

## Guía docente

### Identificación de la asignatura

<b>Asignatura / Grupo</b>	20591 - Ampliación de Estadística: Aprendizaje Estadístico / 9
<b>Titulación</b>	Grado en Matemáticas - Cuarto curso
<b>Créditos</b>	6
<b>Período de impartición</b>	Segundo semestre
<b>Idioma de impartición</b>	

### Profesores

#### Horario de atención a los alumnos

Profesor/a	Hora de inicio	Hora de fin	Día	Fecha inicial	Fecha final	Despacho / Edificio
Arnau Mir Torres <i>Responsable</i> <a href="mailto:arnau.mir@uib.es">arnau.mir@uib.es</a>						Hay que concertar cita previa con el/la profesor/a para hacer una tutoría
Ricardo Alberich Martí <a href="mailto:r.alberich@uib.es">r.alberich@uib.es</a>						Hay que concertar cita previa con el/la profesor/a para hacer una tutoría
Irene María García Mosquera <a href="mailto:irene.garcia@uib.es">irene.garcia@uib.es</a>	14:30	15:30	Martes	09/09/2019	20/12/2019	D-122 / Anselm Turmeda
	10:30	11:30	Martes	01/10/2019	20/12/2019	AT, D-122

### Contextualización

Con el desarrollo de las tecnologías y la era de “Big Data”, los problemas estadísticos se hacen cada vez más retadores y comunes en muchas áreas científicas y en aplicaciones industriales y comerciales. Es por esto que para los matemáticos se hace imprescindible el conocimiento de técnicas estadísticas modernas, ligadas a la informática, que posibiliten extraer patrones y tendencias importantes para comprender “que nos dicen los datos”, es decir, aprender a partir de estos.

En este curso presentamos algunas de las nuevas e importantes ideas de aprendizaje estadístico supervisado y no supervisado: desde la extensión de modelos lineales estudiados en las asignaturas básicas, pasando por métodos de clasificación, remuestreo y árboles de decisión hasta llegar a minería de texto.

Durante el curso se realizarán talleres prácticos, donde se resolverán problemas reales provenientes de diversas áreas de conocimiento y que ilustran cómo implementar cada uno de los métodos utilizando el software estadístico R. Además, se espera que el estudiante aprenda a redactar y presentar un informe profesional junto con el modelo de datos y los algoritmos, con el objetivo de hacer, además, desarrollos reproducibles.

Esta asignatura es optativa y se imparte en el segundo semestre del cuarto curso de Grado de Matemáticas.

### Requisitos



## Guía docente

Haber cursado con aprovechamiento: Análisis de Datos (20582) e Inferencia Estadística (20326).

Dominio de herramientas de informática.

### Recomendables

Dominio a nivel intermedio del programa R.

## Competencias

### Específicas

- \* E40: Desarrollar la capacidad de identificar y describir matemáticamente un problema, de estructurar la información disponible y de seleccionar un modelo matemático adecuado para su resolución
- \* E48: Resolver y analizar problemas básicos de modelos lineales usando la teoría de regresión
- \* E23: Saber manejar las funciones elementales y sus aplicaciones a la modelización de fenómenos tanto continuos como discretos

### Genéricas

- \* TG2: Desarrollar capacidades de análisis y síntesis, de organización y planificación, y de toma de decisiones
- \* TG3: Capacidad para comunicarse de manera oral o escrita con personas con diferentes niveles de conocimientos en matemáticas
- \* TG12: Capacidad de proponer, analizar, validar e interpretar modelos de situaciones reales sencillas

### Básicas

- \* Se pueden consultar las competencias básicas que el estudiante tiene que haber adquirido al finalizar el grado en la siguiente dirección: [http://estudis.uib.cat/es/grau/comp\\_basiques/](http://estudis.uib.cat/es/grau/comp_basiques/)

## Contenidos

La asignatura está compuesta por tres bloques:

### Contenidos temáticos

#### Bloque I. Aprendizaje estadístico

- \* Introducción al aprendizaje estadístico: supervisado y no supervisado.
- \* Extensión de los modelos lineales.
- \* Introducción a la estadística Bayesiana.

