

Job Scrapper With Alert System GUI

¹Satish Chadokar, ²Vishal Chourey, ³Yash Sirsam, ⁴Vani Dhote, ⁵Nikita Pawar

Shri Balaji Institute of Technology and Management

Abstract:

Our project constitutes a comprehensive platform that serves as a unified resource for accessing targeted employment opportunities. Users can input specific keywords, such as "Python internships," and our platform will aggregate and display relevant internship and job listings sourced from diverse internship and job-providing platforms. Designed as a one-stop solution, our platform streamlines the job search process by consolidating information, providing users with a centralized hub for identifying and applying to desired positions. By entering a keyword, such as "Python internships," individuals can efficiently access a curated compilation of relevant opportunities within their skillset from a variety of internship and job-providing platforms.

Keywords: Machine learning, Python, Web scrapping, Job Finder, Beautiful Soup.

I. Introduction

Our project addresses the challenge of aggregating comprehensive information about internships and jobs specifically tailored to individual skillsets. By consolidating data from various internship and job-providing platforms onto a single platform, users can streamline their job search process.

Traditional methods of job hunting on platforms like AngelList, Internshala, and Indeed involve navigating through postings and applying directly on the respective company websites. While this approach eliminates the need for direct company contact, it may not be the most effective method for job seekers. Platforms like AngelList and LinkedIn have limitations for job seekers, such as regional constraints and incomplete company information. Our platform overcomes these limitations by providing access to all jobs within and beyond the user's area. Our platform offers features similar to AngelList and LinkedIn, along with additional benefits like comprehensive company information, cultural insights, career opportunities, and salary data.

Traditional platforms also impose restrictions on the number of visible job postings, making it challenging for users to find the perfect match. Our platform eliminates this constraint by providing access to an extensive range of relevant job posts from different companies, facilitating a more effective and comprehensive job search experience.

II. Literature and Data Review

A. Existing System of Finding Internship and Jobs

In the current scenario, individuals typically navigate across diverse platforms, such as AngelList, LinkedIn, Internshala, Indeed, Naukri.com, to discover job opportunities aligned with their desired skillset. While these platforms are undeniably valuable, instances arise where suitable job listings may not be found. Seeking assistance on forums sometimes yields no responses, especially on less-popular platforms lacking adequate traffic.

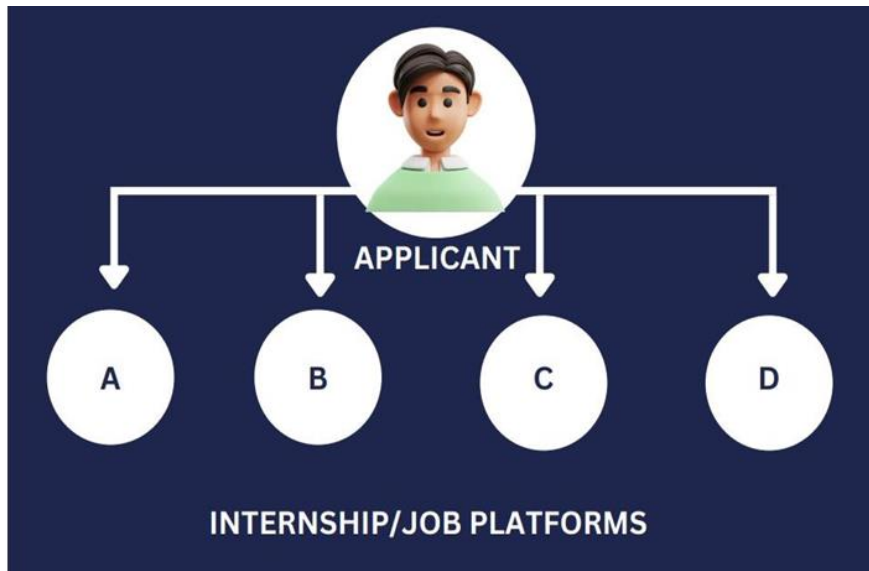


Fig.1 The usual way of finding jobs online

The process of scouring through various sites, like Internshala, for jobs or internships matching specific skills, time flexibility, and work preferences (e.g., remote or in-office) is notably time-consuming. Daily monitoring of these sites becomes burdensome, prompting the need for a more efficient solution. Consequently, we conceived a platform where users can post their job requirements, connecting with suitable profiles nationwide that share similar job aspirations.

