

# Curso Core Networks de Análisis de datos con Python y Machine Learning

**Duración:** 5 Días - 30 horas

## Descripción del curso:

El curso trata de proporcionar los conceptos y habilidades necesarias para que los alumnos puedan desarrollar aplicaciones distribuidas para el análisis de datos con Python y Spark, que forma parte del conjunto de herramientas del ecosistema Big Data. El alumno puede conocer y practicar con los módulos como Streaming, SQL, GraphX o MLib.

## Dirigido a:

Principalmente a desarrolladores, sobre todo para aquellos que tengan conocimientos y experiencia sobre el lenguaje de programación Python.

## Contenido del curso:

### 1. Machine learning como campo de la IA

- Acotando AI
- Machine learning vs Deep learning
- ML y Big Data
- Rol del científico de datos
- Fases en ML
- Datos y DataSets
- Selección de características o features

- Dimensiones etiquetas atributos
- Método supervisado y no supervisado
- Modelos en ML

## 2. Entrada y salida con Python

- Trabajando con archivos
- Manipulación de datos
- Networking: operaciones

## 3. Distribución de paquetes

- Concepto de package
- Instalación con pip
- Anaconda
- Formato Wheel

## 4. Numpy y Scipy

- Qué son y para que sirven
- NumPy datatypes
- NumPy character codes
- n-dimensional arrays: operaciones

## 5. Plotting con Python

- Librerías más populares
- Matplotlib: configuración
- Modificar colores y líneas
- Mostrar varios valores de forma simultanea
- Establecer límites y rejilla
- Agregar etiquetas, leyenda
- Controlar el formato gráfico
- Controlar el estilo de línea
- Control de la escala: ticks
- Tipos de gráficos y su uso

## 6. Pandas

- Qué es y para que sirve
- Las series: operaciones
- Los DataFrames: operaciones

- Carga de datos en Dataframes
- NumPy y Pandas

## 7. Algoritmos Machine learning

- Algoritmos Machine Learning supervisados
- Algoritmos Machine Learning no supervisados
- Elegir el algoritmo correcto
- Ciclos de vida en el desarrollo de Machine Learning
- Clasificando con k-Nearest Neighbors (SL)
- Algoritmo k-Nearest Neighbors
- La tasa de error
- Árboles de decisión (SL)
- Usando árboles de decisión
- Random Forests
- Naive Bayes Classifier (SL)
- Clasificación de documentos con Naive Bayes
- Learning Type no supervisado: Clustering
- K-Means Clustering (UL)
- K-Means Clustering in a Nutshell
- Análisis de regresión
- Tipos de regresión
- Modelo de regresión simple linear
- Linear Regression Illustration
- Least-Squares Method (LSM)
- LSM Assumptions
- Adaptando modelos de regresiónn linear en R
- Ejemplo de uso de la función R's lm()
- Ejemplo de uso de lm() con un Data Frame
- Regresión logística
- Regresión vs Clasificación
- Análisis de Time-Series
- Descomponiendo Time-Series