



Universidad
de Alcalá

GUÍA DOCENTE

Fundamentos de la Ciencia de Datos

Grado en
Ingeniería Informática

Universidad de Alcalá

Curso Académico 2020/2021

4º Curso – 1^{er} Cuatrimestre

GUÍA DOCENTE

Nombre de la asignatura:	Fundamentos de la Ciencia de Datos
Código:	781006
Titulación en la que se imparte:	Grado en Ingeniería Informática
Departamento y Área de Conocimiento:	Departamento Ciencias de la Computación
Carácter:	Optativa de tecnología específica ⁽¹⁾
Créditos ECTS:	6
Curso y cuatrimestre:	Cuarto Curso /Primer Cuatrimestre
Profesorado:	Juan José Cuadrado Gallego Consultar en la página web del departamento
Horario de Tutoría:	Se indicará el primer día de clase
Idioma en el que se imparte:	Español (English Friendly)

1.A PRESENTACIÓN

Los planes de estudios de los Grados en Ingeniería Informática, Ingeniería de Computadores y Sistemas de Información están estructurados en tres bloques de formación: Básica, Obligatoria y Optativa. Los bloques de formación básica y obligatoria cubren el cuerpo de conocimiento principal propuesto por los informes o guías curriculares: *Computing Curricula: Computer Science 2001* [CC 2001],

Computing Curricula Software Engineering 2004 [SE 2004], *Computing Curricula 2005*; de ACM-IEEE, *Computer Engineering* de 2004 [CE 2004]; y la guía IS 2002, de la AIS [AIS 2002].

La materia de Fundamentos de la Ciencia de Datos se encuadra dentro del bloque de Formación Optativa, que consta de 57 ECTS para los Grados en Ingeniería Informática e Ingeniería de Computadores y de 45 ECTS para el grado en Sistemas de Información. En los tres grados, para obtener 45 créditos se podrán obtener con una combinación de módulos y materias optativas, un mínimo de 30 ECTS; prácticas externas, un máximo de 15 ECTS; créditos obtenidos mediante el programa de bonocréditos, un máximo de 9 ECTS. Y para los Grados en Ingeniería Informática e Ingeniería de Computadores los 12 grados restantes se obtendrán a través de materias transversales definidas para toda la universidad. Los módulos y materias optativas, todas de 6 ECTS, entre las que se encuentra Fundamentos de la Ciencia de Datos, están diseñadas para intensificar la formación del alumno en materias específicas que complementan la formación básica y obligatoria.

La asignatura Fundamentos de la Ciencia de Datos está orientada a la intensificación de los conocimientos pertenecientes a la materia de Ingeniería del Conocimiento relacionados con la Recuperación de Información y

Procesamiento automático de la información y transformación en conocimiento. Lo que de forma más general se conoce como Big Data, definición que fundamentalmente comprende el estudio y conocimiento de los fundamentos básicos y las técnicas utilizadas en Datawarehousing y Data Mining.

¹ Las asignaturas optativas de tecnología específica dan lugar a la mención y en este Grado son obligatorias de cursar.

Los contenidos que el alumno verá en esta asignatura son de la máxima actualidad e interés, ya que se enmarcan dentro del fenómeno conocido como el Big Bang Data o la explosión de los datos, que está relacionado con la generación, el procesamiento y, sobre todo, la interpretación o búsqueda de información de interés en grandes bases de datos. La asignatura Fundamentos de la Ciencia de Datos dotará al alumno con los conocimientos básicos que le permitirán, mediante el desarrollo y la profundización de los mismos, convertirse en un científico de datos, perfil para el que, según una de las consultoras más prestigiosas del sector, se prevé que se crearán 1.2 millones de empleos en Europa en los próximos dos años.

Prerrequisitos y Recomendaciones:

Para la asignatura Fundamentos de la Ciencia de Datos se recomienda haber superado las materias básicas y obligatorias: *Estadística y Matemáticas*.

