

## DATOS GENERALES DE LA ASIGNATURA

**ASIGNATURA:** Tecnologías de Almacenamiento de Datos

**PLAN DE ESTUDIOS:** Máster Universitario en Gestión y Análisis de Grandes Volúmenes de Datos: Big Data

**GRUPO:** 1920-01

**CENTRO:** Escuela Politécnica Superior

**CARÁCTER DE LA ASIGNATURA:** Obligatorio

**ECTS:** 6,0

**CURSO:** 1º

**SEMESTRE:** 1º Semestre

**IDIOMA EN QUE SE IMPARTE:**

Castellano, excepto las asignaturas de idiomas que se impartirán en el idioma correspondiente

## DATOS DEL PROFESOR

**NOMBRE Y APELLIDOS:** Abraham Requena Mesa

**EMAIL:** [arequena@uemc.es](mailto:arequena@uemc.es)

**TELÉFONO:** 983 00 10 00

**CV DOCENTE:**

Graduado en ingeniería informática por la Universidad de Sevilla.

Máster en ingeniería del software, especialidad en técnicas inteligentes (Universidad de Sevilla).

Profesor el año pasado de la UEMC en su máster Big Data. Además, colaboraciones con otras instituciones en este mismo ámbito.

**CV PROFESIONAL:**

3 años en una multinacional española trabajando en un proyecto Big Data para Endesa. De esos 3, uno lo he pasado dirigiendo parte de un proyecto anexo, también para Endesa.

Desde Junio de 2017, he ampliado mi rama de conocimiento y actualmente trabajo en un proyecto de Data Science con algoritmos de Machine Learning.

## DATOS ESPECÍFICOS DE LA ASIGNATURA

**DESCRIPTOR DE LA ASIGNATURA:**

Durante el desarrollo de la asignatura de TECNOLOGÍAS PARA EL ALMACENAMIENTO DE DATOS, realizaremos un recorrido por algunas de las herramientas que se sitúan en el ecosistema Big Data.

Comenzaremos con una introducción al mundo del Big Data, su definición y alguna de las aplicaciones que este famoso término tiene el mundo real. Big Data es un término en auge, pero también es un término 'virgen', del cual aún es difícil dar una definición clara y acotar a que llamamos con este nombre. Estudiaremos cuales son las claves y los conceptos básicos en los que se basa el Big Data para funcionar.

Luego, pasaremos a ver algunas bases de datos (BBDD) analíticas, su estructura y la manera en la que podemos gestionarla. Uno de los grandes problemas de hoy día, es gestionar la inmensa cantidad de datos que usamos y creamos a diario. Estos datos, no pueden ser tratados por sistemas o herramientas tradicionales, por lo que tenemos que buscar nuevas herramientas que nos faciliten este trabajo. Estudiaremos bases de datos no relacionales, y como podemos sacar análisis de sus datos de una manera simple aún siendo estos desestructurados.

Después de esto, comenzaremos a estudiar qué algoritmos podemos usar para procesar estos grandes conjuntos de datos. Al fin y al cabo, nuestro objetivo final va a ser analizar esos datos que almacenamos, y además vamos a tener que hacer este análisis en un tiempo prudencial. Por ello, vamos a estudiar que algoritmo podemos utilizar para este fin.

El alumno no es necesario que tenga conocimientos previos acerca del mundo de Big Data, aunque sí es recomendable tener conocimientos acerca de los lenguajes de programación java y scala. Por último, sería también recomendable tener ciertos conocimientos sobre comandos Linux.

