Project Report

Name: Bhargav Gembali

Role: Software Developer – Intern

Company: Roots Innovation Labs Pvt. Ltd. (AIQoD)

Duration: 27th May 2024 – 26th July

---

1. Web Scraper Bot

Objective:
To develop an automated solution for extracting and transforming web data into structured formats (JSON or CSV).

Technologies Used:

- **Node.js**
- **Puppeteer**
- **JSON, CSV**

Project Description:
The Web Scraper Bot was engineered and deployed using Node.js and Puppeteer. The bot is designed to efficiently extract data from web pages, transforming the unstructured web content into structured JSON and CSV formats. This automated solution facilitates data handling and ensures that the extracted data is readily available for further analysis or storage. The bot requires URL of the website to be scraped, action to be performed, next button (text-context), and format desired.

Key Features:

- Automated data extraction from various web pages.
- Handles paginations, by identifying the provided label name or text context of the next page.
- Transformation of data into JSON and CSV formats for ease of use and compatibility.
- Deployment on a Node.js server for continuous and scalable data scraping operations.

Outcome:
The bot has streamlined the data collection process, reducing the time and effort required for manual data gathering. It also ensured accuracy and consistency in the extracted data, which was critical for downstream processing.

---

2. Resume Analyzer

Objective:
To build an application capable of parsing resumes and extracting relevant data using large language models for scoring and extracting data.

Technologies Used:

- **Node.js**
- **Large Language Models (LLM)**
- **CSV Export**

Project Description:
The Resume Analyzer is a Node.js-based application designed for parsing resumes and extracting critical information using LLM-driven data extraction techniques. The extracted data is organized and exported in CSV format, making it easier for HR professionals to review and assess candidate information.

Key Features:

- Automated parsing of resumes to extract key details such as name, contact information, skills, experience, and score.
- Use of LLM (meta llama 3.8 B) to accurately interpret and extract data from various resume formats.
- Capability to export parsed data into CSV files for easy integration with HR management systems.

Outcome:
The Resume Analyzer improved the efficiency of resume processing, allowing for quicker and more accurate assessments of candidate qualifications. It also reduced the workload on HR teams by automating a significant portion of the resume review process.

---

3. LinkedIn Connector

Objective:
To create an automated solution for managing LinkedIn posts and analytics.

Technologies Used:

- **Node.js**
- **OAuth 2.0**
- **Postman (for validation)**

Project Description:
The LinkedIn Connector is an application integrated with OAuth 2.0 for seamless interaction with LinkedIn's API. It automates the process of posting content on LinkedIn and retrieving analytics data, thereby simplifying social media management.

Key Features:

- OAuth 2.0 integration for secure authentication and authorization.
- Automated posting of content on LinkedIn profiles and pages.
- Retrieval and analysis of LinkedIn analytics to measure content performance.

- Validation and testing of API interactions using Postman.

Outcome:
The LinkedIn Connector enhanced the automation of social media management tasks, reducing manual effort and ensuring timely posting and analytics tracking. It also provided a secure and efficient way to handle LinkedIn API interactions.

---

4. Playwright Enhancements

Objective:
To extend Playwright's capabilities for dynamic test generation and recording.

Technologies Used:

- **TypeScript**
- **JavaScript**
- **Playwright Code-gen (Testing)**

Project Description:
This project involved modifying and extending the backend of Playwright, a popular end-to-end testing tool, to handle dynamic variable management during test recording. The enhancements were made in TypeScript/JavaScript, allowing for more flexible and customizable test scenarios.

Key Features:

- Implementation of dynamic variable handling during test recording, enabling more versatile test cases.
- Modifications to Playwright's backend to support these new features.
- Integration of user prompts for managing dynamic variables during test creation.

Outcome:
The enhancements to Playwright significantly improved the flexibility and power of test case generation, making it easier to create and maintain complex test scenarios. This led to better test coverage and more reliable software testing processes.

Conclusion

During the internship at Roots Innovation Labs Pvt. Ltd. (AIQoD), Bhargav Gembali has developed four key projects that significantly enhanced automation and efficiency in web data extraction, resume parsing, LinkedIn social media management, and software testing. By leveraging advanced technologies such as Node.js, Puppeteer, Large Language Models, OAuth 2.0, and Playwright, he engineered solutions that streamlined processes, reduced manual effort, and improved accuracy across various domains. These projects not only demonstrated his technical expertise and problem-solving skills but also contributed valuable tools that can be scaled and adapted for future applications.