1.B PRESENTATION

The curricula of Degrees in Computer Engineering, Computer Engineering and Information Systems are structured in three blocks of training: Basic, Mandatory and Optional. The blocks of basic and mandatory training covering the main body of knowledge proposed by the reports or curriculum guides: Computing Curricula: Computer Science 2001 [CC 2001] Software Engineering Computing Curricula 2004 [SE 2004] Computing Curricula 2005; ACM-IEEE Computer Engineering 2004 [CE 2004]; and IS 2002, AIS [AIS 2002] guide.

Matter Automatic Knowledge Extraction falls within the block Optional Formation, consisting of 57 ECTS for degrees in Computer Engineering and Computer Engineering and 45 ECTS for the degree Information systems. In the three degrees, to obtain 45 credits may be obtained with a combination of modules and electives, a minimum of 30 ECTS; external practices, a maximum of 15 ECTS; Credits earned through the program bonocréditos, a maximum of 9 ECTS. And for Degrees in Computer Engineering and Computer Engineering remaining 12 degrees were obtained through cross materials defined for the whole university. Modules and optional subjects, all of 6 ECTS, including Automatic Extraction is Knowledge, they are designed to enhance students' training in specific areas that complement the basic and mandatory training.

Automatic extraction of knowledge course is aimed intensification of knowledge pertaining to the subject of related Knowledge Engineering and Information Retrieval

Automatic information processing and transformation into knowledge. What more generally it is known as Big Data, definition comprising mainly study and knowledge of the basic fundamentals and techniques used in data mining.

The contents that the student will see in this subject are of utmost current interest because are part of the phenomenon known as the Big Bang Data or the explosion of data, which is related to the generation, processing and above all interpretation or search relevant information in large databases. The Automatic Extraction of Knowledge course will equip students with basic knowledge that will allow, by developing and deepening them, become a data scientist, profile for which, according to one of the most prestigious consulting firms in the sector, it predicts that 1.2 million jobs will be created in Europe over the next two years.

Prerequisites and Recommendations:

Subject to Automatic Knowledge Extraction is recommended to have overcome the basic and mandatory subjects: Statistics and Mathematics.

2. COMPETENCIAS

Competencias generales:

CG8 Conocimiento de las materias básicas y tecnologías, que capaciten para el aprendizaje y desarrollo de nuevos métodos y tecnologías, así como las que les doten de una gran versatilidad para adaptarse a nuevas situaciones.

CG9 Capacidad para resolver problemas con iniciativa, toma de decisiones, autonomía y creatividad. Capacidad para saber comunicar y transmitir los conocimientos, habilidades y destrezas de la profesión de Ingeniero Técnico en Informática.

Competencias específicas:

CC4 Capacidad para conocer los fundamentos, paradigmas y técnicas propias de los sistemas inteligentes y analizar, diseñar y construir sistemas, servicios y aplicaciones informáticas que utilicen dichas técnicas en cualquier ámbito de aplicación.

CC5 Capacidad para adquirir, obtener, formalizar y representar el conocimiento humano en una forma computable para la resolución de problemas mediante un sistema informático en cualquier ámbito de aplicación, particularmente los relacionados con aspectos de computación, percepción y actuación en ambientes o entornos inteligentes.

CC7 Capacidad para conocer y desarrollar técnicas de aprendizaje computacional y diseñar e implementar aplicaciones y sistemas que las utilicen, incluyendo las dedicadas a extracción automática de información y conocimiento a partir de grandes volúmenes de datos.

Resultados del aprendizaje de la asignatura de Fundamentos de la Ciencia de Datos:

A continuación, se enumeran los Resultados del aprendizaje que el estudiante deberá adquirir al cursar la asignatura de Fundamentos de la Ciencia de Datos. Dichas Competencias son:

RA1. Conocer las diferencias entre aprendizaje supervisado y no supervisado además de su contexto de aplicación.

RA2. Ser capaz de modelizar y resolver problemas de aprendizaje mediante paradigmas computacionales (supervisado y no supervisado).

RA3. Ser capaz de explicar los modelos aplicados y resultados obtenidos del aprendizaje, incluyendo sus limitaciones y bondades.

RA4. Ser capaz de entender diferentes tipos de datos, problemas y posibles transformaciones para su aplicación tanto en el aprendizaje como en la evaluación de los resultados.

3. CONTENIDOS

Unidad temática 1. Big data, Data Warehousing

Tema 0. Introducción a Big Data, Data Warehousing y Data mining. 0.1. Big Data.

0.2. Data Warehousing

0.3. Data mining

Tema 1. Adquisición, preprocesamiento y modelado de datos.

1.1. Adquisición de datos.

1.2. Preprocesamiento de datos.

1.3. Modelado de datos.

Unidad temática 2. Data Mining

Tema 2. Búsqueda de patrones

2.1. Bases conceptuales de la búsqueda de patrones.

2.2. Técnicas de asociación.

Tema 3. Clasificación supervisada

3.1. Bases conceptuales de la clasificación supervisada.

3.2. Técnicas de clasificación supervisada.

Tema 4. Clasificación no supervisada 4.1. Bases conceptuales de la clasificación no supervisada.

4.2. Técnicas de detección de clusters.

Tema 5. Identificación de datos anómalos

5.1. Bases conceptuales de la identificación de datos anómalos.

5.2. Técnicas de detección de outliers.

Unidad temática 3. Visualización de datos

Tema 6. Visualización de datos

6.1. Bases conceptuales de la visualización de datos.

6.2. Técnicas de visualización de datos.

Unidad temática 4. Big data, Lenguaje de programación

Tema 7: Lenguaje de programación utilizado en Big Data: R.

7.1. Bases del lenguaje R.

7.2. Aplicación de R en búsqueda de patrones.

7.3. Aplicación de R en clasificación supervisada

7.4. Aplicación de R en clasificación no supervisada

7.5. Aplicación de R identificación de datos anómalos

7.6. Aplicación de R en visualización de datos

Unidades Temáticas	Temas	Total clases de dedicación *	
Big data, Data Warehousing	T0. Introducción a Big Data, Data Warehousing y Data mining. T1. Adquisición, preprocesamiento y modelado de datos.	T0. 1 CA T1. 1 CA	8 h 0,6 ECTS
Data Mining	T2. Búsqueda de patrones T3. Clasificación supervisada T4. Clasificación no supervisada T5. Identificación de datos anómalos	T2. 1 CA T3. 1 CA T4. 1 CA T5. 1 CA	16 h 1,2 ECTS
Visualización de datos	T6. Visualización de datos	T6. 1 CA PEC 1	8 h 0,6 ECTS
Lenguaje de Programación	T7: Lenguaje de programación utilizado en Big Data: R.	T7-T1. 1 CL T7-T2. 1 CL T7-T3. 1 CL T7-T4. 1 CL T7-T5. 1 CL T7-T6. 1 CL PEC 2	28 h 2 ECTS

*Incluye EPC (pruebas de evaluación continua). Acrónimos: CA: Clase Aula // CL: Clase Laboratorio

Cronograma *	
Semana	Contenido
01 ^a	T0. Introducción a Big Data, Data Warehousing y Data mining.
02 ^a	T1. Adquisición, preprocesamiento y modelado de datos
03 ^a	T7: R. Datos
04 ^a	T2. Búsqueda de patrones
05 ^a	T7: R. Aplicación a la búsqueda de patrones
06 ^a	T3. Clasificación supervisada
07 ^a	T7: R. Aplicación a la clasificación supervisada
08 ^a	T4. Clasificación no supervisada
09 ^a	T7: R. Aplicación a la clasificación no supervisada
10 ^a	T5. Identificación de datos anómalos
11 ^a	T7: R. Aplicación a la identificación de datos anómalos
12 ^a	T6. Visualización de datos
13 ^a	T7: R. Aplicación a la visualización de datos
14 ^a	Seminario: Desafíos futuros de la Data Science
15 ^a	PEC 1. T0-T6 // PEC2. T7

*Puede sufrir alguna variación en función las fiestas recogidas en el calendario académico oficial
Acrónimos: PEC: Prueba de Evaluación Continua // PL: Prueba de Laboratorio

4. METODOLOGÍAS DE ENSEÑANZA-APRENDIZAJE.- ACTIVIDADES FORMATIVAS

4.1. Distribución de créditos (especificar en horas)

Número de horas presenciales:	56 horas + 4 horas de examen de evaluación
Número de horas del trabajo propio del estudiante:	90 h
Total horas	150

Acrónimos: CA: Clase Aula // CL: Clase Laboratorio

Al ser una asignatura de 6 ECTS le corresponden 150 horas, de las cuales 50 deben corresponder a presenciales y de ejercicios de evaluación y 100 a no presenciales. La modificación a 10 horas, 0,7 ECTS presenciales y de ejercicios de evaluación más está aprobada por la Junta de Escuela.