#### CONTENIDOS DE LA ASIGNATURA:

##### 1. Tecnologías de almacenamiento de datos

1. Introducción al Big Data y a Hadoop
2. Apache Hadoop I
3. Apache Hadoop II
4. Apache Cassandra I
5. Apache Cassandra II
6. Apache Spark - SparkSQL
7. Apache Spark - Spark Streaming

#### RECURSOS DE APRENDIZAJE:

Los recursos de aprendizaje que se utilizarán en todas las asignaturas de la titulación (salvo las prácticas externas) para facilitar el proceso de enseñanza-aprendizaje, son:

- Campus online de la UEMC (Open Campus)
- Plataforma de Webconference (Adobe Connect)

Las comunicaciones con el profesor serán a través de Open Campus vía Mi correo, Tablón o/y Foro.

Para la asignatura de prácticas externas:

Los estudiantes tendrán acceso a toda la información sobre Prácticas Externas a través del Portal de Empleo y Prácticas de la UEMC (<https://empleo.uemc.es/>) creado a tal efecto y a través del cual se gestionarán todos los procesos administrativos fijados por el Centro de Empleo y Carrera Profesional (CEMCAP): selección, validación, firma de documentación, entrega de memoria, etc.

Además del acceso al Tutor Académico de Prácticas y al de la Empresa tendrán a su disposición al CEMCAP que les proporcionará el asesoramiento necesario para facilitar y realizar las Prácticas externas. Los estudiantes se comunicarán con el CEMCAP a través del correo electrónico ([cemcap@uemc.es](mailto:cemcap@uemc.es)) y el teléfono (983 00 1000 Ext. 12463).

### COMPETENCIAS Y RESULTADOS DE APRENDIZAJE A ADQUIRIR POR EL ALUMNO

#### COMPETENCIAS BÁSICAS:

- CB6. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación
- CB7. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio
- CB8. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios
- CB9. Que los estudiantes sepan comunicar sus conclusiones -y los conocimientos y razones últimas que las sustentan- a públicos especializados y no especializados de un modo claro y sin ambigüedades
- CB10. Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo

**COMPETENCIAS GENERALES:**

- CG01. Capacidad de organización y planificación
- CG02. Comunicación oral y escrita en lengua extranjera
- CG03. Habilidades básicas de informática
- CG04. Capacidad y habilidad para la toma de decisiones
- CG05. Capacidad para trabajar en equipos de carácter interdisciplinar
- CG06. Compromiso ético (saber aplicar la evidencia científica en la práctica profesional y mantener un compromiso ético y de integridad intelectual en el planteamiento de la investigación científica, básica y aplicada)
- CG07. Capacidad de crítica y autocrítica
- CG08. Habilidades interpersonales (tanto con miembros del entorno como con científicos/profesionales de otros centros)
- CG09. Reconocimiento a la diversidad y a la multiculturalidad
- CG10. Desarrollar hábitos de excelencia y calidad en el ejercicio profesional
- CG11. Capacidad para generar nuevas ideas (creatividad)

**COMPETENCIAS ESPECÍFICAS:**

- CE11. Capacidad para definir la arquitectura de una plataforma BigData y reconocer los servicios que conforman el ecosistema de una plataforma
- CE12. Capacidad para determinar el volumen de almacenamiento, memoria y capacidad de procesamiento en la plataforma para realizar un correcto dimensionamiento de los nodos que componen el cluster Big Data.

**RESULTADOS DE APRENDIZAJE:**

El alumno será capaz de:

- El alumno será capaz de utilizar las técnicas y herramientas básicas para almacenar y extraer información estructurada en bases de datos relacionales
- El alumno será capaz de reconocer los servicios que conforman la arquitectura de una plataforma de grandes volúmenes de datos

**BIBLIOGRAFÍA Y RECURSOS DE REFERENCIA GENERALES**

**BIBLIOGRAFÍA BÁSICA:**

- Rafael Caballero, Enrique Martín (2015): Las bases del Big Data. Catarata. ISBN: 978-8490970867
- Boris Lublinsky , Kevin T. Smith, Alexey Yakubovich (2014): Hadoop. Soluciones Big Data. Anaya. ISBN: 978-8441535916

