Guía docente
480093 - TDS - Tratamiento de Datos Socioambientales

Última modificación: 22/05/2024

Unidad responsable: Escuela Técnica Superior de Ingeniería de Caminos, Canales y Puertos de Barcelona
Unidad que imparte: 715 - EIO - Departamento de Estadística e Investigación Operativa.

Titulación: MÁSTER UNIVERSITARIO EN CIENCIA Y TECNOLOGÍA DE LA SOSTENIBILIDAD (Plan 2013). (Asignatura optativa).

Curso: 2024 **Créditos ECTS:** 5.0 **Idiomas:** Inglés

PROFESORADO

Profesorado responsable: CARINA GIBERT OLIVERAS

Otros: Karina Gibert Oliveras
Miquel Sànchez-Marrè

CAPACIDADES PREVIAS

Conocimientos básicos del paquete R
Conocimientos básicos de programación
Estadística Básica

REQUISITOS

Fonaments d'Estadística Aplicada i Mesura de la Sostenibilitat i el Desenvolupament

COMPETENCIAS DE LA TITULACIÓN A LAS QUE CONTRIBUYE LA ASIGNATURA

Específicas:
CE04. Aplicar adecuadamente, y de forma crítica y eficaz, marcos conceptuales, procesos y técnicas de obtención y tratamiento de datos, estadística aplicada, modelización matemática, análisis de sistemas, sistemas de información geográfica, tecnologías de la información y las comunicaciones y la ecología industrial a la solución de retos de la sostenibilidad y desarrollo sostenible.

METODOLOGÍAS DOCENTES

MD1: Lecture or conference (EXP): Sharing knowledge through lectures by professors or by external guest speakers.

MD4: Tutorials of practical or theoretical works (TD): to perform an activity in the classroom, or a theoretical or practical exercise, individually or in small groups, with the advice of the teacher

and

MD6: Extensive project (PA): learning based in the design, planning and realisation in groups of a complex or extensive project or piece of work, applying and extending knowledge and writing a report on this approach and the results and conclusions

OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA

La asignatura da una visión global de la aplicación del Data Science a la resolución de problemas socioambientales reales. El curso prepara para utilizar Minería de Datos en un proceso completo de Descubrimiento del Conocimiento, para extraer información relevante de distintos tipos de datos socioambientales (encuestas, monitorización, data-warehauses, smart sensors...) y dar soporte a la toma de decisiones en fenómenos u organizaciones ambientales de altos niveles de complejidad. El curso se orienta a problemas socio-ambientales reales y a proporcionar los elementos adecuados para diseñar eficiente y correctamente los procesos de Minería de Datos subyacentes, siempre de acuerdo con el problema real de interés en cada aplicación. También se analizan las competencias requeridas en un Data Scientist para tal propósito. Se presentan los principales métodos de minería de datos; se tratan algunos aspectos prácticos importantes, como el efecto de un pre-procesamiento incorrecto, de una incorrecta elección del método de Minería de datos, de una incorrecta interpretación de resultados, de la suposición de falsas hipótesis sobre el proceso analizado, incluyendo la comunicación efectiva del resultados al decisor y el reporting o también las competencias para el trabajo en equipo. Se discutirá sobre casos socioambientales reales; gestión de agua, actividades turísticas sostenibles, polución, uso del suelo, se utilizarán para discutir la versatilidad de la disciplina para proporcionar mejor conocimiento y soporte a la decisión en un amplio espectro de problemas socio-ambientales de gran complejidad.

HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTADO

Tipo	Horas	Porcentaje
Horas aprendizaje autónomo	80,0	64.00
Horas grupo pequeño	9,0	7.20
Horas grupo grande	24,0	19.20
Horas grupo mediano	12,0	9.60

Dedicación total: 125 h

CONTENIDOS

1. Introducción

Descripción:
1.1. Data Science, Data Mining, Knowledge Discovery from Databases and Intelligent decision support.
1.2. Data Mining Pillars: Statistics, Artificial Intelligence, Information Systems, Visualization

Objetivos específicos:
The Data Science and the overall process of Knowledge Discovery from Databases is presented, together with its steps and including Data Mining itself.
The disciplinary pillars of Data Mining are introduced: Statistics and Artificial Intelligence, Information Systems and Data Visualization
Finally, the basic schema of a Knowledge Discovery process is presented.

