

A
Mini-Project Report on

E-Commerce Data Analyzer

Submitted in partial fulfillment of the requirements
for the degree of

BACHELOR OF ENGINEERING

IN

Computer Science & Engineering

Artificial Intelligence & Machine Learning

by

Yash Penkar (22106118)

Prarthana Patil (22106035)

Rutuja Pawar (22106043)

Gauri Ramekar (22106068)

Under the guidance of

Miss. Taruna Sharma



**Department of Computer Science & Engineering
(Artificial Intelligence & Machine Learning)**

A. P. Shah Institute of Technology

G. B. Road, Kasarvadavali, Thane (W)-400615

University Of Mumbai

2023-2024



A. P. SHAH INSTITUTE OF TECHNOLOGY

CERTIFICATE

This is to certify that the project entitled “**E-Commerce Data Analyzer**” is a bonafide work of Yash Penkar (22106118), Prarthana Patil (22106035), Rutuja Pawar (22106043), Gauri Ramekar (22106068) submitted to the University of Mumbai in partial fulfillment of the requirement for the award of **Bachelor of Engineering in Computer Science & Engineering (Artificial Intelligence & Machine Learning)**.

Miss. Taruna Sharma
Mini Project Guide

Dr. Jaya Gupta
Head of Department



A. P. SHAH INSTITUTE OF TECHNOLOGY

Project Report Approval

This Mini project report entitled “**E-Commerce Data Analyzer**” by **Yash Penkar, Prarthana Patil, Rutuja Pawar and Gauri Ramekar** is approved for the degree of *Bachelor of Engineering in Computer Science & Engineering*, (AIML) 2023-24.

External Examiner: _____

Internal Examiner: _____

Place: APSIT, Thane

Date:

Declaration

We declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission hasnot been taken when needed.

Yash Penkar
(22106118)

Prarthana Patil
(22106035)

Rutuja Pawar
(22106043)

Gauri Ramekar
(22106068)

ABSTRACT

E-commerce is the buying and selling of goods and services over the internet. It is a multi-billion dollar industry because it has many advantages for both customers and businesses, such as convenience, variety, lower costs, and global reach. Price comparison website have become today one of the most important sources for the purchase of all kinds of products. Many strategies were developed by analyzing the user's behavior in order to attract more business and engage people. Since there are a lot of e-commerce sites available, it becomes difficult for users to choose the best price for the desired product among those sites. Comparing e-commerce products using web explorer allows users to analyze prices and obtain the desired product at the minimum price. Users can also select several products within the same category to compare their characteristics. To get information about products from e-commerce web crawlers and web scraping techniques are used to get detailed information. The proposed system is based on users' expectations and security. The user can clearly know the exact price of the product and is looking for discount offers in different e commerce sites. The proposed system can be a simple and effective way to know the exact price and the offers provided by the different e-commerce websites. This way aims to provide a solution for online customers to purchase products at cheap prices and save their valuable time, effort and money.

INDEX

Index	Page no.
Chapter-1	
Introduction	1-3
Chapter-2	
Literature Survey	4-8
2.1 History	
2.1 Review	
Chapter-3	
Problem Statement	9-10
Chapter-4	
Experimental Setup	11-15
4.1 Hardware setup	
4.2 Software Setup	
Chapter-5	
Proposed system and Implementation	16-23
5.1 Block Diagram of proposed system	
5.2 Description of Block diagram	
5.3 Implementation	
Chapter-6	
Conclusion	24-26
References	27