Is the time spent applying for jobs online worthwhile? While it can be if approached strategically, selecting the right websites is crucial. The team at GCF Global recommends utilizing well-established platforms with a reputable history to ensure a seamless job search experience without unnecessary redirects to other sites.

B. Our Proposed System for finding Internship and Job

In the contemporary landscape, job seekers often face challenges navigating various job-providing platforms to align opportunities with their skillsets, time flexibility, and work preferences, whether remote or in-office. The proliferation of job platforms exacerbates this complexity.

To mitigate these challenges and streamline the job search process, we have developed a unified platform, serving as a one-stop solution. This platform aims to alleviate time constraints and the cumbersome process of scouring multiple job platforms. It is meticulously designed to deliver the most pertinent and suitable job listings to users.

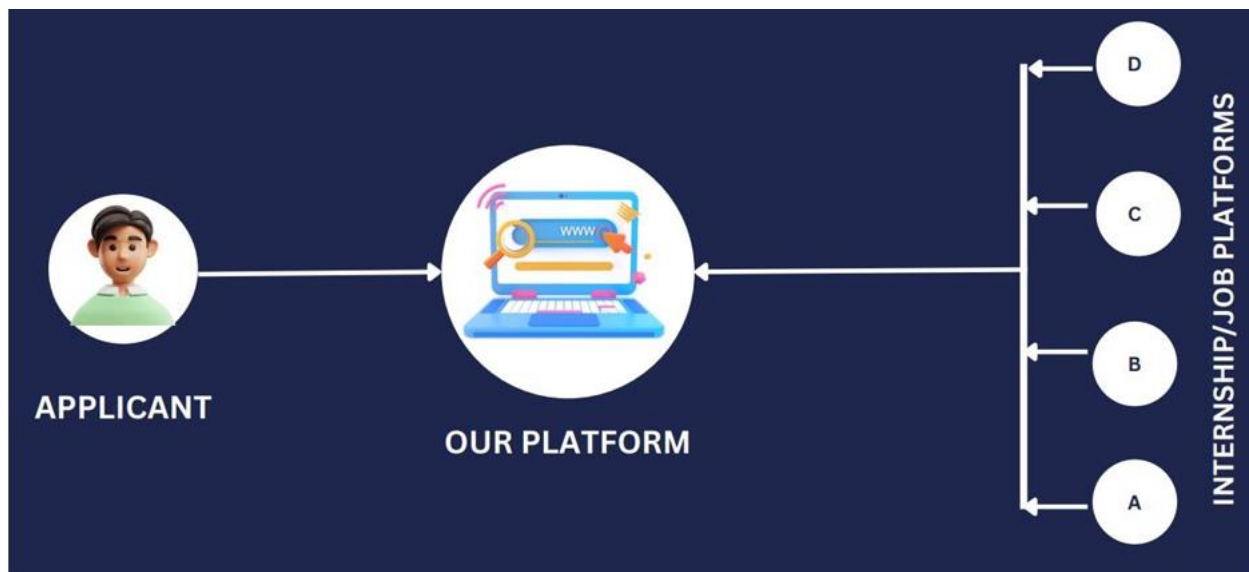


Fig.2 Finding jobs with the help of proposed platform

Key features of the platform include account creation, streamlined job applications, and efficient application management. Job seekers can easily apply for available positions and manage their applications with customized search and viewing options. The platform also consolidates comprehensive information about companies and their job offerings in a centralized location. Employers benefit from a centralized space to post job requirements, eliminating the need to navigate multiple job platforms.

