



¡Bienvenidos al curso maestro de Web Scraping!

ÍNDICE DEL CURSO

Estas a punto de empezar el curso de Web Scraping más completo del internet. El conocimiento que aprenderás aquí es invaluable. Luego de este curso, serás un maestro del Web Scraping.

En este documento detallo brevemente las actividades y los objetivos de cada sección del curso:

1. Aprendiendo a Programar en Python

Objetivo: Esta sección es para aprender a programar desde cero. Si ya tienes fuertes conocimientos de Python, puedes saltar esta sección por completo.

Aprendizajes: Programación en Python (Imperativa y Orientada a Objetos)

2. Introducción al Web Scraping

Objetivo: Aprenderás todas las bases necesarias para empezar a hacer Web Scraping.

Aprendizajes: Fundamentos del Web Scraping | Fundamentos de la programación web | XPATH.

3. NIVEL 1: Una sola página estática

Objetivo: Aprenderás a realizar Web Scraping de una sola página a la vez.

Librerías: requests, lxml, BeautifulSoup, Scrapy.

Prácticas: Wikipedia, Stackoverflow y Diario El Universo.

4. NIVEL 2: Varias páginas del mismo dominio

Objetivo: Aprenderás a realizar Web Scraping de varias páginas al mismo tiempo utilizando la librería Scrapy.

Librerías: Scrapy.

Prácticas: Trip Advisor, Mercado Libre, IGN.



5. NIVEL 3: Páginas dinámicas

Objetivo: Aprenderás a realizar Web Scraping de páginas cuya información carga de manera dinámica. Aprenderás a llenar formularios, manejar varios tabs, iniciar sesión, hacer scrolling, dar clicks y esperar por eventos.

Librerías: Selenium.

Prácticas: OLX, Mercado Libre, Google Places, Twitter.

6. NIVEL 4: APIs y iframes

Objetivo: Aprenderás a realizar Web Scraping de APIs y iframes. Aprenderás que son APIs y como funciona el formato JSON.

Librerías: Requests.

Prácticas: Udemy, W3Schools.

7. NIVEL 5: Autenticación y Captchas

Objetivo: Aprenderás a realizar autenticación antes de realizar una extracción de datos. Aprenderás a resolver CAPTCHAs de manera manual y automática.

Librerías: Requests, Selenium.

Prácticas: Github, Google Demos.

8. Evitando Problemas: Ética

Objetivo: Aprenderás a realizar Web Scraping de manera responsable.

Aprendizajes: User Agents | VPNs | ¿Cómo evitar baneos?

9. NIVEL BONUS: Almacenamiento y actualización de datos

Objetivo: Aprenderás a realizar extraer imágenes. Aprenderás a extraer datos y almacenarlos en MongoDB. Aprenderás a extraer y actualizar datos de manera periódica.

Librerías: Requests, Pillow, Pymongo, Selenium.

Prácticas: OLX, Accuweather.