

Máster Universitario en Big Data para los Negocios

ASIGNATURAS

1º Cuatrimestre

1. Programación para Procesamiento de Big Data con Lenguajes Distribuidos (6 ECTS/1Q)

- Paradigma de computación distribuida para Big Data: conceptos clave (paralelismo, distribución, procesamiento tradicional vs distribuido), batch vs streaming.
- Spark como motor de procesamiento: arquitectura de Spark, RDDs, modelo de transformaciones/acciones, DataFrames y funciones avanzadas.
- Programación con PySpark: programación Spark en Python, DataFrames, transformaciones y acciones.
- Modelos de procesamiento: batch y tiempo real.
- Lectura/escritura desde sistemas externos: conexión con fuentes de datos, integración con NoSQL y Data Lakes, lectura/escritura desde PySpark.

2. Adquisición, Ingesta y Procesamiento de información para Big Data (6 ECTS/1Q)

- Ciclo de vida del dato y arquitectura de ingesta: ciclo de vida del dato, arquitectura de pipelines y workflows.
- Fuentes y tipos de datos: APIs, sensores, logs, bases de datos y archivos; datos estructurados, semiestructurados y no estructurados.
- Formatos y estructuras de intercambio: CSV, JSON, Avro, Parquet, serialización.
- Tecnologías de ingesta: batch, streaming, Kafka, APIs, colas de mensajes y conectores.
- Automatización de la ingesta: scheduling, orquestación y despliegue básico de pipelines.

3. Impacto del Big Data en los Negocios y las Organizaciones (6 ECTS/1Q).

- Big Data y creación de valor: generación de valor, toma de decisiones, evolución del Big Data, BI y analítica avanzada, roles de datos.
- Arquitectura de datos en la empresa: ingesta, almacenamiento (Data Lake, Data Warehouse, Lakehouse), procesamiento ETL/ELT, aprendizaje automático e infraestructura.
- Casos de uso empresariales: marketing, industria 4.0, finanzas, retail, salud y sector público.
- Futuro del Big Data: IA generativa, vector search y automatización.

4. Plataformas Tecnológicas para Big Data: Almacenamiento (6 ECTS/1Q)

- Fundamentos del almacenamiento en Big Data: sistemas distribuidos y almacenamiento masivo.
- Modelos y formatos de almacenamiento: Parquet, ORC y organización de datos.
- Arquitecturas modernas: Data Lake, Data Warehouse y Lakehouse.
- Tecnologías de almacenamiento: SQL, NoSQL, almacenamiento de objetos, arquitecturas híbridas y criterios de selección.
- Gobierno del dato: metadatos, catálogo de datos y capa semántica.



5. Plataformas Tecnológicas para Big Data: Infraestructuras y Arquitectura (6 ECTS/1Q)

- Arquitecturas distribuidas para Big Data: principios de sistemas distribuidos y arquitecturas escalables.
- Cloud Computing (IaaS/PaaS/SaaS): servicios cloud, arquitecturas gestionadas y diseño de sistemas en la nube.
- Infraestructura para sistemas de datos: virtualización, nube vs local, escalabilidad horizontal.
- Contenedores y microservicios: Docker, Kubernetes e Infraestructura como Código (IaC).
- Orquestación y recursos: clústeres, supervisión, monitorización y optimización de recursos.

2º Cuatrimestre

6. Analítica Avanzada de Datos (3 ECTS/2Q).

- Introducción al análisis de datos: aplicación al negocio, estadística vs aprendizaje automático, metodologías de ciencia de datos.
- Análisis descriptivo: calidad, auditoría, análisis univariado y multivariado.
- Análisis supervisado I (Regresión): preparación de datos, regresión lineal y no lineal, evaluación.
- Análisis supervisado II (Clasificación): preparación de datos y modelos de clasificación.
- Análisis no supervisado (Clustering): preparación de datos, algoritmos de agrupamiento y evaluación.

7. Ciberseguridad y Privacidad en Plataformas de Datos (3 ECTS/2Q).

- Seguridad y privacidad en Big Data: ciclo de vida del dato, riesgos y seguridad de sistemas.
- Privacidad y protección del dato: anonimización, privacidad y criptografía.
- Control de acceso: IAM, RBAC, ABAC y cifrado.
- Seguridad en arquitecturas Big Data: detección de amenazas, SIEM y seguridad por diseño.
- Cumplimiento normativo: GDPR, SOC2, HIPAA y gestión de riesgos.

8. Visualización avanzada para Big Data (6 ECTS/2Q).

- Fundamentos de visualización: data storytelling, percepción visual y UX.
- Desafíos con datos masivos: escalabilidad, muestreo, agregación y latencia.
- Herramientas de visualización: Power BI, Looker y visualización programática.
- Dashboards y analítica operativa: KPIs y cuadros de mando.
- Visualización en tiempo real: streaming y dashboards en tiempo real.



9. Gestión de Proyectos Big Data (3 ECTS/2Q).

- Estrategia de datos en la empresa: impacto del dato, estrategia empresarial y roles.
- Gestión de proyectos de datos: tipos de proyectos, problemas típicos y ciclo de vida.
- Metodología Scrum, CRISP-DM y adaptación a proyectos Big Data.
- Gobierno del dato y responsabilidad en IA: oficina de dato, gobierno del dato, sesgos y responsabilidad en IA.

10. Modelo DevOps de Integración Continua (3 ECTS/2Q).

- Introducción a DevOps, DataOps y MLOps: conceptos, objetivos, cultura y organización.
- CI/CD para datos: control de versiones, integración y despliegue continuo.
- Automatización de workflows: orquestación, trazabilidad y monitorización.
- Introducción a MLOps: ciclo de vida de modelos ML, versionado, despliegue y monitorización.

11. Prácticas Académicas Externas (6 ECTS/2Q).

- El contenido de las prácticas externas se basará en el desarrollo laboral en una entidad colaboradora vinculada mediante convenio con la universidad. El tema se concretará antes del inicio de la estancia y podrá estar relacionado con distintos ámbitos profesionales. Las prácticas tendrán carácter presencial y se desarrollarán en el centro de trabajo designado.

12. Trabajo Fin de Máster (6 ECTS/2Q).

- Desarrollo de un trabajo original e individual para verificar la adquisición de competencias generales y específicas del título. El trabajo debe abordar un problema o caso real relacionado con el análisis de datos masivos para los negocios, integrando conocimientos, habilidades y competencias adquiridas durante el máster.



COMPLEMENTOS FORMATIVOS:

Solo a cursar en caso de que el requisito legal de acceso lo requiera.

13. Estructura y análisis de datos. Bases de datos relacionales y lenguaje SQL

- Introducción a las bases de datos relacionales
- Fundamentos de SQL
- Consultas básicas
- Funciones y agregaciones
- Joins y relaciones entre tablas
- Subconsultas y operaciones avanzadas
- Manipulación de datos

14. Fundamentos computacionales para Big Data

- Arquitectura de computadores
- Sistemas operativos
- Redes de comunicación
- Algoritmia y estructuras de datos básicas
- Infraestructura para datos

15. Fundamentos de programación

- Fundamentos de Programación en Python: Sintaxis, estructuras de control y estructuras de datos.
- Python para análisis y manipulación de datos: Lectura/escritura, limpieza básica, funciones, librerías principales (NumPy, Pandas, visualización básica)
- Bases de datos relacionales y SQL: agregaciones, joins, subconsultas y operaciones avanzadas. Insert/Update/Delete y creación/modificación de tablas.

