



UNIVERSIDAD
POLITÉCNICA
DE MADRID

PROCESO DE
COORDINACIÓN DE LAS
ENSEÑANZAS PR/CL/001



E.T.S. de Ingenieros
Industriales

ANX-PR/CL/001-01

GUÍA DE APRENDIZAJE

ASIGNATURA

55000602 - Análisis De Datos

PLAN DE ESTUDIOS

05TI - Grado En Ingeniería En Tecnologías Industriales

CURSO ACADÉMICO Y SEMESTRE

2022/23 - Primer semestre

Índice

Guía de Aprendizaje

1. Datos descriptivos.....	1
2. Profesorado.....	1
3. Conocimientos previos recomendados.....	2
4. Competencias y resultados de aprendizaje.....	3
5. Descripción de la asignatura y temario.....	4
6. Cronograma.....	6
7. Actividades y criterios de evaluación.....	8
8. Recursos didácticos.....	11

1. Datos descriptivos

1.1. Datos de la asignatura

Nombre de la asignatura	55000602 - Análisis de Datos
No de créditos	3 ECTS
Carácter	Optativa
Curso	Cuarto curso
Semestre	Séptimo semestre
Período de impartición	Septiembre-Enero
Idioma de impartición	Castellano
Titulación	05TI - Grado en Ingeniería en Tecnologías Industriales
Centro responsable de la titulación	05 - Escuela Técnica Superior De Ingenieros Industriales
Curso académico	2022-23

2. Profesorado

2.1. Profesorado implicado en la docencia

Nombre	Despacho	Correo electrónico	Horario de tutorías *
Eduardo Caro Huertas (Coordinador/a)	Estadística	eduardo.caro@upm.es	Sin horario. Siempre solicitar cita previamente por email
Jesus Juan Ruiz		jesus.juan@upm.es	Sin horario. COORDINADOR DE LA ASIGNATURA Siempre solicitar cita previamente

por email

* Las horas de tutoría son orientativas y pueden sufrir modificaciones. Se deberá confirmar los horarios de tutorías con el profesorado.

3. Conocimientos previos recomendados

3.1. Asignaturas previas que se recomienda haber cursado

- Calculo li
- Algebra
- Ecuaciones Diferenciales
- Calculo I
- Estadística
- Diseño De Experimentos Y Modelos De Regresion

3.2. Otros conocimientos previos recomendados para cursar la asignatura

- La mayoría de los alumnos han cursado previamente dos asignaturas (Estadística y Diseño de Experimentos & Regresión) con el programa R/RStudio. Estos conocimientos previos son muy útiles para seguir la asignatura de Análisis de Datos.
- Conocimientos básicos de programación

4. Competencias y resultados de aprendizaje

4.1. Competencias

CE25F - Capacidad para elegir y aplicar técnicas de modelado de grandes bases de datos y técnicas de predicción para variables cuantitativas y cualitativas.

CG1 - Conocer y aplicar conocimientos de ciencias y tecnologías básicas a la práctica de la Ingeniería Industrial.

CG2 - Poseer capacidad para diseñar, desarrollar, implementar, gestionar y mejorar productos, sistemas y procesos en los distintos ámbitos industriales, usando técnicas analíticas, computacionales o experimentales apropiadas.

CG3 - Aplicar los conocimientos adquiridos para identificar, formular y resolver problemas dentro de contextos amplios y multidisciplinarios, siendo capaces de integrar conocimientos, trabajando en equipos multidisciplinarios.

CG5 - Saber comunicar los conocimientos y conclusiones, de forma oral, escrita y gráfica, a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CG6 - Poseer habilidades de aprendizaje que permitan continuar estudiando a lo largo de la vida para su adecuado desarrollo profesional.

CG7 - Incorporar nuevas tecnologías y herramientas de la Ingeniería Industrial en sus actividades profesionales.

CG9 - Organización y planificación en el ámbito de la empresa, y otras instituciones y organizaciones de proyectos y equipos humanos.

4.2. Resultados del aprendizaje

RA85 - Utilizar programas de ordenador de análisis estadístico general y de cálculo científico.

RA82 - Identificar problemas que pueden plantearse en términos estadísticos.

RA83 - Interpretar y comunicar los resultados del análisis estadístico con rigor utilizando el lenguaje apropiado.

RA306 - Situarse con actitud crítica y en distintos puntos de vista ante la validez de los cálculos y resultados.

RA84 - Comprender las limitaciones de los modelos estadísticos cuando se trabaja con problemas reales. Evaluar posibles métodos alternativos.