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

According to the rapid increase of using internet-based technologies, e.g., digital sensors, cloud computing, etc., a massive amount of data is being generated and stored. Based on the data variety, velocity, and volume, they are called Big Data. Many companies take advantage of analysing these big data to enhance their business strategy and gain benefits. Moreover, the accelerating use of the internet resulted in activating E-commerce activities. E-commerce is the online process of purchasing products, services, and information or selling and exchanging them. E-commerce companies have taken the benefits of analysing big data to improve their processes, as well as maintaining and increasing their revenues. In recent years, e-commerce has become an integral part of people's lives, providing easy and convenient access to products and services. With the growing trend of online shopping, two of the leading e-commerce giants in India, Snapdeal and Amazon, have emerged as strong competitors. Both platforms offer a wide range of products and services to their customers, including electronics, fashion, grocery, and much more. This comparative analytical study aims to explore the various aspects of Snapdeal and Amazon, including their business model, target audience, marketing strategies, product offerings, customer service, and pricing policies. The study aims to provide insights into the strengths and weaknesses of both platforms, enabling consumers to make informed decisions while choosing an e-commerce platform for their needs. The increasing use of intelligent devices and other mediums has paved the way for users to purchase products virtually anywhere. This has increased the participation of online shoppers in the development of e-commerce businesses. This large number of e-commerce websites places users in a difficult position to search and choose to purchase a single product from several e-commerce websites. The proposed solution helps online users get prices for their products from multiple e-commerce sites across a single web interface. This will save users time, money, and effort in finding the

same product prices on different e-commerce sites. The proposed system utilizes a process of web mining as the part of web mining they are two types of techniques one is web scraper. Web scraper technique to extract data from e-commerce web pages and a web crawler to link products. As the digital economy expands, the role of web scraping becomes ever more important. Web scraping is done for a number of reasons, including: Major search engines like Google will scrape websites to determine relevant search results when users type in keywords. Scraping is an excellent way for retailers to perform market research and check out the competition. This is a popular use of scraping for news websites, as information can be gathered from multiple sources to populate a live feed. Scraping can be used for build data sets for machine learning software, though this has proved controversial as original creators are sometimes not credited for derivative content the programs create. This is another very popular use for scraping (usually with automated tools). It's mostly used to check the prices on products for e-commerce websites.

CHAPTER 2

LITERATURE SURVEY

2. LITERATURE SURVEY

2.1 HISTORY

Most of us have shopped online for something at some point, which means we've taken part in e-commerce. So, it goes without saying that e-commerce is everywhere. But very few people may know that e-commerce has a history that goes back to before the internet began. The history of E-Commerce is quite fascinating. It began in the 1970s when companies used an electronic system called the Electronic Data Interchange (EDI) to transfer documents. However, it wasn't until 1994 that the first online retail transaction took place, involving the sale of a CD through Net Market. E-Commerce, as we know it today, involves conducting business transactions over the internet, including selling information, services, and goods through computer telecommunications networks. It has evolved from simple transactions to complex exchanges involving business-to-consumer, business-to-business, and even consumer-to-consumer interactions. The introduction of the World Wide Web in 1991 and the first browser in 1993 significantly shifted e-commerce to the internet. With the rise of smartphones and broadband connections, e-commerce has further transitioned to mobile devices, expanding its reach and convenience. Among innovations that have contributed to the growth of e-commerce are electronic directories and search engines for finding information on the Web; software agents, or bots, that act autonomously to locate goods and services; systems that recommend products to users based on their profile; and digital authentication services that vouch for identities over the Internet. Those intermediary services facilitate the sale of goods (actually delivering the goods in the case of information), the provision of services such as banking, ticket reservations, and stock market transactions, and the delivery of remote education and entertainment. Today, e-commerce encompasses a wide range of online marketplaces, auction platforms, and

millions of e-retailers, profoundly affecting how we shop and conduct business globally. The industry has gone through so many changes since then, resulting in a great deal of evolution.

2.2 LITERATURE REVIEW

[1] E-commerce Price Comparison Website Using Web Scraping:

International Journal Innovative Research in Engineering & Multidisciplinary Physical Sciences Volume 11, June 2023

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.

[2] E-Commerce Price Comparison With Review Sentimental Analysis:

Sanket Bezalwar et al, International Journal of Computer Science and Mobile Computing, Vol.11 Issue.3, March- 2022

One of the features that would be covered into PriceWar.com is via way of means of permitting customers to set rate cause alert wherein via way of means of purchaser can have the capacity to set a sure rate for a specific object, and as soon as there's vendor gives the rate in shape with the parameter set via way of means of the consumer, PriceWar.com will ship a notification alert thru electronic mail announcing that the products present-day rate has precipitated the alert. This sort of alert is broadly being utilized by the web sites that offer gear and steering for inventory change business. When the rate of a sure inventory has reached a sure cost set via way of means of the consumer, it's going to cause the alert and notify the consumer