**WEBS DE REFERENCIA:**

Web / Descripción

[Apache Hadoop](http://hadoop.apache.org/docs/stable/)(http://hadoop.apache.org/docs/stable/)

Web de Apache Hadoop

[Apache Cassandra](http://cassandra.apache.org/doc/latest/)(http://cassandra.apache.org/doc/latest/)

Web de Apache Cassandra

[Apache Spark](https://spark.apache.org/)(https://spark.apache.org/)

Web de Apache Spark

[Spark SQL](https://spark.apache.org/docs/latest/sql-programming-guide.html)(https://spark.apache.org/docs/latest/sql-programming-guide.html)

Documentación SparkSQL

[Spark Streaming](https://spark.apache.org/docs/latest/streaming-programming-guide.html)(https://spark.apache.org/docs/latest/streaming-programming-guide.html)

Documentación Spark Streaming

[HDFS](http://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-hdfs/hdfsdesign.html)(http://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-hdfs/hdfsdesign.html)

Documentación HDFS

[Scala Language](https://docs.scala-lang.org/tour/tour-of-scala.html)(https://docs.scala-lang.org/tour/tour-of-scala.html)

Tour por Scala

[Java Language](https://docs.oracle.com/javase/tutorial/java/index.html)(https://docs.oracle.com/javase/tutorial/java/index.html)

Tour por Java

[Comandos Linux](https://blog.desdelinux.net/mas-de-400-comandos-para-gnulinux-que-deberias-conocer/)(https://blog.desdelinux.net/mas-de-400-comandos-para-gnulinux-que-deberias-conocer/)  
 Comandos básicos

## PLANIFICACIÓN DEL PROCESO DE ENSEÑANZA-APRENDIZAJE DE LA ASIGNATURA

### METODOLOGÍAS:

La asignatura se imparte con un alto contenido práctico que mantiene un equilibrio entre las sesiones de videoconferencia en directo (que se utilizarán para presentar los contenidos teóricos previstos en cada uno de los temas de la asignatura), y las actividades de evaluación de continua a realizar por el alumno (entrega de trabajos, participación en debates, resolución de ejercicios, etc.) relacionadas directamente con esos contenidos.

### MÉTODO DIDÁCTICO:

El papel del profesor cobra importancia a través de la impartición de clases magistrales en tiempo real por videoconferencia que podrá utilizar para explicar los contenidos teóricos, resolver dudas que se planteen durante la sesión, ofrecer retroalimentación sobre las actividades de evaluación continua o realizar sesiones de tutoría de carácter grupal.

### MÉTODO DIALÉCTICO:

Se caracteriza por la participación de los alumnos en las actividades de evaluación continua de debate y la intervención de éstos a través del diálogo y de la discusión crítica (seminarios, grupos de trabajo, etc.). Utilizando este método el alumno adquiere conocimiento mediante la confrontación de opiniones y puntos de vista. El papel del profesor consiste en proponer a través de Open Campus temas referidos a la materia objeto de estudio que son sometidos a debate para, posteriormente, evaluar el grado de comprensión que han alcanzado los alumnos.

### MÉTODO HEURÍSTICO:

Este método puede desarrollarse de forma individual o en grupo a través de las actividades de evaluación continua (entregas de trabajos, resolución de ejercicios, presentaciones, etc.). El objetivo es que el alumno asuma un papel activo en el proceso de aprendizaje adquiriendo los conocimientos mediante la experimentación y la resolución de problemas.

### CONSIDERACIONES DE LA PLANIFICACIÓN:

Las ACTIVIDADES FORMATIVAS que se realizan en la asignatura son las siguientes:

**Clases teóricas:** Actividad dirigida por el profesor que se desarrollará de forma sincrónica en grupo. Para la realización de esta actividad en Open Campus, la UEMC dispone de herramientas de Webconference que permiten una comunicación unidireccional en las que el docente puede desarrollar sesiones en tiempo real con posibilidad de ser grabadas para ser emitidas en diferido.