Estrategias metodológicas

Clases presenciales	<p>Las clases presenciales se realizarán en el laboratorio. Habrá dos tipos de sesiones:</p> <ul style="list-style-type: none"> ▪ Sin utilización de ordenadores: Se presentarán los contenidos teóricos de la asignatura, mediante herramientas de presentación y se resolverán casos prácticos y ejercicios que complementarán los conocimientos teóricos. ▪ Con utilización de ordenadores: Se resolverán, mediante la utilización del lenguaje de programación R los mismos casos prácticos y ejercicios que se hayan resuelto en la clase previa, en la que no se utilizaban ordenadores. <p>La aplicación práctica, mediante la resolución de cuestiones y ejercicios, tendrá un papel fundamental, como estrategia metodológica, en la impartición de la asignatura. Y en consecuencia se pedirá al alumno un alto grado de implicación y participación en la resolución de las cuestiones y ejercicios, tanto de forma individual, como en grupo. Como parte de dicho trabajo podrán realizarse presentaciones orales por los alumnos, a sus compañeros y a los profesores, del trabajo realizado individualmente o en grupo.</p>
Trabajo autónomo	<p>Dentro del trabajo autónomo se realizarán las siguientes actividades:</p> <ul style="list-style-type: none"> ▪ Estudio de los conceptos teóricos impartidos. ▪ Resolución de ejercicios. Esta será la actividad fundamental que el alumno deberá realizar como trabajo autónomo. Será muy importante que el alumno la dedique suficiente esfuerzo, ya que de su correcta ejecución dependerá en gran medida el aprovechamiento de las clases presenciales. ▪ Resolución de cuestiones con el lenguaje de programación R. ▪ Realización de resúmenes textuales y esquemas gráficos sobre los conceptos impartidos. ▪ Consulta de fuentes y recursos bibliográficos o electrónicos.

Tanto las clases presenciales como el trabajo autónomo estarán apoyados por tutorías, en el horario oficial establecido para la asignatura, que podrán realizarse tanto de forma individual como en grupo. Además, los estudiantes podrán solicitar tutorías fuera del horario oficial establecido para las mismas, en los casos en los que no hayan podido acudir a las mismas.

Materiales y recursos

1. Entorno de programación R

Los alumnos podrán utilizar toda la infraestructura disponible del aula de teoría y del Laboratorio donde dispondrán de un ordenador con el software necesario para realizar los trabajos que se le pidan en la asignatura. Además, el campus online pone a disposición del alumno una de las principales plataformas informática de tele formación Blackboard Learn 9.1, con la que podrá seguir a través de internet el desarrollo de la asignatura.

2. Entorno de programación R

R es una de las principales soluciones de software realizar estudios de Big Data del mundo y ofrece paquetes y métodos de análisis con la que resolver los problemas de Big Data para negocio e investigación. Además permite programar soluciones propias. El entorno estará instalado en los ordenadores del laboratorio y además se enseñará a los a descargarlo e instalarlo. En el Aula Virtual de la asignatura de Fundamentos de la Ciencia de Datos los alumnos encontrarán también manuales de utilización de R, que completarán la información sobre su manejo que se impartirá en las clases de laboratorio.

3. Biblioteca de la Universidad de Alcalá

Los alumnos tendrán a su disposición los recursos bibliográficos de la biblioteca de la Universidad de Alcalá. Se procurará que existan en la misma, especialmente en la Biblioteca Politécnica suficientes copias de todos los libros recomendados. En el caso de no encontrar suficientes copias en la biblioteca de alguno de los libros recomendados, el alumno debe dirigirse al profesor responsable de su grupo o al coordinador de la asignatura, trasladarle el problema y, en el caso de que se vea necesario, se solicitará la compra de nuevas copias del libro.

