iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.60688

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Slickdeal Aggregator: A Web Application using Web Scraping

Aviraj Jagtap[1], Raunak Jaiswal[2], Vaibhav Kedari[3], Nikhil Khajure[4]
[1, 2, 3, 4]*Students, Department of Computer Engineering, Zeal College of Engineering and Research, Pune, Maharashtra*

Abstract: ***An innovative method to facilitate well-informed purchasing decisions by consolidating and analyzing product data sourced from various online platforms. Given the rapid expansion of the e-commerce sector, numerous online retail outlets offer a wide assortment of products, complicating the process of comparing prices and discounts for consumers. By harnessing web scraping technology, we streamline the collection and organization of product information, thereby improving efficiency and accuracy.***

Our approach involves employing ReactJS for frontend interface design and Node.js for backend server development. At the heart of our methodology lies the integration of web scraping capabilities through Puppeteer, enabling automated extraction of product data. The resultant platform provides users with a seamless browsing experience and comprehensive product insights. Beyond its immediate applications, our research contributes to the advancement of web scraping technology within e-commerce domains.

Keywords: Web scraping , Puppeteer

I. INTRODUCTION

In the ever-expanding realm of e-commerce, consumers are faced with an increasingly vast selection of products across a multitude of online platforms. However, amidst this abundance, the challenge of navigating diverse pricing structures and identifying the best deals remains prevalent.

This study aims to tackle this challenge by developing a dynamic price comparison website, leveraging state-of-the-art technologies and methodologies. Through the integration of frontend development practices, backend scripting techniques, and advanced data scraping mechanisms, our objective is to provide users with a streamlined and comprehensive platform to facilitate informed purchasing decisions.

At the forefront of our development strategy lies the adoption of ReactJS for frontend interface construction. By harnessing React's component-based architecture, we prioritize modularity and reusability, allowing for the seamless integration of various features and functionalities.

Additionally, the utilization of Bootstrap ensures that our frontend design is not only visually appealing but also responsive across a range of devices and screen sizes, enhancing user accessibility.

Simultaneously, our backend development approach revolves around the utilization of Node.js as the runtime environment, complemented by Express.js for efficient routing and middleware management. This lightweight web framework facilitates seamless communication between the frontend interface and the underlying server, ensuring robust data transmission. Furthermore, the incorporation of diverse Node.js libraries enhances backend functionalities, enriching the platform with additional capabilities and optimizing overall performance.

A cornerstone of our price comparison website's functionality is the implementation of web scraping techniques using Puppeteer, a potent Node.js library.

Puppeteer empowers us to automate the extraction of product data from various online retailers, encompassing prices, descriptions, reviews, and availability status. By automating these web interactions, we not only ensure timely and accurate data collection but also alleviate the manual effort traditionally associated with such tasks.

In essence, our research endeavors to bridge the gap between consumers and the wealth of product offerings available online. Through the synergistic integration of frontend development, backend scripting, and data scraping technologies, our aim is to empower users with an intuitive and comprehensive platform for navigating the intricacies of e-commerce pricing. Through this endeavor, we seek to contribute to the advancement of both technological innovation and consumer empowerment within the e-commerce landscape.

©IJRASET: All Rights are Reserved

II. MOTIVATION

The explosion of online commerce has made it much easier for buyers to find products; however, the issue of comparing prices, discounts, and product reviews is rather daunting. It is difficult for consumers to know whether they are getting a good deal with so many options and offers available.

This project aims at making shopping easy for everyone. Therefore, we want to build a website that not only compares prices but also collects information on discounts, product reviews and availability from various e-shops around the world in order to save time for the customers.

What inspires us are recent technological advancements towards solving real-life problems. Our plan is to merge front-end and back-end development methodologies with sophisticated data scraping tools so as to create an ecommerce platform that makes online shopping stress-free. Eventually, our intention is to provide customers with necessary information to facilitate effective purchasing by doing away with unnecessary purchases or hustles in finding best deals.

III. OBJECTIVE

1) Designing and implementing a user-friendly frontend interface using ReactJS, prioritizing modular and reusable components to improve user experience.
2) This will involve developing a strong server backend using express.js as well as node.js so that it can facilitate effective data exchange between the front-end interface and database.
3) Advanced web scraping techniques have been integrated through the use of Puppeteer which enables automation for such information like price, discounts, product reviews and availability from various e-commerce platforms.
4) It can best be done through the enablement of user comparisons of prices, discounts and product reviews among online retailers.
5) Consequently, there is a need to disseminate research findings concerning web scraping technology's progressions including methodologies in e-commerce industry.

