

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 3, Issue 2, July 2023

News Aggregation using Web Scraping News Portals

Mr. Mayur Bhujbal¹, Ms. Bhakti Bibawanekar², Dr. Pratibha Deshmukh³

Bharati Vidyapeeth's Institute of Management and Information Technology, Navi Mumbai, India mayurbhujbal3@gmail.com, bibawanekarbhakti@gmail.com, pratibha.deshmukh@bharatividyapeeth.edu

Abstract: In today's fast-paced world, staying informed through reliable and high-quality news is important. However, with an overwhelming number of online publishers and platforms, accessing information from multiple sources can be time-consuming and challenging. To simplify this process, news aggregators have emerged as a solution. These platforms allow users to customize their news consumption by selecting preferred websites and receiving curated articles from those sources in one centralized location. This saves valuable time and effort but also streamlines information gathering in our daily routines. To develop an influential online news aggregator, web scraping for data structuring is crucial. Web scraping involves analyzing a website's HTML structure to extract the desired data. By understanding the underlying structure of web pages, developers can fetch relevant information such as article titles, summaries, authors, and publication dates. People often struggle to find valuable news sources amid the vast sea of information available online. News aggregators address this issue by providing personalized news feeds based on individual interests. Users serve as valuable resources for knowledge tailored to specific preferences. While news aggregators are widely recognized for their popularity and utility, there is still room for improvement in their role in software development. Enhancements could include refining algorithms for more accurate content curation. In addition, Users could include implementing better user interfaces for enhanced user experiences and exploring innovative features like sentiment analysis and topic clustering. By continually iterating and improving upon existing news aggregator platforms, developers can ensure users have access to timely, relevant, and trustworthy news conveniently and efficiently.

Keywords: Web scraping, Online news portals, Information gathering, News Aggregation

I. INTRODUCTION

A web application that aggregates data from multiple websites and presents it in one place is an essential tool for staying informed in today's fast-paced world. With the abundance of magazines and news sites distributing their content across various platforms, the task of visiting 10-20 different sites every day to gather information becomes time-consuming. However, a news aggregator simplifies this process by allowing users to select the websites Users want to follow. The aggregator collects articles from these sources, enabling users to access information from multiple websites with just a click. This not only saves time but also provides users with a competitive advantage by granting them valuable knowledge that others may lack.

Web scraping, also known as data scraping, is a technique used to extract content and information from the internet. It involves retrieving data from websites and storing it in a local file, which can then be manipulated and analysed as needed. While copying and pasting content into an Excel spreadsheet can be considered a form of web scraping on a small scale, the term typically refers to software applications or bots that automate this process. These bots can visit websites, navigate through relevant pages, and extract valuable data efficiently. Web scraping has become increasingly valuable in the digital age due to the abundance of constantly updating and dynamic big data. It finds numerous applications, particularly in the field of data analytics. Market research companies, for instance, use scrapers to extract data from social media or online forums for tasks like customer sentiment analysis. Similarly, companies leverage web scraping to gather data from e- commerce platforms like Amazon or eBay for competitor analysis. Contact scraping is another common application where companies extract contact information from the web for marketing purposes.

Copyright to IJARSCT www.ijarsct.co.in DOI: 10.48175/IJARSCT-12138





International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 3, Issue 2, July 2023

II. LITERATURE SURVEY

News One is a dedicated platform that collects all the latest news updates from multiple national and international sources and summarizes them in short and clear language. This online platform provides service-oriented interactions between users across the web. The main slogan of this application is quick access to news. Get your news straight to you without wasting time searching or loading news. News One uses web scraping/crawling techniques to extract content from various news websites. The main concept deals with news feeds, a manager or sub-directory adds all news URLs to the database. Crawlers retrieve content from RSS feeds at cached URLs. It uses bots that dynamically extract content at regular intervals. This paper describes a classification model that references URLs and extracts useful information for classifying documents into categories. Users can have a flexible experience with this application. [1]

