

TABLA DE CONTENIDOS SESIÓN #63_11/01/2024 PROFESOR LUIS MIGUEL CHICA

Parte 1: Introducción a Python (Duración: 1 hora)

1. Repaso de los Fundamentos de Python (30 minutos)

- Introducción Breve a Python
- Sintaxis Básica
- Variables y Tipos de Datos
- Operadores

2. Estructuras de Control (30 minutos)

- Condicionales (if, elif, else)
- Bucles (for, while)
- Funciones
- Introducción a Clases y Objetos

Parte 2: Repaso de Plataformas GitHub y Hugging Face (Duración: 30 minutos)

1. GitHub

- Conceptos Básicos y Uso
- Gestión de Repositorios

2. Hugging Face

- Características y Aplicaciones
- Integración con Python y Otras Librerías

Parte 3: Librerías para Generación de Datasets (Duración: 1 hora)

1. Pandas (20 minutos)

- Introducción y Casos de Uso
- Lectura y Escritura de Datos
- Operaciones Básicas con DataFrames

2. NumPy (20 minutos)

- Visión General
- Operaciones con Arrays

- Integración con Pandas

3. Librería Datasets de Hugging Face (20 minutos)

- Presentación y Características
- Carga y Manipulación de Datasets
- Subida de Datos a Hugging Face

Parte 4: Práctica con wget y Pandas (Duración: 1:30 horas)

1. Uso de wget para Descargar Datasets (15 minutos)

- Introducción a wget
- Descarga de Archivos CSV

2. Procesamiento con Pandas (45 minutos)

- Limpieza de Datos
- Agregación y Transformación de Datos

3. Integración con la Librería Datasets de Hugging Face (30 minutos)

- Preparación de los Datos para Subida
- Creación de un Repositorio Personal en Hugging Face
- Subida del Dataset Procesado

5 Datasets CSV para Descargar

1. Dataset de Iris (Iris Flower Dataset)

- URL: Kaggle Iris Dataset

2. Datos de Vivienda de Boston (Boston Housing Dataset)

- URL: Kaggle Boston Housing Dataset

3. Dataset de Vinos (Wine Quality Dataset)

- URL: UCI Machine Learning Repository - Wine Quality

4. Dataset de Diabetes (Diabetes Dataset)

- URL: Kaggle Pima Indians Diabetes Database

5. Dataset de Pasajeros del Titanic (Titanic Passenger Dataset)

- URL: Kaggle Titanic Dataset