



CURSO INTERNACIONAL

Curso Data Science, Big Data & Machine Learning con IA



Clases en tiempo real



56 Horas Académicas



Acerca del Programa

El programa Data Science, Big Data & Machine Learning con IA brinda una formación integral en análisis y gestión de datos como activo estratégico. A través de un enfoque teórico-práctico, los participantes aprenderán a usar herramientas como Python, Pandas, Spark y TensorFlow para diseñar e implementar soluciones de analítica avanzada e inteligencia artificial.

El curso prepara a los profesionales para impulsar la transformación digital y optimizar la toma de decisiones mediante el uso ético y estratégico de los datos.

→ **Objetivos:**

El programa busca desarrollar la capacidad de analizar, procesar y modelar grandes volúmenes de datos mediante herramientas y técnicas de ciencia de datos, machine learning y big data, promoviendo buenas prácticas de visualización y gobernanza.

Los participantes aprenderán a:

- Utilizar Python y fundamentos estadísticos para el análisis de datos.
- Comprender las arquitecturas y tecnologías Big Data.
- Diseñar y entrenar modelos de machine learning y deep learning.
- Aplicar NLP e inteligencia artificial generativa en contextos reales.



→ **Beneficios**

- Formación integral y aplicada
- Dominio de herramientas actuales del ecosistema de datos
- Metodología online en vivo con sesiones prácticas
- Preparación para roles estratégicos en análisis de datos
- Acceso para ingresar al curso por un año.

*Sujeto a la programación del año

→ **Criterios de certificación:**

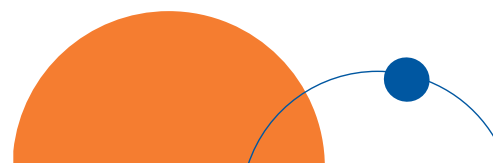
El curso establece como requisitos mínimos para la aprobación que el participante alcance una nota final no menor de doce (12) en la escala vigesimal y acredite una asistencia mínima del setenta y cinco por ciento (75%) a las sesiones programadas. La participación en clase constituye un componente esencial del proceso de aprendizaje.

El cumplimiento de ambos criterios es de carácter obligatorio; en caso contrario, el participante no podrá aprobar el curso, independientemente de las calificaciones obtenidas en las evaluaciones.

→ **Certificación:**



Certificado de participación con validez internacional,
a nombre de New Horizons Corporation





01. Fundamentos de Ciencia de Datos y Programación

Sesión 1: Introducción a la Ciencia de Datos y Python para Data Science

- ¿Qué es la Ciencia de Datos?
- Ciclo de vida de un proyecto de ciencia de datos
- Roles en ciencia de datos (científico de datos, analista, ingeniero, etc.)
- Aplicaciones reales en distintas industrias
- Herramientas y entornos de trabajo
- Python para Data Science
- Instalación y configuración de entorno (Anaconda, Jupyter)
- Sintaxis básica de Python
- Tipos de datos, variables y operadores
- Estructuras de control (condicionales y bucles)
- Funciones y manejo de errores

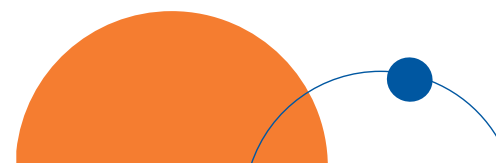
Sesión 2: Librerías clave (NumPy, Pandas, Matplotlib/Seaborn)

- Introducción a NumPy: arrays, operaciones y funciones
- Uso de Pandas: Series, DataFrames, importación/exportación de datos
- Limpieza y transformación de datos
- Introducción a visualización con Matplotlib
- Visualizaciones estadísticas con Seaborn

02. Estadística y Matemáticas Aplicadas

Sesión 3: Estadística y Matemáticas Aplicadas

- Estadística para Ciencia de Datos
- Estadística descriptiva: media, mediana, moda, varianza, desviación estándar
- Visualización de datos estadísticos
- Distribuciones de probabilidad (normal, binomial, etc.)
- Inferencia estadística: estimaciones y pruebas de hipótesis
- Matemáticas para Machine Learning
- Álgebra lineal: vectores, matrices, operaciones básicas
- Cálculo diferencial aplicado a funciones de costo
- Probabilidad y combinatoria
- Conceptos de optimización: gradiente y mínimos locales/globales





03. Fundamentos de Big Data y Arquitecturas

Sesión 4: Fundamentos del Big Data (las 5V y casos de uso)

- Definición de Big Data
- Las 5 V's: Volumen, Velocidad, Variedad, Veracidad y Valor
- Diferencias entre datos estructurados, no estructurados y semiestructurados
- Casos de uso de Big Data en sectores como salud, retail, finanzas, etc.

