

Shivam Dixit

shivam.r.dixit@gmail.com

[linkedin.com/in//shivam-dixit](https://www.linkedin.com/in//shivam-dixit)

github.com/shivvu

Data Engineer, experienced in scraping websites, automating processes and designing CI/CD pipelines.

Skills & Qualifications

- 3+ Years experience with Web scraping, Automation.
- Extensive knowledge in crawling websites.
- Experienced in CI/CD Pipeline & basic knowledge in API.
- 1+ year experience in designing ETL pipelines

Tech Stack

- **Languages:** Python, SQL
- **Crawling Frameworks:** Scrapy, Selenium, Playwright
- **Databases:** Postgres & MySQL
- **Other:** Scrapinghub & AWS

Work History

Data Automation Engineer | forage.ai

10/2021 - Present

Worked on exciting projects including collecting and analyzing data from various sources using the Scrapy framework, as well as developing APIs using our own in-house Flask framework.

- Created more than 200 spiders of varying complexity.
- Designing robust CI pipeline from scratch & has been running successfully from 1 year without any failure.
- Developed an automated spider scheduling pipeline that runs on a scrapy cloud, resulting in 100% increase in efficiency.

Data Scientist | HolidayME

10/2020 - 10/2021

collecting data from different travel websites to provide better deals to customers.

- Developed custom web scraping scripts using tools like Scrapy, or Selenium to meet specific project requirements.
- Conducted data cleaning and preprocessing to ensure accuracy and consistency of scraped data
- 12+ Implemented proxy and user-agent rotation to prevent IP blocking and improve data collection

Data Engineer | Dataweave (Ecommerce)

07/2020 - 10/2020

Using in-house tools to crawl the data at scale & customizing scripts as per requirement.

- Configuring the complex websites & tuning the settings for automating delivery.
- Writing script(plugin) to scrape data from different website across various sections
- Mapping the data according to product variants & ensuring the data quality.