

CURSO de ESPECIALIZACIÓN en: **BIG DATA, DATA SCIENCE E INTELIGENCIA ARTIFICIAL**

MÓDULOS	BIG DATA Y ARQUITECTURA DE DATOS CON POWER BI	4 semanas
	PYTHON PARA CIENCIA DE DATOS Y MACHINE LEARNING	4 semanas
	INTELIGENCIA ARTIFICIAL GENERATIVA Y APLICACIONES	4 semanas
Requisitos	Conocimientos básicos de computación.	
Duración	3 MESES (Martes, Jueves y Viernes)	

MÓDULO 1. BIG DATA y ARQUITECTURA DE DATOS CON POWER BI

OBJETIVO

Comprender los fundamentos de Big Data y las arquitecturas modernas de datos para integrar, almacenar, procesar y transformar información, desarrollando soluciones de Business Intelligence mediante modelos de datos, indicadores y dashboards interactivos.

CONTENIDO

1. Fundamentos de Big Data y Ciencia de Datos

- 1.1. Definición, evolución y características de Big Data
- 1.2. Las 4 V: Volumen, Velocidad, Variedad y Veracidad
- 1.3. Big Data, Ciencia de Datos y Business Intelligence
- 1.4. Tipos y fuentes de datos
- 1.5. Casos de uso de Big Data y analítica empresarial

2. Arquitectura y almacenamiento de datos

- 2.1. Bases de datos relacionales y NoSQL
- 2.2. Data Warehouse, Data Lake y Data Lakehouse
- 2.3. Arquitecturas modernas de datos
- 2.4. Almacenamiento distribuido
- 2.5. Introducción a MongoDB
- 2.6. Almacenamiento de datos en la nube

3. Procesamiento e integración de datos

- 3.1. Conceptos de procesamiento distribuido
- 3.2. Introducción a Apache Spark

- 3.3. Conceptos básicos de PySpark
- 3.4. ETL y ELT
- 3.5. Fuentes y destinos de datos
- 3.6. Diseño de pipelines de datos
- 3.7. Calidad, limpieza y transformación de datos
- 4. Data Warehouse y modelado dimensional**
 - 4.1. Fundamentos de Data Warehouse
 - 4.2. Arquitectura de un Data Warehouse
 - 4.3. Modelo dimensional
 - 4.4. Modelo estrella
 - 4.5. Tablas de hechos y dimensiones
 - 4.6. Relaciones y cardinalidad
 - 4.7. Métricas e indicadores
 - 4.8. Diseño de modelos orientados al análisis
- 5. Business Intelligence con Power BI**
 - 5.1. Fundamentos de Business Intelligence
 - 5.2. Power BI y flujo de trabajo
 - 5.3. Conexión e importación de datos
 - 5.4. Power Query
 - 5.5. Limpieza y transformación de información
 - 5.6. Modelado de datos en Power BI
 - 5.7. DAX, medidas y columnas calculadas
 - 5.8. Indicadores y KPIs
- 6. Visualización y aplicación de Business Intelligence**
 - 6.1. Diseño de dashboards interactivos
 - 6.2. Gráficos, filtros, segmentadores y jerarquías
 - 6.3. Drill-down y navegación
 - 6.4. Publicación y presentación de informes
 - 6.5. Introducción a Looker Studio
 - 6.6. Comparación de herramientas de visualización
 - 6.7. Análisis e interpretación de resultados
 - 6.8. Desarrollo de un dashboard empresarial

MÓDULO 2. PYTHON para CIENCIA DE DATOS y MACHINE LEARNING

OBJETIVO

Aplicar Python para la preparación, análisis y visualización de datos, desarrollando modelos de Machine Learning y Deep Learning para identificar patrones, realizar predicciones y resolver problemas de análisis de datos.

CONTENIDO

1. Python para Ciencia de Datos

- 1.1 Fundamentos de Python
- 1.2 Variables y tipos de datos
- 1.3 Listas, tuplas, diccionarios y conjuntos
- 1.4 Funciones y módulos
- 1.5 Manejo de archivos y datos
- 1.6 Librerías y entornos para Ciencia de Datos

2. NumPy y Pandas

- 2.1 Arrays y operaciones con NumPy
- 2.2 Series y DataFrames
- 2.3 Importación de datos
- 2.4 Selección y filtrado de información
- 2.5 Ordenamiento y agrupación
- 2.6 Limpieza y transformación de datos
- 2.7 Valores nulos y datos inconsistentes
- 2.8 Preparación de datasets para análisis

3. Análisis Exploratorio y Visualización de Datos

- 3.1 Estadística descriptiva aplicada
- 3.2 Distribuciones y comportamiento de los datos
- 3.3 Identificación de patrones y valores atípicos
- 3.4 Correlaciones y relaciones entre variables
- 3.5 Visualización con Matplotlib y Seaborn
- 3.6 Histogramas, boxplots y gráficos de dispersión
- 3.7 Interpretación de resultados

