



Technocation Freelancing Training Institute And Software House

Web Scraping Certification Module

MODULE 1: Introduction to Web Scraping

- Web scraping means collecting data from websites.
- It is used to gather useful information.
- We use simple tools and Python code for scraping.
- It makes searching and learning easy.

MODULE 2: Basics of Websites

- Websites are made of pages.
- Each page has text, pictures, and links.
- Pages are written in HTML language.
- We need to understand this to scrape.

MODULE 3: Tools for Web Scraping

- Python is the main tool for scraping.
- BeautifulSoup is a simple Python library.
- Requests library helps to get pages.
- Tools make scraping easy and fast.

MODULE 4: Understanding HTML Tags

- HTML is like building blocks of a page.
- Tags show where text or image is.
- `<p>` is for paragraph, `` is for image.
- Learning tags helps us pick data.

MODULE 5: Installing Python & Libraries

- First install Python on your computer.
- Then add libraries like BeautifulSoup.

- Install by writing commands in terminal.
- Now you are ready to start scraping.

MODULE 6: Fetching a Webpage

- Use the Requests library to open a page.
- It gives us the page's code.
- We can then read it in Python.
- This is the first step of scraping.

MODULE 7: Parsing the Page with BeautifulSoup

- BeautifulSoup reads the HTML page.
- It helps us find tags inside the page.
- We can pick headings, text, or images.
- It is like searching with a magnifying glass.

MODULE 8: Extracting Text Data

- Text can be inside paragraphs.
- We search for <p> or <h1> tags.
- We collect only useful text.
- This text can be saved in a file.

MODULE 9: Extracting Images & Links

- Images are found in tags.
- Links are inside <a> tags.
- We can collect all links from a page.
- Pictures and links can be saved too.

MODULE 10: Saving Data

- Collected data must be stored safely.
- Save in CSV file like a table.
- Save in JSON for coding use.
- Saving makes data easy to use later.

MODULE 11: Handling Errors

- Sometimes a page may not open.
- Errors happen if internet is slow.
- We write code to try again.
- Handling errors keeps scraping smooth.

MODULE 12: Ethics & Safety in Scraping

- Always respect website rules.
- Never scrape too much at once.
- Use data for learning, not harm.
- Be a responsible web scraper.

MODULE 13: Certification Module

- Build a small web scraper yourself.
- Example: collect book titles or images.
- Submit your project for review.
- Get your **Web Scraping Certificate**.

+92 308 514 5822

www.technocation.pk

info@technocation.pk