5. EVALUACIÓN: Procedimientos, criterios de evaluación y de calificación

La evaluación de la asignatura de Fundamentos de la Ciencia de Datos se realizará siguiendo la *Normativa Reguladora de los Procesos de Evaluación de los Aprendizajes* aprobada en Consejo de Gobierno de 5 de mayo de 2016.

Sobre la *Programación del Proceso de Evaluación*, en lo referido a *Programación y convocatorias*, en esta guía docente, se recogen los siguientes aspectos:

Tipo, características y modalidad de los instrumentos y estrategias que forman parte del proceso de evaluación de la asignatura

Sobre las *Convocatorias*, a continuación se recoge de manera expresa el sistema de evaluación y el tipo de pruebas, tanto de la convocatoria ordinaria como de la extraordinaria, y se especifican los resultados de la convocatoria ordinaria que se ha estimado oportuno, tener en cuenta para la convocatoria extraordinaria:

Sobre la *superación de las prácticas obligatorias presenciales*. En la asignatura de Fundamentos de la Ciencia de Datos la superación de las prácticas obligatorias presenciales no será considerada elemento imprescindible de la evaluación, tanto en la convocatoria ordinaria como en la extraordinaria; pero sí su realización.

Sobre la *evaluación continua*, además la aplicación de todos los aspectos contemplados en la normativa, se establece lo siguiente:

Para poder evaluarse por evaluación continua habrá que haber asistido al menos a 12 clases.

El proceso de evaluación continua utilizará diferentes estrategias y recogerá, a largo del cuatrimestre, evidencias que guardan relación con todo el proceso de enseñanza-aprendizaje durante la impartición de la asignatura. Dichas evidencias consistirán en las siguientes pruebas:

1. Prueba de Evaluación Continua 1 (PEC 1): Consistirá en preguntas teóricas y ejercicios escritos basados en los contenidos de los temas 0, 1, 2, 3, 4, 5 y 6.
2. Prueba de Evaluación Continua 2 (PEC 2): Consistirá en la realización de ejercicios prácticos con el lenguaje de programación R basados en el contenido del tema 7 aplicado a los temas 1, 2, 3, 4, 5 y 6.
3. Prueba de Laboratorio 1 (PL 1): Consistirá en la realización de una práctica, que se realizará con el lenguaje de programación R, basado en los contenidos del tema 1.
4. Prueba de Laboratorio 2 (PL 2): Consistirá en la realización de una práctica, que se realizará con el lenguaje de programación R, basado en los contenidos del tema 2.
5. Prueba de Laboratorio 3 (PL 3): Consistirá en la realización de una práctica, que se realizará con el lenguaje de programación R, basado en los contenidos del tema 3.
6. Prueba de Laboratorio 4 (PL 4): Consistirá en la realización de una práctica, que se realizará con el lenguaje de programación R, basado en los contenidos del tema 4.
7. Prueba de Laboratorio 5 (PL 5): Consistirá en la realización de una práctica, que se realizará con el lenguaje de programación R, basado en los contenidos del tema 5.
8. Prueba de Laboratorio 6 (PL 6): Consistirá en la realización de una práctica, que se realizará con el lenguaje de programación R, basado en los contenidos del tema 6.

Sobre el *desarrollo de las pruebas de evaluación* a continuación se especifica la duración de cada una de las pruebas:

1. PEC 1 y PEC 2: Tendrán una duración máxima de 1h 55' y se realizarán en el aula.
2. PL 1, PL 2, PL 3, PL 4, PL5 y PL6: Serán realizadas como parte del trabajo autónomo del alumno, durante un plazo mínimo de dos semanas.

En lo referido a los *procedimientos de evaluación*, para la asignatura de Fundamentos de la Ciencia de Datos, además la aplicación de todos los aspectos contemplados en la normativa, para la *evaluación final* se establece lo siguiente:

La Prueba de Evaluación Final Ordinaria (PEF-O): Consistirá en preguntas basadas en los contenidos de los 7 temas de la asignatura. Tendrá lugar en la fecha oficialmente fijada por la dirección de los estudios para la realización del examen final.

