

**PROGRAMACIÓN DIDÁCTICA**

**CURSO 2023-2024**

**MÓDULO BIG DATA APLICADO**



**IES  
CO  
MER  
CIO**

**FAMILIA PROFESIONAL INFORMÁTICA Y COMUNICACIONES**

**CURSO DE ESPECIALIZACIÓN EN INTELIGENCIA ARTIFICIAL Y BIG DATA**

**PROFESORES:**

**VESPERTINO: Rafael Juan Alamañac  
Garrido**

## ÍNDICE

1. INTRODUCCIÓN	3
2. OBJETIVOS	3
2.1 Competencia general del Curso	3
2.2 Competencias profesionales, personales y sociales del módulo	3
2.3 Objetivos generales del curso que contribuye a alcanzar el módulo	4
2.4 Objetivos del módulo	5
3. CONTENIDOS Y DISTRIBUCIÓN TEMPORAL	5
3.1 Contenidos básicos	6
3.2 Contenidos actitudinales	7
3.3 Distribución temporal	8
4. UNIDADES DIDÁCTICAS	8
5. METODOLOGÍA	15
5.1 Materiales y recursos didácticos	15
6. EVALUACIÓN	16
6.1 Criterios de evaluación	16
6.2 Instrumentos y procedimientos de evaluación	18
6.3 Criterios de calificación	18
6.4 Criterios de recuperación	20
6.5 Actividades de refuerzo o recuperación	21
6.6 Recuperación de módulos pendientes	21
7. ATENCIÓN AL ALUMNADO CON NECESIDADES ESPECÍFICAS DE APOYO EDUCATIVO	21
8. ACTIVIDADES COMPLEMENTARIAS Y EXTRAESCOLARES	21
9. NECESIDADES Y PROPUESTAS DE FORMACIÓN DEL PROFESORADO	21
10. BIBLIOGRAFÍA DE REFERENCIA	22

## 1. INTRODUCCIÓN

El módulo profesional Big Data Aplicado, cuyo código es 5075, está regulado por el Real Decreto 279/2021, de 20 de abril, por el que se establece el Curso de especialización en Inteligencia Artificial y Big Data y se fijan los aspectos básicos del currículo.

Este módulo tiene asignadas 75 horas a razón de 6 horas semanales y equivale a 8 créditos ECTS.

## 2. OBJETIVOS

### 2.1 Competencia general del Curso

La **competencia general** para el Curso de Especialización de Big Data e Inteligencia Artificial es:

*Programar y aplicar sistemas inteligentes que optimizan la gestión de la información y la explotación de datos masivos, garantizando el acceso a los datos de forma segura y cumpliendo los criterios de accesibilidad, usabilidad y calidad exigidas en los estándares establecidos, así como los principios éticos y legales.*

### 2.2 Competencias profesionales, personales y sociales del módulo

Las competencias profesionales, personales y sociales de este módulo son las que se relacionan a continuación:

b) Desarrollar e implementar sistemas de Inteligencia Artificial que faciliten la toma de decisiones ágiles dentro de un negocio gestionando y explotando datos masivos.

c) Gestionar la transformación digital necesaria en las organizaciones para la consecución de la eficiencia empresarial mediante el tratamiento de datos.

h) Integrar sistemas de explotación de grandes volúmenes de datos aplicando soluciones de Big Data.

i) Implantar las funcionalidades, procesos y sistemas de decisiones empresariales aplicando técnicas de Big Data en ellos.

j) Ejecutar el sistema de explotación de datos según las necesidades de uso y las condiciones de seguridad establecidas asegurando el cumplimiento de los principios legales y éticos.

k) Configurar las herramientas que se usan para construir soluciones Big Data y de Inteligencia Artificial.

l) Gestionar de manera eficiente los datos, la información y su representación para transformarlos en conocimiento.

m) Cumplir la legislación vigente que regula la normativa de los medios de comunicación audiovisual y de la accesibilidad universal.

n) Adaptarse a las nuevas situaciones laborales, manteniendo actualizados los conocimientos científicos, técnicos y tecnológicos relativos a su entorno profesional, gestionando su formación y los recursos existentes en el aprendizaje a lo largo de la vida y utilizando las tecnologías de la información y la comunicación.

ñ) Resolver situaciones, problemas o contingencias con iniciativa y autonomía en el ámbito de su competencia, con creatividad, innovación y espíritu de mejora en el trabajo personal y en el de los miembros del equipo.

o) Organizar y coordinar equipos de trabajo con responsabilidad, supervisando el desarrollo del mismo, manteniendo relaciones fluidas y asumiendo el liderazgo, así como aportando soluciones a los conflictos grupales que se presenten.

p) Comunicarse con sus iguales, superiores, clientela y personas bajo su responsabilidad, utilizando vías eficaces de comunicación, transmitiendo la información o conocimientos adecuados y respetando la autonomía y competencia de las personas que intervienen en el ámbito de su trabajo.

q) Generar entornos seguros en el desarrollo de su trabajo y el de su equipo, supervisando y aplicando los procedimientos de prevención de riesgos laborales y ambientales, de acuerdo con lo establecido por la normativa y los objetivos de la empresa.

r) Supervisar y aplicar procedimientos de gestión de calidad, de accesibilidad universal y de «diseño para todas las personas», en las actividades profesionales incluidas en los procesos de producción o prestación de servicios.

s) Ejercer sus derechos y cumplir con las obligaciones derivadas de su actividad profesional, de acuerdo con lo establecido en la legislación vigente, participando activamente en la vida económica, social y cultural.

### **2.3 Objetivos generales del curso que contribuye a alcanzar el módulo**

Los objetivos generales de que este módulo contribuye a alcanzar son los siguientes:

b) Seleccionar datos relevantes de la empresa u organización para desarrollar e implementar soluciones que faciliten la toma de decisiones.

c) Aplicar técnicas de tratamiento de datos para gestionar la transformación digital en las organizaciones.

h) Utilizar soluciones de Big Data para integrar sistemas de explotación de datos.

i) Analizar y evaluar soluciones Big Data para su implantación en las funcionalidades, procesos y sistemas de decisiones.

j) Determinar la documentación técnica y normativa vigente de los procedimientos de protección de datos para ejecutar el sistema de explotación de datos cumpliendo con los principios legales y éticos.

k) Determinar la solución de Inteligencia Artificial y Big Data para configurar las herramientas y lenguajes específicos.

l) Aplicar técnicas Big Data para gestionar los datos de la organización y obtener conocimiento a partir de ellos.

m) Analizar y utilizar los recursos y oportunidades de aprendizaje relacionados con la evolución científica, tecnológica y organizativa del sector y las tecnologías de la información y la comunicación, para mantener el espíritu de actualización y adaptarse a nuevas situaciones laborales y personales.

n) Desarrollar la creatividad y el espíritu de innovación para responder a los retos que se presentan en los procesos y en la organización del trabajo y de la vida personal.

ñ) Evaluar situaciones de prevención de riesgos laborales y de protección ambiental, proponiendo y aplicando medidas de prevención personal y colectiva, de acuerdo con la normativa aplicable en los procesos de trabajo, para garantizar entornos seguros.

o) Identificar y proponer las acciones profesionales necesarias, para dar respuesta a la accesibilidad universal, al «diseño para todas las personas», así como para evitar posibles sesgos de género en el desarrollo y aplicaciones de Inteligencia Artificial y Big Data.

p) Identificar y aplicar parámetros de calidad en los trabajos y actividades realizados en el proceso de aprendizaje, para valorar la cultura de la evaluación y de la calidad y ser capaces de supervisar y mejorar procedimientos de gestión de calidad.

## 2.4 Objetivos del módulo

Los resultados de aprendizaje son los objetivos del módulo:

1. Gestiona soluciones a problemas propuestos, utilizando sistemas de almacenamiento y herramientas asociadas al centro de datos.
2. Gestiona sistemas de almacenamiento y el amplio ecosistema alrededor de ellos facilitando el procesamiento de grandes cantidades de datos sin fallos y de forma rápida.
3. Genera mecanismos de integridad de los datos, comprobando su mantenimiento en los sistemas de ficheros distribuidos y valorando la sobrecarga que conlleva en el tratamiento de los datos.
4. Realiza el seguimiento de la monitorización de un sistema, asegurando la fiabilidad y estabilidad de los servicios que se proveen.
5. Valida las técnicas de Big Data para transformar una gran cantidad de datos en información significativa, facilitando la toma de decisiones de negocios.

## 3. CONTENIDOS Y DISTRIBUCIÓN TEMPORAL

Los contenidos se organizan en 6 unidades de trabajo

UT 1. Sistemas de Almacenamiento.

UT 2. Acceso a datos.

UT 3. Introducción Hadoop.

UT 4. Ecosistema Hadoop.

UT 5. Administración y monitorización de sistemas.

UT 6. Técnicas Big Data toma decisiones BI.

La relación de unidades didácticas o de trabajo, con sus títulos y las distintas partes de las mismas (objetivos, contenidos, actividades y recursos) viene desarrollada en el punto 4 de la presente programación didáctica. Al ser la duración de algunas de las unidades de trabajo de

más de dos horas, en el momento de impartirlas en clase se dividirán en subunidades con determinados contenidos.

### 3.1 Contenidos básicos

Gestión de soluciones con sistemas de almacenamiento y herramientas del centro de datos para la resolución de problemas:

- Almacenamiento de datos masivo.
- Procesamiento de datos.
- Analítica de Big Data en los ecosistemas de almacenamiento.
- Big Data y Cloud.

Gestión de sistemas de almacenamiento y ecosistemas Big Data:

- Computación distribuida. Computación paralela,
- Sistemas de almacenamiento distribuidos. Tolerancia a fallos.
- Herramientas:
  - Map Reduce.
  - Pig, Hive, Flume.
  - Sqoop, Oozie.
  - Automatización de Jobs.
  - Consultas Pig y Hive.
  - Otras herramientas.

Generación de mecanismos de Integridad de los datos. Comprobación de mantenimiento de sistemas de ficheros:

- Calidad de los datos.
- Comprobación de la integridad de datos de los sistemas de ficheros distribuidos. Sumas de verificación.
- Movimiento de datos entre clusters. Actualización y migración. Metadatos.