[3] Price Comparison Website for Online Shopping: *Department of Computer Science and Engineering, St. Peter's Engineering College, Hyderabad , IJRTI Volume 9, Issue 6 June 2021*

The website provides users with useful information that will help them making informed decision. With this price comparison website, it solves the problems of the working people to check on the price before buying products. 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. 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] Price Comparison for Products in Various E-Commerce Website: *Dept. of Computer Science and Engineering Meenakshi Sundararajan Engineering College, Chennai, India, IJRTI / Volume 8, Issue 5,2023*

The comparison of e-commerce products utilizing web mining is web-based. The based system will assist users in making decisions while. Shopping for products online. This website will support users Analyze the prices that are present in the various e-commerce shopping. Web sites so that they learn to know the lowest price for the product. With a better understanding. The website will also feature the installation of Compare products with any of their specifications that belong to the same Category. This will save shoppers efforts and valuable time.

[5] E-Commerce Product Price Tracker: *Computer Science and Engineering 1,2JSS Science and Technology University, Mysuru, India, JETIR Volume 8, Issue 6, June 2021*

The idea in this article is a real-time program that tracks the website's dynamic price behavior. The customer's pricing will be compared to the current price by crawling the real time price from the website, and if the condition is met, an alert

will be issued to the user telling them of the new price and a link to the product on the website. By hosting the crawler application on the cloud and making it available to all users, the job may be expanded to a larger area. This solution eliminates the need for the crawler to be installed on the local machine. If we use it to track the prices of other websites, it can also be utilized as third-party software.

[6] Comparative Analytical Study Of Snapdeal And Amazon: *Parul University ,TIJSER ,Volume 10,Issue 3,March 2023*

Snapdeal and Amazon are two of the largest e-commerce platforms in India, and both offer a wide range of products, customer care services, and delivery options. While there are some differences between the two platforms in terms of product assortment, pricing, and customer care, they are both highly regarded by Indian consumers. In terms of product availability, both platforms offer a similar range of products, with some differences in specific categories. Amazon has a slightly more extensive assortment of branded products, while Snapdeal has a larger variety of private label products. Both platforms have a similar pricing strategy, with occasional discounts and promotions to attract consumers.

CHAPTER 3

PROBLEM STATEMENT

3. PROBLEM STATEMENT

The rapid growth of e-commerce has led to a surge in the number of online platforms, making it increasingly challenging for users to efficiently compare prices and make informed purchasing decisions. The abundance of e-commerce websites and the diverse range of products they offer create a need for a streamlined solution that empowers users to easily identify the best prices and discount offers across multiple platforms. Existing methods for price comparison often lack user-friendly interfaces and comprehensive data, leaving consumers with a time-consuming and cumbersome process. This project aims to address these challenges by developing a user-centric system that utilizes web mining techniques to gather detailed product information, allowing users to efficiently compare prices and make informed choices, ultimately enhancing their online shopping experience.

CHAPTER 4

Experimental Setup

4. EXPERIMENTAL SETUP

4.1 Hardware Setup

To run the software efficiently and effectively, it's important to have a computer system with the appropriate hardware configuration. Below is the recommended hardware setup to ensure optimal performance of the software:

1. Processor (CPU):

- The software can be run on a wide range of processors, but a modern multi-core processor is recommended for optimal performance.
- Minimum: Dual-core processor
- Recommended: Quad-core processor or higher

2. Memory (RAM):

- Sufficient RAM is essential for smooth operation, especially when dealing with large datasets or performing resource-intensive tasks.
- Minimum: 4 GB RAM
- Recommended: 8 GB RAM or higher

3. Storage (HDD/SSD):

- Adequate storage space is necessary to store the software application, operating system, and any additional data files.
- Minimum: 128 GB solid-state drive (SSD) or hard disk drive (HDD)
- Recommended: 256 GB SSD or higher for faster read/write speeds

4. Display:

- A high-resolution display with adequate size ensures clear visualization of the software interface and any graphical outputs.
- Recommended: Full HD (1920 x 1080) or higher resolution monitor

5. Peripherals:

- Standard peripherals such as a keyboard and mouse are required for input control.
- Additional peripherals such as headphones or speakers may be necessary for audio output, depending on the software's features.
- Optional peripherals such as a printer or scanner may be needed for tasks involving document processing or printing functionality.

