

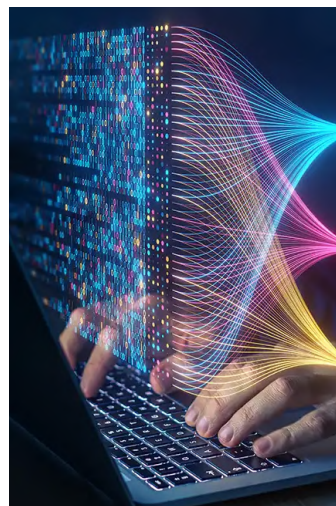
> Especialización en MACHINE LEARNING FOR BIG DATA



¿Por qué llevar esta especialización?

¿Sabías que cuando la data que alimenta a los modelos predictivos escala y se hace masiva, también lo deben hacer las plataformas sobre las que los construimos e inclusive el lenguaje que se emplea?

Por ello, DMC INSTITUTE ha diseñado la Especialización en Machine Learning for Big Data que te permitirá comprender como emplear PySpark, una extensión de Python que opera sobre sistemas distribuidos, para el entrenamiento y puesta en funcionamiento de proyectos robustos de ciencia de datos.



Sobre esta Especialización

08

sesiones

32

horas
académicas

07

talleres
prácticos

¿Cómo impulsamos tu carrera?

- Sesiones 80% enfocadas en la práctica.
- Enfoque en **Casos Aplicados a Negocio**, enfrentando los retos del mercado.
- Énfasis en **habilidades técnicas**.
- **Mentoría especializada** con docentes expertos.
- Acompañamiento **constante**.



Objetivos de la Especialización

OBJETIVO GENERAL:

- Desarrolla algoritmos con el lenguaje PySpark para la implementación de proyectos de Machine Learning, tanto para regresión como para clasificación, que trabajan sobre datos masivos (Big Data).

Requisitos

- Conocimiento en entrenamiento y evaluación de modelos predictivos.
- Manejo de librerías Python para machine learning.
- Mínimo un año de experiencia trabajando en ciencia de datos.
- Contar con una laptop o computadora de escritorio con disponibilidad de micrófono y cámara web.
- Tener instalado los softwares y herramientas señalados en la sección Contenidos.

Dirigido a

- Data Scientist, que actualmente busquen:
 - Evolucionar sus proyectos de ciencia de datos para que puedan operar sobre data masiva (Big Data).
 - Dominar PySpark para el entrenamiento y evaluación de modelos predictivos sobre un cluster Big Data.



Malla Curricular

I. Machine Learning for Big Data

1. Técnicas de clasificación sobre Big Data

- Introducción a PySpark. Características y diferencias funcionales respecto a Python. Overview librerías Spark para machine learning.
- **Taller:** PySpark para un proyecto de clasificación basado en Random Forest.
- **Taller:** PySpark para un proyecto de clasificación basado en Regresión logística monomial.
- **Taller:** PySpark para un proyecto de clasificación basado en Regresión logística multinomial.
- **Taller:** PySpark para un proyecto de clasificación basado en Máquina de soporte vectorial (SVM).
- **Taller:** PySpark para un proyecto de clasificación basado en OneVsRest.

2. Técnicas de regresión sobre Big Data

- **Taller:** PySpark para un proyecto de regresión lineal (Gaussian).
- **Taller:** PySpark para un proyecto de regresión generalizado (GLM).

3. Técnicas avanzadas

- **Taller:** PySpark para un proyecto de regresión basado en Decision-Tree.
- **Taller:** PySpark para un proyecto de regresión basado en RandomForest.
- **Taller:** PySpark para un proyecto de clasificación basado en Gradient-Boost.
- **Taller:** PySpark para un proyecto de regresión basado en Gradient-Boost.
- **Taller:** PySpark para un proyecto de clasificación basado en máquinas de factorización (FM).
- **Taller:** PySpark para un proyecto de regresión basado en máquinas de factorización (FM).

Nuestra Propuesta de Capacitación

Las metodologías que aplicamos



Desarrollo de competencias clave en el mundo de los datos

Analiza • Innova • Transforma



Aprendizaje Secuencial

- Descubre conocimiento de vanguardia
- Explora con la guía del experto
- Aplica lo aprendido



Aprendizaje basado en práctica (Learning by Doing)

- Resuelve retos
- Aprende en base a proyectos
- Analiza casos



¿Qué certificado obtendrás?

- Certificado por aprobación de la Especialización en Machine Learning for Big Data, por un total de 32 horas académicas.



¿Por qué elegirnos?



Somos los primeros en Perú en apostar por el desarrollo de profesionales y empresas en data & analytics con más de 16 años de experiencia.



Las empresas worldclass de Latam confían en nosotros para acompañarlas en su transformación hacia el enfoque data driven.



Nuestros docentes son destacados expertos en data & analytics que lideran equipos de alto rendimiento en las empresas más grandes de Latam.



Nuestra metodología "Aprende haciendo" ha logrado que nuestra comunidad de +25K profesionales en todo Latam mejoren su situación laboral.



Tenemos el portafolio más completo con +150 capacitaciones sincrónicas y asincrónicas que se ajustan a diferentes perfiles y niveles de conocimiento.





Visita nuestra web

www.dmc.pe

Síguenos en:

