GUÍA DOCENTE
2025 - 2026

FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
Nombre completo	Big Data y Tecnologías de Internet/Big Data and Internet Technologies
Código	E000013658
Título	Máster Universitario en Análisis de Negocio / Master in Business Analytics por la Universidad de Deusto y la Universidad Pontificia Comillas
Impartido en	Máster Universitario en Análisis de Negocio/Master in Business Analytics [Primer Curso]
Nivel	Postgrado Oficial Master
Cuatrimestre	Semestral
Créditos	3,0 ECTS
Carácter	Obligatoria
Departamento / Área	Departamento de Telemática y Computación
Responsable	Carlos Miguel Vallez Fernández
Horario	Revisar Horario
Horario de tutorías	Concertar cita con el profesor por email

Datos del profesorado	
Profesor	
Nombre	Carlos Miguel Vallez Fernández
Departamento / Área	Departamento de Telemática y Computación
Despacho	D-401 ICAI
Correo electrónico	cmvallez@icai.comillas.edu

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura

Aportación al perfil profesional de la titulación

La tecnología Big Data es esencial en la disciplina de Analítica avanzada y procesado masivo de datos y una herramienta de uso cada vez más extendido en las empresas como habilitador de los análisis avanzados. De esta manera el objetivo de la asignatura es que el alumno aprenda los conceptos relacionados de Big data, desde sus origen e historia, características y diferencias respecto a sistemas tradicionales, su arquitectura paralela y las diferencias con otros sistemas paralelos y ecosistema del ciclo del dato básico, desde su adquisición del dato, almacenamiento al procesamiento. Se relacionará las capacidades que ofrece esta tecnología con los problemas que para los que está especialmente indicada, las necesidades que cubre y los casos de uso que consiguen aportar valor diferencial a las empresas y la sociedad.

Al finalizar el curso, los alumnos entenderán los bloques habituales de una instalación tipo, el funcionamiento de big data, su utilidad y sabrán en qué casos debe ser aplicado, por qué y para qué, de tal manera que puedan proponer y defender casos concretos de uso.

Luego se aborda Internet y su importancia en la recopilación y distribución de información. Se exploran los diferentes tipos de protocolos de Internet, como HTTP, FTP y DNS.

Por último, se analiza el concepto de IoT (Internet de las cosas), que se refiere a la conexión de dispositivos y sensores a Internet para recopilar datos en tiempo real.

Prerrequisitos

Fundamentos básicos de informática y programación

Competencias - Objetivos

Competencias

Conocimientos o contenidos

CO6	Conocer de manera intuitiva los elementos esenciales y los conceptos de la tecnología vinculada con las bases de datos, el almacenamiento y recuperación de la información, el Big Data, Internet y demás tecnologías conectivas como el #internet de las cosas#, teniendo una visión panorámica, crítica y prospectiva de todos estos temas.

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos

1: Introducción a Big Data

- ¿Qué es Big Data y qué no es Big Data?.
- El valor del dato.
- Volúmenes y fuentes de datos. Información estructurada y no estructurada.
- Utilidad, necesidad y ejemplos de casos de uso Big Data.

2: Ciclo de vida de un proyecto Big Data

- Data Pipeline de un proyecto Big Data. (Ingestión, almacenamiento, procesado/computación y Presentación/visualización).
- Perfiles profesionales.
- Ejemplo básico de ingestión y problemática asociada por recursos disponibles.
- Introduccion a Gobierno, privacidad y protección del dato. Perfil DPO , RGPD y derechos ARCO.

3: Conceptos básicos sobre sistemas

- Sistemas operativos basados en Linux. Introducción y comandos elementales.
- Sistemas Distribuidos. Introducción y conceptos relacionados (Paralelización, Tolerancia a fallos, Balanceo, Disponibilidad, Redundancia)
- Concepto de virtualización. Introducción y plataformas. Máquinas virtuales vs contenedores.
- Datacenters: redes de alta velocidad y computación de alto rendimiento.

4: Ecosistema Hadoop

- Introducción al Procesamiento masivo y Hadoop. Componentes
- Arquitectura Hardware y Software.
- Almacenamiento Masivo
- Sistema de ficheros HDFS.
- Servicios y roles.
- YARN.
- MapReduce.
- Spark.

5: Cloud & Big Data

- Infraestructura on-premise vs Cloud.
- Concepto "as a service".
- Proveedores y plataformas Cloud. Servicios y productos ofrecidos.

6: Internet, IoT y Tecnología Big Data en la empresa

- Internet y su importancia en la recopilación y distribución de información
- Introducción a protocolos de, como HTTP, FTP y DNS.
- Concepto de IoT (Internet de las cosas). Recopilación de datos en tiempo real
- Impacto de Big Data en Analytics y la Inteligencia Artificial/Machine Learning.
- Big Data como tecnología exponencial y disruptiva en el negocio.

.

METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura

La asignatura es de orientación práctica y aplicada. Pero con introducción de conceptos teóricos necesarios.