6. Internet Connectivity:

- A reliable internet connection is recommended for accessing online resources, updates, and any cloud-based services that the software may utilize.

7. Operating System:

- The software is compatible with various operating systems, including Windows, macOS, and Linux distributions.

4.2 Software Setup

In this section, we will outline all the necessary software tools and packages required for setting up and running the software smoothly. Below is a comprehensive list of software components:

1. Python:

- Python is the primary programming language used for developing the software.
- Version: Python 3.x (e.g., Python 3.6, Python 3.7, Python 3.8)

2. Integrated Development Environment (IDE):

- An IDE provides a convenient development environment with features such as code editor, debugger, and project management tools.
- Recommended IDEs: PyCharm, Visual Studio Code, Jupyter Notebook

3. Libraries and Packages:

- Various Python libraries and packages are used for web scraping, data manipulation, GUI development, and more. Install the following packages using pip (Python package manager) or conda (Anaconda distribution) as needed:

- pandas: For data manipulation and analysis
- BeautifulSoup: For web scraping
- requests: For making HTTP requests
- tkinter: For building GUI interfaces
- pyshorteners: For URL shortening
- Pillow: For image processing (required for displaying images in GUI)

You can install these packages using the following commands:

```
pip install pandas beautifulsoup4 requests pyshorteners pillow
```

4. Web Browser:

- A web browser is required to access online resources and view web-based documentation or tutorials related to the software.
- Recommended browsers: Google Chrome, Mozilla Firefox, Microsoft Edge

5. Operating System:

- The software is compatible with various operating systems, including Windows, macOS, and Linux distributions.

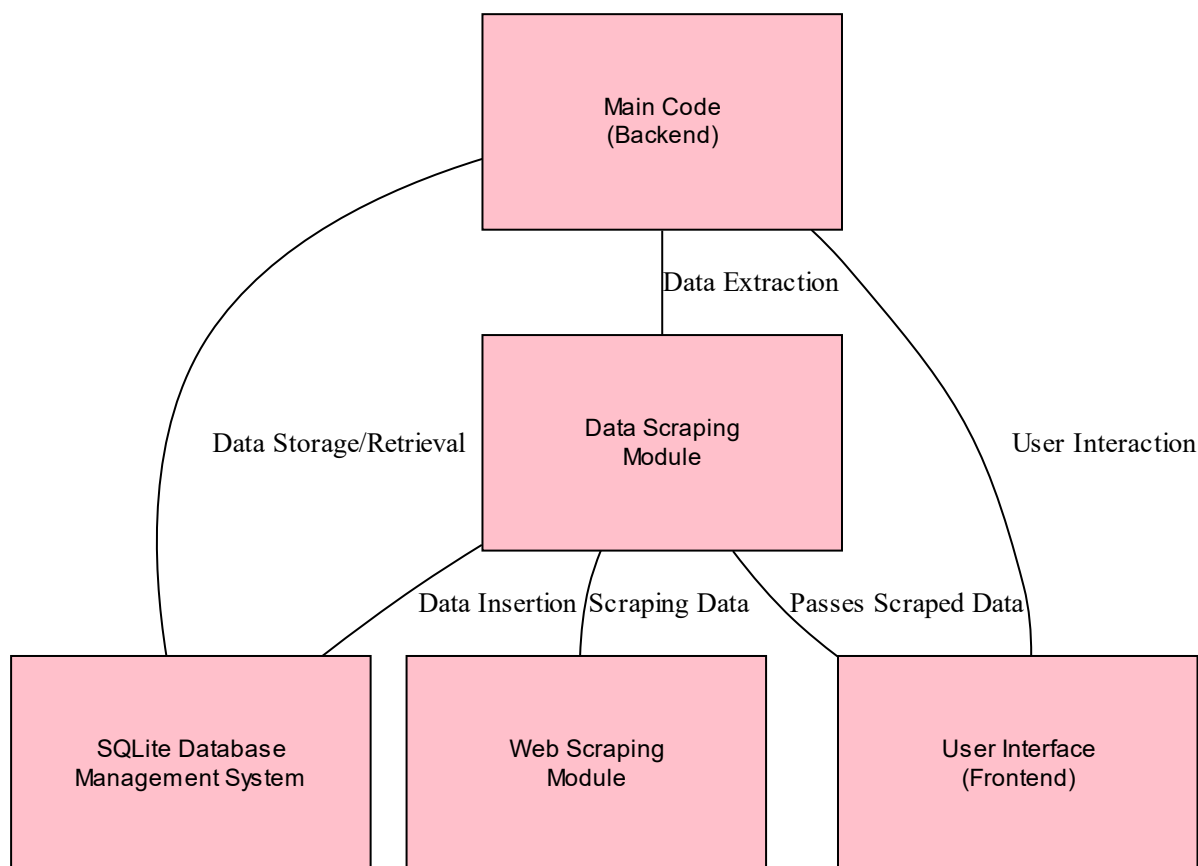
By ensuring that all the necessary software tools and packages are installed and configured correctly, users can set up the software environment and start using the application without any compatibility issues or errors.

