



Universidad de
América[®]
Código SNIES 1715

CURSO TRATAMIENTO DE DATOS E INTRODUCCIÓN A MACHINE LEARNING CON PYTHON

Presencial **40** Horas

Fundación Universidad de América | Vigilada MinEducación



ISO 9001:2015
ISO 45001:2018
ISO 14001:2015
BUREAU VERITAS
Certification





PRESENTACIÓN

Este curso tiene como objetivo proporcionar herramientas avanzadas para el tratamiento de datos, facilitando su posterior análisis mediante técnicas básicas de Machine Learning. Su enfoque será práctico, con ejercicios de programación en Python utilizando los frameworks pandas y sklearn, donde en cada clase se combinará una explicación teórica con la solución de un ejercicio aplicado.

La preparación de datos es una de las principales actividades de los científicos de datos e ingenieros, representando más del 70% del tiempo en proyectos de Machine Learning, ya que los errores en el código suelen derivarse de un tratamiento inadecuado de los datos. Además, durante el análisis exploratorio, se pueden identificar patrones y características inherentes antes de aplicar técnicas más avanzadas.

Este curso está diseñado para dotar a los participantes de herramientas que les permitan preparar conjuntos de datos estructurados para su exploración y posterior análisis con técnicas básicas de Machine Learning, abordando algoritmos clásicos con un enfoque práctico, lo que permitirá obtener resultados de manera rápida y optimizar modelos para seleccionar el que mejor describa el comportamiento del fenómeno analizado.

OBJETIVO GENERAL

Mostrar la preparación de datos como la etapa más importante antes de realizar cualquier análisis o plantear modelos de Machine Learning.

OBJETIVOS ESPECÍFICOS

- Hacer una introducción a los algoritmos de Machine Learning mediante el uso de la librería sklearn.
- Enseñar a los estudiantes a realizar un análisis exploratorio EDA y poder preparar los datos para algoritmos de Machine Learning mediante el framework pandas de Python.

```
}}  
code === 271  
{linklv1List.map((val) => {  
  return (  
    <DropdownItem  
      goToMenu="main"  
      leftIcon={<BoltIcon />}  
      key={val.id}  
      goHref={val.link}
```



PERFIL DEL ASPIRANTE

Este curso está dirigido a profesionales, estudiantes y entusiastas del análisis de datos e Inteligencia Artificial que deseen fortalecer sus habilidades en el tratamiento de datos y en la implementación de técnicas básicas de Machine Learning utilizando Python.

Se recomienda que los participantes tengan conocimientos fundamentales sobre el lenguaje Python, así como experiencia en la gestión de entornos virtuales con herramientas como Anaconda, Virtualenv o Poetry. También es conveniente contar con una comprensión básica de estadística descriptiva para interpretar adecuadamente las métricas de evaluación de modelos.

COMPETENCIAS PREVIAS:

- Conocimientos básicos del lenguaje de programación Python
- Conocimientos básicos en estadística descriptiva

METODOLOGÍA

La metodología será práctica y enfocada en el desarrollo de software, donde en cada clase se verá una explicación teórica del tema a tratar con un ejemplo y su solución en Python.

De igual manera, se propondrá un ejercicio corto para aplicar lo aprendido en la parte teórica: el ejercicio se resolverá de manera participativa por los estudiantes y se resolverán todas las dudas en el mismo espacio.

CONTENIDO TEMÁTICO

MÓDULO 1

6 horas

Repaso de conceptos básicos de Python

- Instalación de entorno Anaconda
- Creación de entornos virtuales y el uso de los mismos
- Programación en notebooks
- Programación de scripts, clases y métodos (conceptos básicos de programación orientada a objetos)
- Tipos de datos en Python

MÓDULO 2

20 horas

Manipulación de datos con pandas

- Pandas como librería para manipulación de datos
- Lectura de archivos desde diferentes fuentes, csv, xls, SQL, txt, json y lectura desde url
- Manipulación básica para poder entender los datos que se leyeron en el apartado anterior
- Visualizaciones básicas para entender el comportamiento estadístico básico de los datos

MÓDULO 3

7 horas

Aprendizaje supervisión

- Regresión lineal
- Regresión lineal múltiple
- Regresión logística
- Árboles de decisión

MÓDULO 4

7 horas

Aprendizaje no supervisado

- Clustering
- Análisis de componentes principales
- Kmeans

La Universidad de América se reserva el derecho de ajustar, cambiar o modificar el contenido del programa.



DOCENTE



ÓSCAR RIOJAS



Científico de datos con experiencia en la construcción de modelos de análisis de datos y tecnología blockchain, así como en la creación de algoritmos para extraer valor de los datos, comprensión de los lenguajes de programación para procesar los datos, como Python, R y SQL.

ROSSEMBER EDÉN CÁRDENAS TORRES



Magíster en Ciencias Matemáticas de la Universidad Nacional de Colombia, con experiencia en programación en Python, inteligencia artificial y automatización de procesos. Ha liderado el Semillero de Inteligencia Artificial en la Universidad de América, impulsando proyectos en IA aplicada. Ha desarrollado cursos en Python, Flowise y Make, con enfoque en la integración de modelos de lenguaje para la automatización y optimización de tareas.

CERTIFICADO

Se hará entrega de un Certificado expedido por la Universidad de América correspondiente al programa, a los participantes que asistan a por lo menos el 80% del tiempo total en horas del curso.

De lo contrario, se expedirá una constancia con el número de horas lectivas.





Universidad de
América[®]

Código SNIES 1715



Más información:

Dirección de Mercadeo y Admisiones
educacion.continua@uamerica.edu.co

 **310 560 1538**

EcoCampus de Los Cerros - Avenida Circunvalar No 20 -53
Tel: (60 1) 3376680 opción 1

Sede Norte - Calle 106 No. 19 - 18
Tel: (60 1) 6580658 opción 1

Bogotá D.C., Colombia.

www.uamerica.edu.co

Fundación Universidad de América | Vigilada MinEducación



ISO 9001:2015
ISO 45001:2018
ISO 14001:2015
BUREAU VERITAS
Certification