The web is now so rich that it has become a leading collection of fact-checkers. These fact-seeking journalists manually scan dozens of sources every day for relevant information. But in the field of data journalism, where data is constantly growing in real time, this manual research becomes impossible and increasingly labor intensive. One of the techniques used in practice to solve this problem is the automatic extraction of information on web pages, which is known as "web scraping". The main purpose of web scraping is to enter the web page. The main goal of web scraping is to extract web pages, specific and highly structured data with less human effort. This paper presents an automated extraction of articles and journalistic claims implemented in 15 news websites in Senegal. [2]

Increasing information in the form of news is one of the main problems of any society. This data is in an unstructured format and is difficult to manage manually and use effectively. In this study, aimed to build a classifier that classifies news articles from online Senegalese websites based on their content. Use web crawling technology to collect news articles from Senegalese online media on Politics, Religion, Justice, Health, Science, Truth Seeking. Use the Ova method to transform the multi-label problem into a binary classification problem and classify online news articles in this way. Various supervised machine learning algorithms measure the performance of selected models using the area under the ROK curve (AUC). [3]

In recent years and especially during the pandemic, everything is online. All connected through the internet. With this rapid growth, cybercrimes are also increasing. The government is trying to resolve the situation, but it is not enough. India suffered over 1,000,000 victims. 1.25 trillion (\$16 billion) due to cybercrime. Online news is a reliable source of information that is always up-to-date and available for free. In this paper, review previous work on web hacking. Web Crawling is a way of gathering information from web pages. With this knowledge, Author created a web scraping system that collects news articles related to cybercrime. From this data, Users can categorize crimes by region and time. This can help law enforcement agencies and create awareness among the general public.[4]

Instagram is the most popular social media and is also used as a medium for advertising and product marketing. One of the most effective ways to sell products is through endorsements. However, most Instagram accounts now have fake followers or accounts, often called fake followers. Based on this problem, Instagram needs a decision support system to select recommended accounts. In this study, the researcher used a simple attachment weighting (SAW) method for web extraction to process parameters and automatically extract data from Instagram accounts. Instagram account parameters used in this research include followers, likes, followers, and posts that are updated regularly. As the results show with 75% accuracy of the system, it is the result of determining the optimal parameters to choose the Endorse account on Instagram. This proves that the selected parameters are the best parameters for selecting Instagram accounts for verification.[5]

It uses Machine Learning on images and creates a feature vector that maps a set of objects to numbers. This technology is used by organizations such as Google and Facebook to create digital profiles for their users. This project proposes to use this technology to detect criminals who have escaped from their previous records. The NCRB (National Crime Records Bureau) report shows that 70% of crimes are committed repeatedly by the same criminals. These criminals can be identified by referring to photo or video frames captured by cameras installed in various places and can also be used to trace missing children. The disadvantage is that the images are often blurred, unclear and unrecognizable to the human eye.[6]

This article explains stock forecasting using Machine Learning. Input parameters include Open, High, Low, Close, Trade Volume, Earnings Ratio, MA, MACD for more accuracy. A Machine Learning algorithm, Random Forest Regression, implemented in the Python programming language is used to predict the stock market. This algorithm is used

Copyright to IJARSCT www.ijarsct.co.in DOI: 10.48175/IJARSCT-12138





International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 3, Issue 2, July 2023

in conjunction with web search techniques used to obtain current market data for stock on historical stock data. An iterative learning model takes forecasted prices as input to predict future long-term stock prices.[7]

Web scraping and semantic scripts play an important role. Web mining allows searching for relevant results from the web and is used to extract meaningful information from search patterns stored on the server. Web usage mining is a form of web mining that extracts information about the access patterns/behavior of users visiting a website. Web Crawling, another technique, is the process of extracting useful information from HTML pages, which can be done using a Prolog-based scripting language known as Prolog Server Pages (PSP).[8]