Actividades vinculadas:
Presentación del proyecto a desarrollar a lo largo del curso y organización de los equipos de trabajo

Dedicación: 2h 30m
Grupo grande/Teoría: 2h 30m

título castellano

Descripción:
contenido castellano

Dedicación: 2h
Grupo grande/Teoría: 2h

título castellano

Descripción:
contenido castellano

Dedicación: 1h
Grupo grande/Teoría: 1h

título castellano

Descripción:
contenido castellano

Dedicación: 1h
Grupo grande/Teoría: 1h

título castellano

Descripción:
contenido castellano

Dedicación: 5h
Grupo grande/Teoría: 5h

título castellano

Descripción:
contenido castellano

Dedicación: 2h 30m
Grupo grande/Teoría: 2h 30m

título castellano

Descripción:
contenido castellano

Dedicación: 2h 30m
Grupo grande/Teoría: 2h 30m

título castellano

Descripción:
contenido castellano

Dedicación: 2h 30m
Grupo grande/Teoría: 2h 30m

título castellano

Descripción:
contenido castellano

Dedicación: 2h 30m
Grupo grande/Teoría: 2h 30m

título castellano

Descripción:
contenido castellano

Dedicación: 2h 30m
Grupo grande/Teoría: 2h 30m

título castellano

Descripción:
contenido castellano

Dedicación: 2h 30m
Grupo grande/Teoría: 2h 30m

título castellano

Descripción:
contenido castellano

Dedicación: 2h 30m
Grupo grande/Teoría: 2h 30m

título castellano

Descripción:
contenido castellano

Dedicación: 2h 30m
Grupo grande/Teoría: 2h 30m

Alcance del proceso de KDD

Descripción:
concepto y fronteras del concepto

Dedicación: 2h
Grupo grande/Teoría: 2h

Ciencia de datos

Descripción:
Proceso, perspectiva histórica, motivación e impacto actual
Diseño de un proceso de descubrimiento de conocimiento

Actividades vinculadas:
Diseñar el proceso KDD de la práctica

Dedicación: 2h
Grupo grande/Teoría: 2h

ACTIVIDADES

Presentación intermedia de proyectos

Descripción:
Presentación oral de la primera parte del proyecto y discusión

Objetivos específicos:
Hito para sincronizar a todos los estuidantes con un plan de trabajo realizable
Se evalúan las habilidades en comunicación y reporting, así como las competencias técnicas y de organización del trabajo en equipo

Dedicación: 2h 30m
Grupo grande/Teoría: 2h 30m

nombre castellano

Dedicación: 2h 30m
Grupo grande/Teoría: 2h 30m

SISTEMA DE CALIFICACIÓN

Se desarrollará un proyecto de envergadura por grupos, aplicando un proceso completo de data science con datos reales, el cual incluirá la aplicación de métodos vistos en clase. El proyecto se realiza bajo la supervisión del profesor.
Una entrega intermedia (D1) ayudará a planificar mejor el trabajo global (D2). Habrá una nota que se asignará de la siguiente forma:

NotaFinal = 0.4D1 + 0.6D2

donde

D1 = 0.4 * calidad del documento escrito + 0.3 * calidad de la presentación oral y discusión + 0.2 * rendimiento individual mostrado en las sesiones de laboratorio

D2 = alfa * (0.4 * calidad del documento escrito + 0.3 * calidad de la presentación oral y discusión + 0.2 * rendimiento individual mostrado en las sesiones de laboratorio)

siendo alfa un factor entre 0.5 y 1.5 que resulta del Proceso de cross-evaluation realizado por los compañeros de grupo de trabajo en fecha de entrega del trabajo final.

AV2. Prueba oral de control de conocimientos (PO).
AV3. Trabajo realizado en forma individual o en grupo a lo largo del curso (TR). Incluye tanto la evaluación de resultados e informes, como la presentación oral de los mismos.
AV4. Asistencia y participación en clases y laboratorios (AP).
AV5. Calidad y rendimiento del trabajo en grupo (TG).