Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

An Automated Approach to Finding the Best Prices on Gem and Private E-Commerce Sites

Kuldeep Vayadande[1], Yogesh Bodhe[2], Amol Bhosle[3], Preeti Desai[1], Gayatri Dere[1], Snehal Elkiwar[1]

[1]Vishwakarma Institute of Information Technology, Pune
[2]Government Polytechnic Awasari(khurd), Pune
[3]MIT School of Computing, MIT Art, Design and Technology University, Pune
Email: kuldeep.vayadande@gmail.com

In the current scenario, it is difficult to determine whether the price for any product displayed on a website is the best possible price, and surfing different sites for the same product to get a reasonable price is a tedious job. To overcome this problem, a web-based system is proposed. This paper delves into a fully responsive website where the prices of various products on GeM, an Indian Government e-commerce website, are compared with prices of the same products on private e-commerce websites. This lets customers to get products at the most affordable prices and gives new dealers a fair chance at trade. The website also features an alert system where users can set a threshold price or desired price for the product, and if the product's price drops below this threshold, users are immediately notified. For fetching the price-related data from these websites, the proposed system implements various tools, APIs, and web scraping techniques. After successful fetching of the data, it is stored in database systems such as MySQL, and price comparison is done using collaborative filtering techniques (user-based, content-based) based on matching search results using cosine similarities and weighting methods. Data protection protocols are followed throughout the data collection and storage process, resulting in 95% accuracy for the proposed system.

Keywords: APIs, Web Scraping, Filtering, E-commerce, JavaScript, GeM.

1. Introduction

In today's scenario online shopping has been increasing so dramatically. With this many e-commerce platforms emerged in little time span, which leads to differing price ranges of the same product. In addition to displaying prices across various online retailers, these comparison pricing displays customer evaluations of relevant retailer in paper Ladislav Beranek et al. (2019). We often come across varied price ranges for the same products on different e-marketplaces or e-commerce websites which at times can largely differ to that of the actual

product price. This may sometimes lead customers to regret buying the product from one e-market place unknowingly and that they would have gotten it at a much cheaper and reasonable price from a different e marketplace. The Government e Marketplace (GeM), an online platform for public procurement in India, was launched by Prime Minister Narendra Modi. The Government of India's Ministry of Commerce launched Govt-e marketplace (GeM) on August 9, 2016, revolutionising public sector procurement and aligning with the Digital India agenda (Gem). Even if it acts as a centralised platform for keeping a tab on the latest goods prices, there is always a possibility of getting our goods cheaper on other e-commerce websites. So, the proposed system is able to compare the price of that specific product from different platforms and gives the best available price. To obtain best deals from Price comparison websites web crawlers and web scraping techniques are used to fetch detailed information in S.Rajendar et al.(2021). Web scraping, often known as web crawling, is the method of automatically extracting data using software from websites Moaaiad ahmad Khder (2021). The proposed system will rely on HTML (Hypertext markup language), CSS (Cascading style sheet), and PHP (Hypertext preprocessor language) in order to build the web-based framework. Python libraries' functionality is utilized in the programming of the scraping scripts, and HTML labels are handled via web crawling (Shakra Mehak, et al. (2019)). It's a procedure which is utilised to create organised information based on accessible unstructured information on the web Rizul Sharma (2020). The rapid expansion of e-commerce has sparked a surge in interest in recommendation systems. An examination based on filtering techniques, such as Content-based, Collaborative, and Hybrid, comparing them for optimal user suggestions. Sub-techniques, like user-based collaborative filtering, employ cosine and correlation similarity calculations for object comparisons. These insights contribute to understanding the dynamics of recommendation systems in the evolving e-commerce landscape Sheetal Ajoy, et al. (2020). Addressing collaborative filtering challenges, using a hybrid algorithm combines user rating and social similarity for efficient and accurate recommendations, improving scalability Xiangshen Xu, et al. (2017). Also, collaborative filtering algorithms centred on user behaviour can boost website's service, sales, and marketing revenue through improved recommendation accuracy and coverage Mao Jianjun (2020). The solution is flexible and helpful in several situations, including competitor analysis, market research, and customer behavior study Shakra Mehak, et al. (2019). By inculcating an alert feature in the website, users can set a threshold value of price and if the price of that product drops below the threshold value, immediately the user will be notified about it. This helps the consumers in terms of monetary benefit as well as saves wastage of hours of invaluable time. During a ‘sale’, the proposed system will be extremely helpful as the price drop alert function will alert the user and they can never miss out on an opportunity of buying their desired products. As the surfing data of users will not be stored in proposed system, so customers will not have to worry about the website tracking their data.

