Guia docent
480093 - TDS - Tractament de Dades Socioambientals

Última modificació: 22/05/2026

Unitat responsable: Escola Tècnica Superior d'Enginyeria de Camins, Canals i Ports de Barcelona
Unitat que imparteix: 715 - EIO - Departament d'Estadística i Investigació Operativa.

Titulació: MÀSTER UNIVERSITARI EN CIÈNCIA I TECNOLOGIA DE LA SOSTENIBILITAT (Pla 2013). (Assignatura optativa).

Curs: 2026 **Crèdits ECTS:** 5.0 **Idiomes:** Anglès

PROFESSORAT

Professorat responsable: CARINA GIBERT OLIVERAS

Altres: Karina Gibert Oliveras
Miquel Sànchez-Marrè

CAPACITATS PRÈVIES

Coneixements bàsics del paquet R
Coneixements bàsics de programació
Estadística Bàsica

REQUISITS

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

COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

Específiques:
CE04. Aplicar adequadament, i de forma crítica i eficaç, marcs conceptuals, processos i tècniques d'obtenció i tractament de dades, estadística aplicada, modelització matemàtica, anàlisi de sistemes, sistemes d'informació geogràfica, tecnologies de la informació i les comunicacions i l'ecologia industrial a la solució de reptes de la sostenibilitat i desenvolupament sostenible.

METODOLOGIES DOCENTS

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

OBJECTIUS D'APRENENTATGE DE L'ASSIGNATURA

L'objectiu principal d'aquesta assignatura és proporcionar una visió global de l'aplicació del Data Science a la resolució de problemes sòcio-ambientals reals.
Es presenta la utilització de tècniques de Mineria de Dades en un procés complet de Descobriment de Coneixement, destinat a extreure informació rellevant de diferent tipus de dades sòcio-ambientals (enquestes, monitorització de processos, data warehouses, smart sensors...) amb l'objectiu de donar suport a la presa de decisions en fenòmens o organizacions amb alts nivells de complexitat. El curs s'orienta a problemes sòcio-ambientals reals i a proporcionar els elements adients per a dissenyar eficientment i correcta els processos de Mineria de Dades, sempre d'acord amb el problema real d'interès en cada aplicació. També s'analitzen les competències requerides en un Data Scientist per a tal propòsit. Es presenten els principals mètodes de mineria de dades; s'entrenen alguns aspectes pràctics importants en aplicacions reals, com l'efecte d'un pre-processament incorrecte, d'una incorrecta selecció del mètode de mineria de dades, d'una incorrecta interpretació de resultats, de l'assumpció d'hipòtesis falses per al procés analitzat; la comunicació efectiva de resultats als decisors o el reporting també s'analitzen. Aquestes qüestions ajuden a garantitzar la validesa i utilitat dels resultats finals, així com l'impacte real de l'anàlisi en el domini d'estudi. Es discutirà sobre casos sòcio-ambiental reals; gestió d'aigua, activitats turístiques sostenibles, pol.lució, ús del sól, s'utilitzaran per mostrar la versatilitat de la disciplina per a proporcionar un millor coneixement i millor suport a la presa de decisions en un ampli espectre de problemes sòcio-ambientals reals i d'alt nivell de complexitat.

HORES TOTALS DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores grup mitjà	10,0	7.99
Hores aprenentatge autònom	88,4	70.66
Hores grup gran	20,0	15.99
Hores grup petit	6,7	5.36

Dedicació total: 125.1 h

CONTINGUTS

1. Introducció

Descripció:
1.1. Ciència de les Dades, Mineria de dades, Descobriment de coneixement a partir de dades i Suport intel.ligent a la presa de decisions.
1.2. Pilars de la Mineria de Dades: Estadística, Intel.ligència Artificial, Sistemes d'informació, Visualització

Objectius específics:
Es presenten les disciplines de ciència de les dades i el procés genèric de Descobriment del Coneixement a partir de dades, amb les seves etapes, incloent la de mineria de dades.
S?introdueixen els pilars disciplinaris de la mineria de dades: Estadística, Intel.ligència artificial, sistemes d?informació i Visualització de dades
Finalment, es presenta l?esquema bàsic d?un procés de descobriment del coneixement.