IV. METHODOLOGY

A. *Web scraping with Puppeteer*

We use Puppeteer, which is a Node.js library for parsing data from e-commerce websites, to ensure web scraping functionalities. Puppeteer performs interaction programs with precision and a high degree of scalability, meaning the program can automate navigation and data scraping on various websites.

We also craft scripts to pull data from other websites. Our scripts seek product information from online stores and pull in a wide range of information, including prices, discounts, reviews and availability. With these scripts, we aim to make the data extraction process smoother and more vivid. By comparing this information, our price comparison website can display the state of products from different online stores in detail.

B. *Data Aggregation and Analysis*

Data collection and API integration algorithms play an important role in the performance of price comparison web applications. These algorithms have the ability to ableto handle dynamic processes and multiple data types, each with its own structure and format.. The data collection process begins with identifying sources, which are typically e-commerce websites, inventory, and other online businesses. These algorithms are responsible for collecting information from various sources and integrating this information into a unified and user-friendly interface

C. *Frontend Development with ReactJS*

The frontend interface, the use of ReactJS, emphasizing its factor-based structure for more suitable scalability.Focus on modularity and reusability of components to make certain flexibility and maintainability of the interface layout.Employ Bootstrap for responsive design and format components to guarantee seamless person experience throughout numerous devices.

D. *Backend Development with Node.js and Express.js*

A robust backend server utilizing Node.js as the primary runtime environment.Express.js, a lightweight web framework for Node.js, to manage routing and middleware for efficient data transmission. Integrate relevant Node.js libraries to augment backend functionalities and optimize server performance.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

V. SYSTEM ARCHITECTURE

A. System Architecture

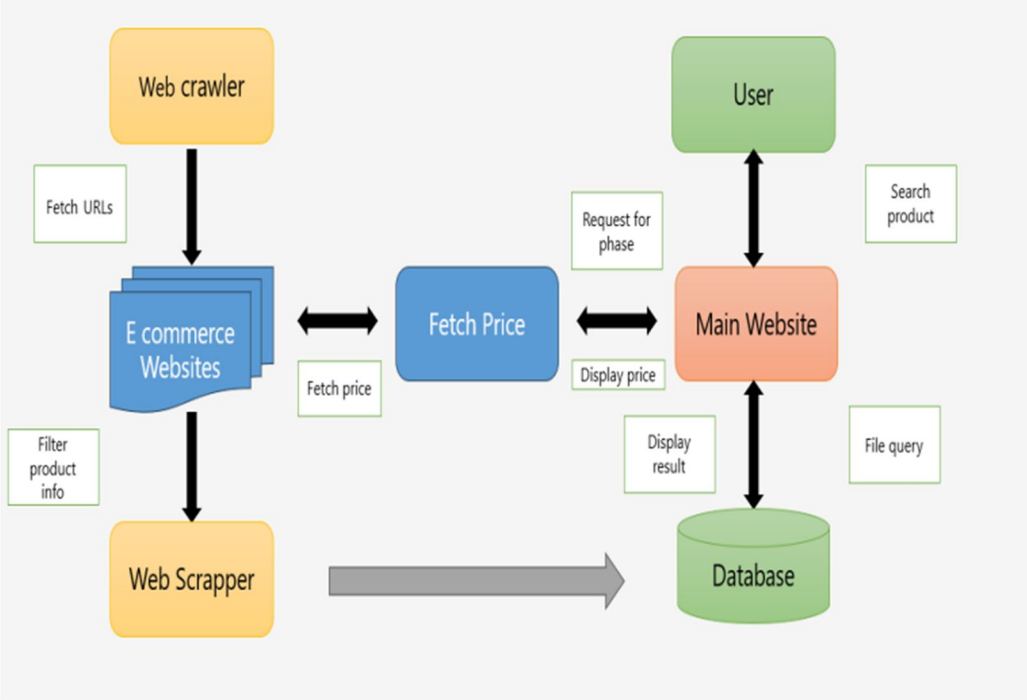
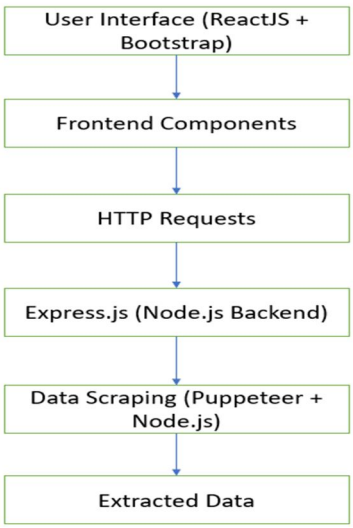