Monitorización, optimización y solución de problemas:

- Herramientas de monitorización: Interfaz web del Jobtracker y Namenode, entre otras.
- Análisis de los históricos.
- Monitorización del clúster: Ganglia, entre otros.

Validación de técnicas Big Data en la toma de decisiones en Inteligencia de negocios BI:

- Modelos de Inteligencia de negocios.
- Proceso del modelo KDD (Knowledge Discovery in Databases).

- Etapas: Selección, limpieza, transformación de datos, minería de datos, interpretación y evaluación de datos.
- Implantación de modelos de inteligencia de negocios BI.
- Técnicas de validación de modelos BI.

### 3.2 Contenidos actitudinales

Los contenidos actitudinales, merecen especial atención ya que son necesarios para poder integrarse en el mundo laboral. Es por ello que existen algunas actitudes asociadas al comportamiento y a la realización de trabajo, de forma individual o en grupo, cuya adquisición se ha de contemplar y fomentar en el desarrollo de todas las unidades didácticas programadas para este módulo:

#### Actitudes Personales:

- Aceptar y cumplir el reglamento interno del Instituto
- Aceptar y cumplir las normas de comportamiento y trabajo establecidas durante el curso
- Utilizar los equipos y programas informáticos cumpliendo las normas establecidas, las de seguridad e higiene y los requisitos legales
- Mantener su puesto de trabajo en perfecto estado
- Respetar y valorar la utilización de técnicas y procedimientos para mantener la seguridad, integridad y privacidad de la información
- Ser puntual
- Participar activamente en los debates y en los grupos de trabajo
- Valorar la evolución de la técnica para adaptarse al puesto de trabajo
- Interesarse por la formación permanente en cuestiones relacionadas con su trabajo
- Perseverar en la búsqueda de soluciones
- Valorar la constancia y el esfuerzo propio y ajeno en la realización del trabajo. Querer aprender y mejorar.
- Demostrar interés, participar, realizar aportaciones y comprometerse con el desarrollo del módulo.
- Mostrar interés por la utilización correcta del lenguaje
- Realizar su trabajo personal de forma autónoma y responsable. No apropiarse del trabajo ajeno.
- Saber rodearse de los materiales necesarios para desarrollar correctamente su trabajo. Traer siempre el material necesario.
- Responsabilizarse de la ejecución de su propio trabajo y de los resultados obtenidos
- Orden y método en la realización de tareas
- El esmero, la pulcritud y la puntualidad en la entrega de actividades. Evitar las faltas de ortografía y cuidar la redacción.
- Demostrar interés por la conclusión total de un trabajo antes de comenzar el siguiente.

**Actitudes relacionales:**

- Respeto por otras opiniones, ideas y conductas. Saber estar en todos los sentidos.
- Tener conciencia de grupo, integrándose en un grupo de trabajo, participando activamente en las tareas colectivas y respetando las opiniones ajenas
- Respetar la ejecución del trabajo ajeno en el grupo, compartiendo las responsabilidades derivadas del trabajo global
- Valorar el trabajo en equipo como el medio más eficaz para la realización de ciertas actividades
- Mantener actitudes de solidaridad y compañerismo.

**3.3 Distribución temporal**

Este módulo tiene asignadas 140 horas anuales, distribuidas a razón de 6 horas semanales.

UT 1. Sistemas de Almacenamiento. (18 horas)

UT 2. Acceso a datos. (32 horas)

UT 3. Introducción Hadoop. (12 horas)

UT 4. Ecosistema Hadoop. (38 horas)

UT 5. Administración y monitorización de sistemas. (26 horas)

UT 6. Técnicas Big Data toma decisiones BI. (14 horas)

**4. UNIDADES DIDÁCTICAS**

POR CADA UNIDAD DIDÁCTICA SE PRECISARÁ:

- o OBJETIVOS A ALCANZAR
- o CONTENIDOS QUE SERÁN DESARROLLADOS
- o ACTIVIDADES DE ENSEÑANZA-APRENDIZAJE Y DE EVALUACIÓN
- o RECURSOS NECESARIOS PARA SU REALIZACIÓN

**UT1. Sistemas de Almacenamiento. (18 horas)****Objetivos:**

- Comprender los fundamentos de las bases de datos.
- Diferenciar entre bases de datos relacionales y no relacionales.
- Dominio de SQL.
- Consultas a BBDD.
- Dominio de las bases de datos NoSQL.
- Modelado de datos para NoSQL.
- Análisis comparativo de tipos de bases de datos.

**Contenidos:**

I. Introducción a las bases de datos

A. Definición de las bases de datos



- B. Importancia de las Bases de Datos en el Big Data
  - C. Tipos de bases de datos (relacionales y no relacionales)
- II. Bases de datos relacionales
- A. Comprensión de los modelos de datos relacionales
  - B. SQL (Lenguaje de Consulta Estructurado)
    - Conceptos básicos de SQL
    - Escritura de Consultas SQL
  - C. Sistemas populares de gestión de bases de datos relacionales (RDBMS)
    - MySQL
    - SQL Server
- III. Bases de datos no relacionales
- A. Introducción a las bases de datos no relacionales (NoSQL)
  - B. Tipos de bases de datos NoSQL
    - Almacenes de documentos (por ejemplo, MongoDB)
    - Almacenes de Clave-Valor (ej. Redis)
    - Almacenes de columnas (por ejemplo, Cassandra)
    - Bases de datos de grafos (por ejemplo, Neo4j)
  - C. Modelado de datos en bases de datos NoSQL
  - D. Consulta de datos en bases de datos NoSQL
- IV. Comparación entre bases de datos relacionales y no relacionales
- A. Ventajas y desventajas de las bases de datos relacionales
  - B. Ventajas y desventajas de las bases de datos no relacionales
  - C. Casos de uso para cada tipo de base de datos

**Actividades:**

- Realización de consultas a distintos sistemas de bases de datos.
- Reconocimiento de ventajas y desventajas de los distintos sistemas de bases de datos.
- Realización y corrección de ejercicios sobre la unidad.

