

De Excel a Machine Learning: Análisis de Datos Inteligente

Duración: 16 horas

Modalidad: Teórico-práctica

Nivel: Introductorio-intermedio

Requisitos previos: Manejo básico de Excel. No se requieren conocimientos de programación.

Descripción del curso

Este curso práctico te guía paso a paso desde el uso de Excel para el análisis de datos hasta la implementación de modelos de *Machine Learning* con Python, con ejecución y visualización integradas en Excel. Aprenderás a explorar, limpiar y visualizar información de manera inteligente, sin necesidad de conocimientos avanzados de programación.

Al finalizar, serás capaz de automatizar tus análisis, generar reportes dinámicos y aplicar técnicas de aprendizaje automático en tus propios proyectos, potenciando la toma de decisiones basada en datos.

Objetivo general

Fortalecer las capacidades de los participantes en el uso combinado de Excel y Python para el análisis inteligente de datos, aplicando técnicas de Machine Learning orientadas a generar conocimiento útil y respaldar decisiones técnicas o productivas basadas en la predicción y clasificación de datos.

Objetivos específicos

- Utilizar herramientas de Excel para realizar estadística descriptiva, preprocesamiento de datos y detección de outliers.
- Importar, transformar y analizar conjuntos de datos en Python mediante las librerías pandas y NumPy.
- Implementar modelos básicos de Machine Learning con scikit-learn para tareas de regresión y clasificación.
- Integrar y automatizar flujos de trabajo entre Excel y Python para generar reportes inteligentes y visualizaciones dinámicas.

Competencia General

El participante será capaz de diseñar y ejecutar flujos de análisis de datos combinando Excel y Python, desde la preparación de la base de datos hasta la generación de modelos predictivos y reportes automatizados, fomentando el uso de herramientas digitales en la toma de decisiones técnicas y productivas.

Resultados de Aprendizaje

Al finalizar el curso, el participante podrá:

- Aplicar métodos de análisis estadístico y limpieza de datos en Excel.
- Configurar y ejecutar modelos básicos de Machine Learning en Python.
- Evaluar modelos mediante métricas cuantitativas.
- Automatizar y visualizar resultados integrando Excel y Python.

Contenido Programático (16 horas)

Sesión 1. Fundamentos del Análisis Inteligente de Datos (4 h)

- Introducción al Machine Learning: tipos de modelos y aplicaciones.
- Conceptos clave: datasets, variables, features, target, entrenamiento y validación.
- Métricas de evaluación: precisión, R^2 , matriz de confusión.
- Detección y limpieza de outliers (método IQR y Tukey).

Sesión 2. Excel para Estadística Descriptiva y Preprocesamiento (4 h)

- Funciones estadísticas y lógicas aplicadas al análisis de datos.
- Tablas dinámicas, filtros y segmentación avanzada.
- Detección y corrección de inconsistencias.
- Exportación de datos limpios a CSV.
- (Desarrollo Plus: visualización de datos en Power BI).

Sesión 3. Python para Procesamiento y Modelado de Datos (4 h)

- Entorno Google Colab y librerías principales: NumPy, pandas, Matplotlib, scikit-learn.
- Limpieza automatizada: valores nulos, duplicados y tipos de variables.
- Implementación de modelos de regresión y clasificación.
- Visualización básica y exportación de resultados a Excel.

Sesión 4. Automatización e Integración Excel–Python (4 h)

- Automatización de reportes con pandas.ExcelWriter y openpyxl.
- Creación de dashboards simples con resultados del modelo.
- Flujo completo: carga, modelado y despliegue de resultados en Excel.
- (Desarrollo Plus: visualización de predicciones en Power BI).

Metodología

El curso se desarrolla mediante una combinación de:

- Clases magistrales demostrativas.
- Ejercicios prácticos guiados paso a paso.
- Aplicaciones con datasets reales.
- Actividades integradoras al final de cada sesión.

Herramientas y Recursos

- **Software:** Microsoft Excel, Google Colab, Python (pandas, NumPy, Matplotlib, scikit-learn, openpyxl).
- **Complementarios:** Power BI (demostrativo).
- **Material de apoyo:** notebooks interactivos, plantillas en Excel y guías paso a paso.

Bibliografía y recursos digitales

McKinney, W. (2022). *Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter* (3rd ed.). O'Reilly Media. <https://wesmckinney.com/book/>

Géron, A. (2023). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (3rd ed.). O'Reilly Media. <https://www.oreilly.com/library/view/hands-on-machine-learning/9781098125974/>

pandas development team. (n.d.). *pandas documentation*. pandas. <https://pandas.pydata.org/docs/>

Pedregosa, F., et al. (n.d.). *scikit-learn user guide*. scikit-learn. <https://scikit-learn.org/stable/>

Kaggle. (n.d.). *Datasets*. <https://www.kaggle.com/datasets>

Duración total del curso

El curso tiene una duración total de 16 horas, distribuidas en cuatro sesiones teórico-prácticas de 4 horas cada una. Cada sesión combina la explicación conceptual con ejercicios guiados y actividades aplicadas en Excel y Python.

Duración total del curso

Claudia Serpa Imbett – Fundamentos y Conceptos, Machine Learning en Python
Raúl Vicuña – Excel y demostraciones en Power BI.