



 **UPB**

Formación Continua



**DIPLOMADO** >>

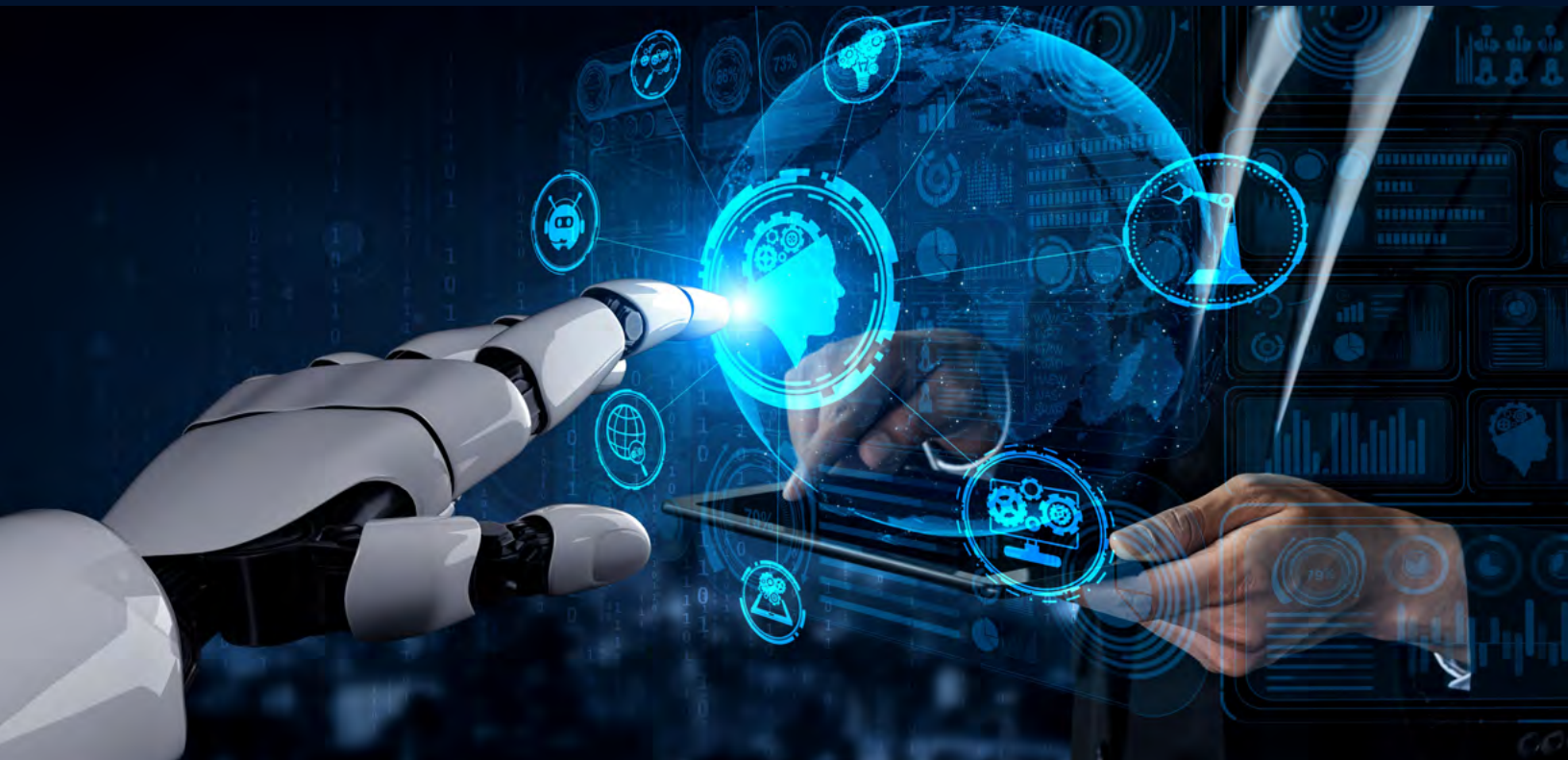
# MACHINE LEARNING Y CIENCIA DE DATOS

---



Modalidad Virtual

La ciencia de datos se ha convertido en una de las habilidades con mayor demanda en los últimos años, dada la disrupción digital actual y dados los grandes volúmenes de datos que manejan las empresas para desarrollar sus principales estrategias. El científico de datos hace que mediante ciertos modelos de machine learning se reduzcan sustancialmente los tiempos de trabajo y se maximicen los resultados de la empresa. Esta implementación debe partir no solo del aprendizaje de técnicas, sino de la implementación de modelos de análisis de datos a la vida real, que permitan agilizar procesos mediante diferentes herramientas analíticas y de programación.



## OBJETIVO GENERAL

Desarrollar herramientas que faciliten el análisis, la predicción y el manejo masivo de datos a través métodos y técnicas relacionadas a la ciencia de datos para optimizar tiempos de trabajo, innovar en procesos y manejo de lenguajes de programación inherentes a la ciencia de datos.

## OBJETIVOS ESPECÍFICOS

Al término del programa los(as) participantes estarán en condiciones de:

- Manejar de forma experta lenguaje R y su entorno R Studio.
- Manejar de forma experta lenguaje Python y todos sus entornos de desarrollo integrado. (Jupyter, Spyder, Google Collaboratory, Pycharm entre otros).
- Realizar modelos predictivos para tomar decisiones de manera óptima en las organizaciones.

- Manejar grandes volúmenes de datos y gestionarlos a través de aplicaciones SQL.
- Trabajar con grandes volúmenes de datos en entornos Hadoop.

## DIRIGIDO A

- Profesionales, consultores y toda persona interesada en aprender sobre ciencia de datos cuantitativos, gestión de grandes volúmenes de datos y automatización de procesos.

## REQUISITOS TÉCNICOS

Todos los participantes deberán contar con:

- Una laptop con procesador i3 o superior.