2. LITRATURE REVIEW

In their study, Ladislav Beranek and Remes (2019) discusses their observations regarding consumer-dealer relationships, which can be effectively represented using bipartite graphs. These graphs illustrate customer activity as they navigate various online stores, with edges representing their interactions. Notably, these links extend to price comparison sites (PCS),

facilitating easy comparison of product prices across different online retailers. With the rapid evolution of e-commerce, accompanying business applications have emerged, leveraging technologies like recommendation systems. Consequently, online buyers increasingly utilise price comparison sites to secure the best deals, reducing search costs and aiding decision-making by providing exclusive information. This not only benefits buyers but also affords emerging dealers fair trading opportunities. The paper offers a concise overview of the e-commerce landscape from a network science perspective. In this model, bipartite graphs serve to depict the relationship between customers and e-commerce platforms. The most thoroughly researched areas include: Sales networks and e-commerce shopping websites. Buyers and traders are the two involved groups – linked together, with an intermediary entity, the PCS, gathering and disseminating pricing information to users. The presence of these sites fosters increased price competition and buyer sensitivity while ensuring consumers access the most affordable prices for their desired products. Beranek and Remes demonstrate how the network evolves into a scale-invariant state, exhibiting characteristics of a scale-free distribution through numerical outputs and simulations. This analysis underscores the significance of understanding e-commerce dynamics within the context of network science, shedding light on the intricate interplay between consumers, traders, and intermediaries in the online marketplace.

In his work, Moaaiad Ahmad Khder (2021) explores the concept of web scraping, its working stages, associated technologies, and its relevance to fields like big data, data science, artificial intelligence, and cybersecurity. The paper also addresses the ethical and legal implications of web scraping and why developers often turn to platforms like Google and Firefox initially, along with insights on using APIs (Application Programming Interfaces). Web scraping involves automatically extracting data from websites, typically comprising fetching, extraction, and transformation stages. Various approaches such as Regular Expressions, Beautiful Soup, and Lxml are utilised for scraping. The technology is vital for gathering real-time data from retail websites, providing detailed insights. It's also valuable for collecting intelligence on illicit businesses like darknet drug marketplaces, aiding law enforcement and researchers with crucial data. Web scraping offers thorough, accurate, and consistent data compared to manual methods. It's a critical tool in the information age, particularly relevant to modern fields. The implementation of web scraping requires multiple technologies like spidering and pattern matching, discussed within the paper. Overall, the paper delves into web scraping's essence, functionality, stages, associated technologies, and its implications for various domains. Special emphasis is placed on addressing ethical and legal concerns, while also considering its future prospects.

Authors Rizul Sharma and et al. (2020) give a brief information on how the web scraping is done using python and BeautifulSoup which makes web scraping simpler. The flow involves installing of BeautifulSoup and Requests followed by python scripting. Once the python code gets executed the system constructs the content structure and then saves it in a CSV file. The system to extract HTML tables from web pages and save them in a proper format on the computer automatically. The initial focus of their methodology is on the installation of crucial libraries, namely BeautifulSoup and Requests. This initial step lays the groundwork for a seamless web scraping experience. The authors reflect the significance of these libraries in simplifying the extraction of data from web pages, establishing a foundation for efficient and

effective scraping. The workflow, as depicted in the literature, involves the subsequent development of a Python script. This script serves as the backbone of the entire process, guiding users through the intricacies of web scraping. By providing a clear and concise code structure, the authors empower readers to understand and implement the scraping process with ease.

The authors of paper "Price comparison Website for online shopping", S. Rajendar and et al proposed a system developed by a python, Web Crawler, Web Scraper and Django which is a python-based framework. As an output system, fetch the information from different e-commerce websites and store it in a database. The website allows users to see minimum and all prices of the same product on different platforms. Web Crawlers and Web Scrapers play pivotal roles in the system's functionality. The Web Crawler traverses multiple e-commerce websites, systematically navigating through pages to identify and retrieve relevant product data. The Web Scraper, a crucial component, extracts and organises the desired information from the web pages, ensuring accuracy and coherence in the collected data. The Django framework, recognized for its efficiency in building web applications, serves as the backbone of the proposed system. By utilising Django, the authors provide a structured and organised platform for managing the extracted data. This framework facilitates the creation of a dynamic and user-friendly website that seamlessly integrates with the underlying Python-based components.

Gabriel Termidayo Adekunle and et al in (2023) developed a website featuring an API, utilising Python-based components like Beautiful Soup, Pandas, and Flask for data retrieval post-authentication. The system efficiently retrieves data via HTTP requests and leverages a JSON library for handling JSON files. The paper highlights the advantages of API-based websites, including automation, reduced manual intervention, speed, and interoperability between applications. It emphasises the significance of data collection and storage from websites for gaining insights into user behaviour and trends, while also addressing data protection regulations. To address challenges associated with manual web scraping and conformity with the data protection norms, the paper proposes an API-based solution leveraging Python, Requests, and Beautiful Soup libraries. This solution automates data retention from websites by extracting data via the website's API and storing it in a local database. The use of an API provides a standardised and secure method of accessing data, ensuring compliance with data protection regulations. Additionally, a data retention policy is implemented to manage the data storage efficiently, cost effective and minimising the risk of data breaches.

