plastic bat , Hard Plastic Bat , Cricket b...	₹338	₹930	800-900 g	3.6
4	Khilaunewala Plastic Bat, Plastic Cricket bat,...	₹379	₹750	0.49 kg	None
5	Strauss Rebel Cricket Bat, Size SH (34 X 4.5 i...	₹399	₹1,199	850-900 g	4.1
6	Tucker Hard full size PVC/Plastic Cricket Bat	₹320	₹1,799	750-850 g	3.9
7	BOOM BOOM master for age 6 to 8 year kids PVC/...	₹229	₹999	400 g	3.6
8	MRF Genius Virat Kohli |Size-6| For 10-13 Year...	₹339	₹495	950-1100 g	3.6
9	Lyzoo Plastic, Hard PVC Heavy duty 15+ Age Pla...	₹399	₹1,350	850-900 g	4
10	Strauss Rebel Cricket Bat, Size SH (34 X 4.5 i...	₹379	₹1,199	850-900 g	4.1

Figure 10: Sample snapdeal

CONCLUSION

Users can access helpful information on the website, which will assist them in making decisions that are in their best interests. It is now possible for working people to check on the price of things before making purchases, as a result of the existence of a website that compares prices. Users of this website will be able to compare costs on a variety of e-commerce shopping websites in order to choose which website offers the best combination of low cost and a good deal on the product they are interested in purchasing. The purchasers are going to unquestionably appreciate the time and effort that this saves them. In the end, this will help buyers shop online by bringing together tactics, the greatest offers and deals from all of the biggest online retailers, and by providing customers with an easier way to shop online.

Users will be able to acquire valuable information from the website, which will assist them in arriving at the best choice. The need for working people to check on the price of things before purchasing them is alleviated by the existence of this website that compares prices. It offers a platform for vendors to promote new products, announce ongoing promotions or deals, and enable customers to purchase products at prices that are more competitive with the market.

APPENDIX (SCREENSHOTS)

