

Guía docente

295102 - 295II012 - Análisis de Datos y Reconocimiento de Patrones

Última modificación: 09/08/2024

Unidad responsable:	Escuela de Ingeniería de Barcelona Este
Unidad que imparte:	723 - CS - Departamento de Ciencias de la Computación. 749 - MAT - Departamento de Matemáticas. 707 - ESAII - Departamento de Ingeniería de Sistemas, Automática e Informática Industrial.
Titulación:	MÁSTER UNIVERSITARIO EN INGENIERÍA INTERDISCIPLINARIA E INNOVADORA (Plan 2019). (Asignatura obligatoria). MÁSTER UNIVERSITARIO EN INGENIERÍA QUÍMICA (Plan 2019). (Asignatura obligatoria). MÁSTER UNIVERSITARIO ERASMUS MUNDUS EN INGENIERÍA DE SISTEMAS SOSTENIBLES (EMSSE) (Plan 2024). (Asignatura optativa).
Curso: 2024	Créditos ECTS: 6.0 Idiomas: Inglés

PROFESORADO

Profesorado responsable: CRISTIAN MATA MIQUEL

Otros: Primer quadrimestre:
GERARD ESCUDERO BAKX - Grup: T11, Grup: T12
CRISTIAN MATA MIQUEL - Grup: T11, Grup: T12
FRANCESC POZO MONTERO - Grup: T11, Grup: T12
BARDIA RAFIEIAN - Grup: T11, Grup: T12

COMPETENCIAS DE LA TITULACIÓN A LAS QUE CONTRIBUYE LA ASIGNATURA

Específicas:

CEMUEQ-01. Aplicar conocimientos de matemáticas, física, química, biología y otras ciencias naturales, obtenidos mediante estudio, experiencia y práctica, con razonamiento crítico para establecer soluciones viables económicamente a problemas técnicos
CEMUEQ-03. Conceptualizar modelos de ingeniería, aplicar métodos innovadores en la resolución de problemas y aplicaciones informáticas adecuadas, para el diseño, simulación, optimización y control de procesos y sistemas
CEMUEII-02. Aplicar técnicas de reconocimiento de patrones, inteligencia artificial y análisis estadístico de datos que permitan tomar decisiones de forma objetiva, cuantitativa y reproducible en problemas de naturaleza multidisciplinar.

Genéricas:

CGMUEQ-04. Realizar la investigación apropiada, emprender el diseño y dirigir el desarrollo de soluciones de ingeniería, en entornos nuevos o poco conocidos, relacionando creatividad, originalidad, innovación y transferencia de tecnología
CGMUEQ-05. Saber establecer modelos matemáticos y desarrollarlos mediante la informática apropiada, como base científica y tecnológica para el diseño de nuevos productos, procesos, sistemas y servicios, y para la optimización de otros ya desarrollados
CGMUEII-01. Participar en proyectos de innovación tecnológica en problemas de naturaleza multidisciplinar, aplicando conocimientos matemáticos, analíticos, científicos, instrumentales, tecnológicos y de gestión.

Transversales:

05 TEQ. TRABAJO EN EQUIPO: Ser capaz de trabajar como miembro de un equipo interdisciplinar ya sea como un miembro más, o realizando tareas de dirección con la finalidad de contribuir a desarrollar proyectos con pragmatismo y sentido de la responsabilidad, asumiendo compromisos teniendo en cuenta los recursos disponibles.
06 URI. USO SOLVENTE DE LOS RECURSOS DE INFORMACIÓN: Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de la especialidad y valorar de forma crítica los resultados de esta gestión.
03 TLG. TERCERA LENGUA: Conocer una tercera lengua, que será preferentemente inglés, con un nivel adecuado de forma oral y por escrito y en consonancia con las necesidades que tendrán las tituladas y los titulados en cada enseñanza.

METODOLOGÍAS DOCENTES

La metodología docente incluye clases de teoría, sesiones de laboratorio y actividades de aprendizaje autónomo mediante la realización de proyectos prácticos y el análisis de aplicaciones reales.

OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA

HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTE

Tipo	Horas	Porcentaje
Horas grupo grande	27,0	18.00
Horas grupo pequeño	27,0	18.00
Horas aprendizaje autónomo	96,0	64.00

Dedicación total: 150 h

CONTENIDOS

Análisis exploratorio de datos

Descripción:

- Visualización de datos (histogramas, diagramas de caja, gráficos cuantil-cuantil, diagramas de dispersión multidimensional, etc.)
- Agrupación de datos (k-medias, aglomerativo, modelos de mezclas gaussianas)
- Reducción de la dimensionalidad y análisis de componentes principales.
- Representación de datos y extracción de características.
- Métricas de datos, distancias, normas, etc.

Actividades vinculadas:

Sesión de laboratorio 1: visualización de datos y técnicas de agrupamiento

Sesión de laboratorio 2: Análisis de componentes principales y reducción de dimensionalidad

Dedicación: 10h

Grupo grande/Teoría: 6h

Grupo pequeño/Laboratorio: 4h

Aprendizaje automático supervisado

Descripción:

- Introducción a los problemas de Clasificación / Regresión.
- Métodos basados en la distancia: kNN y algoritmo centroides
- Métodos probabilísticos: Naïve Bayes y LDA
- Métodos basados en reglas: árboles de decisión y AdaBoost
- Métodos basados en hiperplanos: kernels y SVM

Actividades vinculadas:

Sesión de laboratorio 3: Clasificación supervisada I

Sesión de laboratorio 4: Clasificación supervisada II

Dedicación: 12h

Grupo grande/Teoría: 8h

Grupo pequeño/Laboratorio: 4h

Validación de modelos

Descripción:

- Errores de tipo I y tipo II
- Pruebas de contraste de hipótesis univariadas y multivariadas
- Inferencia estadística y estimación de parámetros (máxima verosimilitud, Bayesiana, bootstrapping)
- Procedimientos de validación: validación cruzada; leave-one-out, etc.

Actividades vinculadas:

Sesión de laboratorio 5: Contraste de hipótesis

Sesión de laboratorio 6: Procedimientos de validación de modelos

Dedicación: 12h

Grupo grande/Teoría: 8h

Grupo pequeño/Laboratorio: 4h

Redes neuronales y aprendizaje profundo

Descripción:

- Introducción a las redes neuronales artificiales (ANNs) y al aprendizaje profundo.
- ANNs feed-forward para clasificación y regresión
- Entrenamiento de ANNs: algoritmo de propagación hacia atrás, etapas de optimización, estrategias avanzadas (complejidad de la red, detención temprana, abandono, regularización del peso)
- Arquitecturas especializadas: redes neuronales recurrentes, autocodificadores, redes adversas generativas, redes neuronales convolucionales.

Actividades vinculadas:

Sesión de laboratorio 7: Redes neuronales artificiales

Sesión de laboratorio 8: Aprendizaje profundo

Dedicación: 10h

Grupo grande/Teoría: 6h

Grupo pequeño/Laboratorio: 4h

Temas avanzados i aplicacions

Descripción:

Seminarios de expertos, proyectos aplicados, estudio publicaciones recientes, temes d'actualitat, etc

Dedicación: 8h

Grupo grande/Teoría: 4h

Grupo pequeño/Laboratorio: 4h

SISTEMA DE CALIFICACIÓN

Examen parcial 30%

Examen final 30%

Proyectos y ejercicios 40%



BIBLIOGRAFÍA

Básica:

- Bishop, Christopher M. Pattern recognition and machine learning. New York: Springer, cop. 2006. ISBN 9780387310732.
- Duda, Richard O; Hart, Peter E; Stork, David G. Pattern classification. 2nd ed. New York [etc.]: John Wiley & Sons, cop. 2001. ISBN 0471056693.

RECURSOS

Otros recursos:

Lenguaje python: <https://www.python.org/> /> Librería de herramientas matemáticas: <http://www.numpy.org/> /> Librería de aprendizaje automático: <https://scikit-learn.org/stable/> /> Representación gráfica: <https://matplotlib.org/> /> Repositorio de datos: <https://archive.ics.uci.edu/ml/index>