Information extraction is a framework used to extract structured insights from text. Among the mining paradigms, event extraction is a technique that answers certain questions (who did what, where, and when). Such responses can be used for the decision-making process. Primary research focuses on extracting the frequency and location of events, including data from news headlines produced by online news channels. Instead, temporal frequency is derived from online news headlines and events related to crime and natural disasters, such as earthquakes. For this purpose, data was collected by scraping two popular news websites. [9]

Data accuracy is an increasingly important concern on the Internet, especially on social media, but web-scale data hampers the ability to identify, evaluate, and correct information on these platforms, or so-called "fake news." In this article, Author suggest how to spot "fake news" and how to use it on Facebook, one of the most popular online social media platforms. This method uses a Naïve Bayes classification model to predict whether a Facebook post will be labeled as REAL or FAKE. The results can be improved by using some of the techniques discussed in the paper. The results obtained show that the problem of detecting fake news can be solved by machine learning methods.[10]

Users are challenged to extract and reuse information from websites by copying or hacking content from websites. Another way is to write a script that retrieves data via API. This allows you to access your data in a cleaner way than hacking. However, APIs are difficult for programmers and almost impossible for non-programmers. This effort allows users to access the API without programming. Create a schema to define how to interact with the data API. Next, develop ScrAPIr. ScrAPIr is a standard user GUI that allows users to get information from existing API specifications and a second GUI that allows users to create specifications for specific APIs. and sharing. Lab tests show that non-programmers can access their APIs using ScrAPIr, and programmers can access their APIs using ScrAPIr on average 3.8 times faster than programmers. [11]

Information is everywhere on the Internet. Searching the web for useful information and information has become a normal activity. Information on the page is stored in tables, articles, comments, various HTML tags, etc. you may find Gathering large amounts of information from the web is not an easy task, but it is a great way to gather as much information as you can. used in further analysis. In this post, it will deal with crunching data from various places on the internet and storing it in a database in order to collect and analyze data in the used car market. [12]

With the rapid proliferation of online shopping websites, people often face difficulties in finding the best product deals and offers. This article presents a web search and search method to identify the best deals from five e-commerce websites. The framework uses HTML and CSS for the front-end and PHP for the back-end. Python library is used for coding and web scripting based on HTML script. In particular, the framework dynamically retrieves and displays results without storing them in a local database, improving storage and processing capabilities. The data mining process achieves 93% accuracy with minimal computational and time requirements.[13]

When the number of available data sources is large and unlimited, the process of finding data becomes very complex. Visiting data sources one by one and comparing data or information from all data sources visited will add a lot of time to the data recovery process. Users need a method that can collect data from multiple sources to facilitate the data retrieval process. This study uses 3 e-commerce websites as data sources. Using the lookup procedure, a new variable will be created that can hold data from the data source, and then the data will be stored in the database. Web browsers work by fetching and parsing HTML tags as needed. Moreover, the data retrieval process can be done easily by collecting data in the database and using a query, this process can be done in less time with 100% recall and 93.9% accuracy level. [14]

Information drives modern business, and the Internet is a data center. Thus, data integration provides users with all different data sources in a unified way. This paper presents a model that provides automated support for multi-domain queries. This model (a) integrates various services into the global ontology (b) formulates the requirements in the global

Copyright to IJARSCT www.ijarsct.co.in DOI: 10.48175/IJARSCT-12138





International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 3, Issue 2, July 2023

ontology and rewrites the requirements in terms of local services (c) provides several web-related services that help deal with XAMPP Dashboard problems.[15]

A stock tape reading monitors stock prices at one point in time and records stock prices. One way to control stock prices is to control the exchange. The time required to record stock quotes varies from three minutes to two hours depending on the experience of each user. The Taiwanese market is open from 09:00 to 13:30. How to divide four hours of work time into reading tapes, the mastery of stock users. Stock price order created by reading tapes can be predicted by stock users. The purpose of reading tapes is to record stock prices, but not to predict them. This study proposes a method for reading tapes using web scraping. This method periodically cancels stock prices and creates a sequence of those stocks in MS Excel format. This application meets the demand of stock users. Since stock users want to get the consistency of the stock price from day to day rather than the format of the stock, it is more intuitive to predict the consistency of the stock price. This search method implements the previous scheme and creates a GUI to query stock prices and order stocks of various companies in seconds. Periodic updates of stock prices and timeshare prices. [16]