Login page

Figure 11: login Page
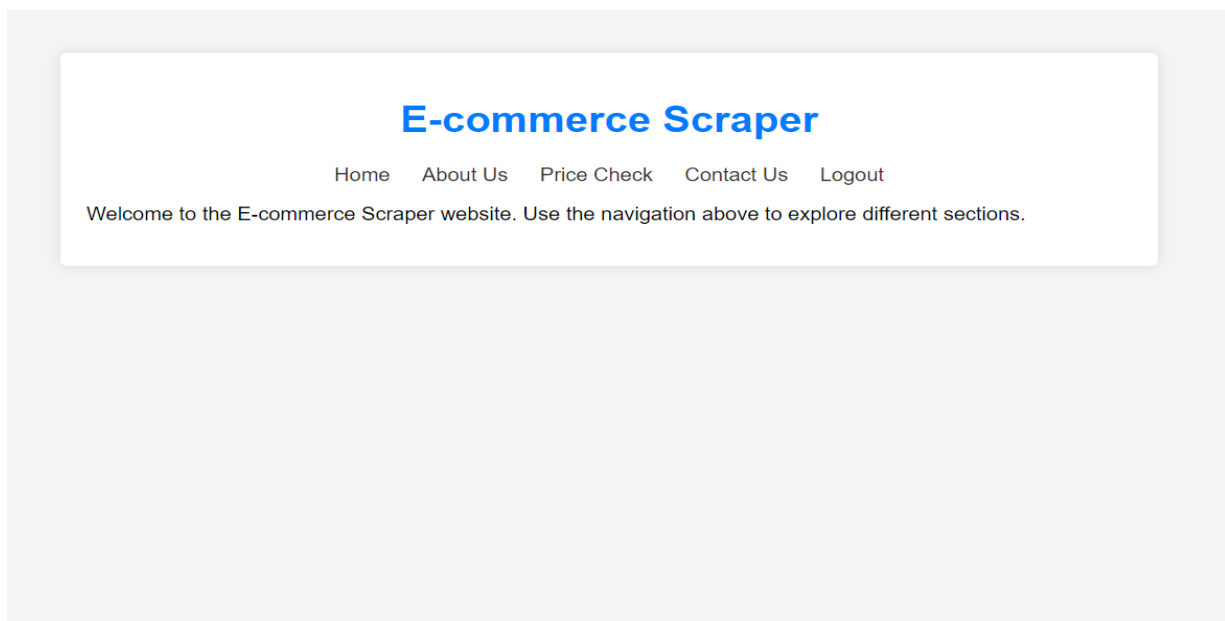
Onload Page

Figure 12: Onload Page

Home Page
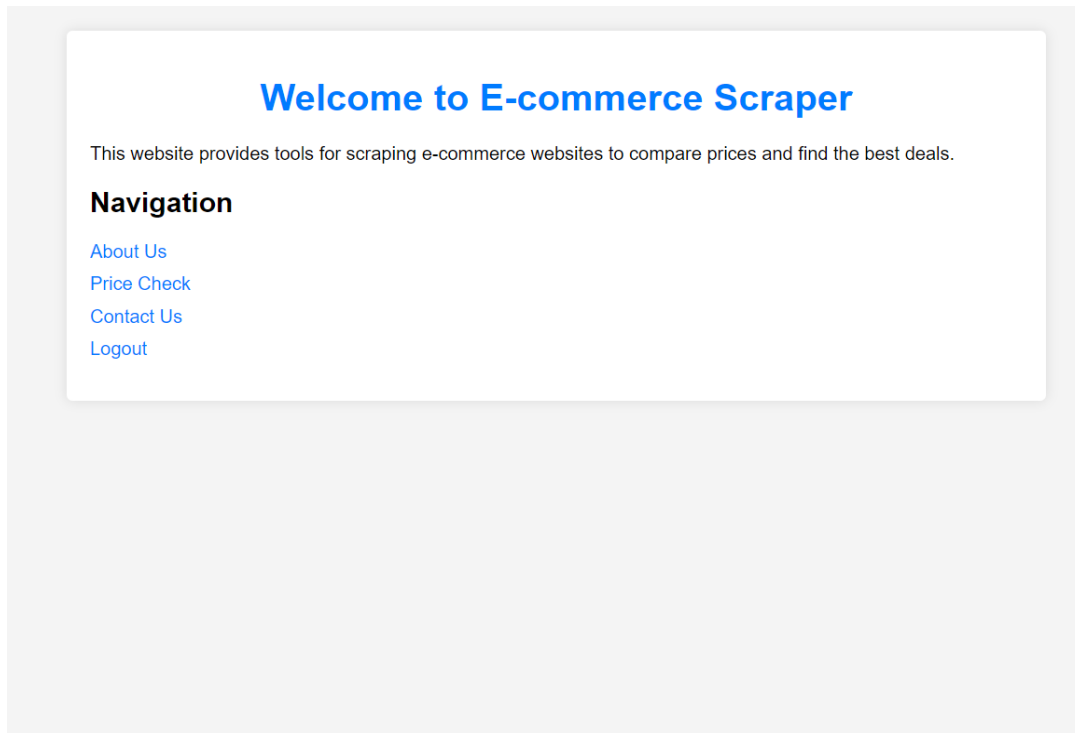
