Abstract. "Personalized Job Opportunity Finder powered by Web Scraping" research presents a comprehensive exploration into the development and implementation of an innovative Job Portal, aiming to redefine the traditional job search experience. The project leverages cutting-edge technologies, including web scraping techniques and integration with prominent communication tools. The experimental setup involves meticulous web scraping from renowned job portals, Google and Apple, using Puppeteer and Cheerio for data extraction. The user interface is thoughtfully designed, featuring an intuitive registration form and a dynamic home page that showcases personalized job recommendations. The research delves into the advantages of scraping job data from top companies, showcasing its efficacy compared to traditional methods such as partnerships and direct job postings. Big companies such as Google, Apple and Meta don't use partnerships for most of its hiring, that makes it difficult for candidates who aspire careers at such companies. Direct partnerships may have delays in updating job listings, whereas scraping allows for more immediate access to new postings.

Keywords. Web Application, Web Scraping, Job Portal, Apply Link, Dashboard, Profile, Puppeteer, Cheerio, Recommendation, Authentication.

1 Introduction

Traditional job search techniques frequently fail to offer each candidate options that are relevant and customized in the modern employment market. Seeing this difficulty, our initiative unveils a cutting-edge career site with the goal of completely changing the employment search process. Using state of the art technologies web scraping and machine learning in particular we aim to build a platform that adapts to the distinct profiles and favourite things of every user.

It may be difficult to find a job, especially for people switching careers, returning to the workforce after a break, or first-time applicants. The goal of this project is to make this difficult undertaking a methodical, approachable, and user-friendly procedure for applicants and companies alike. The objective is to streamline the sometimes laborious process of
locating an appropriate employment portal, irrespective of the job seeker's or the company's geographic location.

With the use of an online web application, our project, "Personalized Job opportunity Finder powered by Web Scraping" acts as a web scraper pulling relevant job postings according to the user's profile. It is intended to be easy to use, effective, convenient, and methodical. Candidates may register on the platform, look through opportunities according to their credentials, and apply with ease.

After user registration and login, our web application employs web scraping techniques to fetch job listings from prominent platforms such as Google, Apple and other companies. The backend of the application, powered by Node.js with Express, utilizes Puppeteer, a headless browser automation tool, to navigate and extract relevant job information from the HTML content of job listing pages.

To achieve this, the application sends HTTP requests to the job portal websites, retrieves the HTML content, and then uses Cheerio, a lightweight and fast HTML parsing library for Node.js, to navigate and extract information from the HTML structure. Upon successful scraping, the web application displays the registered users with potential job options. This notification includes essential information about the specific job, allowing users to stay informed about relevant opportunities.

The automated scraping process significantly expedites job searching, replacing manual efforts with efficient, programmatic approaches. The system stores user information in a MongoDB database, facilitating seamless job applications and providing a centralized platform for managing current and future job opportunities.

In conclusion, "Personalized Job opportunity Finder powered by Web Scraping" offers a user-friendly interface, automatic scraping for real-time updates, and customized job suggestions, all of which combine to provide a revolutionary approach to job hunting. This project intends to improve the efficacy and efficiency of the job search process by fusing web scraping and machine learning technology, ultimately matching individuals with openings in a more dynamic and simplified way.

2 Literature Survey

Our research, "Personalized Job opportunity Finder powered by Web Scraping" investigates how web scraping and machine learning might be used to transform user-centric job suggestions in the rapidly changing job search market. This review of the literature covers important research topics such as customized employment portals, the use of web scraping for data retrieval, improvements in recommendation algorithms, and moral issues. We hope to place our effort within the larger framework of recent advances in the area with this succinct review.

Shubham B. Gulik et al.(2021), This paper discusses about a job search web application called "Scraping of Job Portal" allows job searchers to register and submit applications online. Employers must work hard to identify the best applicant for a vacant position, while job seekers must obtain employment through ads, college fairs, job fairs, etc. These flaws are fixed by this program, which also provides job searchers with an easy-to-use platform for job searching and application. Candidates can use sophisticated search methods to look for employment in any field. Basically, their scraper gathers information from many job portals and presents job seekers with possibilities based on their requirements, including
location, pay, working hours, and technical abilities. They have the option to submit their
resumes and other basic job search data, which is kept in a database. Candidates from
anywhere in the globe can utilize their site without facing any geographical restrictions.
Additionally, the registered users of our platform will receive email alerts from us with basic
job information and an application link.

L. Gangadhara Reddy et al. (2022), This paper speaks about the practice of web scraping job
portals offers insight into the most sought-after talents that hiring businesses are looking for,
the industries that provide more employment chances to job searchers, and other elements
that influence the hiring process, such as a candidate's experience. Through a selection of
employment portals, this study determines which areas in India are offering more career
chances (Naukri, Indeed, LinkedIn). The findings showed that selling abilities are the most
in-demand job needs, followed by business management and IT-programming skills (Java,
Python, JDBC, SQL, Data Management, and CSS). Additionally, Bangalore (20%) offers
the greatest job chances out of all the other cities, followed by the IT-Software (41%) and
Banking (8%), Recruitment (8%), and other industries. Additionally, the survey lists the top
20 job names and roles that recruiters offer to job searchers.