**Actividades prácticas:** Actividades supervisadas por el profesor que se desarrollarán fundamentalmente de forma asíncrona, y de forma individual o en grupo:

- Actividades de debate. Se trata de actividades desarrolladas en el foro de Open Campus, en las que se genera conocimiento mediante la participación de los estudiantes en discusiones alrededor de temas de interés en las distintas asignaturas.
- Entregas de trabajos individuales o en grupo a partir de un enunciado o unas pautas de trabajo que establecerá el profesor.
- Resolución de ejercicios y problemas que el alumno debe realizar a través de Open Campus en un periodo de tiempo determinado. Esta actividad puede ser en formato test de evaluación.

**Tutorías:** Las tutorías podrán tener un carácter sincrónico o asíncrono y podrán desarrollarse de manera individual o en grupos reducidos.

Están previstas dos sesiones de tutoría por videoconferencia, una al inicio y otra al final del semestre. En la primera se presentará la asignatura y la guía docente y en la segunda, en las semanas previas a la evaluación final, se dedicará a la resolución de dudas de los estudiantes.

Además, el docente utiliza el Tablón, el Foro y el Sistema de correo interno de Open Campus para atender las necesidades y dudas académicas de los estudiantes.

#### SESIONES EN TIEMPO REAL

En la asignatura se planifican clases magistrales y tutorías a través de videoconferencias.

La asistencia a las videoconferencias no será obligatoria, pero sí recomendable para un adecuado seguimiento de la asignatura, la comprensión de los materiales y el desarrollo óptimo de las actividades de aprendizaje. En cualquier caso, salvo circunstancias excepcionales, será posible acceder a ellas en diferido a las 48 horas máximo desde su celebración.

#### EVALUACIÓN CONVOCATORIA ORDINARIA:

Evaluación continua 60%

Evaluación final 40%

##### 1. EVALUACIÓN CONTINUA

A lo largo de la planificación de la asignatura el alumno realizará **ACTIVIDADES DE EVALUACIÓN CONTINUA** que forman parte de la calificación de la asignatura con un peso del 60% sobre la nota final.

Para superar la evaluación continua, el alumno debe obtener una media de igual o superior a 5 entre todas las actividades. En el caso de no superar la evaluación continua, se guardan para la convocatoria extraordinaria aquellas actividades aprobadas.

El sistema de evaluación de esta asignatura acentúa el desarrollo gradual de competencias y resultados de aprendizaje y, por tanto, se realizará una evaluación continua a través de las distintas actividades de evaluación propuestas. El resultado de la evaluación continua se calcula a partir de las notas obtenidas en cada actividad teniendo en cuenta el porcentaje de representatividad en cada caso.

Todas las actividades son obligatorias y deberán entregarse en las fechas previstas para ello, teniendo en cuenta:

- Las actividades de evaluación continua (entrega de trabajos) se desarrollarán según se indica y, para ser evaluadas, los trabajos deberán ser entregados en la forma y fecha prevista y con la extensión máxima señalada. No se evaluarán trabajos entregados posteriormente a esta fecha o que no cumplan con los criterios establecidos por el profesor.
- La no entrega de una actividad de evaluación continua en forma y plazo se calificará con un 0 y así computarán en el cálculo de la nota de evaluación continua y final de la asignatura.
- Las actividades presentadas en grupo se evaluarán de forma grupal y por lo tanto la nota establecida será de aplicación a los miembros del grupo que hayan participado en su realización. Si un miembro del grupo no participa, se le evaluará con 0.
- La participación en los foros y debates vinculados a una actividad de evaluación continua se evaluará de forma individual o colectiva según se haya requerido la participación: el valor de las aportaciones, el número de aportaciones y respuestas en debate a las opiniones de los compañeros. Cualquier comentario aportado en el foro que suponga una falta de respeto a las opiniones de compañeros supondrá el suspenso de la evaluación continua de la asignatura. Además, cualquier tipo de copia o plagio por mínimo que sea, supondrá una calificación de 0 en la actividad correspondiente.
- Las actividades de evaluación continua (tipo test) se desarrollarán con anterioridad a la realización de las pruebas de evaluación parcial (solo para el caso de grados) y final de la asignatura