Figure 13: Home Page

About us
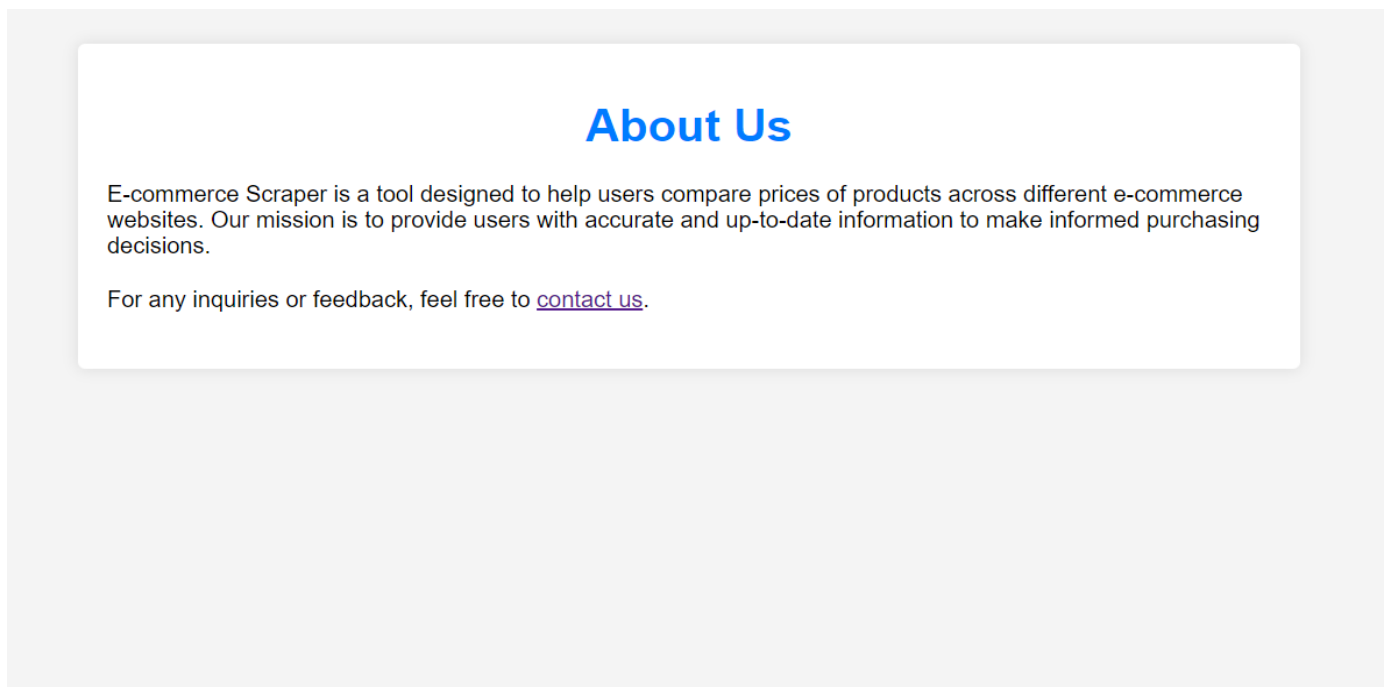
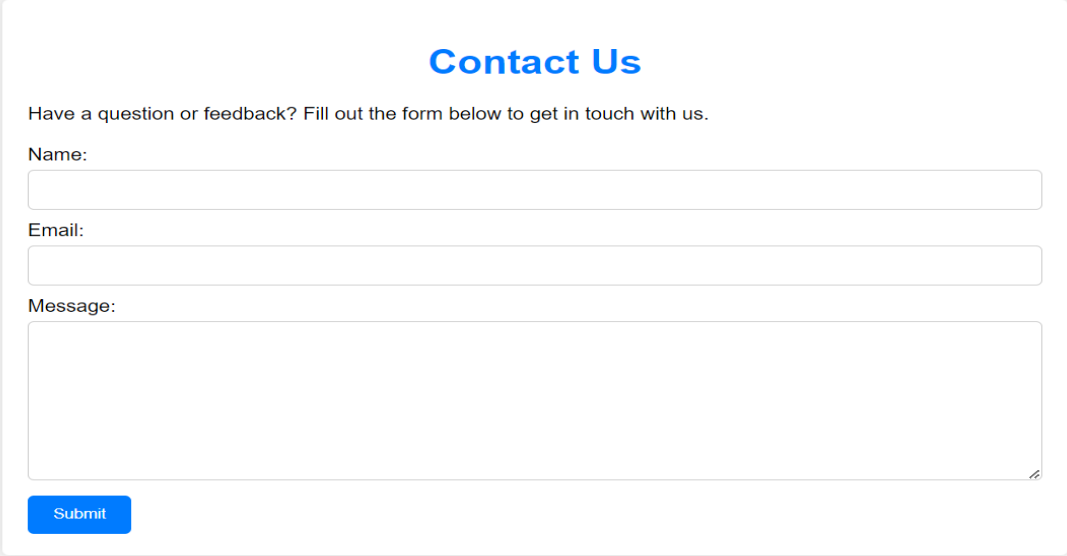*Figure 14*: About us