En la clase presencial se explicarán los contenidos teóricos y conceptos principales, continuando con breves aplicaciones prácticas de los mismos

En las sesiones prácticas se aplicará lo visto y profundizando en las técnicas y conceptos estudiados. Algunas de estas tareas deberán finalizarse como parte del trabajo fuera de las horas de clase por parte de alumno.

Metodología Presencial: Actividades

Lección expositiva: El profesor desarrolla el tema que previamente los alumnos han leído, explicándolo en la pizarra. Una vez desarrollados los conceptos teóricos se aplican a un caso real. Para ello el profesor y/o los propios alumnos, proponen un problema cotidiano al cual se le intenta dar solución con la participación de los alumnos. A continuación, el profesor codifica dicho programa en el ordenador con la participación de los de los alumnos. Este programa codificado por el profesor es facilitado a los alumnos a través de la web de la

asignatura. **Debates grupales, pruebas y resolución de ejercicios**: En estas sesiones se resolverán las dudas surgidas de las lecturas llevadas a cabo por los alumnos de los temas o contenidos propuestos por el profesor. También se realizarán debates grupales y pruebas para facilitar la comprensión de los mismos. Además, se resolverán los pequeños ejercicios propuestos por el profesor. **Prácticas de laboratorio**: A lo largo del curso los alumnos se realizarán prácticas individuales de cada uno de los temas expuestos. En el laboratorio se finalizará el desarrollo de la práctica y se resolverán las dudas pertinentes. **Tutorías**: Se realizarán en grupo e individualmente para resolver las dudas planteadas a los alumnos después de trabajar sobre los distintos temas. También se utilizarán para orientar al alumno en su proceso de aprendizaje.	CO6
Metodología No presencial: Actividades	
1. **Estudio individual del material**. Actividad realizada individualmente por el estudiante para preparar el tema que se discutirá en clase. **(30 horas)**. 2. **Realización de casos y trabajos de aplicación práctica**, empleando datos y programación que plasmen los conocimientos adquiridos en la asignatura	CO6

RESUMEN HORAS DE TRABAJO DEL ALUMNO

HORAS PRESENCIALES	
AF1. Lecciones de carácter expositivo. Sesiones en las que los docentes exponen contenidos concretos, que pueden estar apoyados o no por recursos tecnológicos, en la que existirán periodos de explicación a cuestiones o dudas planteadas por los estudiantes.	AF3. Ejercicios y resolución de casos y de problemas. Sesiones en las que los docentes realizan de manera individual o junto con los estudiantes problemas, casos o cualquier tipo de ejercicio pertinente vinculado con la actualidad de la materia y/o con empresas u organizaciones reales, para la adquisición de competencias de las asignaturas.
10.00	20.00
HORAS NO PRESENCIALES	

AF7. Estudio y lectura organizada. Tiempo de estudio por parte del estudiante dedicado a preparar, profundizar y analizar en los contenidos de las asignaturas o materias, así como realizar ejercicios prácticos.	AF8. Trabajos monográficos y de investigación, individuales o colectivos. Los estudiantes trabajan de manera individual o colectiva un tema objeto de estudio relacionado con una asignatura o materia y/o con empresas u organizaciones reales, mediante la búsqueda de información, el análisis y la elaboración de documentos.	AF3. Ejercicios y resolución de casos y de problemas. Sesiones en las que los docentes realizan de manera individual o junto con los estudiantes problemas, casos o cualquier tipo de ejercicio pertinente vinculado con la actualidad de la materia y/o con empresas u organizaciones reales, para la adquisición de competencias de las asignaturas.	AF6. Sesiones tutoriales. Orientación personal o virtual que el docente realiza con los estudiantes y que consiste en una ayuda particular integrada dentro del proceso formativo, ayudando así al avance en el proceso formativo.
30.00	12.00	12.00	6.00
CRÉDITOS ECTS: 3,0 (90,00 horas)			

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

El uso de IA para crear trabajos completos o partes relevantes, sin citar la fuente o la herramienta o sin estar permitido expresamente en la descripción del trabajo, será considerado plagio y regulado conforme al Reglamento General de la Universidad.

Actividades de evaluación	Criterios de evaluación	Peso
Cuestionario con preguntas de opción única y multiple	Calificación numérica de 0 a 10	40
Quiz cortos en clase durante el curso 20% Trabajo en clase, asistencia y participación 5%	Calificación numérica de 0 a 10 cada una de las partes	25
Prácticas individuales relacionada con la temática de la asignatura	Calificación numérica de 0 a 10	20
Prácticas grupales relacionada con la temática de la asignatura	Calificación numérica de 0 a 10	15

Calificaciones

El uso de IA para crear trabajos completos o partes relevantes, sin citar la fuente o la herramienta (y el resto de elementos que se solicitan y detallan abajo) o sin estar permitido expresamente en la descripción del trabajo, será considerado plagio y regulado conforme al Reglamento General de la Universidad

Se permite el uso de IA única y exclusivamente en la realización de prácticas individuales y/ grupales. Se puede emplear tanto para reformular o refinar textos como para obtener código o ayudar a la resolución de los problemas. Si bien es necesario y obligatorio referenciar los apartados donde ha sido utilizada, referenciar qué IA se ha empleado, incluir el promp empleado y la

respuesta obtenida y finalmente incluir una explicación con vuestras propias palabras que demuestren el entendimento de la respuesta recibida (lo que hace el comando, el significado del párrafo etc)

Queda expresamente prohibido el uso de cualquier IA y/o software de comunicación con una persona o IA durante la realización de los Test/quiz de clase y durante la realización del examen final. El no cumplimiento de esta norma será considerado plagio y regulado conforme al Reglamento General de la Universidad.