Figure 1 System architecture

This diagram shows the layout and workflow of our system. The front of our system has a user-friendly diagram, which can be accessed through the website. This interface provides seamless interaction among users." between the function and the system Leveraging advanced web crawling and scraping techniques " Considering we extract relevant information from e-commerce websites and then the extracted information is securely stored for future reference. Users can use our website to search and request specific products. When our system receives a request for features, it quickly searches local databases for desired features. Then, the site offers complete product descriptions, allowing users to conveniently compare prices across multiple e-commerce platforms. Specifically, our system simplifies the shopping experience by providing a centralized location for users to browse, browse and compare prices on e-commerce websites.

B. Flow chart

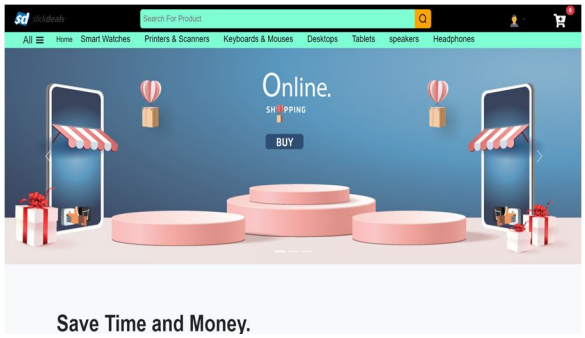


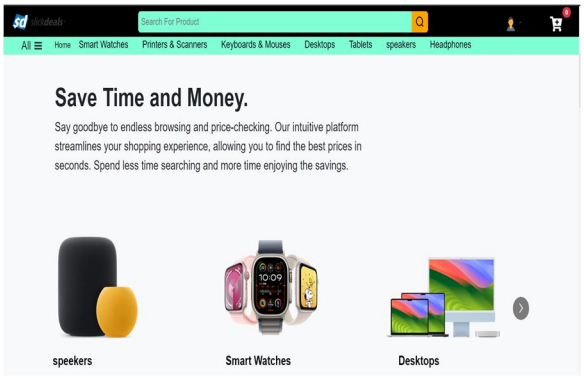
Starting from the Starting layer, the "User Interface" includes frontends constructed in ReactJS and styled with Bootstrap. These capabilities manipulate visual factors and facilitate person interaction. Directly under "Frontend Components" are modules dedicated in particular to UI rendering functions. These objects speak internally and externally thru HTTP requests. The "HTTP Requests" layer that manages the communication among the the front-give up and the again-stop configures the statistics alternate. This entails sending requests from the front quit to the backend server to get hold of or ship statistics. The backend actions are driven by means of "Express.Js", built at the Node.Js framework. It receives incoming HTTP requests from the front, procedures them, and responds consequently. Sitting within the history, the "Middleware" layer acts as an middleman between routing and middle application logic. Middleware services intercept and preprocess requests earlier than they attain the path layer. The "routing" layer directs incoming requests to the relevant operators or controllers inside the backend utility primarily based at the asked endpoint. The "Data Scraping" module at the bottom of the stack makes use of Puppeteer and Node.Js to extract facts from outside resources. This technique can involve automatically enabling network connectivity to retrieve data from networks.Finally, "Extracted Data" represents the statistics acquired via the scraping system. This information can be similarly processed, stored, or presented to customers through an interface.