Contact Us
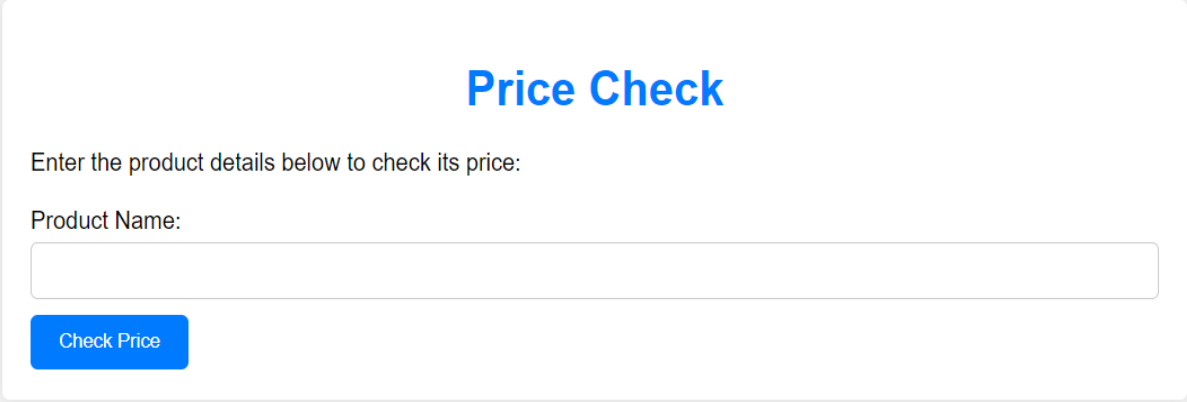
Figure 15: Contact Us

Price-check
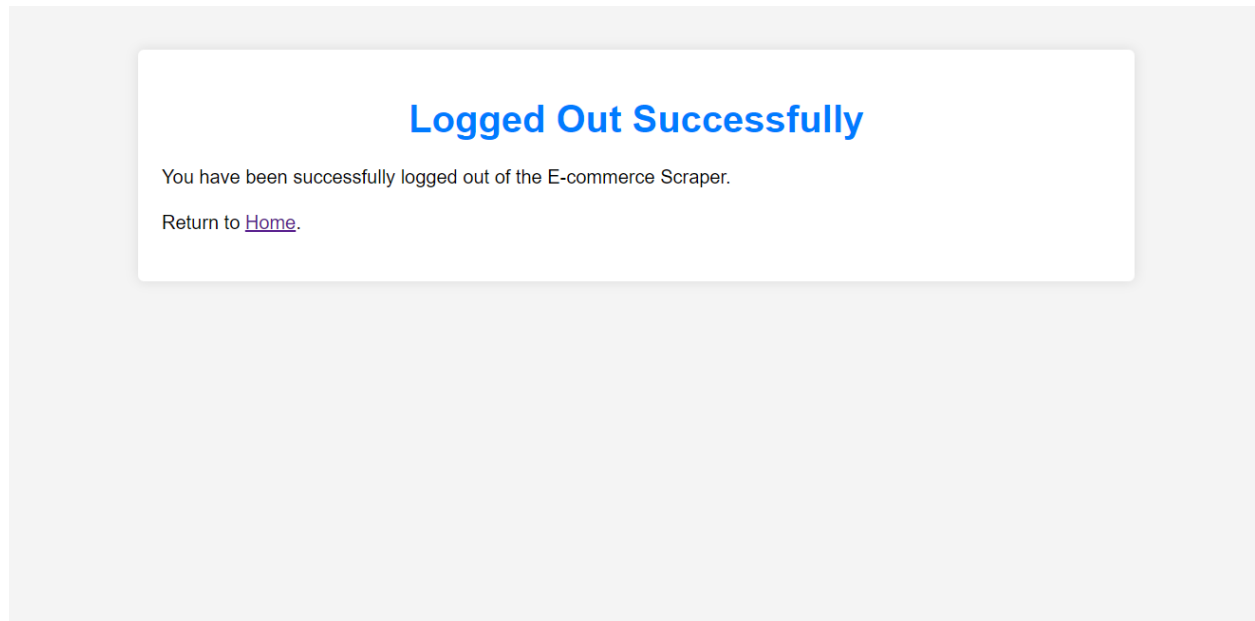

Figure 16: Price-check

Logout

Figure 17: Logout

REFERENCES

1)The use of web scraping in computer parts and assembly price comparison LR Julian, F Natalia - 2015 3rd International Conference on …, 2015 - ieeexplore.ieee.org

2) An overview on web scraping techniques and tools AV Saurkar, KG Pathare, SA Gode - International Journal on Future …, 2018 - ijfrcsce.org

3) Web scraping for unstructured data over web GN Chandrika, S Ramasubbareddy, K Govinda… - Embedded Systems and …, 2020 - Springer

4) Shridevi Swami , Pujashree Vidap *,” Web Scraping Framework based on Combining Tag and Value Similarity*” Proceedings of the IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 6, No 2, November 2013.

