Modelización Predictiva

Curso Académico: (2022 / 2023) **Fecha de revisión: 10/05/2022 14:17:10**

Departamento asignado a la asignatura: Departamento de Estadística

Coordinador/a: GARCIA PORTUGUES, EDUARDO

Tipo: Obligatoria Créditos ECTS : 6.0

Curso : 2 Cuatrimestre : 2

REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)

Cálculo I y II
Algebra Lineal
Programación
Probabilidad y Análisis de Datos
Introducción a la Modelización Estadística
Aprendizaje Estadístico

OBJETIVOS

* Competencias generales

- CG1: Conocimientos y habilidades adecuadas para analizar y sintetizar problemas básicos relacionados con la ingeniería y la ciencia de los datos, resolverlos y comunicarlos eficientemente.
- CG4: Habilidad para resolver problemas tecnológicos, computacionales, matemáticos y estadísticos que puedan surgir en la ingeniería y la ciencia de los datos.
- CG5: Capacidad de resolver problemas matemáticos formulados aplicados a diferentes temas, utilizando algoritmos numéricos y técnicas computacionales.
- CG6: Sintetizar las conclusiones obtenidas de los análisis realizados y presentarlas de forma clara y convincente, tanto por escrito como oralmente.

* Competencias transversales

- CT1: Capacidad de comunicar conocimientos oralmente y por escrito, ante un público especializado y no especializado.

* Competencias específicas

- CE1: Capacidad para resolver problemas matemáticos que pueden surgir en la ingeniería y la ciencia de los datos. Capacidad para aplicar conocimientos sobre: álgebra; geometría; cálculo diferencial e integral; métodos numéricos; algoritmo numérico; estadística y optimización.
- CE2: Identificar correctamente los problemas de naturaleza predictiva correspondientes a determinados objetivos y datos y utilizar los resultados básicos del análisis de regresión como base fundamental de los métodos de predicción.
- CE5: Comprender y manejar los conceptos fundamentales de probabilidad y estadística y ser capaz de representar y manipular los datos para extraer información significativa de ellos.
- CE7: Entender los conceptos básicos de programación y ser capaz de llevar a cabo programas dirigidos al análisis de datos.

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

Este curso está diseñado para dar una visión panorámica de varias herramientas disponibles para la modelización predictiva, a un nivel intermedio-introductivo. Esta visión cubre en profundidad los principales conceptos de los modelos lineales y da una visión general de sus extensiones. Se hace hincapié en proporcionar los principales conocimientos sobre los fundamentos estadísticos/matemáticos de los modelos y en mostrar la aplicación efectiva de los métodos mediante el uso del software

estadístico R.

1. Introducción
   1.1. Resumen del curso
   1.2. Repaso de probabilidad
   1.3. Vectores aleatorios
   1.4. Repaso de inferencia estadística
   1.5. ¿Qué es la modelización predictiva?

2. Regresión lineal simple
   2.1. Formulación y estimación de modelos
   2.2. Supuestos del modelo
   2.3. Inferencia para los parámetros del modelo
   2.4. Predicción
   2.5. ANOVA y ajuste del modelo

3. Regresión lineal múltiple
   3.1. Formulación y estimación de modelos
   3.2. Supuestos de los modelos
   3.3. Inferencia para los parámetros del modelo
   3.4. ANOVA y ajuste del modelo
   3.5. Selección de modelos
   3.6. Manejo de relaciones no lineales
   3.7. Utilización de predictores cualitativos
   3.8. Diagnóstico de modelos y multicolinealidad

4. Extensiones de la regresión lineal
   4.1. Repaso del análisis de componentes principales
   4.2. Regresión de componentes principales
   4.3. Regresión por mínimos cuadrados parciales
   4.4. Modelos lineales regularizados
   4.5. Regresión ridge y lasso

5. Regresión logística
   5.1. Formulación e interpretación del modelo
   5.2. Estimación por máxima verosimilitud
   5.3. Inferencia para los parámetros del modelo
   5.4. Selección del modelo y multicolinealidad
   5.5. Modelos logísticos regularizados

El programa está sujeto a modificaciones menores debido al desarrollo del curso y/o el calendario académico.

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Las lecciones magistrales consisten principalmente en exposiciones teóricas sobre los métodos estadísticos del curso. Estas exposiciones se complementan con ejemplos ilustrativos. Los laboratorios están diseñados para realizar ejercicios y estudios de casos que profundizan en el uso práctico de los métodos vistos. La implementación de los métodos se realiza con el lenguaje estadístico R.

SISTEMA DE EVALUACIÓN

Peso porcentual del Examen/Prueba Final:	60
Peso porcentual del resto de la evaluación:	40

La evaluación continua se hace con una mezcla de:

(a) tres cuestionarios teórico-prácticos;
(b) participación activa en las clases y realización de ejercicios voluntarios.

La Nota de la Evaluación Continua (NEC) es

NEC = min(A + 0.10 * B, 10),

donde:

Peso porcentual del Examen/Prueba Final:	60
Peso porcentual del resto de la evaluación:	40

- A (en la escala 0-10) es la calificación ponderada de los cuestionarios;
- B (en la escala 0-10) es la calificación correspondiente a (b).

Si NEC < 5, el estudiante tendrá que asistir al examen final para superar la asignatura. Para los estudiantes que realicen el examen final, la Nota de la Evaluación Ordinaria (NEO) es

NEO = 0.4 * NEC + 0.6 * NEF,

donde NEF es la Nota del Examen Final. Si NEC >= 5, el estudiante puede optar por (1) realizar el examen final con la anterior ponderación o (2) no realizar el examen final, siendo en este caso NEO = NEC.

La Nota de la Convocatoria Extraordinaria (NCE) se determina por medio de la Nota del Examen Extraordinario (NEE) y la evaluación continua:

NCE = max(0.4 * NEC + 0.6 * NEE, NEE).

Se proporcionan más detalles en Aula Global. La evaluación está sujeta a modificaciones debido al desarrollo del curso y/o al calendario académico.

BIBLIOGRAFÍA BÁSICA

- James, G., Witten, D., Hastie, T. y Tibshirani, R. An Introduction to Statistical Learning, Springer-Verlag, 2013

- Papoulis, A. y Pillai, S. U. Probability, Random Variables, and Stochastic Processes, McGraw-Hill, 2002

BIBLIOGRAFÍA COMPLEMENTARIA

- Hastie, T., Tibshirani, R. y Friedman, J. The Elements of Statistical Learning, Springer, 2013

- Kuhn, M. y Johnson, K. Applied Predictive Modeling, Springer, 2013

- Panaretos, V. M. Statistics for Mathematicians, Springer, 2016

- Peña, D. Regresión y Diseño de Experimentos, Alianza Editorial, 2002

- Seber, G. A. F. Linear Regression Analysis, John Wiley & Sons, 1977

- Wood, S. N. Generalized Additive Models: An Introduction with R, Chapman & Hall/CRC, 2006