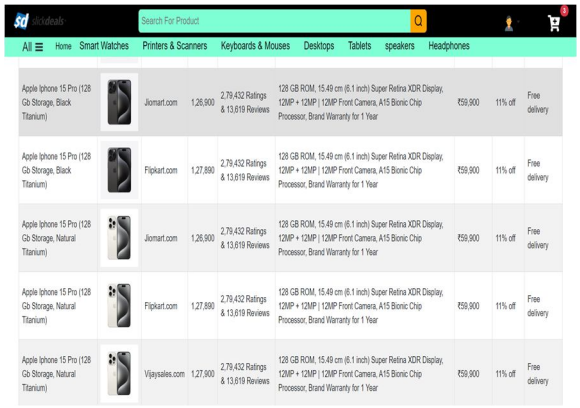
VI. APPLICATIONS

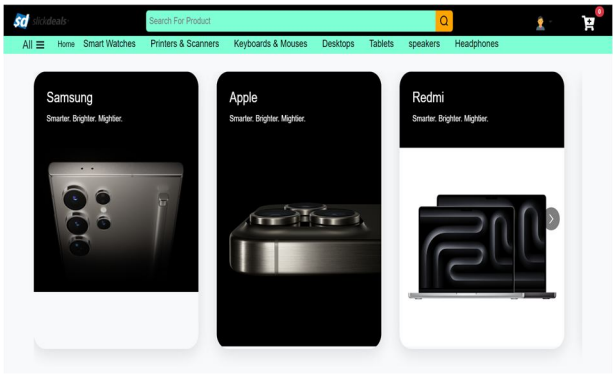
1) *Dynamic Pricing Strategies:* Retailers can make use of real-time fee assessment facts to alter their pricing techniques dynamically. By monitoring competitor fees and marketplace traits in real-time, retailers can optimize their costs to remain aggressive and maximize earnings.
2) *Personalized Recommendations:* Price comparison structures can leverage actual-time information to provide customized product pointers to users based totally on their surfing records, possibilities, and buy conduct. This complements the person experience and increases the likelihood of conversion.
3) *Inventory Management:* E-trade companies can use actual-time price comparison records to inform their inventory control selections. By monitoring product demand and competitor pricing fluctuations in real-time, groups can optimize their inventory ranges and keep away from stockouts or overstock conditions.
4) *Fraud Detection:* Real-time information evaluation on price evaluation platforms can help detect fraudulent activities which include rate gouging or faux discounts. Advanced algorithms can flag suspicious pricing styles and alert platform directors to take important actions to protect consumers.

VII. RESULTS









However, thanks to the web scraping strategies, the desktop development of our website turned to be very successful. We utilize HTML, CSS and React.js front-end technologies to provide a comfortable and smooth navigation of the website. It allows the user to compare products from different websites. Back-end refers to Node.js and Express.js APIs that allow endpoint and data processing functions to operate. The website also adds an effective Puppeteer scraping, which warrants that it will become easy to handle and scrape e-commers product details. All in all, due to our website development, shoppers can choose a better product with a proper price.

VIII. FUTURE SCOPE

1) *Real-Time Notifications:* The platform could implement real-time notifications to alert users about price drops, special promotions, or product availability changes for items on their watchlist. This would help users stay informed and capitalize on time-sensitive deals.
2) *Continuous Performance Optimization:* Regular updates and performance optimizations are essential to ensure that the platform remains fast, reliable, and scalable as user traffic grows. This includes optimizing code, improving server infrastructure, and implementing caching mechanisms to reduce load times.
3) *Enhanced Personalization:* Future updates could focus on enhancing the platform's personalization capabilities, providing users with tailored product recommendations based on their individual preferences, browsing history, and purchase behavior. This could involve the implementation of advanced machine learning algorithms to analyze user data and deliver more relevant content.

IX. CONCLUSIONS

In conclusion, the improvement of a complete fee assessment platform utilising advanced technologies and methodologies holds gigantic capability for revolutionizing the e-trade landscape. Through the combination of frontend development with ReactJS, backend scripting with Node.Js and Express.Js, and complicated internet scraping strategies the usage of Puppeteer, we have effectively created a platform that empowers users with the capability to make informed buying selections. The platform's applications extend past mere price comparison, presenting users enhanced convenience, time financial savings, and personalized shopping reports. By aggregating and reading product facts from numerous on line sources, the platform facilitates seamless contrast of prices, discounts, and critiques across more than one stores, in the long run empowering customers to find the first-class offers effortlessly.