- Memoria RAM de 4GB o más.
- Acceso a internet.
- Auriculares.

# CONTENIDO DEL CURSO

## Módulo 1:

### **Herramientas básicas de visualización de datos.**

- Visualización de datos con Tableau.
  - Introducción a Tableau.
  - Fundamentos de Bases de datos con Tableau.
- Introducción a sistemas integrados de Data Science con Anaconda Navigator.
  - Introducción a PyCharm Professional.
  - Introducción a Jupyter Notebook.
- Introducción a manejo de base de datos con Python.
  - Creación de variables.
  - Importación de bibliotecas.
  - Creación de variables.

## Módulo 2:

### **Estadística aplicada al análisis de datos.**

- Fundamentos de estadística e inferencia avanzada.
  - Distribuciones de probabilidad.
  - Pruebas de hipótesis.
- Estadística Bayesiana.
  - Principio de verosimilitud.
  - Modelos beta-binomiales.

## Módulo 3:

### **Modelos de predicción analítica.**

- Introducción a manejo de datos con RStudio.
  - Presentación lenguaje R y RStudio.
- Predicción con Regresión Lineal en RStudio.
  - Modelado de variables.
- Análisis multivariado para ciencia de datos.
  - Reducción de dimensionalidad.
  - Descomposición de componentes principales.
- Clustering.
  - Agrupaciones significativas, volatilidad y Outliers.
- Regresiones y modelos de predicción en R Studio.
  - Aplicaciones de Corte transversal en RStudio.
  - Aplicaciones de Series temporales en RStudio.
  - Aplicaciones de Datos de panel en RStudio.

## Módulo 4:

### **Fundamentos de Machine Learning.**

- Introducción a ML (Supervisado, no supervisado, semi supervisado).
  - Parcialidad y varianza.
- Fundamentos matemáticos de ML.
  - Álgebra Básica para ML.
  - Álgebra Lineal para ML.
- Gradient Descent (GD).
  - Regresiones lineales con GD.
  - Regresiones logísticas con GD.
- Estadística bayesiana para ML.
  - Naive Bayes Automation.
  - Discretización y Binning.
- Árboles de clasificación y predicción.
  - Método CART.
  - Árboles de clasificación.
  - Árboles de regresión.

