

FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
NombreCompleto	Datos Masivos / Big Data
Código	E000008337
Impartido en	Grado en Análisis de Negocios/Business Analytics y Grado en Derecho [Tercer Curso]
Nivel	Reglada Grado Europeo
Cuatrimestre	Semestral
Créditos	3,0
Carácter	Obligatoria (Grado)
Departamento / Área	Departamento de Telemática y Computación
Responsable	Carlos Morrás Ruiz-Falcó
Horario	Martes de 8 a 10 am.
Horario de tutorías	Lunes de 16:00 a 17:00 y de 19:00 a 21:00, Martes de 10:00 a 11:00 am y viernes de 18:00 a 20:00

Datos del profesorado	
Profesor	
Nombre	Carlos Morrás Ruiz-Falcó
Departamento / Área	Departamento de Telemática y Computación
Correo electrónico	cmorras@icai.comillas.edu

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura
<p>Aportación al perfil profesional de la titulación</p> <p>La tecnología Big Data es esencial en la disciplina de Análítica 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 esta especialmente indicada, las necesidades que cubre y los casos de uso que consiguen aportar valor diferencial a las empresas y la sociedad.</p> <p>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</p>

tal manera que puedan proponer y defender casos concretos de uso.

Prerrequisitos

Ninguno

Competencias - Objetivos

Competencias

Competencias Básicas

CB 02. Saber aplicar e integrar sus conocimientos, la comprensión de éstos, su fundamentación científica y sus capacidades de resolución de problemas en entornos nuevos y definidos de forma imprecisa, incluyendo contextos de carácter multidisciplinar tanto investigadores como profesionales altamente especializados

CB 03. Saber evaluar y seleccionar la teoría científica adecuada y la metodología precisa de sus campos de estudio para formular juicios a partir de información incompleta o limitada incluyendo, cuando sea preciso y pertinente, una reflexión sobre la responsabilidad social o ética ligada a la solución que se proponga en cada caso.

CB 07. Ser capaces de asumir la responsabilidad de su propio desarrollo profesional y de su especialización en uno o más campos de estudio.

Competencias Generales

CG. 1 Capacidad de organización y planificación en la identificación de problemas en el contexto de datos masivos

CG 02. Capacidad de análisis de datos masivos procedentes de diversas fuentes: texto, audio, numérica e imagen.

CG 03. Resolución de problemas y toma de decisiones en un entorno de datos masivos tanto cuantitativos como cualitativos.

CG 07. Capacidad de liderazgo y trabajo en equipo en la sociedad de la información.

CG 11. Capacidad para aprender y trabajar autónomamente en la sociedad de la información.

Competencias Específicas

CE 25. Comprender los conceptos más importantes asociados al Big Data y sus elementos más característicos.

Competencias Específicas de la Optatividad

CEO 24. Conocer y usar técnicas de análisis de información para extraer el conocimiento para la toma de decisiones de negocio.

Resultados de Aprendizaje

CE 25. Comprender los conceptos más importantes asociados al Big Data y sus elementos más característicos

RA1. Conocer el concepto de paralelismo y la importancia de la distribución de datos voluminosos.

RA2. Identificar las necesidades específicas asociadas al manejo, adquisición y almacenamiento de volúmenes masivos de datos.

RA3. Comprender y conocer las tecnologías de comunicaciones (redes) y computación precisas para la distribución y recuperación de información, así como de su análisis.

CEO 24. Conocer y usar técnicas de análisis de información para extraer el conocimiento para la toma de decisiones de negocio.

RA1 Conoce y comprende las herramientas de apoyo a la decisión en el ámbito de la empresa.

RA2 Es capaz de implementar una solución computerizada de apoyo a la decisión.

RA3 Comprende y aplica herramientas de adquisición, representación e inferencia, y aquellas técnicas de aprendizaje automático y basadas en sistemas inteligentes.

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos

Contenidos – Bloques Temáticos

BLOQUE 1: Contenido de la asignatura

1: Big Data Origen e Historia

1. ¿Por qué ahora? La sociedad digital.
2. Las tecnologías exponenciales y las tecnologías disruptivas
3. El problema de Google y Yahoo!.
4. Opensource y el proyecto Nutch.
5. Hadoop y evolución.
6. Tendencias actuales: IoT, el Internet de las Cosas

2: Características fundamentales de Big Data. Las V's

1. Características fundamentales de BD: las 4V

2. Información estructurada y no estructurada. Impacto de BD.

3: Funcionamiento de Big Data

1. Distribución y paralelismo. Características de programación
2. Sistema de ficheros HDFS
3. Redundancia y tolerancia a fallos
4. Arquitectura básica.
 - Herramientas y librerías del ecosistema Hadoop.
 - Procesamiento masivo:
 - YARN.
 - MapReduce.
 - Spark.
 - Administración y monitorización de un cluster.