Consider the case of Ram, a computer science engineering student proficient in Python. In the existing system, Ram would spend considerable time navigating multiple job platforms to find a suitable job. However, our platform offers a more efficient solution, presenting all relevant jobs related to Python and meeting his specified criteria in a consolidated manner. Notably, our platform allows users to simultaneously access multiple sites, overcoming the limitations of other platforms that restrict users to a single site at a time. The interface is intuitively designed, providing users with a seamless browsing experience across multiple platforms.

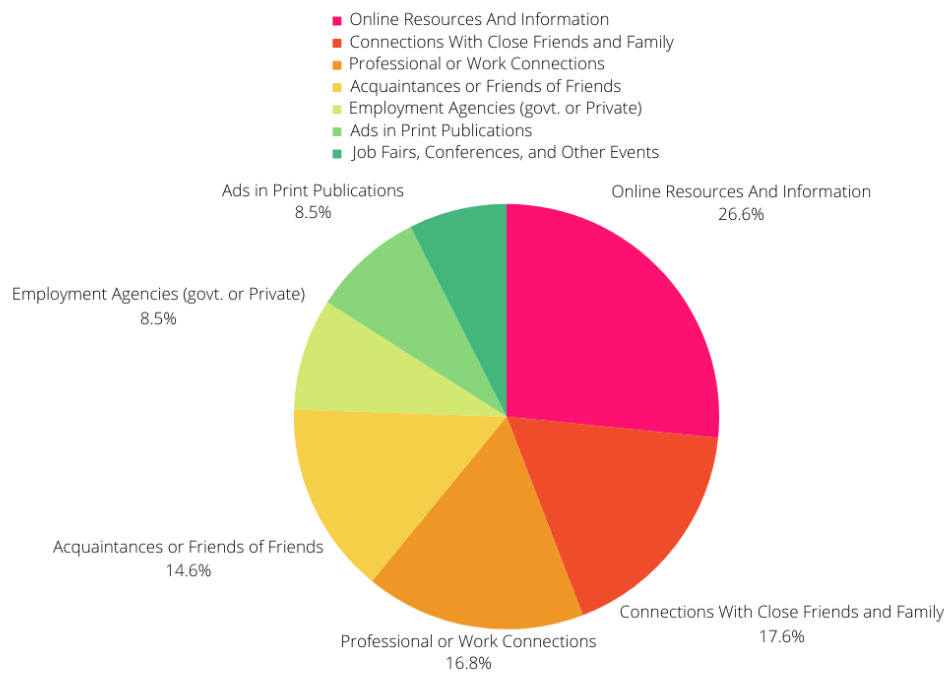


Fig.3 Percentage chart of people seeking jobs through various ways in the recent times

C.Candidates are preferring internet for job hunting

Online job hunting is predominant in today's age, facilitated by platforms like LinkedIn, Indeed, and company websites. Remote work opportunities have increased, and the application process often involves electronic submissions. Job seekers utilize personalized alerts, enhance skills through online courses, and engage in remote interviews. Social media plays a role in job searches, providing insights into companies and industry trends. Overall, the online landscape has transformed how individuals search for, apply to, and secure job opportunities.

- 1) Job Search Platforms: Numerous online job boards and platforms exist, connecting job seekers with employers. Popular platforms include LinkedIn, Indeed, Glassdoor, Monster, and many industry-specific sites. These platforms allow users to create profiles, upload resumes, and apply for jobs easily.
- 2) Company Websites: Many companies post job openings directly on their websites. Job seekers often visit the career sections of company websites to explore available positions, company culture, and application procedures.
- 3) Professional Networking: Platforms like LinkedIn play a crucial role in professional networking and job hunting. Job seekers use these platforms to connect with professionals, follow companies, and stay updated on industry trends. Employers also use LinkedIn to find potential candidates.

4) Remote Job Opportunities: The rise of technology has facilitated an increase in remote work opportunities. Job seekers can now search for and apply to jobs that allow them to work from anywhere, expanding the scope of potential employment.

5) Application Processes: Online job applications typically involve submitting resumes and cover letters electronically. Some platforms also allow for the easy tracking of application statuses. Additionally, some companies use applicant tracking systems (ATS) to streamline the hiring process.

6) Job Alerts and Notifications: Many job platforms offer personalized job alerts based on a user's preferences and skills. These notifications keep job seekers informed about new opportunities that match their criteria.