## Módulo 5:

### **Fundamentos de Deep Learning.**

- Fundamentos de Deep Learning.
  - Fundamentos básicos de redes neuronales.
  - Modelos predictivos de redes neuronales.
- Análisis de discriminante lineal (LDA).
  - LDA con un predictor.
  - LDA con varios predictores.
  - Análisis del discriminante cuadrático.
  - Cálculo de la función discriminante.
  - Normalidad univariante y multivariante.
- Learning Vector Quantization.
  - Algoritmo K-means.
  - K-Vecinos más cercanos.
  - Máquinas de soporte vectorial.

## Módulo 6:

### **Fundamentos de Big Data.**

- Data quality y almacenamiento de datos.
  - Etapas en el proceso de data quality.
- Modelado y minería de datos.
  - Modelos de entidad relación.
  - Bases de datos relacionales.
  - Data Warehouse.
- Arquitectura Big Data.
  - Google Cloud Platform para big data.
  - Cloud Dataflow.
- Integración de datos con bigQuery y DataStudio.
  - Cloud Dataflow.
  - Cloud Pub/Sub.
- Big Data en tiempo real.

## Módulo 7:

### Fundamentos de Cloud Computing.

- Fundamentos Básicos de Cloud Computing.
  - Tipos de cómputo: IaaS, PaaS, SaaS.
  - Laboratorio AWS.
- Entornos Hadoop.
- DockerFile.
  - Elementos Hadoop HDFS.
  - YARN y MapReduce.
  - Flume, Sqoop, Hbase y Hive.
- Otros ecosistemas de cloud computing.

## DIRECTOR ACADÉMICO DEL PROGRAMA Y DOCENTE



**Mauro Delboy**  
Céspedes, Ph.D.c.

Graduado con honores en 2013 de la carrera de economía de la Universidad Privada Boliviana. Obtuvo el grado de Magíster en Administración y dirección de empresas por parte de la universidad Santiago de Chile en titulación conjunta con la UPB. Posee dos diplomados, uno en métodos cuantitativos para el análisis económico, y

el otro en Business Intelligence gerencial. Ese mismo año inició su labor profesional en el sector financiero del país, ocupando cargos como el de asesor de inversiones y de operador de rueda en SUDAVAL S.R.L. También trabajó en el departamento de riesgos del Banco Mercantil Santa Cruz S.A. como analista sectorial, para posteriormente incorporarse como docente tiempo completo a la UPB. En 2019 obtuvo la beca de candidatura doctoral gracias al CEsPI (Centro Studi di Politica Internazionale) para obtener el grado de candidato doctoral en el doctorado de Economía de la Universidad Privada Boliviana conjuntamente con la Università degli Studi Roma Tre. Ha publicado varios artículos científicos en la revista Mondopoli y en EJC Journal of Economics and complexity. Actualmente es director ejecutivo de la maestría UPB - USACH, dictando diferentes cátedras y conferencias tanto a nivel de pregrado como posgrado en el área cuantitativa.

## PLANTEL DOCENTE

El plantel docente está conformado por profesionales de reconocido prestigio en su área, los mismos son seleccionados y evaluados, bajo estrictos procedimientos que toman en cuenta metodología didáctica, facilitación del aprendizaje, transmisión de valores y experiencia laboral en la aplicación de tema.



## METODOLOGÍA

Se aplicará una metodología virtual y participativa de transferencia de conocimientos, combinando actividades sincrónicas (videoconferencias) y asincrónicas (plataforma virtual) con las siguientes características:

### Clases en Salas de Videoconferencias

Se realizarán 3 sesiones de videoconferencia por semana de clases, mediante la aplicación Zoom. Estas salas son de uso sencillo y amigable, permitiendo la interacción entre los alumnos y el docente para el desarrollo de actividades colaborativas.

Todas las sesiones de videoconferencias serán grabadas y puestas a disposición de los alumnos en modo lectura en la plataforma virtual, durante la ejecución de los módulos.