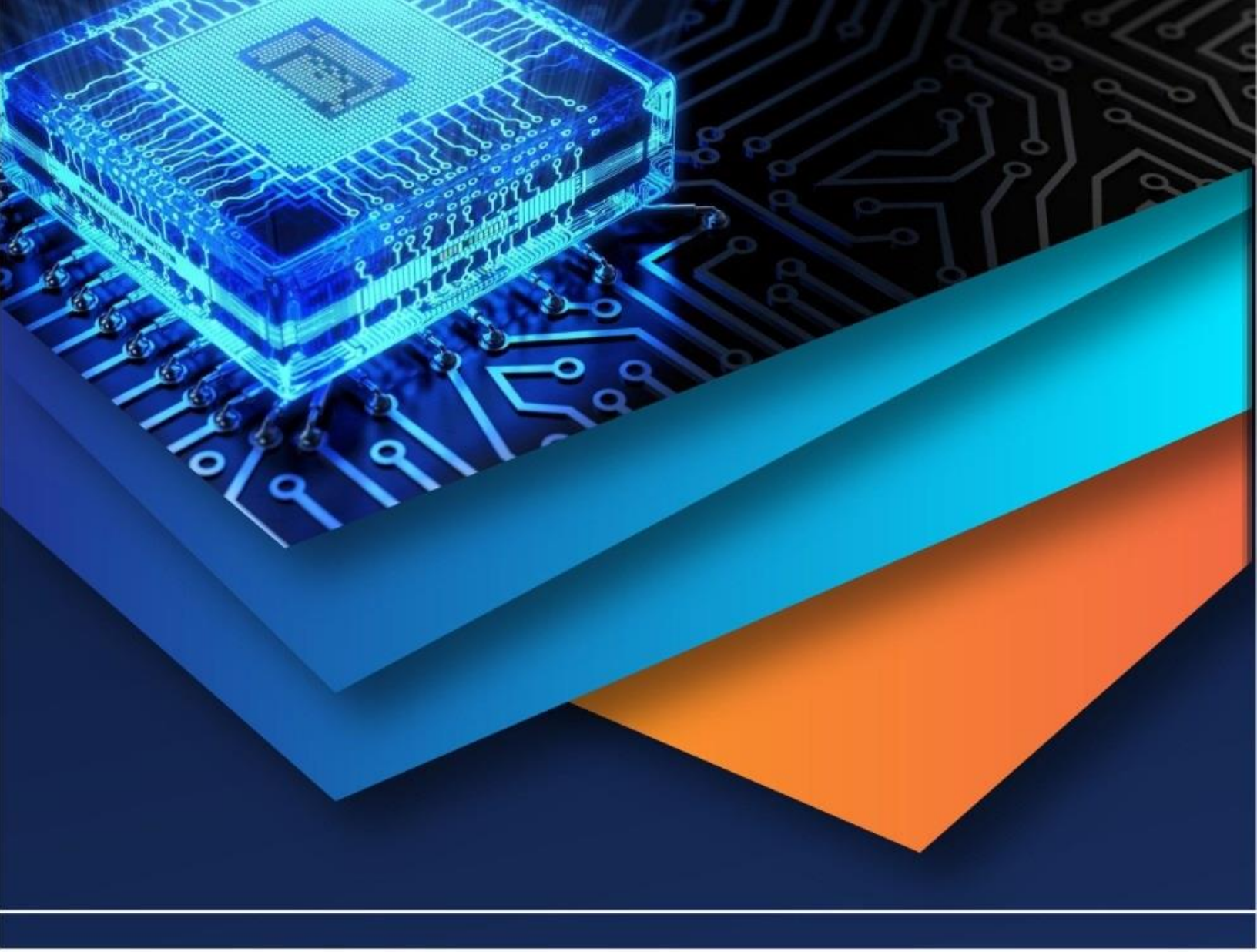
©IJRASET: All Rights are Reserved

REFERENCES

[1] ReactJS: A Comprehensive Analysis of its features, Performance, and Suitability for Modern Web Development Devanshi Bhatt1 , Kunal Parekh2 [2023]
[2] Visualization of Optimal Product Pricing using ECommerce Data N Greeshma, C Raghavendra, K Rajendra Prasad[2019]
[3] Legality and Ethics of Web Scraping; Twenty-fourth Americas Conference on Information Systems, New Orleans, [2018].
[4] Exploiting Filtering approach with Web Scrapping for Smart Online Shopping : Penny Wise: A wise Tool for Online Shopping[2019].
[5] Analysis of E-Commerce and M-Commerce: Advantages, Limitations and Security issues Niranjanamurthy M 1 , Kavyashree N 2, MSRIT, Bangalore, INDIA1
[6] Node.js Challenges in Implementation[2017].
[7] Front-End Development in React: An OverviewSongtaoChen1,UpendarRaoThaduri [2019].
[8] Web scraping technologies in an API world,Daniel Glez-Pen‹a, Ana¤lia Lourenc.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)