Online platforms provide resources for skill development through courses, certifications, and webinars. Job seekers can enhance their skills and qualifications to improve their marketability in their chosen field.

8) Remote Interviews: The interview process has evolved, with many employers conducting initial interviews via video conferencing tools. This allows for a more efficient hiring process, especially for remote or international positions.

9) Research and Reviews: Job seekers can easily research companies and read employee reviews online. This information helps them make informed decisions about potential employers and assess whether a company aligns with their values and career goals.

10) Social Media: Beyond professional networking platforms, social media, such as Twitter and Facebook, are also used by companies to share job opportunities. Job seekers can follow company pages and industry influencers for relevant updates.

While online job hunting offers numerous advantages, it's essential for job seekers to be mindful of privacy and security considerations. Additionally, networking and building a professional online presence are crucial elements for success in today's competitive job market.

III. Methodology and Technology Used

A.Data Scrapping

Data scraping, or web scraping, is an automated process that extracts information from websites. It involves navigating through web pages, parsing HTML or XML content, and extracting specific data elements using techniques like regular expressions or CSS selectors. This method is commonly used for competitive analysis, market research, lead generation, content aggregation, and financial analysis. However, ethical considerations and adherence to legal requirements are essential in its application.

Data retrieved from the internet is showcased on our website through diverse presentation formats. We have the capability to present the data using graphs, tables, and charts, providing our users with a user-friendly means of interpreting the information. For instance, we can employ a pie chart to illustrate the distribution of users across different device types. Additionally, a line graph can be utilized to depict the traffic trend over a specific period.

B.Python

Python is widely used in data scraping due to its rich ecosystem of libraries and tools that simplify the process. Here are some key aspects of Python's role in data scraping:

1) Beautiful Soup and Requests: These are fundamental libraries in Python used for web scraping. Requests facilitates sending HTTP requests to retrieve web pages, while Beautiful Soup aids in parsing and navigating HTML or XML content, making it easier to extract relevant data.

```

import requests
from bs4 import BeautifulSoup

# The URL of the website you want to scrape
url = 'YOUR_TARGET_URL'

# Send a GET request to the website
response = requests.get(url)

# Check if the request was successful
if response.status_code == 200:
    # Parse the HTML content of the page with BeautifulSoup
    soup = BeautifulSoup(response.text, 'html.parser')

    # Now you can navigate the HTML tree and extract the data you need
    # For example, to extract all text within <p> tags:
    paragraphs = soup.find_all('p')
    for p in paragraphs:
        print(p.get_text())
else:
    print(f'Failed to retrieve the webpage. Status code:
{response.status_code}')

```

Fig.4 Code for scrapping data from a particular site using BeautifulSoup

Here's the algorithm for the above code for data scraping from any site-

1. Import the required libraries: requests for making HTTP requests and BeautifulSoup for parsing HTML.
2. Define the URL of the website you want to scrape.
3. Send a GET request to the website using `requests.get(url)`.
4. Check if the request was successful (HTTP status code 200).
5. If the request was successful:
 - a. Parse the HTML content of the page using `BeautifulSoup(response.text, 'html.parser')`.
 - b. Navigate the HTML tree to extract the desired data.
 - c. For example, to extract all text within `<p>` tags, use `soup.find_all('p')`.
 - d. Iterate over the result set and print the text content of each `<p>` tag using `p.get_text()`.
6. If the request was not successful (status code other than 200), print an error message.

2) Scrapy: Scrapy is a powerful and extensible framework specifically designed for web scraping. It provides a high-level API for efficiently extracting data from websites and can handle large-scale scraping projects.

3) Pandas: Once data is extracted, Pandas is often used for data manipulation, cleaning, and analysis. It provides convenient data structures and functions for working with structured data.

4) Regular Expressions (re): Python's re module is useful for pattern matching and extracting specific information from text data during the scraping process.

Python's versatility, ease of use, and the availability of these powerful libraries make it a popular choice for data scraping tasks, whether for small-scale projects or large-scale web scraping operations.