**Recursos:**

- Aula de informática con equipos con un mínimo de 8 GB de RAM, conectados en red y a Internet.
- Proyector y pizarra
- Docker

**UT2. Acceso a datos. (32 horas)****Objetivos:**

- Conectividad de Bases de Datos
- Operaciones con Bases de Datos Relacionales usando Python
- Operaciones con Bases de Datos No Relacionales usando Python

**Contenidos:**

- I. Introducción a la conexión de bases de datos usando Python
  - A. APIs de Bases de Datos en Python (DB-API)
  - B. Librerías comunes de Bases de Datos en Python (e.g., SQLAlchemy, psycopg2, PyMongo)
  - C. Configuración de Conexiones a Bases de Datos
- II. Operaciones con Bases de Datos Relacionales en Python
  - A. Conexión a Bases de Datos Relacionales
  - B. Ejecución de Consultas SQL desde Python
  - C. Obtención y Manipulación de Datos
  - E. Gestión de Errores en Operaciones de Bases de Datos
- III. Prácticas con Bases de Datos Relacionales
  - A. Trabajando con MySQL en Python
  - B. Interacción con PostgreSQL en Python
  - C. Conectividad de Bases de Datos Oracle con Python
- IV. Operaciones de Bases de Datos No Relacionales en Python
  - A. Conexión a Bases de Datos No Relacionales
  - B. Operaciones CRUD (Crear, Leer, Actualizar, Borrar) en Bases de Datos NoSQL
  - C. Manejo de Estructuras de Datos Complejas
  - D. Manejo de Errores en Operaciones de Bases de Datos NoSQL

**Actividades:**

- Integración de MongoDB y Python
- Operaciones de Base de Datos Redis en Python
- Conectividad de Base de Datos Cassandra con Python
- Interacciones de Bases de Neo4j con Python

**Recursos:**

- Aula de informática con equipos con un mínimo de 8 GB de RAM, conectados en red y a Internet.
- Proyector y pizarra

**UT3. Introducción Hadoop. (12 horas)****Objetivos:**

**Contenidos:**

I. Almacenamiento y procesamiento en Hadoop.

II. HDFS.

III. Arquitectura.

A. Namenode

B. Secondary Namenode

C. Datanode

IV. Funcionamiento (lectura y escritura).

V. YARN

VI. MapReduce

**Actividades:**

- Lectura y escritura de datos en HDFS.
- Manejo de YARN
- Programar Jobs con MapReduce en Java.
- Interactuar BBDD y MapReduce.

**Recursos:**

- Aula de informática con equipos con un mínimo de 8 GB de RAM, conectados en red y a Internet.
- Proyector y pizarra

**UT4. Ecosistema Hadoop. (38 horas)****Objetivos:**

- Comprender el variado panorama de herramientas y marcos disponibles para el manejo de big data.
- Dominar el uso de Apache Flume para la ingestión de datos de diversas fuentes.
- Dominar Apache Sqoop para transferir datos entre bases de datos relacionales y Hadoop.
- Aprender a utilizar Apache Pig para la transformación y procesamiento de datos.
- Desarrollar habilidades de consulta y manipulación de datos con Apache Hive.
- Adquirir experiencia en Apache Oozie para la coordinación de flujos de trabajo y la programación de trabajos en el ecosistema de big data.
- Comprender los fundamentos de Apache Spark y su papel en el procesamiento de big data.
- Explorar Resilient Distributed Datasets (RDDs) y la API Spark DataFrames para el procesamiento de datos distribuidos.
- Aprende a realizar agregaciones de datos utilizando Spark DataFrames.
- Obtenga experiencia práctica en el procesamiento de datos en tiempo real con Spark Streaming.

**Contenidos:**

- I. Introduction to Big Data Ecosystem
- II. Data Ingestion and Integration Tools
  - A. Apache Flume
  - B. Apache Sqoop
- III. Data Processing and Querying Tools
  - A. Apache Pig
  - B. Apache Hive
- IV. Workflow and Job Scheduling
  - A. Apache Oozie
- V. Introduction to Apache Spark
  - A. Overview of Apache Spark
  - B. Resilient Distributed Datasets (RDD)
  - C. Spark DataFrames API
  - D. DataFrames Aggregations
  - E. Spark Streaming

**Actividades:**

- Configurar Apache Flume para ingerir datos de varias fuentes en un entorno Hadoop.
- Practicar la transferencia de datos entre bases de datos relacionales y Hadoop utilizando Apache Sqoop.
- Transformar datos utilizando Apache Pig.
- Consulta y manipulación de datos con Apache Hive.
- Crear flujos de trabajo Oozie para coordinar y programar trabajos Hadoop.
- Trabajar con Spark RDDs y la API DataFrames creando tareas de procesamiento de datos, analizar datos creando DataFrames.
- Realizar tareas de agregación de datos utilizando Spark DataFrames.
- Procesar datos en tiempo real utilizando Spark Streaming.