#### Bloque II. Métodos no lineales

- \* Métodos de remuestreo.
- \* Introducción a los métodos de clasificación. Árboles de decisión.
- \* Splines, modelos aditivos y regresión polinomial

#### Bloque III. Minería de texto

- \* Introducción a la minería de texto.

## Guía docente

\* Análisis de sentimientos.

\* n-gramas y topic models.

### Metodología docente

Actividades de carácter presencial:

Asistir a las clases de teoría, realización de talleres y proyectos, presentar resultados.

Actividades de trabajo no presencial:

Resolución y preparación de la exposición de los ejercicios, de las presentaciones, lectura de la bibliografía recomendada, búsqueda en la web de material relevante, experimentación informática, elaboración de programas e informes asociados al proyecto y su presentación.

Con el propósito de favorecer el trabajo del alumno, se ha solicitado que la asignatura forme parte de Aula Digital. Mediante esta plataforma, el alumno tendrá a su disposición un calendario con noticias de interés, documentos electrónicos, problemas y proyectos propuestos tanto para el trabajo individual como en grupo y maneras de comunicarse con los profesores de la asignatura y sus compañeros de clase.

Actividades de trabajo presencial (2,4 créditos, 60 horas)

Modalidad	Nombre	Tip. agr.	Descripción	Horas
Clases teóricas	Introducción a los temas propuestos	Grupo grande (G)	Introducción teórica a los temas del curso. Se presentarán los conceptos, resultados y procedimientos básicos de las técnicas, así como ejemplos detallados de aplicaciones mediante la exposición magistral. Se trabajarán todas las competencias específicas y las genéricas TG2 y TG12.	24
Seminarios y talleres	Talleres de prácticas Bloque II	Grupo mediano (M)	Los alumnos resolverán, redactarán y expondrán la solución de ejercicios del Bloque II: Métodos no lineales, tanto individual como en grupo. Se trabajarán todas las competencias específicas y genéricas.	10
Seminarios y talleres	Talleres de prácticas Bloque III	Grupo mediano (M)	Los alumnos resolverán, redactarán y expondrán la solución de ejercicios del <i>Bloque III: Minería de texto</i> , tanto individual como en grupo. Se trabajarán todas las competencias específicas y genéricas.	10
Seminarios y talleres	Talleres de prácticas Bloque I	Grupo mediano (M)	Los alumnos resolverán, redactarán y expondrán la solución de ejercicios del <i>Bloque I: Introducción al aprendizaje estadístico</i> , tanto individual como en grupo. Se trabajarán todas las competencias específicas y genéricas.	10
Otros	Exposición de trabajos	Grupo grande (G)	Los alumnos harán una exposición de sus trabajos en forma individual. Se trabajarán todas las competencias específicas y la general TG3.	6

Al inicio del semestre estará a disposición de los estudiantes el cronograma de la asignatura a través de la plataforma UIBdigital. Este cronograma incluirá al menos las fechas en las que se realizarán las pruebas de evaluación continua y las fechas de entrega de los trabajos. Asimismo, el profesor o la profesora informará a los estudiantes si el plan de trabajo de la asignatura se realizará a través del cronograma o mediante otra vía, incluida la plataforma Aula Digital.

## Guía docente

### Actividades de trabajo no presencial (3,6 créditos, 90 horas)

Modalidad	Nombre	Descripción	Horas
Estudio y trabajo autónomo individual o en grupo	Tareas personales o en grupo	Estudio individual y resolución de ejercicios. Realización de los trabajos propuestos, tanto individual como en grupo. El alumno debe dedicar tiempo tanto al estudio de la teoría como a la resolución de problemas y/o proyectos a lo largo del curso.  Se trabajarán todas las competencias específicas y genéricas.	90

### Riesgos específicos y medidas de protección

Las actividades de aprendizaje de esta asignatura no conllevan riesgos específicos para la seguridad y salud de los alumnos y, por tanto, no es necesario adoptar medidas de protección especiales.

