

Investigación operativa

Código: 100125
Créditos ECTS: 6

Titulación	Tipo	Curso	Semestre
2500149 Matemáticas	OT	4	0

La metodología docente y la evaluación propuestas en la guía pueden experimentar alguna modificación en función de las restricciones a la presencialidad que impongan las autoridades sanitarias.

Contacto

Nombre: Juan Ramón González Ruíz

Correo electrónico: JuanRamon.Gonzalez@uab.cat

Uso de idiomas

Lengua vehicular mayoritaria: catalán (cat)

Algún grupo íntegramente en inglés: No

Algún grupo íntegramente en catalán: Sí

Algún grupo íntegramente en español: No

Prerequisitos

Esta asignatura supone que el alumno ha obtenido los conocimientos que se imparten en diferentes asignaturas sobre los siguientes temas:

- Cálculo en varias variables.
- Probabilidad.
- Modelos lineales.
- Programación en R.

Objetivos y contextualización

Esta asignatura pretende familiarizar al alumno con diferentes métodos de aprendizaje automático aplicando el punto de vista utilizado cuando se dispone de grandes cantidades de datos.

Competencias

- Ante situaciones reales con un nivel medio de complejidad, recabar y analizar datos e información relevantes, proponer y validar modelos utilizando herramientas matemáticas adecuadas para, finalmente, obtener conclusiones.
- Demostrar de forma activa una elevada preocupación por la calidad en el momento de argumentar o hacer públicas las conclusiones de sus trabajos.
- Que los estudiantes hayan demostrado poseer y comprender conocimientos en un área de estudio que parte de la base de la educación secundaria general, y se suele encontrar a un nivel que, si bien se apoya en libros de texto avanzados, incluye también algunos aspectos que implican conocimientos procedentes de la vanguardia de su campo de estudio.
- Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía.

- Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.
- Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.
- Reconocer la presencia de las Matemáticas en otras disciplinas.
- Utilizar aplicaciones informáticas de análisis estadístico, cálculo numérico y simbólico, visualización gráfica, optimización u otras para experimentar en Matemáticas y resolver problemas.
- Utilizar eficazmente bibliografía y recursos electrónicos para obtener información.

Resultados de aprendizaje

1. Adquirir el dominio y seguridad en el manejo de software científico específico para la resolución de problemas con datos reales y para realizar la simulación.
2. Conocer rudimentos de logística y otros campos en los que se aplica la investigación operativa en el ámbito tecnológico e industrial.
3. Demostrar de forma activa una elevada preocupación por la calidad en el momento de argumentar o hacer públicas las conclusiones de sus trabajos.
4. Distinguir, de un problema, lo que es importante de cara a la construcción del modelo matemático y su resolución de lo que no lo es.
5. Dominar los conceptos básicos de la teoría y ser capaz de combinarlos y utilizarlos para resolver problemas.
6. Encontrar modelos de la realidad científica o tecnológica relativa a un problema de toma de decisiones y expresarla con el lenguaje matemático de los problemas de optimización con programación dinámica o con colas estocásticas.
7. Evaluar la dificultad de hacer un cálculo de probabilidades analítico en situaciones complejas y saber distinguir cuando se pueden realizar estos cálculos y cuando se debe recurrir a la simulación estocástica.
8. Extraer conclusiones adecuadas a partir del resultado del modelo.
9. Que los estudiantes hayan demostrado poseer y comprender conocimientos en un área de estudio que parte de la base de la educación secundaria general, y se suele encontrar a un nivel que, si bien se apoya en libros de texto avanzados, incluye también algunos aspectos que implican conocimientos procedentes de la vanguardia de su campo de estudio.
10. Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía.
11. Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.
12. Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.
13. Saber generar y manipular modelos de simulación de la realidad para establecer y comprobar hipótesis en el estudio de problemas o realidades más complejas.
14. Utilizar eficazmente bibliografía y recursos electrónicos para obtener información.