4: Ecosistema Big Data

1. Ecosistema básico de Hadoop
2. Adquisición de datos. Scoop flume y Kaffka
3. Búsqueda y procesamiento de datos. Hbase y Hive
4. Análisis de datos

5: Big Data y Cloud/la Nube

1. La nube. Definición y principales actores (AWS, Azure y Google)
2. La nube vs on premise

6: El impacto y uso del Big Data

1. Impacto de big data en Analytics y la inteligencia artificial/Machine Learnig. Causas y efectos.
2. Palancas de uso de Big Data empresarial
3. Casos de uso de Big Data
 - Marketing
 - Comercio y Retail
 - Banca
 - Industria.
 - Otros.

BLOQUE 2:Prácticas de la asignatura

Prácticas

1. Arquitectura Hadoop con Big Data: configuración y monitorización de servidores
2. HDFS: Sistema de Almacenamiento en Hadoop

METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura

Metodología Presencial: Actividades

Metodología Presencial: Actividades

Competencias

1. **Clase magistral y presentaciones generales.** Exposición de los principales conceptos y procedimientos mediante la explicación por parte del profesor. Incluirá presentaciones dinámicas, pequeños ejemplos prácticos y la participación reglada o espontánea de los estudiantes. Además, el profesor propondrá la lectura de temas preparados por el profesor, artículos científicos y divulgativos que permitirán introducir o profundizar en un tema, con el objetivo de aportar ideas y generar debate **(24 horas). CE 25,CG 03, CG 02 CG 01, CB 03, CB 02**
2. **Resolución en clase de problemas prácticos.** Resolución de unos primeros problemas para situar al alumno en contexto. La resolución correrá a cargo del profesor y los alumnos de forma cooperativa **(2 horas). CE 25,CG 03, CG 02 CG 01, CB 03, CB 02**
3. **Prácticas.** Cada alumno realizará de forma aislada o en grupo una serie de prácticas regladas. Las prácticas finalizarán con la redacción de un informe o la redacción de las distintas experiencias **(4 horas). CE 25,CG 03, CG 02 CG 01, CB 03, CB 02**

Metodología No presencial: Actividades

Metodología No presencial: Actividades

Competencias

1. **Estudio individual del material.** Actividad realizada individualmente por el estudiante para preparar el tema que se discutirá en clase. **(60 horas). CE 25,CG 03, CG 02 CG 01, CB 03, CB 02**
2. **Resolución de problemas prácticos** a resolver fuera del horario de clase por parte del alumno. El alumno debe utilizar e interiorizar los conocimientos aportados en la materia. La corrección a la clase se realizará por parte de alguno de los alumnos o el profesor según los casos. La corrección individualizada de cada ejercicio la realizará el propio alumno u otro compañero según los casos (método de intercambio) **(10 horas).CE 25,CG 03, CG 02 CG 01, CB 03, CB 02**
3. **Prácticas de la asignatura.** Realización de dos prácticas que plasme los conocimientos adquiridos en la asignatura. **(10 horas).CE 25,CG 03, CG 02 CG 01, CB 03, CB 02**

RESUMEN HORAS DE TRABAJO DEL ALUMNO

RESUMEN HORAS DE TRABAJO DEL ALUMNO

HORAS PRESENCIALES

Lección magistral 24

Debates y resolución de problemas 2

Prácticas 4

Evaluación 2

HORAS NO PRESENCIALES

Trabajo autónomo sobre contenidos teóricos 10

Trabajo autónomo sobre contenidos prácticos 10

Práctica final de la asignatura 10

Estudio 30

CRÉDITOS ECTS:

3 (90 horas)

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

Actividades de evaluación

Criterios de evaluación

PESO

Realización de exámenes:

- Examen final.
- Comprensión de conceptos mediante pruebas abiertas, enfocado a casos prácticos.
- Aplicación de los conceptos adquiridos
- Presentación y comunicación escrita.