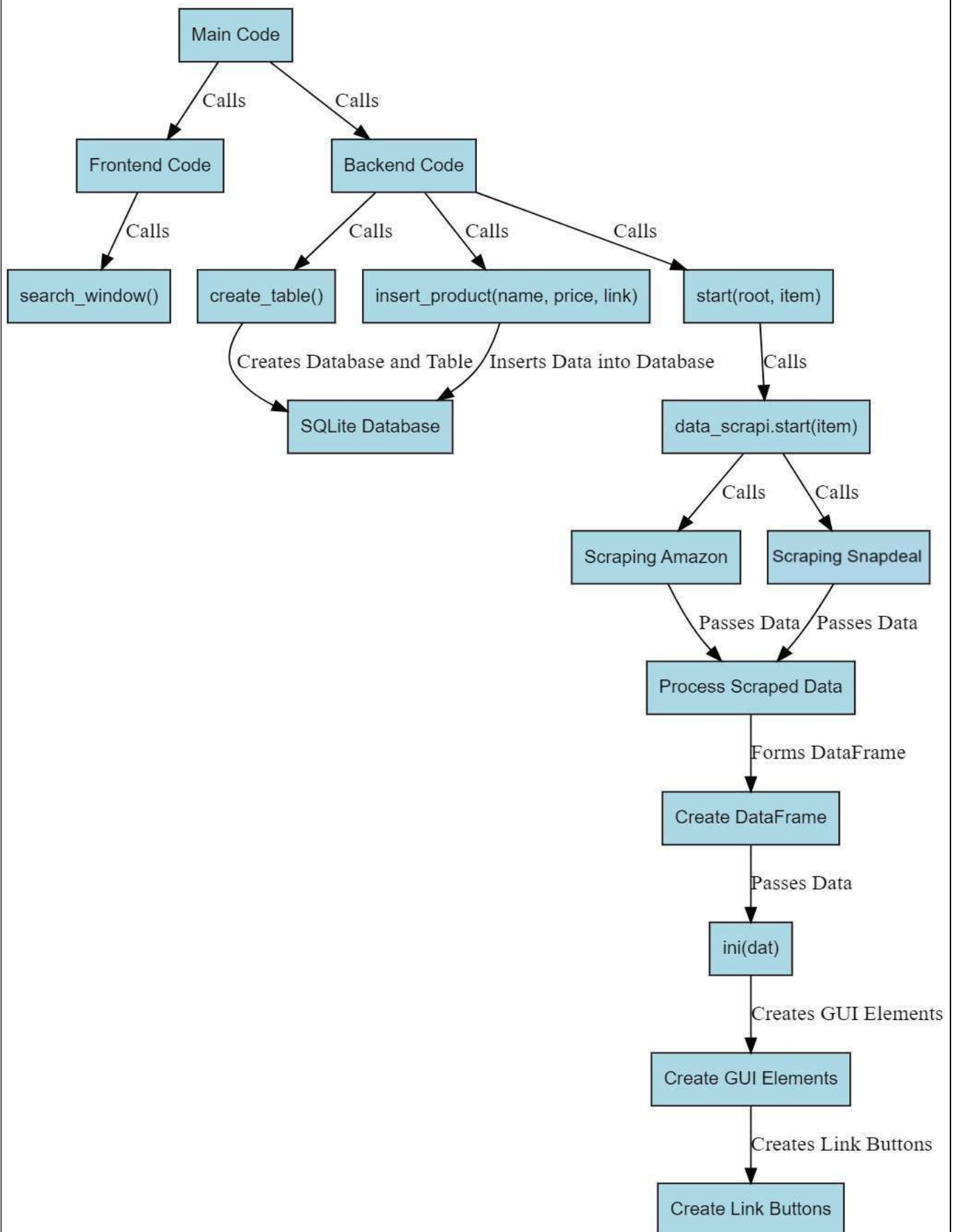
CHAPTER 5

Proposed System & Implementation

5.PROPOSED SYSTEM & IMPLEMENTATION

5.1 Block diagram of proposed system





5.2 Description of block diagram

The block diagram illustrates the system flow of the provided code, which includes the following main components:

1. Main Code (Backend): This component represents the main backend code responsible for coordinating the overall system functionality.

2. Data Scraping Module: This module handles the extraction of data from various sources, such as websites or databases.

3. SQLite Database Management System: It manages the storage and retrieval of data in the SQLite database, serving as the backend data repository.

4. User Interface (Frontend): The frontend component responsible for interacting with users, displaying information, and receiving user input.

5. Web Scraping Module: This module specializes in scraping data from web pages, facilitating the extraction of relevant information.

- Main Code interacts with the Database for data storage and retrieval operations.
- Main Code also communicates with the Data Scraping Module to extract data.
- Extracted data is then inserted into the Database by the Data Scraping Module.
- The Data Scraping Module passes the scraped data to the User Interface, which is responsible for displaying it to the user.
- Additionally, the Main Code facilitates user interaction by communicating with the User Interface.
- The Data Scraping Module may also interact with the Web Scraping Module to scrape data from web pages.

5.3 Implementation

Front End:

1. Imports necessary modules like ``sqlite3``, ``pandas``, ``tkinter``, ``webbrowser``, ``PIL``, and a custom module ``data_scrap``.
2. Defines a function ``create_table()`` to create a SQLite database named ``product_details.db`` if it doesn't exist and a table named ``products`` with columns ``id``, ``name``, ``price``, and ``link``.
3. Defines a function ``insert_product(name, price, link)`` to insert product details into the SQLite database.