Sesión 5: Arquitectura de Big Data (Hadoop, Spark, Data Lakes vs Warehouses)

- Componentes del ecosistema Hadoop
- Introducción a Apache Spark: ventajas y casos de uso
- Diferencias entre Data Lakes y Data Warehouses
- Modelos de procesamiento: batch vs streaming

04. Machine Learning

Sesión 6: Aprendizaje Supervisado y no supervisado

Aprendizaje No Supervisado

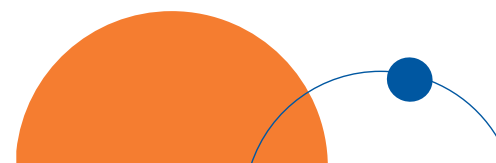
- Concepto de aprendizaje supervisado
- Regresión lineal y regresión logística
- Árboles de decisión y random forests
- Preparación de datos para modelos supervisados

Aprendizaje No Supervisado

- Concepto de aprendizaje no supervisado
- Algoritmos de clustering (K-means, DBSCAN)
- Reducción de dimensionalidad (PCA)
- Casos de uso en segmentación de clientes, análisis exploratorio

Sesión 7: Evaluación de Modelos

- Conjuntos de entrenamiento, validación y prueba
- Métricas de evaluación: accuracy, precision, recall, F1-score, ROC-AUC
- Overfitting vs underfitting
- Validación cruzada
- Ajuste de hiperparámetros (GridSearchCV, RandomizedSearchCV)





05. Deep Learning y Redes Neuronales

Sesión 8: Fundamentos de Redes Neuronales

- ¿Qué es el deep learning?
- Perceptrón y redes neuronales básicas
- Funciones de activación
- Propagación hacia adelante y retropropagación (backpropagation)

Sesión 9: Arquitecturas Avanzadas (CNN, RNN, LSTM)

- Convolutional Neural Networks (CNN): estructura y aplicación en imágenes
- Recurrent Neural Networks (RNN): estructura y uso en datos secuenciales
- LSTM: mejora de RNN para secuencias largas
- Casos de uso prácticos (visión computacional, series temporales)
- "Frameworks populares (TensorFlow, PyTorch, Keras)"

06. Procesamiento de Lenguaje Natural (NLP)

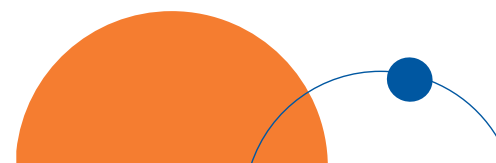
Sesión 10: Procesamiento de Lenguaje Natural (NLP)

- Introducción al NLP y sus aplicaciones
- Preprocesamiento de texto: tokenización, stemming, lematización
- Representación de texto: Bag of Words, TF-IDF, Word Embeddings
- Modelos básicos para clasificación de texto
- Introducción conceptual a transformers y BERT

07. Inteligencia Artificial Aplicada

Sesión 11: Inteligencia Artificial Aplicada (IA generativa, chatbots, casos de uso)

- Diferencias entre Machine Learning, Deep Learning e Inteligencia Artificial
- Introducción a la IA generativa: definición y aplicaciones
- Funcionamiento básico de modelos generativos (p. ej. GPT, DALL·E)
- Construcción y aplicación de chatbots
- Casos de uso reales en distintas industrias





08. Visualización y Business Intelligence

Sesión 12: Herramientas de Visualización y BI

- Principios de visualización efectiva de datos
- Herramientas BI: Power BI, Tableau (introducción y comparativa)
- Creación de dashboards interactivos
- Conexión a fuentes de datos
- Métricas e indicadores clave (KPIs)

09. Ética y Gobernanza en IA

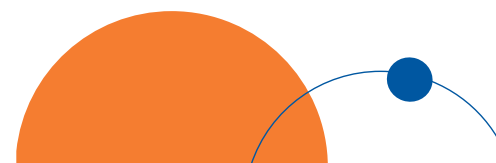
Sesión 13: Ética, Gobernanza y IA Responsable

- Principios éticos en el desarrollo de IA
- Sesgos algorítmicos y sus riesgos
- Marco legal y regulaciones (conceptos generales)
- Transparencia algorítmica y explicabilidad (XAI)
- Gobernanza de datos: privacidad, seguridad y transparencia algorítmica
- Diseño de IA responsable y centrada en el ser humano

10. Proyecto Final Integrador

Sesión 14: Desarrollo de Proyecto

- Definición del problema o caso a resolver
- Recolección, análisis y preparación de datos
- Aplicación de modelos aprendidos
- Visualización de resultados y conclusiones
- Presentación y defensa del proyecto final



BENEFICIOS DE CLASES ONLINE EN VIVO



Online Live

Clases en tiempo real (conéctate desde el lugar que estés)



Acceso a las clases grabadas

Podrás ver las clases grabadas hasta por 90 días



Certificado Internacional

A nombre de New Horizons Corporation



Capacidad

Máximo 20 alumno



Discusiones

Con sus compañeros y el instructor en tiempo real



Informes e inscripciones:



www.newhorizons.edu.pe
940 068 987
Info@newhorizons.edu.pe

New Horizons Perú
RUC: 20306532201
Av. Santa Cruz 870, Miraflores