 5) Dr. Rajendra Nath ,Khyati Chopra,” *Web Crawlers: Taxonomy, Issues & Challenges*” Proceedings of the International Journal of Advanced Research in Computer Science and Software Engineering , Volume 3, Issue 4, April 2013, pp. 944-948.

6) Jos´e Ignacio Fern´andez-Villamor, Jacobo Blasco-Garc´ıa, Carlos ´A. Iglesias, Mercedes Garijo "A Semantic Scrapping Model for Web Resources" Spain. .

7)Richard K. Lomotey, Ralph Deters*,” RSenter: Tool for Topics and Terms Extraction from Unstructured Data Debris*”, Proceeding of the IEEE International Congress on Big Data, 2013.

8) Web and android application for comparison of e-commerce products A Ambre, P Gaikwad, K Pawar, V Patil - no, 2019 - academia.edu

9) Rahul Dhawani, Mrudav Shukla, Priyanka Puvar, Bhagirath Prajapati,” *A Novel Approach to Web Scraping Technology*” Proceeding of the International Journal of Advanced Research in Computer Science and Software Engineering ,Volume 5, Issue 5, MAY 2015.

10) E-Commerce Web-Crawling to Facilitate Consumers for Economical Choices S Saeed, M Naqvi, M Memon - International Journal of …, 2020 - journal.scientiaca.org

110Future Comparisons: What's next for Price Comparison Websites? – *Webcredible* (2009)

12)Knight, E. (2009) : "*The Use of Price Comparison Sites in the UK General Insurance Market.*

13)Moraga-Gonzalez, J.L. , Wildenbeest, M.R. (2011) :
Comparison Sites.
        On the Horizon (2009), May.

Internetworldstat.com - "*World's Internet Usage Statistics*"
[retrieved on June 20, 2012] at http://www.internetworldstats.com/stats.htm

14)Grant, K.B (2008) [retrieved on June 20, 2012] – *"Which Price Comparison Shopping is The Best?"* at http://www.smartmoney.com/spend/family-money/whichcomparison-shopping-site-is-best/

15)Ong K. (2011) : *"Online Shopping Statistics in Malaysia"*
[retrieved on June 21, 2012] at
http://humanwebsite.com.my/blog/online-shopping-statisticsin-malaysia.html

16)Malaysia Crunch website : *"Malaysia's E-Commerce Statistic"* [retrieved on June 25, 2012] at http://www.malaysiacrunch.com/2009/09/malaysias-ecommerce-statistics.html

17)Donna L. : *"New Website Allow U.S Consumers To Find Local Grocery Deals"*[retrieved on July 5, 2012] at http://couponing.about.com/od/groceryzone/qt/mgdpr.htm

18)PaySult.com – *"Malaysian Internet Users spend 20 hours a Week Online, 54% Purchase Online"* [retrieved on July 10, 2012] at http://www.paysult.com.my/malaysian-internetusers-spend-20-hours-a-week-online-54-purchase-online/

19)One-minute Marketing - *"Consumer Buying Behavior: Online vs Offline Shopping"* [retrieved on Oct 3, 2012] at http://blog.jbamedia.com/2010/09/13/consumer-buyingbehavior-online-vs-offline/

20)Rueter, T , July 31, 2012: *"Consumers Plan More Web Research before Buying School Supplies"* [retrieved on Oct 4, 2012] at http://www.internetretailer.com/2012/07/31/consumers-planmore-web-research-buying-school-supplies

21)Olson, M , October 11,2010 : *"Majority of Consumers Research Products Online before Purchase"*[retrieved on Oct 2, 2012] at http://janrain.com/blog/majority-consumersresearch-products-online-purchase/

22*)*Constantinides, E (2004) : *"Influencing the Online Consumer's Behavior : the Web Experience"*

23)Basu, R. : *"Implementing Six Sigma and Lean – a Practical Guide to Tools and Techniques"* [retrieved on Oct 2, 2012] at http://www.scribd.com/doc/67871803/37/D1-IPODiagram

24) Java Stack Overflow at https://stackoverflow.com