**Recursos:**

- Aula de informática con equipos con un mínimo de 8 GB de RAM, conectados en red y a Internet.
- Proyector y pizarra

**UT 5. Administración y monitorización de sistemas. (26 horas)****Objetivos:**

- Aseguramiento de la calidad de los datos.
- Integridad de los datos en sistemas de archivos distribuidos.
- Checksums para la detección de errores.
- Movimiento y migración de datos.
- Uso de herramientas de monitorización: Jobtracker, Namenode y otras.

- Monitorización de clústeres con Grafana y Ganglia.

**Contenidos:**

## I. Generación de mecanismos de integridad de datos

- A. Calidad de los datos
- B. Comprobación de la integridad de los datos en sistemas de archivos distribuidos
- C. Checksums
- D. Movimiento de Datos entre Clusters
- E. Actualización y migración
- F. Gestión de metadatos

## II. Supervisión, optimización y resolución de problemas

- A. Introducción a las herramientas de supervisión
  - Interfaz web de Jobtracker
  - Monitorización de Namenode
  - Otras herramientas de monitorización relevantes
- B. Análisis de datos históricos
- C. Monitorización de clústeres
  - Uso de Grafana
  - Utilización de Ganglia

**Actividades:**

- Evaluar la calidad de los conjuntos de datos proporcionados, identificando y documentando problemas como valores faltantes, duplicados e inconsistencias.
- Calcular sumas de comprobación en archivos de datos y los utilizan para detectar y corregir errores.
- Planificar, ejecutar y supervisar la transferencia de datos en un cluster, garantizando al mismo tiempo la integridad de los mismos.
- Explorar estrategias para actualizar y migrar datos en sistemas de archivos distribuidos, incluyendo versionado, cambios de formato de datos y evolución de esquemas.
- Usar herramientas de monitorización como la interfaz web Jobtracker y la monitorización Namenode para monitorizar el rendimiento del cluster, diagnosticar problemas y optimizar la ejecución de trabajos.
- Configurar Grafana y Ganglia para el monitoreo de clústeres en tiempo real.

**Recursos:**

- Aula de informática con equipos con un mínimo de 8 GB de RAM, conectados en red y a Internet.
- Proyector y pizarra

**UT 6. Técnicas Big Data toma decisiones BI. (14 horas)**

**Objetivos:**

- Desarrollar una comprensión del papel de BI en la toma de decisiones informadas dentro de las organizaciones.
- Adquirir conocimientos sobre los diferentes tipos de modelos de BI, incluyendo modelos descriptivos, de diagnóstico, predictivos y prescriptivos.
- Comprender el proceso KDD y su importancia en la extracción de conocimientos valiosos a partir de grandes conjuntos de datos.
- Explorar y dominar cada etapa del proceso KDD, desde la selección de datos hasta la evaluación de modelos.
- Aprenda cómo implementar de forma práctica los modelos de BI y sus casos de uso en el mundo real.
- Comprender la importancia de validar los modelos de BI para garantizar la fiabilidad y precisión de los resultados de la toma de decisiones.
- Adquirir habilidades en el uso de diversas técnicas de validación y métricas de evaluación para evaluar el rendimiento de los modelos de BI.
- Aprenda a aplicar metodologías de validación cruzada y pruebas para validar eficazmente los modelos de BI.

**Contenidos:**

- I. Introducción a la inteligencia empresarial en la toma de decisiones
  - A. Papel del Business Intelligence (BI) en la toma de decisiones
  - B. Integración de Big Data y BI
- II. Modelos de Inteligencia de Negocio
  - A. Visión general de los modelos de BI
  - B. Tipos de Modelos BI (Descriptivo, Diagnóstico, Predictivo, Prescriptivo)
- III. Proceso de modelos de descubrimiento de conocimientos en bases de datos (KDD)
  - A. Comprensión del proceso KDD
  - B. Fases clave del KDD
- IV. Etapas del Proceso KDD
  - A. Selección de datos
  - B. Limpieza y preprocesamiento de datos
  - C. Transformación de datos
  - D. Minería de datos
  - E. Interpretación de resultados
  - F. Evaluación de los modelos de minería de datos
- VI. Técnicas de Validación de Modelos BI
  - A. Importancia de la Validación de Modelos

- B. Técnicas de Validación de Modelos BI
- C. Métricas de Evaluación para Modelos BI
- D. Validación cruzada y pruebas

**Actividades:**

- Realizar la validación utilizando varias técnicas como la validación cruzada y la validación de espera.
- Validar un modelo BI real utilizando datos reales de un escenario empresarial.
- Utilizar estas herramientas para validar los modelos de BI y comparar los resultados.