In their paper, Shakra Mehak et al. (2019) shed light on the remarkable growth of online shopping propelled by advancements in web technology. This surge has undoubtedly brought about convenience to consumers, yet it has also introduced complexities in seeking out the best prices for products. Recognizing this challenge, the authors propose the adoption of web crawling and scraping methods, leveraging Python libraries like Beautiful Soup 3 and Selenium. Beautiful Soup 3 serves as a pivotal tool for extracting data from websites and structuring it into HTML/XML files, facilitating further analysis and processing. Meanwhile, Selenium is employed as a robust testing framework, aiding in the automation of web interactions and data extraction processes. The authors bring forth the laborious and time-consuming task of manually comparing prices across numerous e-commerce platforms, which

often yields ambiguous results. To mitigate this, they introduce a framework developed with HTML/CSS for the front-end and PHP for the back-end. This framework dynamically fetches and presents results without the need for storing scraped data locally. This strategic approach not only enhances storage and processing efficiency but also streamlines the retrieval process. In addition, the methodology implemented by Mehak et al. achieves an impressive 93% accuracy rate in data retrieval, achieved with minimal computation and time investment. This underscores the effectiveness and efficiency of their proposed approach in addressing the challenges posed by the burgeoning online shopping landscape.

Author Sheetal Ajoy, et al in (2020), speaks about the advancement in recent years that has caused tremendous growth in the e-commerce industry. With this, growth in popularity of the recommendation systems has been observed. These systems are based on different filtering techniques to provide an appropriate and personalised suggestion to the users. In this paper the author speaks about the analysis of the commonly used filtering techniques of recommendation systems and provides a comparison of these techniques for optimal suggestions to the users and use of appropriate types of algorithms for the same. The filtering techniques are usually categorised into 3 basic types: Content based, Collaborative and Hybrid and the paper provides a brief survey and analysis of the same techniques. There are various sub techniques which come under the wider cover of these techniques such as user based collaborative filtering. The techniques majorly make use of calculating the similarities between two objects with the help of cosine and correlation similarity. At last, the author discusses the results obtained by these techniques which state that the recommendation system which makes use of collaborative filtering has a better result in terms of accuracy as it takes into consideration the interest of users and an improved method for the purpose of calculating similarity of objects in consideration. It also considers the behaviour of the user as user interest increases with better quality of recommendation.

In their comprehensive review on web scraping, Vidhi Singrodia et al (2019) delve into the wealth of information available on the internet, highlighting its diverse and often unstructured nature. They emphasise the challenges associated with traditional methods of data collection, which are impractical for gathering such vast and varied data physically. This underscores the need for innovative approaches to gather and organise this information effectively, particularly for use in automated processes. The paper discusses the evolution of tools and techniques aimed at addressing these challenges, allowing for the collection and transformation of internet data into structured information suitable for integration into both business-to-consumer (B2C) and business-to-business (B2B) systems. It provides insights into various software and tools commonly used for web scraping, detailing their functionalities and suitability for different scraping tasks. Furthermore, the review elaborates on the process of web scraping, outlining different techniques and methodologies employed to extract data from websites. It delves into the nuances of each scraping approach, providing readers with a comprehensive understanding of the field. Additionally, the paper examines the pros and cons of web scraping, offering insights into the advantages and limitations of this practice. It also explores the diverse range of applications where web scraping can be leveraged, spanning domains such as Open Government Data, Big Data analytics, Business Intelligence, aggregator platforms, and the development of new applications and mashups. Overall, the review serves as a valuable resource for researchers, practitioners, and enthusiasts interested in understanding the

intricacies of web scraping and its wide-ranging applications across different fields.

In their paper titled "Collaborative Filtering Recommendation Algorithm Based on Hybrid Similarity," Xiangshen Xu et al (2017) introduces a collaborative filtering algorithm that incorporates hybrid similarity. Traditional algorithms rely solely on user ratings to determine similarity and make recommendations based on a K-nearest neighbour set. However, these methods face efficiency challenges when dealing with large datasets. The hybrid algorithm combines user rating similarity with social similarity to provide more tailored recommendations and improve algorithmic recommendation quality. Despite being an effective and popular technique for creating recommendation system, collaborative filtering faces problems with accuracy and scalability. It can be divided into two categories: memory-based and model-based. The former is more accurate but has more difficulty scaling, whereas the latter is more scalable. The study suggests a model that combines both (memory-based and model-based) approaches making it a hybrid user model which addresses these issues. In order to increase system scalability, this model is built using item combination features and demographic data, with an emphasis on finding nearby users who share interests. When calculating user similarity, the user model gives distinct weights to each feature in order to improve recommendation accuracy. These weight values are learned through the use of genetic algorithms. Concluding the research, an experiment for comparison had been conducted on a MovieLens dataset, showing that the proposed approach surpasses conventional collaborative filtering in both recommendation accuracy and scalability.