Activitats vinculades:
Presentació del projecte a desenvolupar al llarg del curs i organització dels equips de treball

Dedicació: 2h 30m
Grup gran/Teoria: 2h 30m

2. Abast, el procés de KDD.

Descripció:
2.1. Tipus de problemes susceptibles de Ciència de Dades
2.2. Dominis poc estructurats
2.3 Coneixement a priori; Coneixement implícit. Causes i conseqüències
2.4. Principals softwares de Mineria de dades (R, weka, RapidMiner)

Objectius específics:
Es discuteixen problemes sòcio-ambientals de diferents naturaleses i llur different nivells de complexitat, d'acord amb la classificació proposada per Simpson. S'introdueix la gestió dels dominis poc estructurats i del coneixement implícit, causes i conseqüències.
Es presenten algunes eines software per desenvolupar tasques de mineria de dades, amb un focus especial en el sistema R.

Dedicació: 2h
Grup gran/Teoria: 2h

3. Formalització del problema de Ciència de Dades i disseny d'un procés complet de Descobriment de Coneixment

Descripció:
Es treballen les passes del procéss de Ciència de Dades i del procés de descobriment de coneixement

Activitats vinculades:
Definir el projecte a desenvolupar, identificar les fonts de dades

Dedicació: 1h
Grup gran/Teoria: 1h

4. Estructures de dades

Descripció:
4.1 Principals fonts de dades sòcio-ambientals
4.2 Representació de dades i metadades

Objectius específics:
S'estudien les principals estructures de dades analitzades en mineria de dades per camps sòcio-ambientals
Importància de les metadades. Formats i continguts

Activitats vinculades:
Construir el fitxer de metadades de les dades del projecte

Dedicació: 1h
Grup gran/Teoria: 1h

5. Pre-processament de dades

Descripció:
5.0 Metodologia de referència per a pre-processament
5.1 Qualitat de les dades
5.2 Filtratge i mostreig
5.3. Tractament de dades mancants
5.4 Tractament de dades anòmales
5.5 Transformació de les dades i dades derivades
5.6 Feature Weighting i reducció de la dimensionalitat

Objectius específics:
Es discuteix la importància de la qualitat de les dades i les conseqüències de la manca de qualitat. S'introdueixen els aspectes més rellevants de la fase de preparació de dades: Dades mancants, detecció de dades anòmales, variables derivades, transformacions, filtratge, mostreig, selecció i rellevància de les variables, reducció de la dimensionalitat (selecció de variables i mètodes factorials), tots crítics per a garantir la validesa del procés d'anàlisi. Es proporcionaran guies de bones pràctiques, així com una metodologia general de referència

Activitats vinculades:
Pre-processar les dades del projecte

Dedicació: 5h
Grup gran/Teoria: 5h

6. Selecció del mètode de mineria de dades

Descripció:
6.1. L'aproximació orientada al problema
6.2. Criteris que determinen el mètode de mineria de dades adient
6.3. El mapa conceptual de mètodes de mineria de dades (mapa DMMDM)

Objectius específics:
El curs segueix un enfoc orientat al problema, on la naturalesa intrínseca del poblema determina el procés d'anàlisi i no al revés. Es discuteix sobre els factors que determinen la correcta elecció del mètode de mineria de dades a emprar en un procés concret de Ciència de Dades. El mapa DMMCM presenta una tipologia de mètodes de mineria de dades que s'aporta com a base conceptual per a la selecció.

Dedicació: 2h 30m
Grup gran/Teoria: 2h 30m

7. Fase de mineria de dades: Mètodes descriptius

Descripció:
7.1. Mètodes descriptius
clustering: Mètodes de particions, jeràrquics. Escalabilitat. Mètodes híbrids, introducció de coneixement expert a priori. Explicitació de coneixement
7.2 Caracterització de les classes

Objectius específics:
S'estudien mètodes d'identificació i caracterització de perfils

Activitats vinculades:
Fer clustering de les dades del projecte

Dedicació: 2h 30m
Grup gran/Teoria: 2h 30m

8. Fase de Mineria de dades: Mètodes associatius

Descripció:
8.1. Regles d'associació
8.2. Mètodes factorials
8.3. Xarxes baiessianes

Objectius específics:
Aquest tema introdueix mètodes per a descobrir les relacions d'associació entre les variables de la base de dades

Activitats vinculades:
Aplicar algun mètode associatiu a les dades del projecte

Dedicació: 2h 30m
Grup gran/Teoria: 2h 30m

9, Fase de mineria de dades: Mètodes discriminants

Descripció:
9.1 Arbres de decisió
9.2 Inducció de regles de classificació
9.3 Màquines de vector de suport
9.4 anàlisi discriminant
9.5 mètodes d'ensemble i bagging
9.6 mètodes híbrids

Objectius específics:
Mètodes per predir una variable de classe o una variable qualitativa. Se'n veuran almenys 3.