Chaimaa Lotfi et al. (2021), This paper speaks about how organizations may now analyze
massive data volumes and obtain fresh insights thanks to big data analytics. It assists in
providing basic answers to inquiries concerning the performance and operations of
businesses. It also aids in the discovery of obscure patterns in sizable datasets or
combinations of them. Big data approaches are becoming more and more important to use
and analyze for organizational success in today's data-driven environment. More precisely,
given the abundance of data available on the Internet from social media to websites, online
portals, and platforms, to mention a few organizations must be able to mine that data in
order to extract valuable insights. In this sense, web scraping is a basic strategy. Therefore,
in order to better provide academics and managers with useful information on how to mine
online data most efficiently, this work intends to present an updated literature review
regarding the most sophisticated Web Scraping methodologies. The first section of the
article presents the fundamental layout of a web scraper as well as the many industries and
fields in which web scraping is used. The various techniques and technologies for Web
scraping are then discussed. The report concludes with a proposal for a technique to create
Web scraping using various technologies.

Vishnu Priya N P et al. (2023), This paper introduces about the online networks known as
job portals help connect job searchers with possible employers. Information from job portals
is frequently gathered via web scraping, which is the process of obtaining data from
websites, in order to obtain insights about the state of the employment market. Software
programs or bots can extract data such as job names, industries, necessary skills, and
locations of job vacancies via web scraping employment sites. By analyzing this data,
important insights into the labour market may be obtained, such as the most in-demand skill
sets, the top hiring industries, and the areas with the most employment prospects. Job
searchers may also obtain helpful information on job titles and roles given by recruiters, as
well as the abilities necessary for a certain job, via web scraping employment portals. Job
searchers may use this information to gain a better understanding of the labour market and
the skills they need to develop in order to become more employable. For example, job
seekers might enroll in classes to enhance their programming abilities and boost their
chances of landing a position in the IT business if web scraping indicates that programming
skills are in high demand. It's crucial to remember, though, that web scraping employment
websites without authorization or in violation of the terms of service of the portal may be
unethical and unlawful. Data scraping may be prohibited by certain employment platforms, and any effort to do so might result in legal action. As such, it is imperative that web scraping of employment websites be carried out in an ethical and responsible way, making sure that any data extraction is carried out within the bounds of the law and with the necessary licenses.

Tariqul Islam et al. (2023), This paper speaks about Picking up valuable data from the internet is the most critical concern for semantic web discernment. This can be accomplished in a few ways in which web mining, web scraping plays an vital part. A wealthy source of genuine and current information is the web, where information are regularly displayed and put away in a structure that needs a few wrangling and change before being prepared for examination. Web mining enables the search for relevant results from the internet and is utilized to extract important data from disclosure designs put away on servers. Web utilization mining could be a sort of web mining that burrows into the root/etiquette data of clients who visit websites. Web scraping, another strategy, recovers valuable data from HTML pages that can be executed employing a scripting dialect known as Introduction Server Pages (PSP). Semantics explanations are strategies that make it possible to combine semantics and a formal structure into unrecognized content records, an vital angle of semantic data recovery that can be performed by a tool known as KIM (knowledge information management). This article clarifies how web scratching works and how it can be implemented in an academically and actually executable way at different levels of insights and information mining educational module. Too, we return to, explore, and discuss a few of the data recovery methodologies on the internet, such as digging for web utilization, web scraping, and outlining with illustrations for better or more proficient data extraction on the web.

Vivek Kumar Sehgal et al. (2013), This paper tells that the gaining knowledge and specialized work skills are now recognized as the primary goals for university students. Making educated decisions requires knowledge, particularly in situations of urgency. Any organization needs knowledge and knowledge management (KM) to provide it a competitive advantage in the demanding and globalised world of today. The authors of this paper have presented an online recruiting system architecture that enables companies to submit job adverts, which job seekers may access during their job search. Based on industry demands, this employment site can record work requirements.

Vidhi Singrodia et al. (2019), This paper speaks about Web grants a wide scope of realities and information source built up by people. In spite of the fact that, it might comprise of an gigantic grouping of disparate and sickly organized information, challenging in collection in a physical implies and problematical for its utilization in mechanical forms. Since the later past, methods along with different outfits have been created to allow information gathering and modification into organized data to be fulfilled by B2C and B2B frameworks. This paper will center on different viewpoints of web scratching, starting with the fundamental presentation and a brief discourse on different software's and devices for web rejecting. We had too clarified the method of web scratching with an elaboration on the different sorts of web scratching procedures and at last concluded with the masters and cons of web scratching and an in detail depiction on the different areas where it can be connected. The openings taking an advantage of these information are various which might incorporate fields concerning Open Government Information, Enormous Information, Commerce Insights, aggregators and comparators advancement of modern applications and mashups among formers.
4 Proposed Approach

In our research paper, we present a holistic approach to building an advanced job portal that prioritizes user experience, efficiency, and security. The platform aims to provide personalized job recommendations by considering individual profiles, including graduation year, course, CGPA, experience, and skills.