**Recursos:**

- Aula de informática con equipos con un mínimo de 8 GB de RAM, conectados en red y a Internet.
- Proyector y pizarra

## 5. METODOLOGÍA

La L.O.E. y el R.D. 1147/2011, de 29 de julio, por el que se establece la ordenación general de la formación profesional del sistema educativo disponen que “la metodología didáctica de las enseñanzas de Formación Profesional integrará los aspectos científicos, tecnológicos y organizativos con el fin de que el alumnado adquiera una visión global de los procesos propios de la actividad profesional correspondiente”.

Se utilizará un método activo y participativo. Al alumno se le proporcionarán una serie de conceptos y procedimientos, junto con la documentación y las indicaciones necesarias, fomentando preguntas y debates para la asimilación de los mismos. Una vez comprendidos, el alumno pasará a adquirirlos mediante la realización de actividades prácticas, completando así el proceso de aprendizaje. Por tanto, las clases serán fundamentalmente prácticas.

Asimismo, se utilizará el espacio polivalente del ATecA (Aula de Tecnología Aplicada) y el AEP (Aula de Emprendimiento Profesional) para trabajar en él con nuevas metodologías didácticas como ABP (aprendizaje basado en proyectos o retos) y trabajo colaborativo. Al tratarse de un espacio polivalente y dotado de medios tecnológicos avanzados hace posible que se puedan abordar en él actividades, trabajos y proyectos en parejas o en pequeño grupo, contribuyendo así al desarrollo de las competencias personales, sociales y profesionales del alumnado.

### 5.1 Materiales y recursos didácticos

Los materiales y recursos necesarios para el correcto desarrollo del módulo serán los siguientes:

- Pizarra blanca de rotulador.
- Un ordenador para cada alumno y uno para el profesor con doble pantalla.

- Instalación de red, con acceso a Internet.
- Vídeo proyector.
- Apuntes proporcionados por la profesora.
- Cuaderno/folios para tomar apuntes.
- Pendrive o similar para almacenar las prácticas.
- Software necesario para el seguimiento del aula y realización de prácticas.
- Aula Virtual de Educación del Gobierno de La Rioja para publicación de apuntes y actividades por parte de la profesora y entrega de las mismas por parte de los alumnos: <https://aulavirtual-educacion.larioja.org/>

Además, para trabajar en determinados proyectos en pequeño grupo o en pareja se intentará usar el espacio polivalente del ATecA y la AEP.

## 6. EVALUACIÓN

La evaluación es continua y según el momento en que se realice recibe el nombre de inicial, formativa, sumativa y final, por lo que a lo largo del curso se irá recogiendo toda la información para sistematizar periódicamente y presentarla en las sesiones de evaluación del equipo docente del grupo de alumnos.

Se realizarán tres sesiones de evaluación durante el curso, en donde se medirá la superación de contenidos conceptuales, procedimentales y actitudinales por parte del alumno.

Por cada evaluación se realizarán uno o más exámenes teóricos y prácticos individuales, de los contenidos impartidos hasta la fecha.

El Proyecto Educativo de Centro recoge la asistencia mínima a clase como requisito para no perder el derecho a evaluación continua. La falta justificada o no a un número superior al 10% de las clases hace que el alumno pueda perder el derecho a evaluación continua.

Mediante la evaluación se determina si la planificación del curso ha permitido alcanzar los objetivos propuestos o es necesario reestructurar todo el proceso.

### 6.1 Criterios de evaluación

**Gestiona soluciones a problemas propuestos, utilizando sistemas de almacenamiento y herramientas asociadas al centro de datos.**

- a) Se ha caracterizado el proceso de diseño y construcción de soluciones en sistemas de almacenamiento de datos.
- b) Se han determinado los procedimientos y mecanismos para la ingestión de datos.
- c) Se ha determinado el formato de datos adecuado para el almacenamiento.
- d) Se han procesado los datos almacenados,
- e) Se han presentado los resultados y las soluciones al cliente final en una forma fácil de interpretar.

**Gestiona sistemas de almacenamiento y el amplio ecosistema alrededor de ellos facilitando el procesamiento de grandes cantidades de datos sin fallos y de forma rápida.**



a) Se ha determinado la importancia de los sistemas de almacenamiento para depositar y procesar grandes cantidades de cualquier tipo de datos rápidamente.

b) Se ha comprobado el poder de procesamiento de su modelo de computación distribuida.

c) Se ha probado la tolerancia a fallos de los sistemas.

d) Se ha determinado que se pueden almacenar tantos datos como se desee y decidir cómo utilizarlos más tarde.

e) Se ha visualizado que el sistema puede crecer fácilmente añadiendo módulos.

**Genera mecanismos de integridad de los datos, comprobando su mantenimiento en los sistemas de ficheros distribuidos y valorando la sobrecarga que conlleva en el tratamiento de los datos.**

a) Se ha valorado la importancia de la calidad de los datos en los sistemas de ficheros distribuidos.

b) Se ha valorado que a mayor volumen de tratamiento de datos corresponde un mayor peligro relacionado con la integridad de los datos.

c) Se ha reconocido que los sistemas de ficheros distribuidos implementan una suma de verificación para la comprobación de los contenidos de los archivos.

d) Se ha reconocido el papel del servidor en los procesos previos a la suma de verificación.