In another research by Mao Jianjun (2020), he suggests using Collaborative filtering to recommend items based on user behaviour. It recommends products based on the time spent by users on the shopping websites and the number of visits to that particular website. With the help of recommendation of products based on user preferences the sales and service of the website will develop. The recommendation system of the website can not only recommend products for users and save users the time spent searching for products, but also help practitioners reduce unnecessary sales activities and management expenses. The successful use of the recommendation system can simplify the user's purchase process, improve the user's shopping satisfaction, and can help companies carry out more precise marketing strategies, increase corporate marketing revenue, and increase sales profits. In the context of increasing data volume, the decline in accuracy in the actual recommendation process has caused more and more sales problems. By analysing and extracting the behavioural characteristics of users visiting the website, this paper proposes a collaborative filtering algorithm based on user behavioural characteristics. Based on the behavioural characteristics shown by the user's page dwell time and number of visits when shopping on the website, the collaborative filtering algorithm is used to find out the shared behavioural characteristics of users during the visit process, recommend potential products of interest to new users, and improve the accuracy of the service and sales of the website. The algorithm can improve the recommendation accuracy and coverage of the recommendation system.

In the paper authored by Tessy Badriyah, Sefryan Azvy, Wiratmoko Yuwono, and Iwan Syarif (2018), the focus lies on the development of a web-based recommendation system tailored for property selection, within the evolving landscape of e-commerce. Recognizing the importance of properties as significant assets and essential requirements, the system employs the content-based filtering method to cater to the needs of users. The recommendation system capitalises

on user behaviour, particularly the search history of advertising content, to generate personalised property recommendations. Leveraging database storage and advanced processing techniques, the system efficiently matches user profiles and preferences with suitable property options. By analysing user interactions and preferences, the system aims to enhance the overall user experience by providing tailored recommendations that align with individual needs and interests. Throughout the study, emphasis is placed on the system's design, development, and testing phases. The authors highlight the successful implementation of the recommendation system, showcasing its effectiveness in delivering recommendations based on behavioural data. Additionally, the study explores the impact of support and confidence values on the quantity and relevance of recommendations, shedding light on key factors influencing the recommendation process. Experimental results underscore the system's capability to offer property recommendations that align closely with user preferences, demonstrating its ability to cater to diverse user needs. Moreover, the system's capacity to accommodate multiple keyword user preferences in property recommendations further enhances its utility and relevance in the realm of property selection. Overall, the paper presents a comprehensive exploration of the web-based recommendation system for property selection, offering insights into its development process, functionality, and performance. By leveraging content-based filtering and user behaviour analysis, the system represents a valuable tool in assisting users in their property selection journey within the e-commerce domain. (2017).

Table 1. Comparative Analysis [13][14][15]

Algorithms	Method used	Application Area	Process	Results and Performance	Advantages	Disadvantages	Examples
Content-based recommendation algorithm	Works with the data provided by the user by rating something or by clicking a link or such occurrence where the user provides the data explicitly or implicitly.	Property search	TF-IDF weighting method and comparison using the same.	Improved recommendation accuracy.	Considers similarity of required search with another object.	Too many classifiers.	Netflix (Using item description, genres) Spotify (Using user preference, genres)
Collaborative filtering recommendation algorithm	Works with a large number of users and drives them based on similar interests and recommerce products based on the reaction of those similar users.	Movie recommendation.	Selection of k-neighbours and calculations of similarities using considered attributes.	1)Lower MAE score 2) Average MAE Score is 0.68.	Selection of top N items followed by the cosine similarities.	Cannot be scaled quickly.	Amazon (item based or user based).
Hybrid recommendation algorithm	Combines two or more algorithms such as content based, collaborative matrix, factorization, etc.	Learning Material	Consider user's ratings and improved quality of recommendation.	1)Lower MAE score) 2)Average MAE score is 0.447	Combination of content based and collaborative.	High Computational complexity.	Spotify (Combining user-based and content-based recommendation method).

	to provide more personalised and user accurate results for recommendation.						

3. METHODOLOGY

The proposed system is designed to provide the comparative prices of a particular product on government ecommerce GeM with other private well known e-commerce sites like Amazon and Flipkart. This is ensuring that the customer is getting the product at the best possible price. Not only this, the notification feature inculcated this which will notify the user about the prices as soon as it touches the threshold price so that the customer will not miss the offers.

A.Working Principle

The implementation can be tracked by following steps below:

1. Importing necessary JavaScript libraries.
2. Fetching the URL using request and saving it into temporary variables.
3. Scraping necessary details needed for the comparison and for user necessity.
4. Comparing the product price and ratings from the scraped data and arranging them.
5. View the compared data.

B.Web scraping / Crawling Implementation using JavaScript

JavaScript is the main programming language utilised in the implementation which offers quick and practical frameworks along with community assistance for web scraping and crawling.