4. Modifies a function ``start(root, item)`` to utilize the ``data_scrap`` module to fetch product details and insert them into the SQLite database.
5. Defines a function ``shorten(title)`` to shorten the product name for display.
6. Defines a function ``ini(dat)`` to initialize the graphical user interface (GUI) using ``tkinter``. It displays product details like name, price, and a link to the website where the product is available. It also creates clickable links to open the website in a web browser.
7. Defines a function ``search_window()`` to create a search window where users can enter the product they want to search for.
8. Calls the ``create_table()`` function to create the database table.
9. Calls the ``search_window()`` function to start the search interface.

Overall, this program fetches product details from a website, stores them in a SQLite database, and provides a GUI for users to search and view product details.

Back End:

Import Statements:

- ``pandas``: Used for data manipulation and analysis.
- ``BeautifulSoup`` from ``bs4``: A library for pulling data out of HTML and XML

files.

- ``requests``: A library for making HTTP requests in Python.
- ``re``: A module providing support for regular expressions.
- ``pyshorteners``: A library to shorten URLs.

2. Function ``generate_url(part1, part2, search_for, ch)``:

- Takes four parameters: ``part1``, ``part2``, ``search_for``, and ``ch``.
- Constructs a URL by replacing spaces in the search query (``search_for``) with the character specified by ``ch``.

3. Function ``start(search_for)``:

- Takes a search query (``search_for``) as input.
- Generates URLs for Amazon and Snapdeal using the ``generate_url`` function.
- Sends HTTP requests to the generated URLs and retrieves the HTML content of the web pages.
- Parses the HTML content using ``BeautifulSoup``.
- Extracts product information such as product name, price, and link to the site.
- Constructs a dictionary (``data``) containing the extracted information.
- Converts the dictionary into a pandas DataFrame.
- Shortens the URLs using the ``pyshorteners`` library.
- Returns the DataFrame containing the product information.

