#	Task Description		3-Nov	10-Nov	17-Nov	24-Nov	1-Dec	8-Dec	15-Dec	22-Dec	29-Dec	5-Jan
			W1	W2	W3	W4	W5	W6	W7	W8	W9	W10
1	Define Problem	Plan	▓	▓								
		Actual	▓	▓	▓							
2	Review Literature	Plan		▓	▓	▓						
		Actual		▓	▓	▓						
3	Design System	Plan				▓	▓	▓				
		Actual										
4	Develop Prototype	Plan						▓	▓	▓		
		Actual										
5	Validate Prototype	Plan								▓	▓	▓
		Actual										
6	Write Report	Plan		▓		▓		▓		▓		▓
		Actual	▓	▓	▓	▓						

Problem Definition Report

Project Name: Applying Web Scraping with Machine Learning for Job Searching

Team Member:

Asif 011212032

Shahriar Rahman 011212067
Sabrina zaman 011212137
Safrin Sanzida Islam 011212124
Nafisa Zaman 011212120
Abubakkar Siddique. 011201329

Supervisor:

Prof. Dr. Hasan Sarwar
Department of CSE & DEAN

Course Title : Final Year Design Project - I   **Section :** C

Course Teacher : Prof. Dr. Hasan Sarwar   **Date :** 10- 11- 2024

Problem Definition

The job market is overwhelmed with scattered and unstructured job data across various online platforms, making it difficult for job seekers to find relevant opportunities, for employers to understand market trends, and for researchers to analyze labor dynamics. This fragmentation creates inefficiencies, limits decision-making, and results in missed opportunities for all stakeholders.

#	Task Description		3-Nov W1	10-Nov W2	17-Nov W3	24-Nov W4	1-Dec W5	8-Dec W6	15-Dec W7	22-Dec W8	29-Dec W9	5-Jan W10
1	Define Problem	Plan	▓	▓								
		Actual	▓	▓	▓							
2	Review Literature	Plan		▓	▓	▓						
		Actual		▓	▓							
3	Design System	Plan				▓	▓	▓				
		Actual										
4	Develop Prototype	Plan						▓	▓	▓		
		Actual										
5	Validate Prototype	Plan								▓	▓	▓
		Actual										
6	Write Report	Plan		▓		▓		▓		▓		▓
		Actual	▓	▓	▓	▓						

Motivation

In today's fast-changing world, jobs and skills are constantly evolving. Staying updated about job trends, salary ranges, and in-demand skills is essential for people making career choices, students planning their education, or organizations shaping policies. However, finding this information can be time-consuming and overwhelming.

This project simplifies that process by automatically collecting and analyzing data from job postings across the internet. It helps job seekers find opportunities that match their skills, educators understand what to teach for future jobs, and policymakers get insights into the economy and workforce needs. By turning raw data into clear, easy-to-understand information, the project empowers everyone to make smarter decisions about work and careers, shaping a brighter future for individuals and society as a whole.

Complex Engineering Problem

1. **WK2:** Conceptually-based mathematics, numerical analysis, statistics, and formal aspects of computer and information scienceThe project extensively uses numerical and statistical analysis for insights such as job trends, salary comparisons, and skill analysis. Formal concepts from information science are integral to database design, data processing, and search algorithms.

3. **WK6: Knowledge of engineering practice (technology) in practice areas.** The project demonstrates practical application by using technologies like Python, databases (e.g., MySQL, MongoDB), and visualization tools (e.g., Plotly, Tableau). It involves deploying user interfaces using frameworks such as Flask or React, showcasing hands-on engineering and technology application.

#	Task Description		3-Nov W1	10-Nov W2	17-Nov W3	24-Nov W4	1-Dec W5	8-Dec W6	15-Dec W7	22-Dec W8	29-Dec W9	5-Jan W10
1	Define Problem	Plan	▓	▓								
		Actual	▓	▓	▓							
2	Review Literature	Plan		▓	▓	▓						
		Actual		▓	▓	▓						
3	Design System	Plan				▓	▓	▓				
		Actual										
4	Develop Prototype	Plan						▓	▓	▓		
		Actual										
5	Validate Prototype	Plan								▓	▓	▓
		Actual										
6	Write Report	Plan		▓		▓		▓		▓		▓
		Actual	▓	▓	▓	▓						

2. **WK4:** Engineering specialist knowledge providing theoretical frameworks for accepted practices. The project relies on established frameworks in data scraping and engineering practices (e.g., Scrapy, BeautifulSoup, Selenium) and incorporates machine learning techniques for advanced trend analysis and job recommendations.

4. **WK7:** Comprehension of the role of engineering in society, including ethics and professional responsibility. Ethical considerations, including compliance with website terms of service, the responsible handling of data, and adherence to privacy laws, are central to the project. These highlight the role of engineering in maintaining professional integrity and addressing societal impacts.

	Task Description		3-Nov	10-Nov	17-Nov	24-Nov	1-Dec	8-Dec	15-Dec	22-Dec	29-Dec	5-Jan
			W1	W2	W3	W4	W5	W6	W7	W8	W9	W10
1	Define Problem	Plan	�	▓								
		Actual	▓	▓	▓							
2	Review Literature	Plan		▓	▓	▓						
		Actual		▓	▓	▓						
3	Design System	Plan				▓	▓	▓				
		Actual										
4	Develop Prototype	Plan						▓	▓	▓		
		Actual										
5	Validate Prototype	Plan								▓	▓	▓
		Actual										
6	Write Report	Plan		▓		▓		▓		▓		▓
		Actual	▓	▓	▓	▓						

```python
import matplotlib.pyplot as plt

# Data for the pie chart
labels = ['High Computational Machine', 'Development and Labour Costs',
        'Web Scraping Tools & Software', 'Miscellaneous Costs']
sizes = [150000, 50000, 20000, 10000]  # Corresponding costs for each category
colors = ['#ff9999', '#66b3ff', '#99ff99', '#ffcc99']  # Colors for each section
explode = (0.1, 0, 0, 0)  # Exploding the first slice (High Computational Machine) slightly

# Create the pie chart
plt.pie(sizes, labels=labels, colors=colors, explode=explode, autopct='%1.1f%%', startangle=140)

# Set equal aspect ratio to ensure the pie is a circle
plt.axis('equal')

# Title
plt.title("Cost Breakdown for Web Scraping Job Search Platform")

# Show the plot
plt.show()
```