### Evaluación del aprendizaje del estudiante

Al inicio del semestre estará a disposición de los estudiantes el cronograma de la asignatura a través de la plataforma UIBdigital. Este cronograma incluirá al menos las fechas en las que se realizarán las pruebas de evaluación continua y las fechas de entrega de los trabajos. Asimismo, los profesores informarán a los estudiantes si el plan de trabajo de la asignatura se realizará a través del cronograma o mediante otra vía, incluida la plataforma Aula Digital.

### Fraude en elementos de evaluación

De acuerdo con el artículo 33 del Reglamento Académico, "con independencia del procedimiento disciplinario que se pueda seguir contra el estudiante infractor, la realización demostrablemente fraudulenta de alguno de los elementos de evaluación incluidos en guías docentes de las asignaturas comportará, a criterio del profesor, una minusvaloración en su calificación que puede suponer la calificación de «suspense 0» en la evaluación anual de la asignatura".

### Talleres de prácticas Bloque II

Modalidad	Seminarios y talleres
Técnica	Informes o memorias de prácticas ( <b>no recuperable</b> )
Descripción	Los alumnos resolverán, redactarán y expondrán la solución de ejercicios del Bloque II: Métodos no lineales, tanto individual como en grupo. Se trabajarán todas las competencias específicas y genéricas.
Criterios de evaluación	Los alumnos resolverán ejercicios y problemas con métodos no lineales, tanto de forma individual como en grupo.

Porcentaje de la calificación final: 25%

## Guía docente

### Talleres de prácticas Bloque III

Modalidad	Seminarios y talleres
Técnica	Informes o memorias de prácticas ( <b>no recuperable</b> )
Descripción	Los alumnos resolverán, redactarán y expondrán la solución de ejercicios del Bloque III: Minería de texto, tanto individual como en grupo. Se trabajarán todas las competencias específicas y genéricas.
Criterios de evaluación	Los alumnos resolverán ejercicios y problemas de minería de texto, tanto de forma individual como en grupo.
Porcentaje de la calificación final:	25%

### Talleres de prácticas Bloque I

Modalidad	Seminarios y talleres
Técnica	Informes o memorias de prácticas ( <b>no recuperable</b> )
Descripción	Los alumnos resolverán, redactarán y expondrán la solución de ejercicios del Bloque I: Introducción al aprendizaje estadístico, tanto individual como en grupo. Se trabajarán todas las competencias específicas y genéricas.
Criterios de evaluación	Los alumnos resolverán ejercicios y problemas de aprendizaje estadístico, tanto de forma individual como en grupo.
Porcentaje de la calificación final:	25%

### Exposición de trabajos

Modalidad	Otros
Técnica	Trabajos y proyectos ( <b>no recuperable</b> )
Descripción	Los alumnos harán una exposición de sus trabajos en forma individual. Se trabajarán todas las competencias específicas y la general TG3.
Criterios de evaluación	El trabajo final de la asignatura se realizará de manera individual y consistirá en una pequeña memoria y en una presentación oral.
Porcentaje de la calificación final:	25%

### Recursos, bibliografía y documentación complementaria

Adicional a los libros recomendados, los profesores irán publicando en Aula Digital los apuntes y listas de problemas de cada tema.

#### Bibliografía básica

- \* **An Introduction to Statistical Learning with Applications in R.** Gareth James, Daniela Witten, Trevor Hastie y Robert Tibshirani. Springer. 2013.
- \* **The Elements of Statistical Learning: Data Mining, Inference and Prediction.** Jerone Friedman, Trevor Hastie y Robert Tibshirani. Springer Series in Statistics. Segunda edición. 2009.
- \* **R for Data Science.** Garrett Golemund, Hadley Wickham. Editorial O'Reilly, primera edición. January 2017.