4. Printing Logo:

- Prints a logo before executing the function ``start(search_for)``.

5. Main Execution:

- Calls the ``start(search_for)`` function with the provided search query.

tk

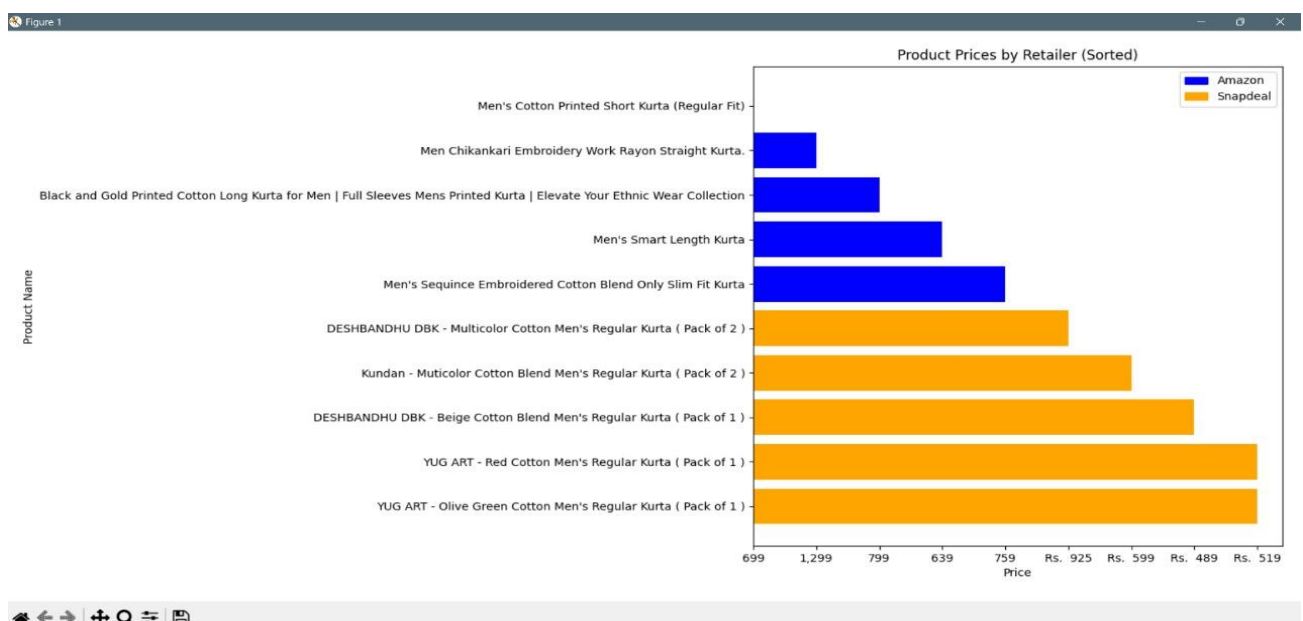
Enter the product

kurta for mens

Search

I agree to the Terms & Conditions

amazon.com			snapdeal		
Product Name	Price	Link To Site	Product Name	Price	Link To Site
Men's Cotton Printed Short Kurta	Rs 699	link	DESHBANDHU DBK - Multicolor Cotton	Rs Rs. 925	link
Men Chikankari Embroidery Work	Rs 1,299	link	Kundan - Muticolor Cotton Blend	Rs Rs. 599	link
Black and Gold Printed Cotton Long	Rs 799	link	DESHBANDHU DBK - Beige Cotton Blend	Rs Rs. 489	link
Men's Smart Length Kurta	Rs 639	link	YUG ART - Red Cotton Men's Regular	Rs Rs. 519	link
Men's Sequence Embroidered Cotton	Rs 759	link	YUG ART - Olive Green Cotton Men's	Rs Rs. 519	link



5.4 Advantages

Competitive Pricing Strategy: E-Commerce Price Comparison allows businesses to analyze competitor pricing strategies comprehensively. By understanding how competitors' price their products, companies can adjust their own pricing strategy to remain competitive while maximizing profitability.