60%

Evaluación del rendimiento teórico continuo:

Realización de pruebas de seguimiento, debates y ejercicios.

- Comprensión de conceptos mediante pruebas objetivas y one minute papers.
- Debate y análisis de los documentos técnico/científicos.
- Análisis e interpretación de los resultados obtenidos en la resolución de problemas mediante debates.

30%

Evaluación del rendimiento práctico continuo:

- Casos a desarrollar en la clase de prácticas.
- Comprensión de conceptos.
- Aplicación de conceptos al desarrollo de casos sobre los nuevos conceptos planteados.
- Aplicación de todos los conocimientos adquiridos durante el curso.

10%

Calificaciones

Criterios de Calificación

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

- Un 60% la calificación del examen final.
- Un 30% será la calificación de las pruebas de seguimiento y debates.
- Un 10% será la calificación de las prácticas.

La calificación en la **convocatoria extraordinaria**:

- Un 60% la calificación del examen final.
- Un 30% será la calificación de las pruebas de seguimiento y debates.
- Un 10% será la calificación de la práctica final.

Para aprobar la asignatura los alumnos tienen que tener al menos 5 puntos sobre 10 en el examen final, y en las prácticas, tanto en la convocatoria ordinaria como en la extraordinaria.

PLAN DE TRABAJO Y CRONOGRAMA

Actividades	Fecha de realización	Fecha de entrega
<p>Actividades Presenciales y No presenciales</p> <p>Fecha de realización</p> <p>Fecha de entrega</p> <ul style="list-style-type: none"> • Lectura y estudio de los contenidos teóricos Después de cada clase • Resolución de los problemas propuestos o debates Semanalmente • Preparación de las prácticas Semanas 		

8, 9 y 10		
Semanas		
11		
<ul style="list-style-type: none">• Preparación del examen final		

BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

- Notas preparadas por el profesor disponibles en Moodle.

Bibliografía Complementaria

- Big Data: A Revolution That Will Transform How We Live, Work, and Think. [Viktor Mayer-Schonberger](#) (Autor), [Kenneth Cukier](#) (Autor)
- George Orwell. 1984 . ISBN: 9788499890944
- Michael Lewis. Moneyball: The Art of Winning an Unfair Game Paperback – March 17, 2004. ISBN: 978-0393324815 or the film 'Moneyball (2011)'. Sony Pictures Director: Bennett Miller

FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
Nombre	Big Data
Código	DTC-
Titulación	Grado en E3-Analytics
Curso	Tercero
Cuatrimestre	1º
Créditos ECTS	3
Carácter	Obligatorio/ formación Básica
Departamento	Telemática y Computación
Área	
Coordinador	Carlos Morrás Ruiz-Falcó

Datos del profesorado	
Profesor	
Nombre	Carlos Morrás Ruiz-Falcó
Departamento	Telemática y Computación
Área	Ingeniería Telemática
Despacho	
e-mail	cmorras@comillas.edu
Teléfono	
Horario de Tutorías	Cita previa con el profesor

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura
Aportación al perfil profesional de la titulación
<p>La tecnología Big Data es esencial en la disciplina de Analítica avanzada y procesamiento 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 esta especialmente indicada, las necesidades que cubre y los casos de uso que consiguen aportar valor diferencial a las empresas y la sociedad.</p> <p>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.</p>

Prerrequisitos

Ninguno

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos

BLOQUE 1: Contenido de la asignatura

Tema 1: Big Data Origen e Historia

- 1.1. ¿Por qué ahora? La sociedad digital.
- 1.2. Las tecnologías exponenciales y las tecnologías disruptivas
- 1.3. El problema de Google y Yahoo!.
- 1.4. Opensource y el proyecto Nutch.
- 1.5. Hadoop y evolución.
- 1.6. Tendencias actuales: IoT, el Internet de las Cosas

Tema 2: Características fundamentales de Big Data. Las V's

- 2.1. Características fundamentales de BD: las 4V
- 2.2. Información estructurada y no estructurada. Impacto de BD.

Tema 3: Funcionamiento de Big Data

- 3.1. Distribución y paralelismo. Características de programación
- 3.2. Sistema de ficheros HDFS
- 3.3. Redundancia y tolerancia a fallos
- 3.4. Arquitectura básica.
 - Herramientas y librerías del ecosistema Hadoop.
 - Procesamiento masivo:
 - YARN.
 - MapReduce.
 - Spark.
 - Administración y monitorización de un cluster.