A continuación se especifica la duración de la Prueba de Evaluación Final Ordinaria (PEF-O): Tendrá una duración máxima de 3h 50' y se realizará en las aulas oficialmente asignadas por la dirección de los estudios.

En lo referido a los *procedimientos de evaluación*, para la asignatura de Fundamentos de la Ciencia de Datos, además la aplicación de todos los aspectos contemplados en la normativa, para la *convocatoria extraordinaria* se establece lo siguiente:

La convocatoria extraordinaria (PEF-E) tendrá el mismo número de pruebas y con el mismo formato que la evaluación final. Aquellos alumnos que hayan aprobado las pruebas PL o PEC durante la evaluación continua no tendrán que realizar todos los ejercicios de R la prueba, en el caso de que los hubiera.

La convocatoria extraordinaria tendrá lugar en la fecha oficialmente fijada por la dirección de los estudios.

Los resultados de las Pruebas de Laboratorio obtenidos en la convocatoria ordinaria, tanto continua como final, se tendrán en cuenta para la convocatoria extraordinaria.

Criterios de evaluación

Los criterios de evaluación que se aplicarán en la asignatura de Fundamentos de la Ciencia de Datos, están vinculados a las competencias que se han definido para la misma, incluidas las competencias actitudinales, y estarán orientados a recoger tanto el nivel de comprensión de los contenidos e ideas básicas, como la capacidad de aplicar los conceptos en la resolución de problemas, sin y con la ayuda de herramientas software.

Además se tendrá en cuenta la actitud de colaboración y proactiva en la realización de las tareas tanto individuales como de grupo y el grado de interés mostrado en la realización y cumplimiento de las actividades de clase y de trabajo autónomo, por lo que será fundamental la entrega en tiempo y formato de todas las actividades de prácticas. Para superar la asignatura hay que obtener todas las competencias de la misma.

La evaluación de la adquisición de competencias tendrá en cuenta, entre otros los siguientes criterios de evaluación:

1. CE1, Dominio y aplicación de los contenidos y conceptos básicos: 40%
2. CE2, Resolución de problemas prácticos propuestas y manejo de herramientas software: 40%
3. CE3, Interés y motivación en la realización de las tareas y prácticas: 20%

Los resultados de evaluación obtenidos estarán fundamentados sobre las siguientes rúbricas:

Sobresaliente: Un excelente dominio de la comprensión y aplicación de los contenidos y conceptos básicos, al igual que en la resolución de problemáticas prácticas relativas a la Fundamentos de la Ciencia de Datos informática. Muestra un elevado interés en la realización y entrega de las tareas y prácticas.

Notable: Un alto manejo de la comprensión y aplicación de los contenidos y conceptos básicos, al igual que en la resolución de problemáticas prácticas relativas a la Fundamentos de la Ciencia de Datos informática. Interés notable en la realización y entrega de las tareas y prácticas.

Aprobado: Suficiente manejo de la comprensión y aplicación de los contenidos y conceptos básicos, al igual que en la resolución de problemáticas prácticas relativas a la Fundamentos de la Ciencia de Datos informática. Discreto interés en la realización y entrega de las tareas y prácticas.

Suspense: Escaso dominio comprensivo y de aplicación de los contenidos y conceptos básicos, al igual que en la resolución de problemáticas prácticas relativas a la Fundamentos de la Ciencia de Datos informática. Bajo interés en la realización y entrega de las tareas y prácticas.

