

# Machine Learning Aplicado con Python

Código: MLG-02

**Propuesta de Valor:** BIG DATA

**Duración:** 40 Horas



El aprendizaje automático tiene una amplia gama de aplicaciones, incluyendo motores de búsqueda, diagnósticos médicos, detección de fraude en el uso de tarjetas de crédito, análisis del mercado de valores, clasificación de secuencias de ADN, reconocimiento del habla y del lenguaje escrito, juegos y robótica.

En este curso se aplicará algoritmos de Machine Learning de forma simple y eficiente. Se utilizará herramientas para análisis y minería de datos implementando esta librería en los proyectos.



## AUDIENCIA

- Este curso está dirigido a:
  - Desarrolladores con experiencia en Programación y Manejo de Datos.



## PRE REQUISITOS

- Conocimientos sólidos en algoritmos y estructura de datos.
- Manejo de Python.
- Conocimientos sólidos en álgebra.



## OBJETIVOS

- Los objetivos del curso, es brindar al participante las herramientas para que logre:
  - Comprender de forma aplicada los modelos de Machine Learning
  - Conocer y entender el API de Scikit-Learn
  - Explorar y entender familias de modelos de Machine Learning
  - Conocer el manejo de Pre-procesamiento
  - Conocer Performance de Modelos (Evolución, mejoras e iteraciones)

 **CERTIFICACIÓN DISPONIBLE**

- Certificado emitido por COGNOS.

 **CONTENIDO**

## 1. INTRODUCCIÓN

- 1.1. INTRODUCCIÓN AL CURSO DE MACHINE LEARNING APLICADO CON PYTHON
- 1.2. MONTAR UN AMBIENTE DE TRABAJO DE MACHINE LEARNING

## 2. CÓMO DEFINIR UN PROBLEMA DE MACHINE LEARNING

- 2.1. IMPORTANCIA DE DEFINIR EL PROBLEMA EN MACHINE LEARNING
- 2.2. PREDECIR EL INGRESO DE PELÍCULAS DE IMDB
- 2.3. TERMINOLOGÍA DE MACHINE LEARNING

## 3. EL CICLO DE INGENIERÍA DE MACHINE LEARNING

- 3.1. EL CICLO DE MACHINE LEARNING

## 4. MONTAR UN AMBIENTE DE TRABAJO PYDATA

- 4.1. CONFIGURACIÓN DEL AMBIENTE DE TRABAJO CON DOCKER
- 4.2. CONSTRUCCIÓN DE CONTENEDORES DE DOCKER
- 4.3. QUÉ ES Y CÓMO SE UTILIZA NUMPY
- 4.4. ARRAYS EN NUMPY
- 4.5. OPERACIONES ARITMÉTICAS EN NUMPY

## 5. PREPARACIÓN DE LOS DATOS

- 5.1. CARGAR LOS DATOS NECESARIOS PARA EL PROYECTO
- 5.2. INSPECCIÓN DE LOS TIPOS DE DATOS
- 5.3. INSPECCIÓN CUANTITATIVA Y DE SALUD DE LOS DATOS
- 5.4. LIMPIAR LOS DATOS
- 5.5. MANEJO DE DATOS FALTANTES

## 6. MODELACIÓN Y EVALUACIÓN

- 6.1. EL OBJETO ESTIMADOR DE SCIKIT-LEARN
- 6.2. IMPLEMENTAR UN MODELO DE REGRESIÓN (LASSO)
- 6.3. AJUSTANDO MODELOS DE MACHINE MEARNING, UNDERFITTING Y OVERFITTING
- 6.4. EVALUANDO EL MODELO

## 7. INGENIERÍA DE CARACTERÍSTICAS

7.1. FEEDBACK DEL MODELAMIENTO

7.2. ANÁLISIS EXPLORATORIO

7.3. CREACIÓN DE CARACTERÍSTICAS

7.4. SELECCIÓN DE CARACTERÍSTICAS Y LA MALDICIÓN DE LA DIMENSIONALIDAD

## 8. MODELOS Y EVALUACIÓN AVANZADA

8.1. VALIDACIÓN CRUZADA

8.2. SELECCIÓN DE MODELOS

8.3. CURVAS DE APRENDIZAJE

8.4. INTRODUCCIÓN A ENSEMBLES Y ÁRBOLES DE DECISIÓN

8.5. RANDOM FOREST Y GRADIENT BOOSTING TREES

8.6. OPTIMIZACIÓN DE HIPERPARÁMETROS

---

## ★ BENEFICIOS

- Al finalizar el curso, los estudiantes aplicará algoritmos de Machine Learning de forma simple y eficiente. Utilizando herramientas para análisis y minería de datos implementando esta librería en los proyectos.