Los alumnos accederán a través de Open Campus a las calificaciones de las actividades de evaluación continua en un plazo no superior a 15 días naturales desde la fecha límite de entrega de la actividad conforme al calendario, excepto causas de fuerza mayor en cuyo caso se informará al alumno a través del Tablón.

##### 2. EVALUACIÓN FINAL

La evaluación continua se complementará con una **EVALUACIÓN FINAL** que se realizará al finalizar el periodo lectivo en cada asignatura

periodo lectivo en cada asignatura.

#### A. TITULACIONES DE GRADO:

La prueba constará de parte práctica y teórica, suponiendo un 40% de la calificación sobre la nota final.

La evaluación final de la asignatura se desarrollará del siguiente modo:

- A mitad de cada semestre se ofrece al alumno realizar de forma voluntaria un parcial para eliminar materia.
- Para eliminar la materia es necesario que el alumno lo supere al menos con un 5. En este caso, se le guardaría la nota del parcial hasta la convocatoria extraordinaria. El alumno sólo podrá presentarse a la segunda parte de la asignatura bien en convocatoria ordinaria o extraordinaria.
- En convocatoria ordinaria, la prueba final constará de dos exámenes (primera y segunda parte de la asignatura)
- En el caso de que el alumno hubiera superado y eliminado materia con el primer parcial, sólo se presentará a la segunda parte. Para superar la asignatura se hará la media siempre que en la segunda parte se obtenga al menos un 4 y la media supere el 5.
- En el caso de que el alumno no hubiera superado el primer parcial, se podrá presentar a ambas partes. Para superar la asignatura se hará la media de ambas partes siempre que se obtenga al menos un 4 en cada una y la media supere el 5.
- El alumno tendrá la posibilidad, siempre dentro de los tres días siguientes a la publicación de las notas, a renunciar a su calificación, y presentarse en la siguiente convocatoria
- El aplazamiento concedido por la Universidad para la realización de una evaluación final tendrá lugar en los periodos de evaluación previstos y comunicados el aula de asesoría personal de la titulación.

La nota final se corresponderá con la media aritmética del resultado obtenido en cada una de las partes. En caso de no superación, se guarda la parte aprobada para la convocatoria extraordinaria.

#### B. TITULACIONES DE MÁSTER

La evaluación continua se complementará con una **evaluación final online** que se realizará al finalizar el periodo lectivo en cada asignatura. Los exámenes serán eminentemente prácticos, de manera que los alumnos puedan disponer de los apuntes.

La prueba supondrá un 40% de la calificación sobre la nota final.

- El alumno tendrá la posibilidad, siempre dentro de los tres días siguientes a la publicación de las notas, a renunciar a su calificación, y presentarse en la siguiente convocatoria
- El aplazamiento concedido por la Universidad para la realización de una evaluación final tendrá lugar en los periodos de evaluación previstos y comunicados el aula de asesoría personal de la titulación.

#### 3. NOTA GLOBAL

La **NOTA GLOBAL** de la asignatura se obtiene ponderando la calificación de la evaluación continua y de la evaluación final según los siguientes porcentajes, y debiendo tener aprobadas ambas partes, continua y final, para superar la asignatura.

Si un alumno no se presenta a la prueba de evaluación final, su calificación en la convocatoria será de “No presentado”, con independencia de que haya realizado alguna actividad de evaluación continua.