A sophisticated matching algorithm is proposed to analyze both job descriptions and user profiles, ensuring precise alignment for optimal job-user fit. The implementation includes a mechanism for daily job updates to offer users the latest opportunities. Additionally, the platform enables users to conveniently track their job applications and allows customization of profiles, preferences, and eligibility criteria for a personalized job search.

Interactive features, such as prompts to confirm job applications, are incorporated to enhance user engagement. Machine learning techniques, coupled with user feedback, are utilized to refine recommendation algorithms continually. The backend, powered by Node.js with Express, employs Puppeteer for web scraping from platforms like Google and Apple, automating the extraction of relevant job information from HTML content.

The automated scraping process expedites job searching, and user data is stored securely in a MongoDB database, providing a centralized platform for seamless job applications and future opportunities. This approach integrates cutting-edge technologies to revolutionize the job search experience, prioritizing user satisfaction, efficiency, and data security. The subsequent sections of the research paper will delve into detailed implementation, results, and the potential impact on the job search landscape.

5 Experimental Setup

This section provides a detailed overview of our Job Portal project's setup, with a focus on the technologies, approaches, and essential elements used to efficiently retrieve job information and facilitate user engagement.

5.1 Implementation of Web Scraping

In order to obtain employment information from Google, Apple, and other job sites, our solution makes use of web scraping techniques. The following are the main elements of our web scraping implementation:

jobScraper.js:

This script is specifically designed to retrieve task data using Cheerio and Puppeteer. Headless browser interactions are automated with Puppeteer, and HTML parsing is made quick with Cheerio. We extract job information, including title, location, and application links, to create the core dataset for our site. You may access more extensive information,
including minimum and desirable credentials, by navigating to the application link for each position.

5.2 User Interaction components

The Job Portal facilitates user-to-user engagement through clearly defined components:

Views /register.html:
The user registration form is organised by this HTML file, which also includes input areas for professional, academic, and personal information. Several input formats, such as text, date, radio buttons, and password, improve how comprehensive user profiles are.

Views /home.ejs:
This EJS file, which represents the user's home page after logging in, dynamically loads job details using job data that has been scraped and the user's profile. Title, address, contact information, and a "Apply Now" button that connects to the appropriate employment application are all displayed on each job card.

5.3 Server Side Logic and Database Management

The Express.js framework is used to encapsulate the server-side code in the app.js file. Important elements consist of:

Express application setup, including middleware configuration for session management. Establishing a link with the MongoDB database to hold user information, such as registration specifics and customized task suggestions. Defining the paths for job data retrieval, registration, user authentication, and home page display.

5.4 Presentation and Style

Our portal's appearance and design are improved by the use of CSS styling in the styles.css file. To ensure a captivating user experience, this styling is used to the main page, navigation bar, and registration form.

5.5 External Dependencies

For smooth operation, the following external dependencies are necessary for our project:
Puppeteer and Cheerio: Needed for effective online scraping of employment data. The web application is built using Express.js and EJS, which are also used to generate dynamic content.

6 Results and Discussion

The "Personalized Job opportunity Finder powered by Web Scraping " initiative has changed the way that people think about job searches in a profound way. We have successfully developed personalized job suggestions with our algorithmic technique, which
incorporates individual variables such as graduation year, course, CGPA, experience, and talents. Users have praised our approach for its increased relevance. The employment pool is diversified by the addition of daily job updates, which guarantee instant access to the newest chances. Application tracking improves consumer comfort and organizational effectiveness by streamlining the employment application process. When combined with interactive features, customization options such as personalized profiles and preferences help to increase user engagement. Employing web scraping methods with Puppeteer and Cheerio speeds up job searches considerably, saving users time and requiring less human labour. Strong data privacy safeguards, including safe storage in a MongoDB database, increase user confidence. In conclusion, our project has effectively combined state-of-the-art technology with user-focused features, changing the way people look for work by offering a safe, secure, and personalized platform.

Fig. 1. Login Page

Fig. 2. Details of Applicant
7 Conclusion

In summary, the "Personalized Job opportunity Finder powered by Web Scraping " initiative has revolutionized the conventional job search experience, surpassing expectations. A platform that caters to the various requirements and preferences of job searchers has been created through the combination of daily updates, application monitoring that is simplified, personalized job suggestions, and an advanced matching algorithm. The ability to customize and interact with the platform has improved user engagement, while web scraping techniques have significantly improved the effectiveness of job searches. Most importantly, consumers’ trust and faith in the platform are guaranteed by our dedication to strict data protection procedures, which include safe storage in a MongoDB database. The initiative is still evolving, but the good reaction that has been received highlights how it has the potential to completely change how people approach and negotiate the job market. "Personalized Job opportunity Finder powered by Web Scraping " is proof of the revolutionary potential of technology in improving user experience and offering a safe, effective, and customized response to the problems associated with job hunting.

References

Multiple Clients” 2013, First International Conference on Artificial Intelligence, Modelling & Simulation.