### Plataforma Virtual para el Aprendizaje

La plataforma UPB Virtual está diseñada para crear espacios de enseñanza virtual, administrar, distribuir y controlar todas las actividades asincrónicas.

Para acceder al material digital y cumplir con diferentes actividades asignadas por el docente, el participante podrá ingresar a la plataforma virtual, sin restricción de horarios.

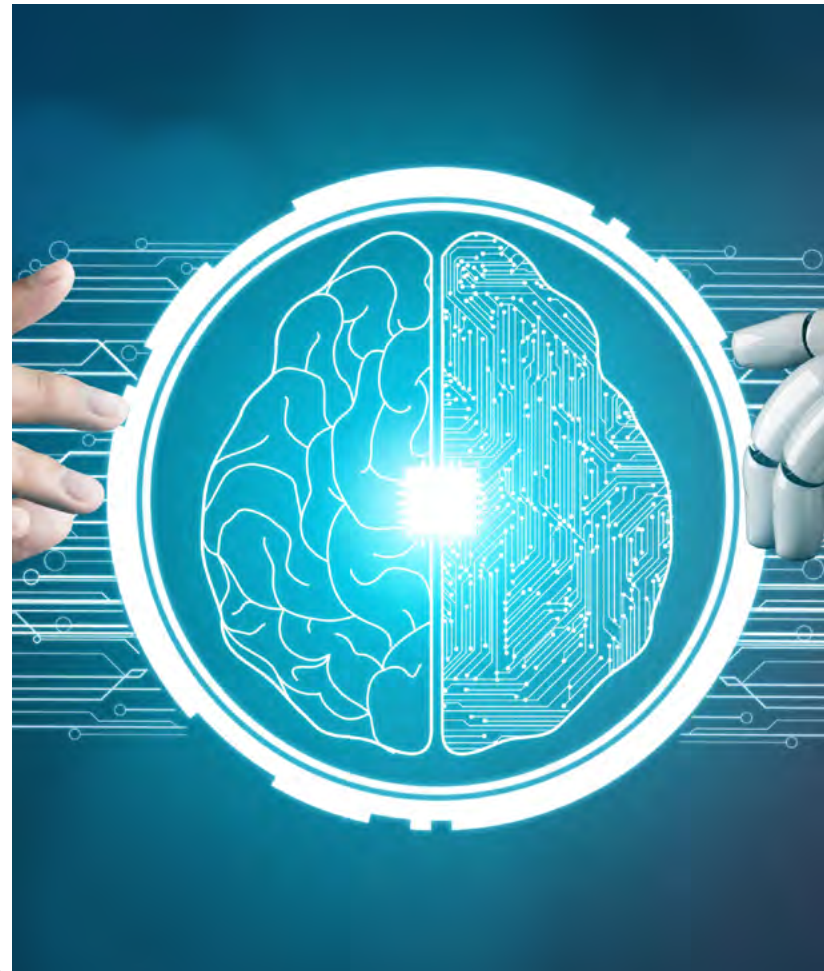
Para la transferencia de conocimientos el docente podrá hacer uso de distintos recursos: foros de trabajo colaborativo, foros de discusión temática, foros de preguntas y respuestas, mensajería interna, trabajos, cuestionarios, tests y otros servicios orientados a lograr un aprendizaje efectivo.

## CARACTERÍSTICAS DEL PROGRAMA

**Duración:** Cinco meses. 450 horas académicas.

**Titulación:** Al cumplir con los requisitos académicos de aprobación de los módulos el(la) participante obtendrá el certificado de "Diplomado en Machine Learning y Ciencia de Datos".

Solicite al contacto regional información actualizada, referida a precio, descuentos, planes de pago, resumen de la experiencia del plantel docente, horarios, cronograma tentativo y proceso de inscripción.



Management  
System  
ISO 9001:2015  
www.tuv.com  
ID 3108623133

Patrocinada por:



### Informaciones e Inscripciones:

☎ 675 88819

✉ mzambrana@upb.edu

### CONTACTO NACIONAL:

✉ formacioncontinua@upb.edu

📘 UPB Formación Continua Central

🌐 www.upb.edu/ceu