III. OBJECTIVES OF THE STUDY

- Understanding web scraping and its application to obtaining news data from various news portals. .
- To identify key elements to scrape from different news portals based on their structure and content. In addition . to headlines, summaries, and authors, publication dates are also included in these elements.
- Automating news article extraction from multiple news portals using web scraping techniques and tools.
- To evaluate and compare different scraping methods and technologies for their efficiency, accuracy, and . scalability in aggregating news data.
- To explore data processing and integration techniques to organize and present scraped news data in a coherent • and user-friendly manner on an aggregated news platform.





Firstly, the URL of the webpage is required to be accessed to fetch the content. URLs were stored in a python dictionary. In order to fetch the content of a webpage for web scraping, the first step is to obtain the URL of the webpage. In this specific case, the URLs are stored in a Python dictionary. A dictionary is a data structure in Python that allows for storing key-value pairs. Each URL is associated with a specific key in the dictionary, which can be used to access the URL when required. By storing the URLs in a dictionary, it becomes easier to manage and retrieve the appropriate URL for scraping the desired webpage.

For getting the HTML structure of the website, HTTP connection has to be established to the web server. Establishing an HTTP connection to the web server is essential for obtaining the HTML structure of a website during web scraping. This connection allows the client (web scraper) to send a request to the server and retrieve the HTML content of the web page. The HTTP connection is established using the appropriate protocol (HTTP or HTTPS) and involves sending a request header to the server, which includes information such as the desired webpage URL, method (GET or POST), and any additional parameters.

Web pages of news articles typically have a standardized HTML structure that can be accessed using Python libraries. By utilizing these libraries, such as BeautifulSoup or lxml, it becomes possible to extract the required data from the HTML structure. This includes retrieving specific elements like article titles, publication dates, authors, and content. Copyright to IJARSCT DOI: 10.48175/IJARSCT-12138 278

www.ijarsct.co.in

ISSN 2581-9429 IJARSCT



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 3, Issue 2, July 2023

With the help of these Python libraries, web scraping techniques can be applied effectively to navigate and extract targeted information from news article web pages.

Python is a popular, high-level programming language known for its readability and versatility. It supports various programming paradigms and incorporates automatic garbage collection. Python's extensive standard library makes it a "batteries-included" language. Beautiful Soup is a Python package used for parsing HTML and XML documents and facilitating web scraping tasks. It generates a parse tree for extracting data from imperfect markup. Requests is an essential library for sending HTTP requests in Python, crucial for working with REST APIs and web scraping. The dateutil package enhances the DateTime module in Python, providing features like relative delta calculations and accurate time zone conversions. HTML is the standard markup language for creating web documents, often complemented by CSS to define the visual appearance and layout. With HTML and JavaScript, CSS shapes the aesthetics of web pages.

V. PROCESS OF NEWS AGGREGATION USING WEB SCRAPING NEWS PORTALS INPUT

Our system operates in two stages. Initially, gather user preferences, followed by data collection from well-known news portals such as Times of India, NDTV News, India Today, The Indian Express, The Hindu, News 18, First Post, and Business Standard. The personalized news is then presented on the user's news feed. The user selects the desired news portals and categories from the available options, and subsequently receives customized news content on their news feed.



DOI: 10.48175/IJARSCT-12138

Copyright to IJARSCT www.ijarsct.co.in





International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 3, Issue 2, July 2023

Collecting the information about tags and which contains information needed to display on a website. Example: Image URL, News heading and Detailed News link.