Criterios de calificación

Evaluación Continua

Los criterios de calificación de las pruebas de evaluación continua de la asignatura de Fundamentos de la Ciencia de Datos serán los siguientes:

1. PEC 1: 20% de la calificación final de la asignatura. Se evalúa el dominio y aplicación de los contenidos y conceptos básicos.
2. PEC 2: 20% de la calificación final de la asignatura. Se evalúa el dominio y aplicación de los contenidos y conceptos básicos.
3. PL 1: 10% de la calificación final de la asignatura. Se evalúa la resolución de problemas prácticos propuestos y manejo de herramientas software.
4. PL 2: 10% de la calificación final de la asignatura. Se evalúa la resolución de problemas prácticos propuestos y manejo de herramientas software.
5. PL 3: 10% de la calificación final de la asignatura. Se evalúa la resolución de problemas prácticos propuestos y manejo de herramientas software.
6. PL 4: 10% de la calificación final de la asignatura. Se evalúa la resolución de problemas prácticos propuestos y manejo de herramientas software.
7. PL 5: 10% de la calificación final de la asignatura. Se evalúa la resolución de problemas prácticos propuestos y manejo de herramientas software.
8. PL 6: 10% de la calificación final de la asignatura. Se evalúa la resolución de problemas prácticos propuestos y manejo de herramientas software.

Participación e implicación del alumno en el proceso enseñanza-aprendizaje, que será valorada en la calificación final de la asignatura. Dicha participación e implicación del alumno en el proceso enseñanza-aprendizaje se valorará fundamentalmente a través de los resultados obtenidos por el alumno en un conjunto de ejercicios que se le irán proponiendo durante el curso. Se evalúa el interés y motivación en la realización de las tareas y Prácticas

Competencias	Resultado Aprendizaje	Criterio de Evaluación	Instrumento de Evaluación	Porcentaje
CG8, CG9, CC4, CC5, CC7	RA1, RA2, RA3, RA4	CE1	PEC1	20
		CE1, CE2	PEC2	20
		CE2, CE3	PL1	10
		CE2, CE3	PL2	10
		CE2, CE3	PL3	10
		CE2, CE3	PL4	10
		CE2, CE3	PL5	10
		CE2, CE3	PL5	10

Evaluación Final y Convocatoria Extraordinaria.

Los criterios de calificación de las pruebas de evaluación final y convocatoria extraordinaria, de la asignatura de Fundamentos de la Ciencia de Datos, serán los siguientes:

PEF-O y Prueba de Evaluación Final Extraordinaria (PEF-E): Se evalúa el dominio y aplicación de los contenidos de la asignatura y conceptos básicos mediante la respuesta a preguntas teóricas.

Convocatoria final y extraordinaria:

Competencias	Resultado Aprendizaje	Criterio de Evaluación	Instrumento de Evaluación	Porcentaje
CG8, CG9, CC4, CC5, CC7	RA1, RA2, RA3, RA4	CE1	PEF-O, PEF-E	50
		CE2	PEF-O, PEF-E	50

6.BIBLIOGRAFÍA

Bibliografía Básica

La bibliografía básica recomendada está compuesta por 4 libros, 2 sobre Big Data y 2 sobre R. Los libros son:

- J. Han et al., *Data Mining. Concepts and Techniques*, Morgan Kaufman, 2012. ISBN: 978-0-12-381479-1. Se trata de uno de los libros de referencia más importantes sobre Big Data. Permite al lector introducirse en todos los conceptos que configuran la materia.
- M. Crawley, *The R Book*. John Wiley & Sons, 2007. Un libro muy interesante para profundizar en los conceptos de R. Además tiene la ventaja de poder leerse online en la web de la Biblioteca de la Universidad de Alcalá.

Bibliografía Complementaria

- J. Leskovec, *Mining of Massive Datasets*. Cambridge University Press, 2011. Un libro muy interesante para profundizar en los conceptos de Big Data. Además tiene la ventaja de poder descargarse en pdf. Estará a disposición de los alumnos en Blackboard.
- P. Teetor, *R Cookbook*. O'Reilly Media, Inc, 2011. ISBN: 978-0-596-80915-7. Libro muy útil para introducirse en el entorno de R. Además tiene la ventaja de poder leerse online en la web de la Biblioteca de la Universidad de Alcalá.

NOTA INFORMATIVA

La Universidad de Alcalá garantiza a sus estudiantes que, si por exigencias sanitarias las autoridades competentes impidieran la presencialidad total o parcial de la actividad docente, los planes docentes alcanzarían sus objetivos a través de una metodología de enseñanza-aprendizaje y evaluación en formato online, que retornaría a la modalidad presencial en cuanto cesaran dichos impedimentos.