La calificación en la convocatoria **ordinaria** de la asignatura se obtendrá como:

- 40 % Examen final (teórico y pequeños ejercicios prácticos en cuestionario)
- 20 % Pruebas prácticas durante el curso
- 20 % Practicas individuales
- 15 % Prácticas en grupo
- 5 % Trabajo en clase y participación

Es necesario tener entregadas las prácticas para poder optar a aprobar la asignatura.

Nota: Para hacer media los alumnos tienen que tener al menos 4 puntos sobre 10 en el examen final de la asignatura además de presentar y obtener al menos un 5 en todas las prácticas

La calificación en la convocatoria **extraordinaria**:

- 40 % Examen final (teórico y pequeños ejercicios prácticos en cuestionario)
- 20 % Pruebas prácticas durante el curso
- 20 % Practicas individuales
- 15 % Prácticas en grupo
- 5 % Trabajo en clase y participación

Es necesario tener entregadas las prácticas para poder optar a aprobar la asignatura.

Nota: Para hacer media los alumnos tienen que tener al menos 4 puntos sobre 10 en el examen final de la asignatura además de presentar y obtener al menos un 5 en todas las prácticas

El uso de IA para crear trabajos completos o partes relevantes, sin citar la fuente o la herramienta o sin estar permitido expresamente en la descripción del trabajo, será considerado plagio y regulado conforme al Reglamento General de la Universidad

PLAN DE TRABAJO Y CRONOGRAMA

Actividades	Fecha de realización	Fecha de entrega
Sesion 1 Sesión 2		

Sem 1	Tema 0 Introducción a la asignatura Tema 1 Introducción al Big Data y Tema 2 Big Data Definition y Características	Tema 2 Big Data Definición y Características (cont) Tema 3.1 Sistemas distribuídos		
Sem 2	Tema 3.1 Sistemas distribuídos (cont) Tema 3.2 Datacenters Tema 4 Tolerancia a fallos en sistemas distribuidos	Tema 5.1 Virtualización Tema 5.2 Consolidación		
Sem 3	Tema 6 Introducción a comandos Linux	Práctica Grupal 1 día 1 Mv's y Linux		
Sem 4	Práctica Grupal 1 día 2 Mv's y Linux	Tema 7 Hadoop I		
Sem 5	Tema 8 Hadoop II Pract Indiv 1 Map Reduce	Pract Grupal 2 Map Reduce / Hadoop		
Sem 6	Tema 9 Servicios Cloud y Roles	Tema 10 Spark		
Sem 7	Práctica individual final Pract Grupal 4 Spark	Pract Grupal 3 Spark		
Sem 8	Tema 11 Introducción a las Tecnologias de Intenernet e IoT			
Sem 9	EXAMEN			

BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

Notas preparadas por el profesor disponibles en Moodle.

Bibliografía Complementaria

- Mayer-Schönberger, V., & Cukier, K. (2013). Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt.
- P. Zikopoulos, C. Eaton, D. DeRoos, T. Deutsch,&G. Lapis,(2012) Understanding Big Data, USA:McGraw-Hill Books
- White, T. (2009). Hadoop. The Definitive Guide. Vol. 1, OReilly Media. Inc, San Jose, CA, USA
- L. Joyanes (2013), Big Data: Análisis de grandes volúmenes de datos en organizaciones, Editorial Alfaomega
- Chambers, B., & Zaharia, M. (2018). Spark: The definitive guide: Big data processing made simple. " O'Reilly Media
- Holmes, A. (2012). Hadoop in practice. Manning Publications Co..
- Shotts, W. (2019). The Linux command line: a complete introduction. No Starch Press.
- Karau, H., Konwinski, A., Wendell, P., & Zaharia, M. (2015). Learning spark: lightning-fast big data analysis. " O'Reilly Media, Inc.".
- M. van Steen and A.S. Tanenbaum, Distributed Systems, 3rd ed., distributed-systems.net, 2017.
- Noah Iliinsky ad Julie (2011) Steele Data Visualizations O'Reilly

- Nathan Yau (2011) Visualize This: The FlowingData Guide to Design, Visualization, and Statistics
- Lutz, M. (2009). Python Pocket Reference: Python in Your Pocket" O'Reilly Media, Inc.".
- Lutz, M. (2013). Learning python: Powerful object-oriented programming. " O'Reilly Media,