De igual modo si el alumno no entrega ninguna actividad de evaluación continua, obtendrá la calificación de “No presentado”, con independencia de que haya aprobado la prueba de evaluación final, en cuyo caso, se le guardaría su calificación para la convocatoria extraordinaria

Por otro lado, si un alumno se presenta tanto a la evaluación continua, como a la evaluación final, y suspende una de las dos partes, la nota final de la convocatoria será la nota de la parte suspendida, con independencia de la nota obtenida en la parte aprobada.

#### EVALUACIÓN CONVOCATORIA EXTRAORDINARIA:

Evaluación continua 60%

Evaluación final 40%

### 1. EVALUACIÓN CONTINUA<br/>

Los estudiantes que no hayan superado la asignatura en la convocatoria ordinaria, porque hayan suspendido la nota final de las actividades de evaluación continua, podrán presentarse a las pruebas establecidas por el profesor en la convocatoria extraordinaria.<br/>

En convocatoria extraordinaria, el alumno solo podrá entregar las actividades de evaluación continua no superadas, guardándose la calificación de las aprobadas, siempre y cuando no haya obtenido una nota igual o superior a 5 en el conjunto de las actividades de evaluación continua, en cuyo caso, se guardarán las notas de todas las actividades.<br/>

### 2. EVALUACIÓN FINAL<br/>

Los estudiantes que no hayan superado la asignatura en la convocatoria ordinaria, porque hayan suspendido la evaluación final, podrán presentarse a la prueba de evaluación final establecida por el profesor en la convocatoria extraordinaria.<br/>

#### A. TITULACIONES DE GRADO:<br/>

- En convocatoria extraordinaria, la prueba final también constará de dos exámenes (primera y segunda parte de la asignatura)<br/>
- En el caso de que el alumno hubiera superado el parcial (al menos un 5) o una de las partes en convocatoria ordinaria (al menos un 5), esta calificación se mantiene para la extraordinaria, presentándose el alumno sólo a lo suspenso. Para superar la asignatura se hará la media entre lo aprobado en ordinaria y la calificación que haya sacado en extraordinaria siempre que se obtenga al menos un 4 y la media supere el 5.<br/>
- En el caso de que el alumno tuviera que presentarse a ambas partes, para superar la asignatura se hará la media siempre que se obtenga al menos un 4 en cada parte y la media supere el 5.<br/>

#### B. TITULACIONES DE MÁSTER<br/>

La evaluación continua se complementará con una **evaluación final online** que se realizará al finalizar el periodo lectivo en cada asignatura. Los exámenes serán eminentemente prácticos, de manera que los alumnos puedan disponer de los apuntes.<br/>

La prueba supondrá un 40% de la calificación sobre la nota final.<br/>

### 3. NOTA GLOBAL<br/>

En la convocatoria extraordinaria, la **nota global** de la asignatura se obtiene ponderando la calificación de la evaluación continua y de la evaluación final, de la misma forma que en la convocatoria ordinaria.<br/>

Al igual que en la convocatoria ordinaria, en la convocatoria extraordinaria es necesario superar tanto la evaluación continua como la evaluación final para aprobar la asignatura.<br/>

Si un alumno no se presenta a la prueba de evaluación final, su calificación en la convocatoria será de “No presentado”, con independencia de que haya realizado alguna actividad de evaluación continua.<br/>

De igual modo si el alumno no entrega ninguna actividad de evaluación continua, obtendrá la calificación de “No presentado”, con independencia de que haya aprobado la prueba de evaluación final.<br/>

Por otro lado, si un alumno se presenta tanto a la evaluación continua, como a la evaluación final, y suspende una de las dos partes, la nota final de la convocatoria será la nota de la parte suspensa, con independencia de la nota obtenida en la parte aprobada.<br/>

## SISTEMAS DE EVALUACIÓN:

SISTEMA DE EVALUACIÓN	PORCENTAJE (%)
Pruebas de ejecución de tareas reales y/o simuladas	15%
Pruebas de respuesta corta	20%
Pruebas de respuesta larga, de desarrollo	20%
Pruebas objetivas	10%
Trabajos y proyectos	35%