**Realiza el seguimiento de la monitorización de un sistema, asegurando la fiabilidad y estabilidad de los servicios que se proveen.**

a) Se han aplicado herramientas de monitorización eficiente de los recursos.

b) Se han recogido métricas, procesamiento y visualización de los datos.

c) Se han generado alertas para detectar un riesgo o mal funcionamiento.

d) Se ha comprobado que las herramientas usadas ofrecen un rendimiento elevado con rapidez.

e) Se ha comprobado la fiabilidad de los datos según respuestas.

f) Se ha analizado la estabilidad de servicios.

**Valida las técnicas de Big Data para transformar una gran cantidad de datos en información significativa, facilitando la toma de decisiones de negocios.**

a) Se han seleccionado gran cantidad de datos estructurados y no estructurados para reforzar la función de BI.

b) Se ha realizado la limpieza y transformación de datos en base a los objetivos predeterminados.

c) Se ha comprobado que el *Big Data* multiplica la relevancia y la utilidad del BI para el negocio.

d) Se han conjugado dentro de un modelo de empresa datos de clientes, financieros de ventas, de productos, de marketing, de redes sociales, de la competencia, entre otros, para extraer un análisis valioso y efectivo para el negocio.

e) Se ha evaluado e interpretado la información extraída de los datos y su influencia en el triunfo de diferentes negocios.

f) Se ha simulado la implantación de un modelo de Inteligencia de negocios BI.

## 6.2 Instrumentos y procedimientos de evaluación

Durante cada periodo de evaluación se le propondrán al alumno una serie de prácticas y ejercicios para desarrollar en clase de obligada ejecución y una o varias pruebas de carácter teórico-práctico en las que se incluyen los conceptos y procedimientos desarrollados hasta el momento.

El Proyecto Educativo de Centro recoge la asistencia mínima a clase como requisito para no perder el derecho a evaluación continua. La falta justificada o no a un número superior al 10% de las clases puede hacer que el alumno pierda el derecho a evaluación continua.

Mediante la evaluación se determina si la planificación del curso ha permitido alcanzar los objetivos propuestos o es necesario reestructurar todo el proceso.

Se utilizarán los siguientes procedimientos de evaluación:

- La observación directa de los alumnos cuando estén realizando las actividades o prácticas tanto individualmente como en grupo o la participación en la clase.
- La valoración de los trabajos realizados (actividades, prácticas, proyectos,...)
- La realización de pruebas (tests, preguntas cortas, resolución de casos prácticos,...)
- En cada evaluación se realizan uno o más exámenes que engloban una o varias unidades de trabajo.

## 6.3 Criterios de calificación

Estos criterios de calificación se aplicarán a los **alumnos que no hayan perdido el derecho a evaluación continua**. Para ello, deben asistir al 90 % de las clases.

La calificación del alumno se realizará considerando los resultados de la evaluación continua en las diferentes actividades de evaluación realizadas como exámenes, ejercicios, prácticas o proyectos.

La **calificación de cada evaluación** se realizará con la siguiente cuantificación:

Exámenes	75 %
Prácticas/proyecto	25 %

Para obtener la calificación de cada evaluación se realizará la media ponderada de los resultados obtenidos en cada uno de los exámenes y prácticas realizadas hasta el momento de la evaluación. La nota de evaluación resultará del truncamiento de esta media ponderada

De acuerdo con la normativa vigente la calificación del módulo profesional es numérica entre 1 y 10, sin decimales. Se considerarán positivas las calificaciones iguales o superiores a cinco puntos y negativas las restantes.

Serán calificados con un 0 aquellos exámenes y/o pruebas en los que se detecte que ha habido copia entre diferentes compañeros y/o se hayan utilizado medios no autorizados por el profesor.

La siguiente tabla muestra **la distribución temporal de las unidades** y el peso de los **exámenes** programados a lo largo del curso.

EVALUACIÓN	UD	HORAS	EXÁMENES PROGRAMADOS	PESO SOBRE EL TOTAL DE EXÁMENES (75 %)	PESO SOBRE LA NOTA FINAL (100%)
1ª	UD 1	18	Examen 1ª evaluación [temas 1-2]	40%	30%
	UD 2	32			
2ª	UD 3	12	Examen 2ª evaluación [temas 3-4]	40%	30%
	UD 4	38			
3ª	UD 5	26	Examen 3ª evaluación [temas 5-6]	20%	15%
	UD 6	14			

La siguiente tabla muestra **el peso de las prácticas** realizadas a lo largo del curso.

EJERCICIOS Y PRÁCTICAS	PESO SOBRE EL TOTAL DE PRÁCTICAS (25%)	PESO SOBRE LA NOTA FINAL (100%)
UD 1	10%	2,5%
UD 2	20%	5%
UD 3	10%	2,5%
UD 4	40%	10%
UD 5	10%	2,5%
UD 6	10%	2,5%
<b>TOTAL</b>	<b>100 %</b>	<b>25 %</b>

Si en una unidad didáctica hay más de una práctica se realizará la media aritmética entre todas las prácticas realizadas en cada unidad.