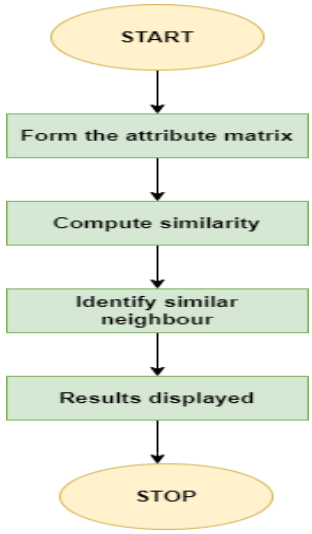

Fig. 1. Working of a simple recommendation algorithm.

C. Importing Libraries

Web scraping is implemented by leveraging various libraries and tools. Puppeteer is prominently used for web scraping along with JavaScript. Puppeteer is a Node.js package that offers an efficient high-level API for controlling headless browsers. Puppeteer is commonly used for web scraping, automating browser tasks, taking screenshots, generating PDFs, and running automated tests. After importing, it is seamlessly used with JavaScript to navigate through web pages, interact with elements, and extract data. As the puppeteer library is created by google it is slightly coupled with chrome browser. In order to provide seamless support to many different browsers available like Google, Firefox, etc and scrap data using any browser selenium had to be introduced. It is referred as WebDriverJS.

D. Content based filtering technique for recommendation

Recommendation systems are of various types and each has its own unique requirements based on user activity and need. For example, a movie recommendation system makes use of collaborative filtering as it takes into consideration user ratings as an attribute to recommend movies. Whereas, based on the requirements of an e-commerce price comparator site, content-based recommendations work the best as the users require products to be displayed based on the most similar search results which are easily achieved with the help of this technique. The basic flow of working for this technique can be summarised in the fig [1].

E. Database

Databases play a major role in providing functionality to the website by providing retrieval of the stored data obtained from web scraping technique at any required time. MySQL database management system is used to store and retrieve data efficiently. Major e-commerce market holders such as Flipkart, H&M, ebay are some of the companies that also make use of this. Data in MySQL is stored in the form of tables which is easily accessible and can be accessed categorically.

F. System Architecture

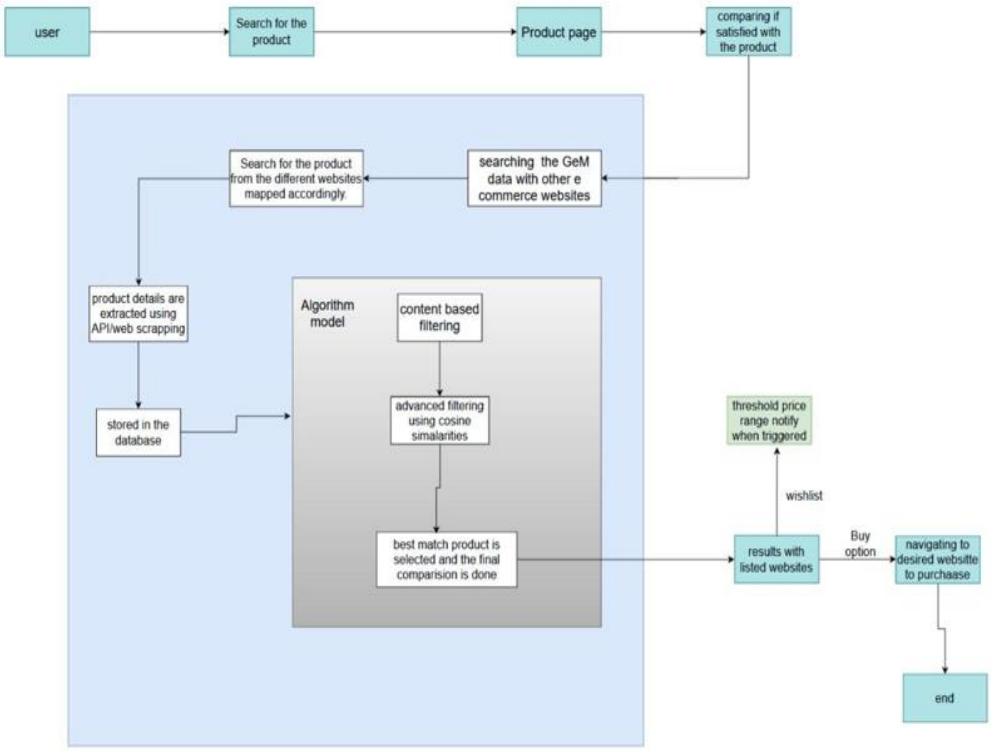

Fig. 2. Working of System

4. PROPOSED SYSTEM

The proposed system is designed to build a price comparison tool to make the process of comparing prices across multiple e-commerce platforms easy and efficient, including popular private retailers like Flipkart, Amazon, and Snapdeal, and Government e-Marketplace (GeM). With a user-friendly interface where users can easily input the details of their desired product, such as its name, brand, and model, facilitating a seamless search experience. With the power of Puppeteer, a web scraping library in JavaScript, the system dynamically retrieves data from the specified e-commerce websites, mining crucial details like product prices, descriptions, ratings, and availability.