Tema 4: Ecosistema Big Data

- 4.1. Ecosistema básico de Hadoop
- 4.2. Adquisición de datos. Scoop flume y Kaffka
- 4.3. Búsqueda y procesamiento de datos. Hbase y Hive
- 4.4. Análisis de datos

Tema 5: Big Data y Cloud/la Nube

- 5.1. La nube. Definición y principales actores (AWS, Azure y Google)
- 5.2. La nube vs on premise

Tema 6: El impacto y uso del Big Data

- 6.1. Impacto de big data en Analytics y la inteligencia artificial/Machine Learnig. Causas y efectos.

6.2. Palancas de uso de Big Data empresarial

6.3. Casos de uso de Big Data

- Marketing
- Comercio y Retail
- Banca
- Industria.
- Otros.

BLOQUE 2:Prácticas de la asignatura

Prácticas

Se realizarán dos prácticas de Big Data donde aplicarán lo aprendido en clase en el ordenador, .

Competencias – Resultados de Aprendizaje

Competencias

Competencias Básicas

CB 02. Saber aplicar e integrar sus conocimientos, la comprensión de éstos, su fundamentación científica y sus capacidades de resolución de problemas en entornos nuevos y definidos de forma imprecisa, incluyendo contextos de carácter multidisciplinar tanto investigadores como profesionales altamente especializados

CB 03. Saber evaluar y seleccionar la teoría científica adecuada y la metodología precisa de sus campos de estudio para formular juicios a partir de información incompleta o limitada incluyendo, cuando sea preciso y pertinente, una reflexión sobre la responsabilidad social o ética ligada a la solución que se proponga en cada caso.

CB 07. Ser capaces de asumir la responsabilidad de su propio desarrollo profesional y de su especialización en uno o más campos de estudio.

Competencias Generales

CG. 1 Capacidad de organización y planificación en la identificación de problemas en el contexto de datos masivos

CG 02. Capacidad de análisis de datos masivos procedentes de diversas fuentes: texto, audio, numérica e imagen.

CG 03. Resolución de problemas y toma de decisiones en un entorno de datos masivos tanto cuantitativos como cualitativos.

CG 07. Capacidad de liderazgo y trabajo en equipo en la sociedad de la información.

CG 11. Capacidad para aprender y trabajar autónomamente en la sociedad de la información.

Competencias Específicas

CE 25. Comprender los conceptos más importantes asociados al Big Data y sus elementos más característicos.

Resultados de Aprendizaje

Al final de curso los alumnos deben ser capaces de:

Comprenderá los conceptos más importantes asociados al Big Data y sus elementos más característicos. En particular comprenderá el funcionamiento, problemática que resuelve, impacto de su uso y los casos reales de negocio en los que emplearlo.

RA1 Conoce y comprende las herramientas de apoyo a la decisión en el ámbito de la empresa.

RA2 Es capaz de implementar una solución computerizada de apoyo a la decisión.

RA1 Conoce y comprende la problemática del Big Data y las técnicas más avanzadas de computación y almacenamiento usadas en la actualidad.

RA3 Comprende y aplica las técnicas de MapReduce y otras técnicas avanzadas de computación específicas para el análisis de datos voluminosos.

CEO 27. Conocer y comprender las herramientas más extendidas de gestión de datos.

METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura	
Metodología Presencial: Actividades	Competencias
<p>1. Clase magistral y presentaciones generales. Exposición de los principales conceptos y procedimientos mediante la explicación por parte del profesor. Incluirá presentaciones dinámicas, pequeños ejemplos prácticos y la participación reglada o espontánea de los estudiantes. Además, el profesor propondrá la lectura de temas preparados por el profesor, artículos científicos y divulgativos que permitirán introducir o profundizar en un tema, con el objetivo de aportar ideas y generar debate (24 horas).</p>	<p>CE 25,CG 03, CG 02 CG 01, CB 03, CB 02</p>
<p>2. Resolución en clase de problemas prácticos. Resolución de unos primeros problemas para situar al alumno en contexto. La resolución correrá a cargo del profesor y los alumnos de forma cooperativa (2 horas).</p>	<p>CE 25,CG 03, CG 02 CG 01, CB 03, CB 02</p>
<p>3. Prácticas. Cada alumno realizará de forma aislada o en grupo una serie de prácticas regladas. Las prácticas finalizarán con la redacción de un informe o la redacción de las distintas experiencias (4 horas).</p>	<p>CE 25,CG 03, CG 02 CG 01, CB 03, CB 02</p>