<img alt="'Destroyed Mentally': Rafael Nadal After Early Australian Open Exit" title="'Destroyed Mentally': Rafael Nadal
n Exit" class="imgbrd" src="https://c.ndtvimg.com/2023-01/iqkjft rafael-nadal 625x300 18 January 23.jpg?downsize=105:79"
shadow-region-image="0"> == \$0

<4 href="https://sports.ndtv.com/australian-open-2023/destroyed-mentally-rafael_al-reacts-to-early-australian-open-exit-3702773#pfrom=home-ndtv toos
tories" onclick="javascript;GAP_Event_tracking("top-stories",""Destroyed Mentally": Rafael Nadal After Early Australian Open Exit");" class="item-ti
tle" ping="/rn-events.taboola.com/newsroom/1.0/ndtv/notify-click?view.id=427754201140270126&page.template=news-home&page.dashboard=news-home&pa

Installation of Needed Libraries of Beautiful soap and requests.



Install the Django using following command

pip install django-import-export

Process:

News aggregators scrape the data from different news portals in different categories and store them into a database. The objective of web scraping is to scrape the data from identified websites and convert it into a form which can be stored into traditional databases.

A news aggregator works in three phases, it scrapes the web for the news articles. Then it stores the image, link, and title of the article in the database. The stored objects in the database are served to the news feed. The client gets information on his news feed.



News scraping is a form of web scraping that focuses on extracting information from public online media websites. It involves automatically retrieving news updates and releases from news articles and websites. Beautiful Soup, a versatile module, is ideal for parsing and traversing HTML code. It enables efficient targeting of various elements such as div, table, td, tr, class, id, and more.



Django (sometimes stylized as django) is a free and open-source, Python-based web framework that follows the model-template-views (MTV) architectural pattern.

Import the libraries which are going to need to display our data on the website.

Copyright to IJARSCT www.ijarsct.co.in DOI: 10.48175/IJARSCT-12138





International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 3, Issue 2, July 2023

import app *as* app from django.shortcuts import render

Python Model classifies the news using some of the popular.

This model finds these keywords in those headlines or bulletins and shows the types of news on the web application. The user chooses the topics according to their preferences on the web application and the model shows the news according to their preferences.



Fig 1. Details of process

News scraping involves extracting news articles and data from a news website and displaying it on a client's website. The process includes identifying the target website, accessing its HTML structure using tools like Requests and BeautifulSoup, parsing the HTML to extract relevant information, cleaning and formatting the data, storing it in a structured format, and integrating it into the client's website. Automation enables clients to keep their sites updated with the latest news, improving user engagement. However, it's crucial to respect website terms, adhere to ethical guidelines, and respect copyright laws when scraping news websites. (See Fig.1)

VI. CLIENT APPLICATION



Fig 2. Functions of News Aggregation Using Web Scraping News Portal

Copyright to IJARSCT www.ijarsct.co.in

DOI: 10.48175/IJARSCT-12138





International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 3, Issue 2, July 2023

VII. RESULT/OUTPUT

Output from Beautiful soap for particular news

title="Bajrang Punia accuses WFI president of abusing, hitting wrestlers" href="/sports/other-sports/story/bajrang-punia-accuses-wfi-president-of-abusing-hitting-wrestlers-2323206-2023-01-18" >Bajrang Punia accuses WFI president of abusing, hitting wrestlers</a

title="Bajrang Punia accuses WFI president of abusing, hitting wrestlers"

src="https://akm-img-a-in.tosshub.com/indiatoday/images/story/202301/bajrang_punia_pti-sixteen_nine.jpeg

In this web application, users can see news with its types. Aggregate the news according to news portal. Example: India Today



Bajrang Punia accuses WFI president of abusing, hitting wrestlers



Tripura Assembly election to be held on February 16, results on March 2



From Bharat Jodo to BJP chalo: Punjab Congress leader quits party, joins saffron camp

VIII. CONCLUSION

People often spend a significant amount of time trying to find valuable news sources. This issue can be resolved through the use of news aggregators, which provide news tailored to individual interests in one convenient location. While it is widely acknowledged that news aggregators are popular and valuable resources for obtaining knowledge, there is a lack of insight into their role in software development and potential areas for improvement.

