Universidad de Navarra

Big data & cloud computing (Gr. IA)

Guía docente 2026-27

PRESENTACIÓN

Breve descripción:

- **Titulación (Módulo/Materia)**: Big Data & Cloud Computing
- Ingeniería en Inteligencia Artificial (Fundamentos de Computación / Almacenamiento y procesamiento de datos)
- **ECTS**: 6 ECTS
- **Curso, semestre**: 3.º curso, 1.º semestre
- **Carácter**: Obligatorio
- **Idioma**: Español

Profesores de la asignatura:

- Fernández Duval, Gonzalo Hernán / Profesor Ayudante Doctor
- Hernantes Apezetxea, Josune / Catedrático

RESULTADOS DE APRENDIZAJE (Competencias)

R7 - Aplicar la computación en la nube para el desarrollo de soluciones y la gestión de datos masivos (Big Data)

PROGRAMA

Tema 1: Fundamentos de Big Data

- Las 5V del Big Data y casos de uso reales en IA
- Arquitecturas de referencia: Data Warehouse vs. Data Lake, Lambda vs. Kappa
- Entorno de trabajo: contenedores con Docker

Tema 2: Computación distribuida y MapReduce

- Paralelismo de datos vs. de tareas y teorema CAP
- El paradigma MapReduce: map, shuffle y reduce

Tema 3: Hadoop

- Arquitectura HDFS y YARN
- Limitaciones de Hadoop y motivación para Spark

Tema 4: Apache Spark

- RDDs: lazy evaluation, DAG y transformaciones
- Spark SQL y DataFrames
- Spark MLlib: pipelines de Machine Learning distribuido

Tema 5: Integración de la arquitectura Big Data local

- Integración de Hadoop y Spark en un pipeline de datos completo
- Buenas prácticas de diseño de pipelines

Universidad de Navarra

Tema 6: Fundamentos de Cloud Computing

- Modelos IaaS, PaaS y SaaS; despliegue público, privado e híbrido
- Almacenamiento en la nube, gestión de costes e IAM básico

Tema 7: Cómputo en la nube y Azure Databricks

- Cómputo elástico y servicios de Big Data gestionado
- Azure Databricks: workspace y notebooks

Tema 8: Migración de arquitecturas Big Data a la nube

- Spark local vs. gestionado: configuración y costes
- Migración de pipelines a Databricks + ADLS

Tema 9: Big Data para IA: ML y visualización en la nube

- Pipelines de datos para ML: feature stores y calidad de datos
- MLOps: entrenar, registrar, desplegar y monitorizar modelos
- Visualización de datos en la nube

ACTIVIDADES FORMATIVAS

La dedicación de 150 h (6 ECTS) a la asignatura se divide en las siguientes actividades formativas:

- Clases presenciales teóricas: 24 horas
- Clases presenciales prácticas, laboratorios o talleres: 20 horas.
- Trabajos dirigidos (proyecto: sesiones en clase + trabajo fuera de ella): 50 horas
- Estudio personal: 50 horas
- Evaluación (exámenes, presentaciones): 6 horas

Metodologías docentes:

- Clases expositivas
- Trabajo individual o en grupo, resolución de problemas y laboratorios técnicos.
- Estudio del alumno basado en documentación técnica y materiales del curso
- Trabajo en equipo sobre un proyecto propio con seguimiento docente
- Realización de pruebas evaluadas y presentación

Las clases expositivas presentan los fundamentos teóricos de cada bloque (Big Data, Hadoop, Spark, Cloud Computing) y se complementan con demostraciones guiadas de las herramientas del curso (Docker, HDFS, PySpark, Azure Databricks). Los laboratorios están dedicados a la aplicación práctica de estos conceptos sobre conjuntos de datos masivos. El trabajo en equipo se estructura en torno a un proyecto propio que evoluciona a lo largo del curso: parte de una arquitectura Big Data local, migra a una arquitectura en la nube y finalmente incorpora componentes de Machine Learning y visualización de datos. El estudio personal se apoya en la documentación oficial de las herramientas y en los materiales entregados en cada sesión.

EVALUACIÓN

CONVOCATORIA ORDINARIA

La evaluación de la asignatura se distribuye en dos partes: pruebas escritas (50%) y proyecto en equipo (50%).

Pruebas escritas (50%):

A lo largo del curso se realizarán dos exámenes parciales:

- Examen parcial 1 (P1): 25%
- Examen parcial 2 (P2): 25%

Si el alumno aprueba ambos exámenes parciales (nota P1 # 5,0 Y nota P2 # 5,0), esta parte de la nota queda cerrada con el promedio ponderado de ambos.

Si suspende alguno de los dos parciales (o ambos), deberá presentarse a un examen final, que evaluará únicamente el contenido de la(s) parte(s) suspendida(s):

- Si suspende solo un parcial, el examen final evalúa únicamente esa parte y mantiene su ponderación original (25%). La nota del parcial aprobado se conserva.
- Si suspende ambos parciales, el examen final evalúa ambas partes, con una ponderación conjunta del 50%.