5. Descripción de la asignatura y temario

5.1. Descripción de la asignatura

La asignatura quiere presentar una visión panorámica de los métodos estadísticos de análisis de datos y las técnicas fundamentales de machine learning. En los últimos años ha nacido una nueva profesión, científico de datos (data scientist), transversal a disciplinas como matemáticas, estadística, ingeniería e informática. Las empresas con gran determinación han creado nuevos departamentos en el área de Business Intelligence y las universidades y escuelas de negocio han puesto en el mercado nuevos títulos académicos relacionados con el análisis de datos. El slogan que más se repite, tanto en ámbitos académicos como empresariales es que las decisiones han de tomarse basándose en datos.

La asignatura pretende enseñar de forma muy práctica los modelos y las técnicas más importantes en el área de análisis de datos utilizando R/Rstudio como herramienta computacional. Los alumnos que hayan cursado el segundo curso en la Escuela tienen una buena formación en Estadística, Diseño de Experimentos y Modelos de Regresión y han aprendido a usar R/Rstudio. Ese será el punto de partida. Una buena guía para el curso es el libro "An introduction to statistical learning (Gareth James, Daniela Witten, Trevor Hastie y Robert Tibshirani)". Algunos autores del libro son profesores de la Universidad de Stanford, y han contribuido de forma muy relevante al desarrollo teórico y práctico de los métodos modernos de análisis estadístico que enseñaremos en la asignatura.

Imprescindible la asistencia a clase: todos los alumnos tendrán que entregar las prácticas realizadas en las sesiones del curso para aprobar la asignatura.

5.2. Temario de la asignatura

1. Estadística descriptiva multivariante
2. Modelo de regresión múltiple
3. Componentes principales y análisis factorial
4. Análisis discriminante
5. Análisis Cluster
6. Árboles de Regresión y Clasificación
7. Random Forest
8. Otras técnicas de "Machine Learning"

6. Cronograma

6.1. Cronograma de la asignatura *

Sem	Actividad en aula	Actividad en laboratorio	Tele-enseñanza	Actividades de evaluación
1	Descriptiva multivariante con R Duración: 02:00 LM: Actividad del tipo Lección Magistral			
2	Modelos de regresión múltiple Duración: 02:00 LM: Actividad del tipo Lección Magistral			
3	Modelos de regresión múltiple Duración: 02:00 LM: Actividad del tipo Lección Magistral			
4	Componentes Principales Duración: 02:00 LM: Actividad del tipo Lección Magistral			Práctica de Regresión EP: Técnica del tipo Examen de Prácticas Evaluación continua Presencial Duración: 02:00
5	Componentes Principales Duración: 02:00 LM: Actividad del tipo Lección Magistral			
6	Análisis Discriminante Duración: 02:00 LM: Actividad del tipo Lección Magistral			
7	Análisis Discriminante Duración: 02:00 LM: Actividad del tipo Lección Magistral			Práctica de Multivariante EP: Técnica del tipo Examen de Prácticas Evaluación continua Presencial Duración: 02:00
8	Exámen PEC1 Duración: 004:00 OT: Otras actividades formativas			
9	Análisis Cluster Duración: 02:00 LM: Actividad del tipo Lección Magistral			
10	Análisis Cluster Duración: 02:00 LM: Actividad del tipo Lección Magistral			
11	Árboles de Regresión y Clasificación Duración: 02:00 LM: Actividad del tipo Lección Magistral			
12	Random Forest 1 Duración: 02:00 LM: Actividad del tipo Lección Magistral			Práctica de A. Cluster EP: Técnica del tipo Examen de Prácticas Evaluación continua Presencial Duración: 02:00

13	Random Forest 1 Duración: 02:00 LM: Actividad del tipo Lección Magistral			
14	Árboles de Clasificación Duración: 00:00 LM: Actividad del tipo Lección Magistral			
15	Random Forest para Clasificación Duración: 02:00 LM: Actividad del tipo Lección Magistral			
16	Exámen PEC2 Duración: 004:00 OT: Otras actividades formativas			Práctica de Árboles de Regresión y Clasificación EP: Técnica del tipo Examen de Prácticas Evaluación continua Presencial Duración: 02:00
17	Exámen Duración: 004:00 OT: Otras actividades formativas			Evaluación EP: Técnica del tipo Examen de Prácticas Evaluación continua y sólo prueba final Presencial Duración: 04:00

Para el cálculo de los valores totales, se estima que por cada crédito ECTS el alumno dedicará dependiendo del plan de estudios, entre 26 y 27 horas de trabajo presencial y no presencial.

* El cronograma sigue una planificación teórica de la asignatura y puede sufrir modificaciones durante el curso derivadas de la situación creada por la COVID-19.