In conclusion, employing web scraping technology for news aggregation from multiple news portals has proven to be an efficient and effective method for collecting and organizing news content. By utilizing automation, web scraping allows for the retrieval of real-time information from various sources, thereby offering users a diverse range of news articles in a consolidated format.

Web scraping empowers news aggregators to simultaneously gather data from different news portals, eliminating the need for manual browsing and searching. This automated approach saves users time and effort, while also ensuring users have access to a broad spectrum of news topics, sources, and perspectives. As a result, users can develop a more comprehensive understanding of current events.

Moreover, web scraping technology enables news aggregators to deliver timely updates and breaking news to their users. By continuously monitoring and extracting data from news portals, aggregators can promptly provide the latest headlines and articles, ensuring users remain well-informed and up to date with the rapidly evolving news landscape.

REFERENCES

- [1]. K. Sundaramoorthy, R. Durga and S. Nagadarshini, "NewsOne An Aggregation System for News Using Web Scraping Method," 2017 International Conference on Technical Advancements in Computers and Communications (ICTACC), 2017, pp. 136-140, doi: 10.1109/ICTACC.2017.43.
- [2]. E. N. SARR, O. SALL and A. DIALLO, "FactExtract: Automatic Collection and Aggregation of Articles and Journalistic Factual Claims from Online Newspaper," 2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), 2018, pp. 336-341, doi: 10.1109/SNAMS.2018.8554421.

Copyright to IJARSCT www.ijarsct.co.in

DOI: 10.48175/IJARSCT-12138





International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 3, Issue 2, July 2023

