

Curso Académico: ( 2021 / 2022 )

Fecha de revisión: 03-06-2021

Departamento asignado a la asignatura: Departamento de Estadística

Coordinador/a: GALEANO SAN MIGUEL, PEDRO

Tipo: Obligatoria Créditos ECTS : 3.0

Curso : 1 Cuatrimestre : 1

**REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)**

Modelos estadísticos para análisis de datos  
Matemáticas aplicadas para análisis de datos

**OBJETIVOS****Competencias Básicas**

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

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

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

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**

Aplicar los fundamentos teóricos de las técnicas de recogida, almacenamiento, tratamiento y presentación de información, especialmente para grandes volúmenes de datos, como base para el desarrollo y adaptación de dichas técnicas a problemas concretos

Identificar diferentes técnicas para almacenar, replicar y distribuir grandes cantidades de datos, y diferenciarlas en función de sus características teóricas y prácticas

Identificar las técnicas de análisis de datos más adecuadas para cada problema y saber aplicarlas para el análisis, diseño y solución de los mismos

Obtener soluciones prácticas y eficientes para problemas de tratamiento de grandes volúmenes de datos, tanto individualmente como en equipo

Sintetizar las conclusiones obtenidas de estos análisis y presentarlas de manera clara y convincente en un entorno bilingüe (español e inglés) tanto por escrito como oralmente

Ser capaz de generar nuevas ideas (creatividad) y de anticipar nuevas situaciones, en los contextos del análisis de datos y de la toma de decisiones

Utilizar habilidades para el trabajo en equipo y para relacionarse con otros de forma autónoma

**Competencias Específicas**

Emplear los resultados básicos de inferencia y regresión como fundamento para métodos avanzados de predicción y clasificación

Identificar y seleccionar las herramientas software adecuadas para el tratamiento de grandes

cantidades de datos

Utilizar procedimientos estadísticos avanzados para el tratamiento de grandes volúmenes de datos en áreas como la estimación, la inferencia, la predicción o la clasificación, así como la manera de aplicarlos de forma eficiente

Identificar correctamente el tipo de problema estadístico correspondiente a unos objetivos y unos datos determinados

Saber diseñar sistemas para el procesamiento de los datos, desde la obtención y filtrado inicial de los mismos, su análisis estadístico, hasta la presentación de los resultados finales

Utilizar técnicas y herramientas de investigación operativa utilizables con datos masivos en procedimientos para su análisis, visualización de sus resultados o dentro de sistemas de apoyo a decisiones

## DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

1. Datos multidimensionales
  - 1.1 Conjuntos de datos multidimensionales.
  - 1.2 Visualización de conjuntos de datos multidimensionales.
  - 1.3 Un breve resumen de las variables aleatorias multivariadas.
  - 1.4 Inferencia con la matriz de datos.
  - 1.5 Imputación de valores perdidos.
2. Técnicas de reducción de dimensión
  - 2.1 Técnicas de reducción de la dimensión.
  - 2.2 Análisis de componentes principales.
  - 2.3 Análisis de componentes independientes.
3. Clasificación no supervisada
  - 3.1 Clasificación no supervisada.
  - 3.2 Agrupación particional.
  - 3.3 Agrupación jerárquica.
  - 3.4 Agrupación basada en modelos.
4. Clasificación supervisada
  - 4.1 Clasificación supervisada.
  - 4.2 K vecinos más cercanos (KNN).
  - 4.3 Métodos basados  $\lambda$  en el teorema de Bayes.
  - 4.4 Regresión logística.

## ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Actividades formativas

Clase teórica

Clases prácticas

Prácticas de laboratorio

Trabajo individual del estudiante

Metodologías docentes

Exposiciones en clase del profesor con soporte de medios informáticos y audiovisuales, en las que se desarrollan los conceptos principales de la materia y se proporciona la bibliografía para complementar el aprendizaje de los alumnos. Lectura crítica de textos recomendados por el profesor de la asignatura: Artículos de prensa, informes, manuales y/o artículos académicos, bien para su posterior discusión en clase, bien para ampliar y consolidar los conocimientos de la asignatura.

Resolución de casos prácticos, problemas, etc. planteados por el profesor de manera individual o en grupo

Elaboración de trabajos e informes de manera individual o en grupo

## SISTEMA DE EVALUACIÓN

Trabajo en grupo (50%)

Tareas individuales (40%)

Presentaciones en clase (10%)

**Peso porcentual del Examen Final:**

0

**Peso porcentual del resto de la evaluación:**

100

#### BIBLIOGRAFÍA BÁSICA

- Daniel Peña Análisis de datos multivariantes, McGraw Hill, 2002