4. Fundamentos de Machine Learning

- 4.1 Conceptos y aplicaciones de Machine Learning
- 4.2 Aprendizaje supervisado y no supervisado
- 4.3 Flujo de trabajo de un proyecto de Machine Learning
- 4.4 Preparación de datos para modelos
- 4.5 División de datos para entrenamiento y evaluación

- 4.6 Selección de características
- 4.7 Entrenamiento y validación de modelos

5. Modelos de Machine Learning

- 5.1 Regresión lineal
- 5.2 Regresión logística
- 5.3 Algoritmos de clasificación
- 5.4 Árboles de decisión
- 5.5 Evaluación y comparación de modelos
- 5.6 Interpretación de resultados

6. Redes Neuronales y Deep Learning

- 6.1 Fundamentos de redes neuronales artificiales
- 6.2 Neuronas, capas y funciones de activación
- 6.3 Procesos de entrenamiento
- 6.4 Introducción al Deep Learning
- 6.5 Arquitecturas básicas de redes neuronales
- 6.6 Preparación de datos para Deep Learning
- 6.7 Evaluación de modelos de Deep Learning
- 6.8 Aplicaciones de Deep Learning

7. Aplicación práctica

- 7.1 Preparación y exploración de un dataset
- 7.2 Desarrollo de un modelo de Machine Learning
- 7.3 Desarrollo de un modelo básico de Deep Learning
- 7.4 Evaluación de resultados
- 7.5 Comparación de modelos
- 7.6 Interpretación y presentación de resultados

MÓDULO 3. INTELIGENCIA ARTIFICIAL GENERATIVA Y APLICACIONES

OBJETIVO

Comprender las arquitecturas de datos en la nube y utilizar plataformas modernas para el almacenamiento, procesamiento y análisis de información, integrando capacidades de Inteligencia Artificial y modelos generativos en soluciones orientadas a datos.

CONTENIDO

1. Cloud Computing aplicado a Datos

- 1.1 Conceptos de computación en la nube
- 1.2 Modelos IaaS, PaaS y SaaS
- 1.3 Arquitecturas Cloud para datos

- 1.4 Servicios Cloud para almacenamiento y procesamiento
- 1.5 Servicios de datos en AWS, Azure y Google Cloud
- 1.6 Seguridad y gestión de datos en entornos Cloud
- 2. Data Lakehouse y arquitecturas modernas**
 - 2.1 Evolución de Data Warehouse a Data Lakehouse
 - 2.2 Arquitectura Data Lakehouse
 - 2.3 Componentes y capas de una plataforma de datos
 - 2.4 Integración de datos estructurados y no estructurados
 - 2.5 Procesamiento y almacenamiento en Lakehouse
 - 2.6 Calidad y gobernanza de datos
 - 2.7 Preparación de datos para analítica e Inteligencia Artificial
- 3. Databricks y procesamiento de datos**
 - 3.1 Introducción a Databricks
 - 3.2 Entorno de trabajo y notebooks
 - 3.3 Ingesta y procesamiento de datos
 - 3.4 Apache Spark en Databricks
 - 3.5 Delta Lake
 - 3.6 Construcción de un Lakehouse
 - 3.7 Transformación y análisis de datos
 - 3.8 Preparación de datos para Machine Learning e IA
- 4. Fundamentos de Inteligencia Artificial**
 - 4.1 Historia, definición y evolución de la Inteligencia Artificial
 - 4.2 Inteligencia Artificial, Machine Learning y Deep Learning
 - 4.3 Tipos y áreas de aplicación de la IA
 - 4.4 IA aplicada al análisis de datos
 - 4.5 Automatización inteligente
 - 4.6 Casos de uso empresariales
- 5. Inteligencia Artificial Generativa**
 - 5.1 Concepto y características de la IA Generativa
 - 5.2 Modelos fundacionales
 - 5.3 Modelos de lenguaje de gran escala (LLM)
 - 5.4 Generación y transformación de texto
 - 5.5 Generación y análisis de información
 - 5.6 IA multimodal
 - 5.7 Aplicaciones empresariales de IA Generativa
- 6. Integración de IA con Plataformas de Datos**

- 6.1 IA aplicada al procesamiento y análisis de datos
- 6.2 Integración de modelos de IA con plataformas Cloud
- 6.3 IA aplicada a Data Lakehouse
- 6.4 IA aplicada a Business Intelligence
- 6.5 IA aplicada a Ciencia de Datos
- 6.6 Automatización inteligente de procesos
- 6.7 Integración de IA en flujos de trabajo
- 6.8 Evaluación y validación de resultados
- 6.9 Ética, privacidad, seguridad y uso responsable

7. Aplicación práctica

- 7.1 Integración de datos en un entorno Cloud
- 7.2 Construcción de una solución básica en Databricks
- 7.3 Preparación de datos para analítica e IA
- 7.4 Integración de capacidades de IA Generativa
- 7.5 Automatización de una tarea basada en datos
- 7.6 Presentación y evaluación de resultados