Contenido

Estos son los contenidos de la asignatura*

- Introducción a Tidyverse
- Introducción al aprendizaje automático
- Regresión lineal y logística
- Tractamiento de Big Data con R
- La librería caret
- Métodos de aprendizaje automático
K-vecinos más próximos

- Análisis discriminante
 - Máquinas de soporte vectorial
- Métodos para tratar datos no balanceados
- Árboles de decisión
 - Árboles de clasificación
 - Árboles de regresión
 - Bagged trees
 - Random Forest
- Boosting
 - AdaBoost
 - GBM clásico
 - GBM estocástico
 - XGBoost
 - Otros

**A menos que las restricciones impuestas por las autoridades sanitarias obliguen a una priorización o reducción de estos contenidos.*

Metodología

La asignatura tiene programadas, a la semana, dos horas de teoría y dos horas de prácticas.

- Clase de teoría: se definen y se explican los diferentes métodos con sus características particulares y se muestran ejemplos concretos.

- Clase de prácticas: se trabajan los métodos explicados en clase de teoría con diversos conjuntos de datos utilizando el lenguaje de programación R.

Se considera que, para cada hora de teoría y prácticas, el alumno deberá dedicar una hora adicional para la preparación y/o finalización de la sesión. Se realizarán preguntas de autoevaluación en el Moodle para consolidar los conocimientos aprendidos en clase

NOTA: La metodología docente propuesta puede experimentar alguna modificación en función de las restricciones a la presencialidad que impongan las autoridades sanitarias.

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

Actividades

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Sesiones de prácticas	50	2	1, 7, 2, 3, 4, 5, 8, 9, 10, 13, 6, 14
Tipo: Supervisadas			
Sesiones de teoría	50	2	1, 7, 2, 3, 4, 5, 8, 9, 10, 11, 12, 13, 6
Tipo: Autónomas			
Trabajos semanales + autoevaluación	50	2	1, 7, 2, 4, 5, 8, 9, 10, 11, 12, 13, 6, 14

Evaluación

La evaluación de la asignatura se llevará a cabo con un examen (final) una serie de trabajos semanales y preguntas de autoevaluación. La nota final se calculará con la fórmula:

$$NF = 0,5 * NE + 0,4 * NP + 0,1 * NA$$

donde NP es la nota promedio de los trabajos semanales, NA la nota promedio de las preguntas de autoevaluación y NE la nota del examen que tiene que ser un 5 como mínimo.

A final de curso habrá una prueba de recuperación para aquellos alumnos tales que NE sea menor que 5 y/o NF menor que 5. En este caso, la nota final se calculará con la fórmula:

$$NF = 0,7 * NR + 0,3 * NP$$

donde NR es la nota del examen de recuperación.

NOTA: La evaluación propuesta puede experimentar alguna modificación en función de las restricciones a la presencialidad que impongan las autoridades sanitarias.

Actividades de evaluación

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Autoevaluación	10%	0	0	1, 3, 4, 5, 9, 10, 11, 12, 14
Examen final	50%	0	0	7, 2, 3, 5, 9, 10, 11, 12, 13, 6, 14
Prácticas	40%	0	0	1, 2, 3, 4, 5, 8, 9, 11, 12, 13, 6, 14

Bibliografía

Bibliografía básica:

- An Introduction to Statistical Learning with Applications in R - Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani

- El bookdown de la asignatura: https://isglobal-brge.github.io/Aprendizaje_Automatico_1/

Bibliografía complementaria:

- The Elements of Statistical Learning: Data Mining, Inference, and Prediction - Trevor Hastie, Robert Tibshirani and Jerome Friedman

- Data Science from Scratch - Joel Grus

- Computer Age Statistical Inference: Algorithms, Evidence and Data Science - Trevor Hastie and Bradley Efron

Software

Tanto la teoría como la parte práctica se hará con R.