Activitats vinculades:
Aplicar almenys dos mètodes discriminants a les dades del projecte

Dedicació: 2h 30m
Grup gran/Teoria: 2h 30m

10. Fase de mineria de dades: Mètodes predictius

Descripció:
10.1 Regressió, model.lització estadística en general
10.2 Mètodes temporals
10.3 Xarxes neuronals artificials
10.4 Swarm intelligence

Objectius específics:
Mètodes per predir una variable numèrica. Se'n veuran almenys 2

Activitats vinculades:
Predicció d'una o vàries variables numèriques

Dedicació: 2h 30m
Grup gran/Teoria: 2h 30m

11. Mineria de dades espai-temporal

Descripció:
11. Models espaitemporals

Objectius específics:
s'introdueixen eines per tractar dades espai-temporals

Activitats vinculades:
Revisió general del projecte

Dedicació: 2h 30m
Grup gran/Teoria: 2h 30m

12. Post-processament i validació

Descripció:
12.1 Eines de post-processament
12.2. Validació de models
12.3. Validació de resultats

Objectius específics:
Eines de post-processament i validació per models i resultats adaptades als diferents mètodes de mineria de dades.

Activitats vinculades:
Validació dels models del projecte

Dedicació: 2h 30m
Grup gran/Teoria: 2h 30m

13. Reporting i comunicació de resultats

Descripció:
13.1. Informes, informes automàtics
13.2. Comunicació de resultats

Objectius específics:
Crucial per garantir que els resultats del procés de ciència de dades aportin suport efectiu a la presa de decisiona i que l'anàlisi tingui impacte real en el domini objecte d'interès

Activitats vinculades:
Revisió de l'informe del projecte

Dedicació: 2h 30m
Grup gran/Teoria: 2h 30m

Abast del procés de KDD

Descripció:
concepte i fronteres del concepte

Dedicació: 2h
Grup gran/Teoria: 2h

Ciència de dades

Descripció:
Procés, perspectiva històrica, motivació i impacte actual
Disseny d'un procés de descobriment de coneixement

Activitats vinculades:
Dissenyar el procés KDD de la pràctica

Dedicació: 2h
Grup gran/Teoria: 2h

ACTIVITATS

Presentació intermèdia de projectes

Descripció:
Presentació oral de la primera part del projecte i discussió
Entregable escrit

Objectius específics:
Fita per sincronitzar tots els estuidants amb un pla de treball realizable
S'avaluen les habilitats en comunicació i reporting, així com les competències tècniques i d'organització del treball en equip

Dedicació: 2h 30m
Grup gran/Teoria: 2h 30m

Presentació final de projectes

Descripció:
Presentació oral i entrega escrita del projecte complet. Discussió general i individual amb els professors

Dedicació: 2h 30m
Grup gran/Teoria: 2h 30m

SISTEMA DE QUALIFICACIÓ

Qualsevol acte de frau acadèmic, plagi, ús o tinença a l'abast de mitjans no autoritzats en qualsevol activitat d'avaluació comportarà la qualificació de zero (0) en la prova o lliurament afectat. A més, d'acord amb la normativa de la Universitat, la possible derivació dels fets per a l'obertura d'un expedient disciplinari implicarà que l'assignatura quedi en l'estat provisional de "pendent d'avaluació" fins a la resolució de l'expedient. La gestió d'aquestes incidències es desplega d'acord amb el **Marc d'actuació per a la integritat acadèmica en l'avaluació de la UPC**.

Es desenvoluparà un projecte d'envergadura per grups, aplicant un procés complet de data science a dades reals, el qual inclourà l'aplicació de mètodes vistos a classe. El projecte es realitza sota la supervisió del professor.
Una entrega intermèdia (D1) ajudarà a planificar millor el treball global (D2). Hi haurà una nota que s'assignarà de la següent forma:

NotaFinal= 0.4D1+0.6D2

on

D1= 0.4*qualitat del document escrit+0.3*qualitat de la presentació oral i discussió + 0.2* rendiment individual mostrat a les sessions de laboratori

D2= alfa * (0.4*qualitat del document escrit+0.3*qualitat de la presentació oral i discussió + 0.2* rendiment individual mostrat a les sessions de laboratori)

essent alfa un factor entre 0.5 i 1.5 que resulta del procéss de cross-evaluation realitzat pels companys de grup de treball a data d?entrega del treball final.

AV2. Prova oral de control de coneixements (PO).
AV3. Treball realitzat en forma individual o en grup al llarg del curs (TR). Inclou tant l'avaluació de resultats i informes, com la presentació oral dels mateixos.
AV4. Assistència i participació en classes i laboratoris (AP).
AV5. Qualitat i rendiment del treball en grup (TG).