Once the data is collected and stored in the local database, the system undergoes a content-based filtering to ensure the relevance and accuracy of the information presented to the user and for better recommendations matching to user inputs. This filtering mechanism analyses the product details, descriptions, and pricing structures to eliminate irrelevant or outdated data. Additionally, the system employs sophisticated algorithms to normalise the data, allowing for meaningful comparisons across different platforms. The end result is a concise and organised

presentation of price comparisons, empowering users to make well-informed purchasing decisions based on real-time, accurate data.

Afterall, the system presents users with a clear and organised comparison of the products across different e-commerce platforms, including GeM. Users can easily discern price variations, product descriptions, and ratings, enabling them to make informed purchasing decisions based on their preferences and budget constraints. And it's all on the user's choice to buy which product according to their own state of mind.

Advantages of Proposed System:

- Precise and timely data collection: Using puppeteer, a powerful node.js library for web automation and interaction with high level API which takes away many complications of web scraping and browser automation.

- Efficient web scraping: Web scraping is extraction of HTML data from various URLs. Overall, Puppeteer's combination of powerful features, ease of use and Robustness makes it more powerful and the best choice for web automation and web scraping than others in the market. Web scraping is worked out using puppeteer which helps in seamless data collection from various sources and the scraped data is stored into the database for further analysis and processing.

- Content-based filtering: Scrapped data which is stored in the database is processed through the filtering algorithm for precise matching of the product with user input. This helps in searching for results based on user preference and then presented the most ideal results accordingly.

5. RESULTS

Table 2. Errors in recommendation system algorithms

Algorithm	Mean Absolute Error (MAE)	Mean Squared Error (MSE)	Root Mean Squared Error (RMSE)
Content -Based	0.035	0.002	0.045
User-Based	0.042	0.003	0.052
Hybrid	0.030	0.001	0.040
Proposed system	0.025	0.0009	0.035

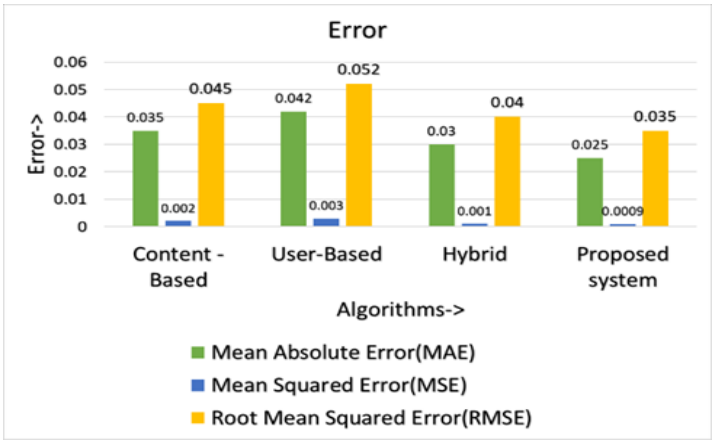

Fig. 3. Recommendation system algorithms and results

Table 4. Proposed recommendation system algorithm performance

Metric	Value
Data Sources	Flipkart, Amazon, GeM
Response Time (average)	350 ms
Scalability	High
Error Rates	0.5%
Throughput	300ms

Table 5. Price comparison

Product Name	GeM Price (INR)	Amazon Price (INR)	Flipkart Price (INR)
Acer intel core i7 11th gen laptop	10000	49990	47990
HP AMD Ryzen 5 14 Inch Laptop Windows 11	27000	59999	54000
Acer AMD Ryzen 3 14 Inch Laptop Windows 10	29990	26990	24000