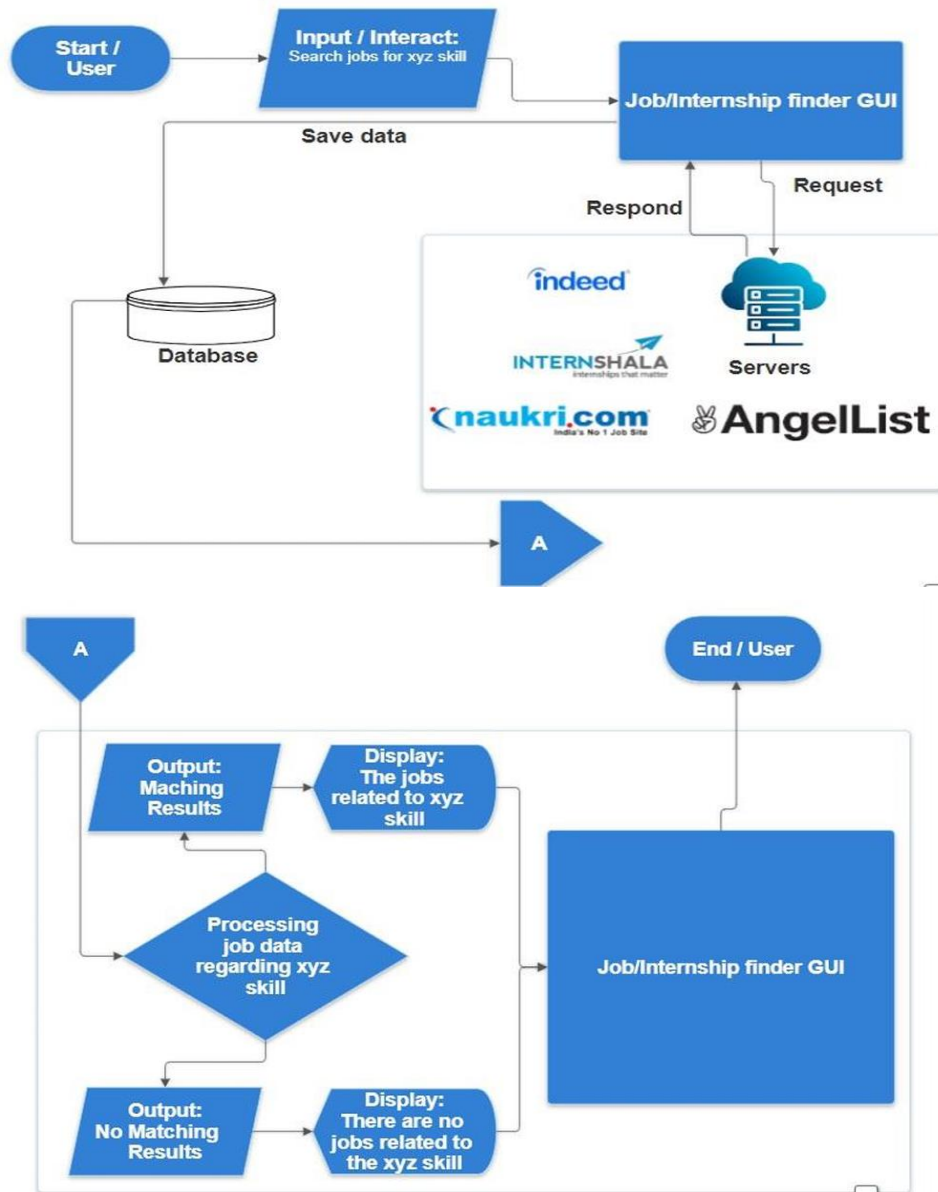


Fig.5 Flow chart of our proposed platforms working for finding jobs

IV. Output and Conclusion

This platform stands as a robust and comprehensive resource catering to job seekers actively exploring employment opportunities. It grants users access to a diverse array of job openings, spanning internships, entry-level positions, and more advanced roles, consolidating them into a single, user-friendly interface. The platform alleviates the need for job seekers to navigate across various platforms, offering a centralized solution for their employment search.