Metodología No presencial: Actividades	Competencias
<p>1. Estudio individual del material. Actividad realizada individualmente por el estudiante para preparar el tema que se discutirá en clase. (60 horas).</p>	<p>CE 25,CG 03, CG 02 CG 01, CB 03, CB 02</p>
<p>2. Resolución de problemas prácticos a resolver fuera del horario de clase por parte del alumno. El alumno debe utilizar e interiorizar los conocimientos aportados en la materia. La corrección a la clase se realizará por parte de alguno de los alumnos o el profesor según los casos. La corrección individualizada de cada ejercicio la realizará el propio alumno u otro compañero según los casos (método de intercambio) (10 horas).</p>	<p>CE 25,CG 03, CG 02 CG 01, CB 03, CB 02</p>
<p>3. Prácticas de la asignatura. Realización de dos prácticas que plasme los conocimientos adquiridos en la asignatura. (10 horas).</p>	<p>CE 25,CG 03, CG 02 CG 01, CB 03, CB 02</p>

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

Actividades de evaluación	Criterios de evaluación	PESO
<p><u>Realización de exámenes:</u></p> <ul style="list-style-type: none"> Examen final. 	<ul style="list-style-type: none"> - Comprensión de conceptos mediante pruebas abiertas, enfocado a casos prácticos. - Aplicación de los conceptos adquiridos - Presentación y comunicación escrita. 	60%
<p><u>Evaluación del rendimiento teórico continuo:</u></p> <p>Realización de pruebas de seguimiento, debates y ejercicios.</p>	<ul style="list-style-type: none"> - Comprensión de conceptos mediante pruebas objetivas y one minute papers. - Debate y análisis de los documentos técnico/científicos. - Análisis e interpretación de los resultados obtenidos en la resolución de problemas mediante debates. 	30%
<p><u>Evaluación del rendimiento práctico continuo:</u></p> <ul style="list-style-type: none"> Casos a desarrollar en la clase de prácticas. 	<ul style="list-style-type: none"> - Comprensión de conceptos. - Aplicación de conceptos al desarrollo de casos sobre los nuevos conceptos planteados. - Aplicación de todos los conocimientos adquiridos durante el curso. 	10%

Criterios de Calificación

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

- Un 60% la calificación del examen final.
- Un 30% será la calificación de las pruebas de seguimiento y debates.
- Un 10% será la calificación de las prácticas.

La calificación en la **convocatoria extraordinaria**:

- Un 60% la calificación del examen final.
- Un 30% será la calificación de las pruebas de seguimiento y debates.
- Un 10% será la calificación de la práctica final.

Para aprobar la asignatura los alumnos tienen que tener al menos 5 puntos sobre 10 en el examen final, y en las práctica, tanto en la convocatoria ordinaria como en la extraordinaria.

RESUMEN PLAN DE LOS TRABAJOS Y CRONOGRAMA

Actividades Presenciales y No presenciales	Fecha de realización	Fecha de entrega
• Lectura y estudio de los contenidos teóricos	Después de cada clase	
• Resolución de los problemas propuestos o debates	Semanalmente	
• Preparación de las prácticas	Semanas 8, 9 y 10	Semanas 11
• Preparación del examen final		

RESUMEN HORAS DE TRABAJO DEL ALUMNO			
HORAS PRESENCIALES			
Lección magistral	Debates y resolución de problemas	Prácticas	Evaluación
24	2	4	2
HORAS NO PRESENCIALES			
Trabajo autónomo sobre contenidos teóricos	Trabajo autónomo sobre contenidos prácticos	Práctica final de la asignatura	Estudio
10	10	10	30
CRÉDITOS ECTS:			3 (90 horas)

BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

- Notas preparadas por el profesor disponibles en Moodle.
-

Bibliografía Complementaria

- Big Data: A Revolution That Will Transform How We Live, Work, and Think. [Viktor Mayer-Schonberger](#) (Autor), [Kenneth Cukier](#) (Autor)
- George Orwell. 1984 . ISBN: 9788499890944
- Michael Lewis. Moneyball: The Art of Winning an Unfair Game Paperback – March 17, 2004. ISBN: 978-0393324815 or the film 'Moneyball (2011)'. Sony Pictures Director: Bennett Miller