- [3]. T. Tona Landu et al., "Machine Learning Algorithm for Text Categorization of News Articles from Senegalese Online News Websites," 2022 17th Iberian Conference on Information Systems and Technologies (CISTI), 2022, pp. 1-8, doi: 10.23919/CISTI54924.2022.9820408.
- [4]. Authors, Joel Christian, Sharada Valiveti & Swati Jain, Department of Computer Science and Engineering, Institute of Technology, Nirma University, Chharodi, Ahmedabad, 382481, Gujarat, India
- **[5].** M. I. Akrianto, A. D. Hartanto and A. Priadana, "The Best Parameters to Select Instagram Account for Endorsement using Web Scraping," 2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia, 2019.
- [6]. S. Ayyappan and S. Matilda, "Criminals and Missing Children Identification Using Face Recognition and Web Scrapping," 2020 International Conference on System, Computation, Automation and Networking (ICSCAN), 2020, pp. 1- 5, doi: 10.1109/ICSCAN49426.2020.9262390.
- [7]. B. B. P. Maurya, A. Ray, A. Upadhyay, B. Gour and A. U. Khan, "Recursive Stock Price Prediction with Machine Learning and Web Scrapping for Specified Time Period," 2019 Sixteenth International Conference on Wireless and Optical Communication Networks (WOCN), 2019, pp. 1-3, doi: 10.1109/WOCN45266.2019.8995080.
- [8]. S. K. Malik and S. Rizvi, "Information Extraction Using Web Usage Mining, Web Scrapping and Semantic Annotation," 2011 International Conference on Computational Intelligence and Communication Networks, 2011, pp. 465-469, doi: 10.1109/CICN.2011.97.
- [9]. A.-U. -N. Fatima, H. Ahmad, M. Ahmad, W. Ahmad and N. Faisal, "Extraction of Temporal Events' Frequency from Online News Channels," 2020 30th International Conference on Computer Theory and Applications (ICCTA), 2020, pp. 109-116, doi: 10.1109/ICCTA52020.2020.9477659.
- [10]. Jain and A. Kasbe, "Fake News Detection," 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), 2018, pp. 1-5, doi: 10.1109/SCEECS.2018.8546944.
- [11]. Alrashed T, Almahmoud J, Zhang AX, Karger DR (2020) Scrapir: making web data apis accessible to end users. In: Proceedings of the 2020 CHI conference on human factors in computing systems.
- [12]. D. Petrović and I. Stanišević, "Web scrapping and storing data in a database, a case study of the used cars market," 2017 25th Telecommunication Forum (TELFOR), 2017, pp. 1-4, doi: 10.1109/TELFOR.2017.8249451.
- [13]. S. Mehak, R. Zafar, S. Aslam and S. M. Bhatti, "Exploiting Filtering approach with Web Scrapping for Smart Online Shopping: Penny Wise: A wise Tool for Online Shopping," 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, 2019, pp. 1-5, doi: 10.1109/ICOMET.2019.8673399.
- [14]. D. Kurniawati and D. Triawan, "Increased information retrieval capabilities on e-commerce websites using scraping techniques," 2017 International Conference on Sustainable Information Engineering and Technology (SIET), Malang, Indonesia, 2017, pp. 226-229, doi: 10.1109/SIET.2017.8304139.
- [15]. D. Mishra and N. Pujari, "Cross-domain query answering: Using Web scrapper and data integration," 2011 2nd International Conference on Computer and Communication Technology (ICCCT-2011), Allahabad, India, 2011, pp. 27-32, doi: 10.1109/ICCCT.2011.6075193.
- [16]. F. Lin and S. Chih Yang, "Web Scraping to Implement Tape Reading on Taiwan Stock Periodically with GUI," 2021 IEEE 3rd Eurasia Conference on IOT, Communication and Engineering (ECICE), Yunlin, Taiwan, 2021, pp. 112-116, doi: 10.1109/ECICE52819.2021.9645633.
- [17]. Deshmukh, P., Avinash, S., Gonsai, A.M., Sonawane, S.S., Khan, T. (2023). Ordering Services Modelling in Blockchain Platform for Food Supply Chain Management. In: Rathore, V.S., Tavares, J.M.R.S., Piuri, V., Surendiran, B. (eds) Emerging Trends in Expert Applications and Security. ICE-TEAS 2023. Lecture Notes in Networks and Systems, vol 681. Springer, Singapore. https://doi.org/10.1007/978-981-99-1909-3_37
- [18]. G, M., Deshmukh, P., N. L., U. K. , Macedo, V. D. J. , K B, V. , N, A. P. , & Tiwari, A. K. (2023). Resource Allocation Energy Efficient Algorithm for H-CRAN in 5G. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3s), 118–126. https://doi.org/10.17762/ijritcc.v11i3s.6172

Copyright to IJARSCT www.ijarsct.co.in DOI: 10.48175/IJARSCT-12138





International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 3, Issue 2, July 2023

- [19]. K. S. Gill, A. Sharma, V. Anand, R. Gupta and P. Deshmukh, "Influence of Adam Optimizer with Sequential Convolutional Model for Detection of Tuberculosis," 2022 International Conference on Computational Modelling, Simulation and Optimization (ICCMSO), Pathum Thani, Thailand, 2022, pp. 340-344, doi: 10.1109/ICCMSO58359.2022.00072.
- [20]. Avinash, S., Deshmukh, P., Jamsandekar, P., Kumbhar, R.D., Kharade, J., Rajendran, R. (2023). Specifying the Virtual Reality Approach in Mobile Gaming Using Unity Game Engine. In: Rathore, V.S., Piuri, V., Babo, R., Ferreira, M.C. (eds) Emerging Trends in Expert Applications and Security. ICETEAS 2023. Lecture Notes in Networks and Systems, vol 682. Springer, Singapore. https://doi.org/10.1007/978-981-99-1946-8 26
- [21]. Sakhare, Mr & Deshmukh, Pratibha. (2022). Blockchain in Petroleum Industry. International Journal of Advanced Research in Science, Communication and Technology. 733-737. 10.48175/IJARSCT-5751.
- [22]. Prerana Patil and Dr. Pratibha Deshmukh (2022). Case Study of Indian Students Going to Abroad for Medical Studies. International Journal of Scientific Research in Engineering and Management (IJSREM).