Otros aspectos a considerar sobre las prácticas son:

1. Sólo se calificarán las prácticas entregadas en plazo.
2. Aquellos proyectos, trabajos, prácticas o ejercicios de diferentes alumnos en los que haya una manifiesta similitud entre ellos o con otros de años anteriores o de Internet serán calificados con un 0.
3. Podrá requerirse la exposición oral de las prácticas por parte de los alumnos.

Si alguna de las unidades didácticas no se ha impartido durante el curso por falta de tiempo, se repartirá su porcentaje asignado al resto de unidades didácticas de forma proporcional a las horas realizadas.

#### 6.4 Criterios de recuperación

El profesor debe facilitar el éxito de sus alumnos, por lo que deben establecerse unos criterios para permitir la recuperación de las partes no superadas. Al ser la evaluación individualizada, las actividades de recuperación podrán variar en función del alumno y de los conocimientos y capacidades que sean objeto de recuperación.

Los procedimientos de recuperación son detección de las carencias del alumno, realización de tareas específicas que refuercen la carencia detectada y reevaluación de los conocimientos y/o capacidades no superadas.

La reevaluación de los conocimientos y/o capacidades no superadas se realizará en diferentes momentos:

- En las **evaluaciones ordinarias** (primera y segunda), que serán realizadas por:
  - **Alumnos que no hayan superado la materia a lo largo del curso.**

Estos alumnos realizarán un examen de todos los contenidos del módulo. La nota final del módulo se calculará con los porcentajes establecidos en los criterios de calificación.

CALIFICACIÓN	1ª ORDINARIA	2ª ORDINARIA
Exámenes	75%	75%
Prácticas	25%	25%

Los alumnos con prácticas o proyecto suspensos deben poder recuperar esta parte mediante la nueva realización de las mismas o a través de la incorporación en el examen final de los contenidos de las prácticas realizadas durante el curso.

- **Alumnos que han perdido el derecho a evaluación continua.** Se llevará a cabo la evaluación de toda la materia impartida en la asignatura en un examen una vez acabada la docencia. En este tipo de evaluación, la nota de los exámenes corresponderá con la calificación obtenida en el examen final. [Este examen podrá tener contenidos de las prácticas realizadas durante el curso.]

CALIFICACIÓN	1ª ORDINARIA	2ª ORDINARIA
Examen	100%	100%

#### 6.5 Actividades de refuerzo o recuperación

Con el fin de afianzar las competencias y conocimientos que se han trabajado durante las clases, se propondrán actividades de refuerzo que podrán ser entregadas para su corrección al profesor.

Además, para aquellos alumnos que no superen la primera evaluación final ordinaria, se continuará trabajando con ellos en clase hasta la segunda convocatoria ordinaria suministrándose actividades para que puedan alcanzar los objetivos del módulo.

## 6.6 Recuperación de módulos pendientes

En el caso de que haya algún alumno realizando la FCT con este módulo pendiente deberá realizar una prueba teórico-práctica que recoja los contenidos del módulo y superarla. Asimismo deberá realizar las actividades, trabajos y prácticas indicadas por la profesora.

## 7. ATENCIÓN AL ALUMNADO CON NECESIDADES ESPECÍFICAS DE APOYO EDUCATIVO

La **atención a la diversidad** en este módulo se centra en prestar apoyo a aquellos alumnos que sufran algún retraso en la adquisición de determinados contenidos del curso ya que no hay ningún alumno con necesidades educativas específicas que exija la adaptación de las pruebas.

Se atenderá a cada alumno de forma individual y en todo caso se procurará aclarar cuantas dudas surjan por parte de los alumnos de forma que aquellos que no hayan alcanzado los conocimientos y procedimientos mínimos harán **actividades de apoyo o de refuerzo** para que cubran las lagunas que tengan y puedan seguir el curso en mejores condiciones, mientras que otros alumnos pueden hacer **actividades de profundización o ampliación**.

## 8. ACTIVIDADES COMPLEMENTARIAS Y EXTRAESCOLARES

Mediante la coordinación con todo el equipo educativo del grupo, se planificarán actividades extraescolares de visita a empresas relacionadas con las competencias profesionales del curso de especialización.

Asimismo, se buscará que los alumnos participen en jornadas tecnológicas o talleres que estén relacionados con el módulo y se promoverá la participación en concursos de nivel autonómico o nacional relacionados con el mismo.

## 9. NECESIDADES Y PROPUESTAS DE FORMACIÓN DEL PROFESORADO

Debido a la rapidez en la **evolución de las tecnologías de la información y la comunicación** es necesaria una **velocidad de reciclaje** en nuevos conocimientos por parte del profesorado muy alta. Los alumnos son los principales interesados en recibir los conocimientos más recientes posibles para una mejor adaptación al mercado laboral.

Por ello, el profesorado necesita cursos para adquirir estos nuevos conocimientos que surgen rápidamente, a través de por ejemplo, **píldoras formativas**, **estancias en empresas** y través de **cursos especializados** de manera que se personalicen a las necesidades de formación del profesorado.

## 10. BIBLIOGRAFÍA DE REFERENCIA

Como bibliografía de referencia para el desarrollo del módulo, se utilizarán, además de los apuntes proporcionados en el aula virtual (<https://aulavirtual-educacion.larioja.org>), libros de consulta y prensa especializada, vídeos explicativos, manuales de Internet etc.