Para aprobar la asignatura mediante el examen final, es obligatorio obtener una **nota mínima de 3,0 en cada una de las partes evaluadas** en dicho examen, independientemente de que la nota final resultante sea igual o superior a 5,0.

Proyecto (50%):

El proyecto se desarrolla en equipo a lo largo del semestre y se evalúa en las siguientes etapas:

- Entrega 1: 15% (repositorio de código + documento)
- Entrega 2: 15% (repositorio de código + documento)
- Presentación oral: 5% (Entrega 1 y 2)
- Evaluación individual: 15%, compuesta por:
  - *Preguntas individuales sobre el proyecto* (10%): se realizan en una clase posterior a las presentaciones finales.
  - *Puntaje asignado por el equipo* (5%): cada estudiante recibe una cantidad de puntos (según el número de integrantes de su equipo) que debe repartir entre sus compañeros de grupo (sin poder asignárselos a sí mismo) según el mérito y el trabajo realizado por cada uno.

Nota final:

La nota final de la asignatura se calcula como la suma ponderada de pruebas escritas (50%) y el proyecto en equipo (50%):

Nota final = Pruebas escritas × 0,5 + Proyecto × 0,5

Condición para aprobar la asignatura:

Para aprobar la asignatura se deben cumplir ambas condiciones:

- La Nota final debe ser igual o superior a 5,0.
- Si el alumno realiza el examen final, debe obtener además una nota mínima de 3,0 en cada una de las partes evaluadas en dicho examen, independientemente de que la Nota final resultante sea igual o superior a 5,0.

CONVOCATORIA EXTRAORDINARIA

Si un alumno suspende la asignatura en la convocatoria ordinaria, deberá presentarse a un examen final extraordinario (EFE), que equivale al 100% de la nota final.

En esta convocatoria únicamente se tendrá en cuenta la nota obtenida en el EFE.

AVISOS Y NORMATIVA DE EVALUACIÓN

ESTUDIANTES CON NECESIDADES EDUCATIVAS ESPECIALES

Los estudiantes con necesidades educativas especiales deberán ponerse previamente en contacto con la Coordinación de Estudios de la Escuela para obtener la autorización correspondiente a las adaptaciones (por ejemplo, disponer de más tiempo en los exámenes). Dicha autorización deberá ser enviada al profesor. Se recomienda realizar esta gestión al comienzo del cuatrimestre.

NORMATIVA DE USO DE ASISTENTES DE INTELIGENCIA ARTIFICIAL Y AYUDA EXTERNA

Las pruebas escritas se realizarán de forma individual y sin conexión a internet, sin ningún tipo de asistencia.

Para la realización de los trabajos se puede utilizar la ayuda externa que se desee. La incorporación al trabajo del resultado de dicha ayuda es responsabilidad del/os autor/es del trabajo por lo que deben ser conocedores de forma completa de dicho contenido y justificar su contenido como autor; en el caso de código de programación además se debe poder explicar qué hace dicho código a nivel de cada término y de algoritmo.

AVISO IMPORTANTE

Se recuerda que cualquier intento de fraude, copia, plagio u otro comportamiento irregular supone una infracción grave tal y como está contemplado en el título IV "Normas de disciplina académica de los estudiantes" dentro del Sistema de normas sobre la convivencia en la Universidad de Navarra .

HORARIOS DE ATENCIÓN

Dra. Josune Hernantes Apezetxea (jhernantes@unav.es)

- Despacho o101 Edificio Urdaneta Planta 1ª
- Horario de tutoría: enviar un mail con antelación para concertar cita.

Dr. Gonzalo Hernán Fernández Duval (ghfernandezd@unav.es)

- Despacho o110 Edificio Urdaneta Planta 1ª
- Horario de tutoría: enviar un mail con antelación para concertar cita.

BIBLIOGRAFÍA Y RECURSOS

Bibliografía básica:

- Kleppmann, M., & Riccomini, C. (2026). Designing Data-Intensive Applications (2.ª ed.). O'Reilly Media Localízalo en la biblioteca (online)
- Damji, J. S., Wenig, B., Das, T., & Lee, D. (2020). Learning Spark: Lightning-Fast Data Analytics (2.ª ed.). O'Reilly Media. Localízalo en la biblioteca (online)
- White, T. (2015). Hadoop: The Definitive Guide (4.ª ed.). O'Reilly Media. Localízalo en la biblioteca (online)

- Erl, T., Puttini, R., & Mahmood, Z. (2013). Cloud Computing: Concepts, Technology & Architecture. Prentice Hall. Localízalo en la biblioteca
- Documentación oficial de Microsoft Azure: https://learn.microsoft.com/azure/

Bibliografía complementaria:

- Huyen, C. (2022). Designing Machine Learning Systems: An Iterative Process for Production-Ready Applications. O'Reilly Media. Localízalo en la biblioteca (online)
- Knaflic, C. N. (2015). Storytelling with Data: A Data Visualization Guide for Business Professionals. Wiley. Localízalo en la biblioteca online / Localízalo en la biblioteca físico y en castellano