LinkedIn Job Scraper - Project Report

Abstract

This project automates the process of collecting job listings from LinkedIn using Selenium and BeautifulSoup. The scraper logs into LinkedIn, searches for jobs with specific keywords and locations, extracts job titles, company names, and posting dates, and saves the data into a CSV file. Additionally, the project includes a data visualization step to show the top companies hiring for the specified role.

Introduction

LinkedIn is a widely used platform for job seekers and recruiters. However, manually searching and collecting job data can be time-consuming. This project was developed to automate job data scraping from LinkedIn, making the process faster and more efficient.

Tools Used

The following tools and libraries were used in building this project: - Selenium: For browser automation and LinkedIn login. - BeautifulSoup: For parsing HTML content and extracting job details. - Pandas: For data storage and CSV export. - Matplotlib: For visualization of job frequency by company. - Python-dotenv: For managing sensitive credentials securely.

Steps Involved

The key steps involved in building the project are: 1. Setting up Selenium WebDriver and logging into LinkedIn. 2. Navigating to the LinkedIn Jobs page with specific keywords and locations. 3. Scrolling the page to load job listings dynamically. 4. Extracting job titles, company names, and posting dates using BeautifulSoup. 5. Saving the scraped data into a CSV file. 6. Removing duplicate job entries. 7. Visualizing job frequencies by company using a bar chart.

Conclusion

This project demonstrates how web automation and scraping techniques can be used to collect and analyze job market data from LinkedIn. It provides a structured dataset for further analysis and insights into hiring trends. The automation significantly reduces manual effort and enables quick updates of job listings.