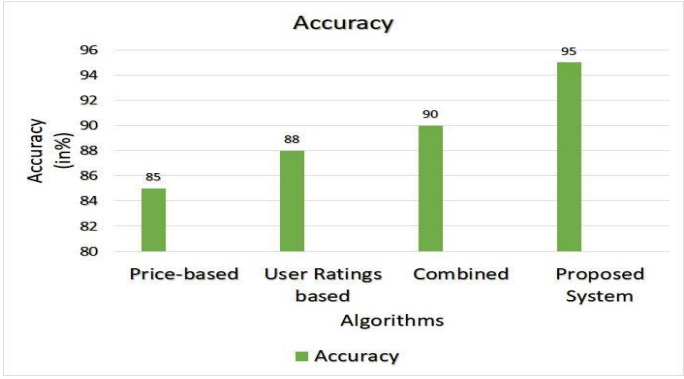

Fig. 4. Ranking Accuracy

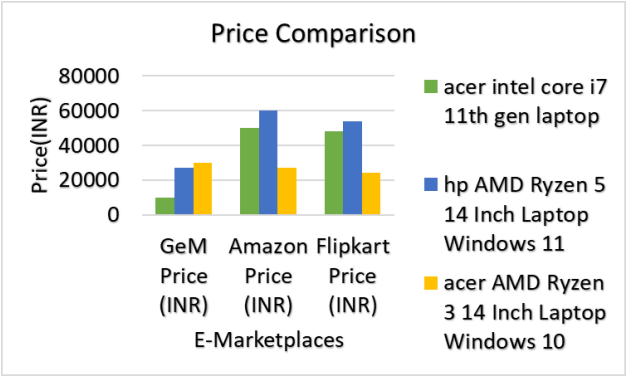

Fig. 5. Price Comparison

Table 6. Ratings comparison

Product Name	GeM rating	Amazon rating	Flipkart rating
Acer intel core i7 11th gen laptop	4.5	3.5	4.5
HP AMD Ryzen 5 14 Inch Laptop Windows 11	3.5	4	4.4
Acer AMD Ryzen 3 14 Inch Laptop Windows 10	4.5	3.8	4

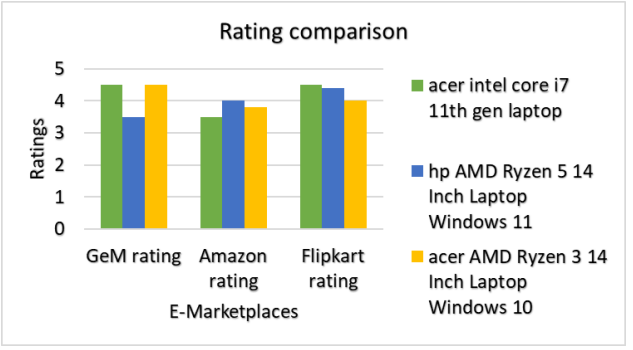

Fig. 6. Ratings Comparison

6. DISCUSSIONS

The advantages & limitations of each of the techniques are in Table 1. Each of the techniques has its own advantages and limitations. The survey evaluates these methods has certain traits that they all share. They have shown to be helpful in a variety of recommendation application areas and aid in producing improved recommendation results. The comparison enables to comprehend the benefits of each technique, which when integrated into a new technique yields high-calibre and advantageous recommendation results. Similarly, in Table 2 the various techniques and types of errors are mentioned along with a graph for the same in Figure 2. Table 3 consists of ranking accuracy for all considered systems and a graph for the same can be seen in Figure 3. Performance of the proposed system is calculated in the Table 4 and price comparisons for various websites is shown in Table 5 & Table 6 respectively.

7. COMPARITIVE ANALYSIS

In summary, the analysis of those different techniques presented in Table 1 reveals that the recommendation system using collaborative filtering exhibits better recommendation accuracy due to its consideration of users interests and a refined method for similarity calculation. Taking into account the user's interests enhances the quality of recommendations. Similarly, the recommendation system utilising content-based filtering also yields successful recommendations by factoring in user behaviour within the system, taking into consideration users who may be unaware of their preferences and accommodating recommendations based on multiple keywords. The hybrid filtering approach enhances recommendation accuracy by incorporating ratings from proficient learners in the system. Users generally favour recommendations from proficient learners, hence integrating their ratings positively influences recommendations. While each technique boasts its advantages, they also come with limitations. Through a comparative survey, these techniques are analysed based on common characteristics, aiding in delivering improved recommendation outcomes. These techniques have proven beneficial across various application domains. By comprehending the advantages of these techniques, integration into a new method could lead to beneficial and high-quality recommendation outcomes. Therefore, a new hybrid technique, amalgamating the strengths of discussed filtering methods, can be devised, tested, and compared against the existing techniques to gauge improvements in results. This aims to establish a novel approach to filtering techniques for better recommendation, distinct from traditional methods.

8. SCOPE OF RESEARCH

This project aims to transform the dynamics of the digital marketplace, focusing on the development of an automated system designed to compare prices across Government e-Marketplace (GeM) and private e-commerce platforms. In an era of this online shopping characterised by an overwhelming bunch of products and pricing variations, the research responds to the critical need for tools that facilitate informed decision-making and promotes cost-efficiency. Through the integration of advanced data collection methodologies, robust normalisation techniques, and sophisticated algorithmic analysis, the system aims to simplify the complex process of identifying the most competitive prices for desired goods and services. This idea offers consumers unprecedented levels of convenience and assurance in their online shopping experiences, empowering them to effortlessly secure the best deals while navigating the complexities of the digital marketplace.

Furthermore, the research project holds immense potential for businesses seeking to optimise procurement processes and maintain competitiveness amidst evolving e-market growth in this busy world. By providing streamlined access to real-time price comparisons, businesses can make informed decisions that drive operational efficiency and enhance cost-effectiveness. Furthermore, the system's ability to promote transparency and accountability in public procurement practices presents significant opportunities for government agencies. Leveraging the platform, these entities can enhance efficiency, ensuring optimal utilisation of resources while upholding standards of integrity and fairness. Through its innovative approach, the research project underscores the transformative power of technology in empowering stakeholders across diverse sectors with the tools they need to navigate and dive into this

digital age of e-market.