7. Actividades y criterios de evaluación

7.1. Actividades de evaluación de la asignatura

7.1.1. Evaluación (progresiva)

Sem.	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
4	Práctica de Regresión	EP: Técnica del tipo Examen de Prácticas	Presencial	02:00	10%	5 / 10	CG6 CG9 CG1 CG3
7	Práctica de Multivariante	EP: Técnica del tipo Examen de Prácticas	Presencial	02:00	10%	5 / 10	CG5 CG7 CG2 CE25F
12	Práctica de A. Cluster	EP: Técnica del tipo Examen de Prácticas	Presencial	02:00	10%	5 / 10	CG5 CG9
16	Práctica de Árboles de Regresión y Clasificación	EP: Técnica del tipo Examen de Prácticas	Presencial	02:00	10%	5 / 10	CG1 CG2 CG3 CG5 CG6 CG7 CG9 CE25F
17	Evaluación	EP: Técnica del tipo Examen de Prácticas	Presencial	04:00	60%	5 / 10	CG3 CG5 CG1 CG2 CG6 CG7 CG9 CE25F

7.1.2. Prueba evaluación global

Sem	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
-----	-------------	-----------	------	----------	-----------------	-------------	------------------------

17	Evaluación	EP: Técnica del tipo Examen de Prácticas	Presencial	04:00	60%	5 / 10	CG3 CG5 CG1 CG2 CG6 CG7 CG9 CE25F
----	------------	--	------------	-------	-----	--------	--

7.1.3. Evaluación convocatoria extraordinaria

No se ha definido la evaluación extraordinaria.

7.2. Criterios de evaluación

EVALUACIÓN DE LA ASIGNATURA

A efectos de evaluación, la asignatura se divide en dos bloques:

- Bloque 1: Descriptiva, Regresión, Componentes Principales y Análisis Discriminante
- Bloque 2: Análisis Cluster, CART, Random Forest y Otras técnicas

Pruebas de evaluación intermedia:

La evaluación continua de cada bloque está compuesta por dos trabajos (10% cada uno) y un examen (30% cada uno). Es decir: toda la evaluación continua consiste en cuatro trabajos y dos exámenes (PEIs).

- Los trabajos serán "evaluados por pares". Estas tareas serán corregidas por los alumnos utilizando el sistema implementado en Moodle. Según este sistema cada alumno evaluará tres ejercicios y su ejercicio será evaluado por tres compañeros. La nota de la tarea será el promedio de las tres calificaciones, siempre y cuando los profesores no detecten ninguna evaluación incorrecta u otra anomalía. Los profesores supervisarán el proceso de evaluación intentando respetar al máximo la evaluación realizada por los alumnos.
- Los dos exámenes online (PECs) consistirán en unos ejercicios de análisis de datos, a realizar mediante el uso del ordenador. La fecha y hora de estas dos pruebas aparece recogida en el POD.

Cálculo de la nota final mediante evaluación continua:

- Nota Bloque 1: nota tarea 1 (10%) + nota tarea 2 (10%) + PE1 (30%)
- Nota Bloque 2: nota tarea 3 (10%) + nota tarea 4 (10%) + PE2 (30%)
- Nota final de la asignatura: nota Bloque 1 + nota Bloque 2

Para aprobar cada bloque, es obligatorio entregar los dos trabajos del bloque, presentarse a la PE, y que la nota de ese bloque sea superior a 5 sobre 10.

Para aprobar la asignatura es necesario tener aprobados los dos bloques.

Examen Final Ordinario (Enero):

El alumno que no haya aprobado alguna de las dos partes tendrá que realizar el examen final ordinario para aprobar la asignatura. El examen final tendrá dos partes, correspondientes a cada parte de la asignatura. El alumno se examinará de las partes de la asignatura que no haya aprobado en los exámenes parciales. Para aprobar la asignatura será necesario tener aprobadas las dos partes (bien en los exámenes parciales o en el examen final).

Cada parte del examen consistirá en unos ejercicios de análisis de datos, a realizar mediante el uso del ordenador. El alumno se llevará su propio portátil. La fecha y hora del examen aparece recogida en el POD.

Examen Final Extraordinario (Julio):

Un examen único para todos los alumnos. Será necesario obtener una calificación igual o superior a cinco para aprobar la asignatura. Las calificaciones previas obtenidas en los exámenes parciales y final no serán tenidas en cuenta en la evaluación de este examen.

El examen consistirá en unos ejercicios de análisis de datos, a realizar mediante el uso del ordenador. El alumno se llevará su propio portátil. La fecha y hora del examen aparece recogida en el POD.

8. Recursos didácticos

8.1. Recursos didácticos de la asignatura

Nombre	Tipo	Observaciones
An introduction to Statistical Learning with applications in R. G. James, D. Witten, T. Hastie y R. Tibshirani. Springer	Bibliografía	
En Moodle se dispondrá la siguiente documentación: apuntes, ejercicios resueltos, diapositivas, etc.	Otros	