Market Insights: Price comparison tools provide valuable insights into market trends and consumer behavior. By analyzing pricing data across various platforms and regions, businesses can identify emerging trends, demand patterns, and price sensitivities, enabling them to make informed decisions.

Optimized Product Positioning: Understanding how competitors' price similar products helps businesses position their offerings effectively in the market. By identifying pricing gaps and opportunities, companies can differentiate their products based on value proposition and adjust pricing accordingly.

Enhanced Customer Experience: Price transparency is crucial for building trust and satisfaction among customers. By offering price comparison features on their e-commerce platforms, businesses can empower customers to make informed purchasing decisions, leading to higher satisfaction levels and repeat business.

Dynamic Pricing: E-Commerce Price Comparison enables businesses to implement dynamic pricing strategies. By monitoring competitor prices in real-time and adjusting prices accordingly, companies can optimize revenue and maximize margins based on demand fluctuations and market conditions.

CHAPTER 6

Conclusion

6.CONCLUSION

The price comparison project between Snapdeal and Amazon presents a valuable resource for consumers seeking the best deals in online shopping. By offering a straightforward comparison of prices, it empowers users to make informed purchasing decisions. In essence, the price comparison project serves as a valuable resource for consumers navigating the vast landscape of online shopping, empowering them with the information needed to make confident and well-informed purchasing decisions while promoting fairness and transparency in the e-commerce ecosystem.

FUTURE SCOPE:

The software's capabilities are just the beginning, and there are several potential avenues for further development and improvement:

- **Enhanced Product Matching:** Improving product matching accuracy is crucial for providing users with relevant and reliable comparison results.
- **Personalized Recommendations:** Tailoring product recommendations based on user preferences enhances user experience and increases engagement.
- **Price Drop Alerts:** Alerting users about price drops helps them make timely purchasing decisions and save money.
- **Geographical Expansion:** Expanding to international markets increases the tool's reach and potential user base.
- **Integration with Voice Assistants:** Voice-enabled functionality adds convenience and accessibility, improving user engagement.
- **In-depth Product Reviews:** Providing comprehensive product reviews assists users in making well-informed decisions beyond price considerations.
- **API Integration:** Providing APIs for developers increases the tool's versatility and potential for integration into various platforms and applications.

REFERENCES

- [1] “E-commerce Price Comparison Website Using Web Scraping”
International Journal of Innovative Research in Engineering & Multidisciplinary
Physical Sciences Volume 11(3),June 2023
- [2] “E-Commerce Price Comparison With Review Sentimental Analysis”, Sanket
Bezalwar, Vikas Bhandekar, Sagar Kumbhare, Rushikesh Rebhankar , Prof.
Prajakta Singam , Dept. Of Computer Technology, Rashtrasant Tukadoji Maharaj
Nagpur University (RTMNU), KDK College Of Engineering, Nagpur,
India,Volume 8, 2023 IJCSMC, Vol. 11, Issue. 3, March 2022
- [3]“Price Comparison Website for Online Shopping” ,S. Rajendar, K. Manikanta,
M. Mahendar, Assistant Prof. (Mrs.) K. Madhavi Department of Computer
Science and Engineering, St. Peter’s Engineering College, Hyderabad. IJCRT |
Volume 9, Issue 6 June 2021
- [4] “Price Comparison for Products in Various E-Commerce Website” Mrs. M.
SOWMIYA, SRINANDHAN CS, MUGESH RAJA M, SUDHEEKSHAN
KUMAR Dept. Of Computer Science & Engineering Meenakshi Sundararajan
Engineering College, Chennai, India. IJRTI | Volume 8, Issue 5
- [5] “E-Commerce Product Price Tracker”, Dr. G Madhusudhan, Nitin
Gopalakrishna Bhat, Sahana Venkatraman Patgar, Chandan N A, Bharath S V,
Computer Science and Engineering 1,2JSS Science and Technology University,
Mysuru, India, JETIR , Volume 8, Issue 6,June 2021
- [6] “Comparitive Analytical Study Of Snapdeal And Amazon”,Visha Patel,(Mba
(Ba) Student) Maazahmed Shaikh,(Mba (Ba) Student) Tushar
Pradhan,(Professor) Parul University Tijer || Issn 2349-9249 || Volume 10, Issue
3, March 2023