

Lenguajes y Paradigmas de Programación

Proyecto

En esta clase discutiremos el proyecto a realizar en el semestre y la app 3 (phyton)

App 3 : 5 de Junio, 23:59 hrs

Proyecto Final: 19 de Junio, 23:59

(se acordará bien en la clase del 13/05)

Proyecto final: Web Scraper

¿Qué es?

Método para recopilar información de Internet → simular la navegación web humana para recopilar información de diferentes sitios web.

Pasos a paso

- Encuentra la URL que deseas extraer.
- Inspeccionando la página.
- Encuentre los datos que desea extraer.
- Escribe el código.
- Ejecute el código y extraiga los datos.
- Almacene los datos en el formato requerido.

Proyecto

Realizar un web scraper que permita:

- Extraer precios de medicamentos de las grandes cadenas farmacéuticas (Salco Brand, Cruz Verde y Farmacias Ahumadas).
- Procesar la información de los medicamentos.
- Comparar las ofertas.

Idea de estructura

- Scraping: paradigma procedural (25% nota)
- Estructura de los medicamentos: orientado a objetos (25% nota)
- Procesamiento de datos: funcional (25% nota)
- Presentación (25% nota)

Trabajar

La idea es que se metan a sitios de existentes de las cadenas farmacéuticas y extraigan la estructura de presentación de un medicamento y la modelen como clases usando orientación a objetos.

Se revisará en la clase del 27 de mayo para consultas.

Recursos de apoyo

- [Introducción al Web Scraping](#)
- [Web Scraping en Python: SCRAPY SPIDER](#)
- [Cómo hacer scrape de sitios web con Python 3](#)
- [Una introducción práctica al web scraping en Python](#)
- [Cómo hacer Web Scraping con Python](#)