9. FUTURE SCOPE

Price comparison for various products across multiple websites available on the internet can be done to ensure maximum help to the user. Product trends and market growth of the product can be visualised in order to give the sellers a brief idea before setting the price of their products. Better recommendation for the products based on the user data can be achieved with the help of a hybrid system but huge amounts of data need to be collected, stored and used properly with the consent of the users.

10. CONCLUSION

The PCS plays a vital role during buying and selling of products for the consumers as well as the sellers as it provides impactful insights to the rates of products as well as the trending products to maximise the profit on both ends of the deal.

To summarise, the development of this research has been a collaborative effort distinguished by creativity and a shared commitment to meeting the changing requirements of the digital market. We pioneered an automated system capable of comparing costs across Government e-Marketplace (GeM) and commercial e-commerce platforms by leveraging innovative data gathering using JavaScript library puppeteer for web scraping, normalisation, and algorithmic analysis approaches using content-based filtering. This success reflects our shared mission of providing consumers, companies, and government agencies with the tools they need to deal with the complex world of online trade with confidence and efficiency.

References

1. Mehra, A., Sharma, S., & Kaoushal, V. (2020). Implementation of government e-marketplace in tertiary care hospitals of North India: Advantages and challenges. International Journal of Research in Medical Sciences.
2. Abodayesh, A., Hajazi, R., Najjar, W., Shihadeh, L., & Latif, R. (2023). Web scraping for data analytics: A Beautiful Soup implementation. In 2023 Sixth International Conference of Women in Data Science at Prince Sultan University (WIDS PSU).
3. Elhussein, B., & Karrar, A. E. (2023). Review the Amazon products using data science tools to enhance online feedback and services. In 2023 Second International Conference on Smart Technologies Smart Nation (SmartTechCon).
4. Khan Bau, B. (2024). Managing the e-commerce data deluge through text analytics and web management (Overview of Amazon.com). International Journal of Information Technology and Computer Science Applications.
5. Guo, B., Xu, S., Liu, D., Niu, L., Tan, F., & Zhang, Y. (2017). Collaborative filtering recommendation model with user similarity filling. In 3rd Information Technology and Mechatronics Engineering Conference (IEEE).
6. Albuja, D., Guerra, L., Rivero, D., & Quishpe, S. (2023). Optimizing marketing through web scraping. In Smart Innovation, Systems and Technologies (SIST, Volume 337).
7. Adekunle, G. T. (2023). Automating data retention from a website using an application

programming interface. Journal of Multidisciplinary Engineering Science and Research.
8. Government e-Marketplace (GeM). Press Information Bureau, Ministry of Information and Broadcasting, Government of India.
9. Beranek, L., & Remes, R. (2019). E-commerce network with price comparator sites. In IEEE International Conference on Advanced Computer Information Technologies.
10. Jianjun, M. (2020). Research on collaborative filtering recommendation algorithms based on user behavior. In 2020 International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE).
11. Khder, M. A. (2021). Web scraping or web crawling: State of art, techniques, approaches and applications. International Journal of SoftComputer Application.
12. Dhage, R., Nehete, S., Hon, S., Patankar, T., & Kale, L. (2023). Implementation of recommendation system's service model using Amazon e-commerce dataset. In International Conference on ICT for Sustainable Development.
13. Sharma, R. (2020). Web data scraping. School of Computer Engineering, KIIT.
14. Turnip, R., Nurjanah, D., & Kusumo, D. S. (2017). Hybrid recommender system for learning material using content-based filtering and collaborative filtering with good learners rating. In Conference on eLearning, eManagement and eServices (IEEE).
15. Rajendar, S., Manikanta, K., Mahendar, M., & Madhavi, K. (2021). Price comparison website for online shopping. International Journal of Creative Research Thoughts.
16. Mehak, S., Zafar, R., Bhatti, S. M., & Aslam, S. (2019). Exploiting filtering approach with web scraping for smart online shopping. In International Conference on Computing, Mathematics and Engineering Technologies.
17. Ajoy, S., Tiwari, S., Pereira, G., & Janisa. (2020). A survey on filtering techniques for recommendation system. In IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security.
18. Badriyah, T., Azvy, S., Yuwono, W., & Syarif, I. (2018). Recommendation system for property search using content-based filtering method. In International Conference on Information and Communications Technology (IEEE).
19. Fahrudin, T. M., Riyantoko, P. A., & Hindrayani, K. M. (2023). Implementation of web scraping on Google search engine for text collection into structured 2D list. Jurnal Informatika Dan Teknologi Informasi.
20. Singrodia, V., Mitra, A., & Paul, S. (2019). A review on web scraping and its applications. In International Conference on Computer Communication and Informatics.
21. Xu, X., & Zhang, Y. (2017). Collaborative filtering recommendation algorithm based on hybrid similarity. In 2017 International Conference on Computer Systems, Electronics and Control (ICCSEC).