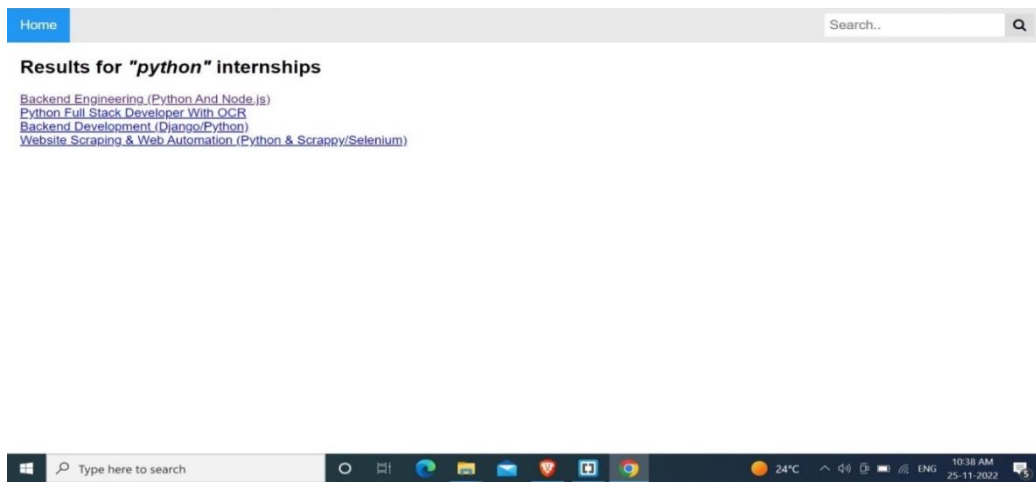


Fig.6 Search results of Python related jobs and internships

Advantages:

A) Unified Access to Opportunities

As a one-stop platform, it streamlines the job search process by presenting a multitude of opportunities in one location. Whether users are seeking internships, entry-level positions, or roles requiring greater experience, the platform ensures a seamless and efficient search experience.

B) Efficient Job Search Tools

The platform goes beyond mere aggregation by providing additional tools and resources. Notably, job alerts keep users informed about the latest postings, ensuring they remain up-to-date with relevant opportunities. This proactive approach enhances the overall efficiency of the job-seeking process.

C) Convenience and Simplicity

The value proposition of this platform lies in its simplicity and convenience. Job seekers can navigate through a variety of opportunities without the need to switch between multiple platforms. The user-centric design promotes an intuitive experience, facilitating a more focused and effective job search.

In essence, this platform serves as an encompassing solution, offering a centralized hub for job seekers to discover and apply for positions that align with their career aspirations. Its multifaceted approach, coupled with efficient tools and resources, positions it as a go-to destination for individuals striving to realize their professional ambitions.

Key Highlights:

1) **Comprehensive Solution:** The platform serves as a one-stop destination for both internship and job seekers, catering to a broad spectrum of skillsets and career stages.

2) **Streamlined Job Search:** By consolidating opportunities, the platform eliminates the need for job seekers to navigate diverse platforms, streamlining the search process and saving valuable time.

In conclusion, this platform stands as an indispensable tool for job seekers, offering not only a rich array of opportunities but also the efficiency and convenience needed to navigate today's competitive job market successfully.

REFERENCES:

- [1] Ryan Mitchell, "Web Scraping with Python: Collecting More Data from the Modern Web", O'Reilly Media, Inc. 2018.
- [2] Richard Lawson, "Web Scraping with Python", Packt Publishing Ltd, 2015.
- [3] Harshit Nigam and Prantik Biswas, "Web Scraping: From Tools to Related Legislation and Implementation Using Python", Springer, Singapore, 2021.
- [4] David Mathew Thomas, Sandeep Mathur, "Data Analysis by Web Scraping using Python", IEEE, 2019.
- [5] Anish Chapagain, "Hands-On Web Scraping with Python: Perform advanced scraping operations using various Python libraries and tools such as Selenium, Regex, and others", Packt Publishing Ltd, 2019.