

Guia docent

240346 - 240E0065 - Ciència de Dades

Última modificació: 03/07/2024

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

Titulació: MÀSTER UNIVERSITARI EN ENGINYERIA INDUSTRIAL (Pla 2014). (Assignatura optativa).
MÀSTER UNIVERSITARI EN ENGINYERIA D'ORGANITZACIÓ (Pla 2021). (Assignatura obligatòria).
MÀSTER UNIVERSITARI EN RECERCA EN ENGINYERIA MECÀNICA (Pla 2024). (Assignatura optativa).

Curs: 2024 **Crèdits ECTS:** 5.0 **Idiomes:** Anglès

PROFESSORAT

Professorat responsable: Puig De Dou, Ignacio

Altres: Marta Castellano Palomino

REQUISITS

Àlgebra Lineal i Càlcul Diferencial i Integral a nivell dels cursos del GETI.
Coneixements de Probabilitat i Inferència a nivell d'Estadística del GETI.
Fonaments de Programació.

COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

Específiques:

CEEORG2. Dissenyar, desenvolupar i aplicar mètodes analítics (mètodes quantitius, estadística, models i eines de decisió) per a la presa de decisions estratègiques, tàctiques i operatives en les organitzacions. (Competència específica associada a l'especialitat en Organització Industrial).

CEETI3. Analitzar dades de gran volum mitjançant models estadístics i algoritmes d'intel·ligència artificial. (Competència específica associada a l'especialitat en Tecnologies de la Informació per a la Indústria).

CEETI4. Determinar la fiabilitat i seguretat en sistemes complexos mitjançant mètodes de manteniment predictiu i seguretat ciberfísica. (Competència específica associada a l'especialitat en Tecnologies de la Informació per a la Indústria).

Transversals:

CT4. ÚS SOLVENT DELS RECURSOS D'INFORMACIÓ: Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació de l'àmbit d'especialitat, i valorar de forma crítica els resultats d'aquesta gestió.

Bàsiques:

CB 6. Tenir i comprendre coneixements que aportin una base o oportunitat de ser originals en el desenvolupament i/o aplicació d'idees, sovint en un context d'investigació

CB 9. Que els estudiants sàpiguin comunicar les seves conclusions i coneixements (i darrers raonaments que els sustentin), a públics especialitzats i no especialitzats de manera clara i sense ambigüitats.

CB10. Que els estudiants poseixin les habilitats d'aprenentatge que els permetin continuar estudiant d'una manera que sigui en gran mesura autodirigida o autonòma.

METODOLOGIES DOCENTS

Presentació de fonaments teòrics de les diferents metodologies en classes magistrals d'hora i mitja. Implementació pràctica de les metodologies amb exemples en sessions en aula d'ordinadors d'hora i mitja.

Anàlisi i presentació d'una problemàtica més complexa a través d'un cas pràctic

OBJECTIUS D'APRENENTATGE DE L'ASSIGNATURA

L'alumne obtindrà un coneixement de les característiques bàsiques dels models d'aprenentatge estadístic i màquina i sabrà quin model emprar en funció del problema a solucionar.

Entendrà les metodologies d'entrenament i validació dels models d'aprenentatge automàtic i serà capaç de planificar de manera robusta el procés d'entrenament d'un model

Coneixerà els diferents models disponibles d'aprenentatge automàtic i serà capaç d'implementar-los en casos reals.

Coneixerà i sabrà emprar programari R per implementar i validar models d'aprenentatge automàtic.

HORES TOTS DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores grup petit	22,5	18.00
Hores grup gran	22,5	18.00
Hores aprenentatge autònom	80,0	64.00

Dedicació total: 125 h

CONTINGUTS

Introducció a l'Aprenentatge Màquina amb R

Descripció:

Data Science en la Indústria.

Conceptes bàsics de l'Aprenentatge Màquina

Introducció a R:

- anàlisi descriptiu
- comandes i programació
- data wrangling

Dedicació: 20h

Grup mitjà/Pràctiques: 10h

Aprenentatge autònom: 10h

Aprenentatge Supervisat: el problema de Regressió i Classificació

Descripció:

Regressió Multivariant Lineal i no-Lineal

Regressió Logística com a mètode de classificació

Bondat d'ajust en models de classificació.

Selecció i ajust de models. Remostreig

Dedicació: 30h

Grup mitjà/Pràctiques: 15h

Aprenentatge autònom: 15h

Mètodes no supervisats

Descripció:

Tècniques de reducció de la dimensionalitat
Mètodes de clusterització basats en distàncies: jeràrquic i aglomeratiu.

Dedicació: 20h

Grup mitjà/Pràctiques: 10h

Aprenentatge autònom: 10h

Mètodes d'Aprenentatge Màquina Avançats: basats en arbres i xarxes neuronals

Descripció:

Mètodes basats en arbres: bagging, random forest i boosting.
Xarxes neuronals

Dedicació: 20h

Grup mitjà/Pràctiques: 10h

Aprenentatge autònom: 10h

SISTEMA DE QUALIFICACIÓ

0.30 examens d'avaluació continuada + 0.25 treball de curs + 0.45 examen final.

Hi haurà tres examens en horari de classe habitual com avaluació continuada sobre els temes tractats fins aquell moment. La nota del final substituirà la del parcial si és superior.

NORMES PER A LA REALITZACIÓ DE LES PROVES.

Apliquen les normes que estableix l'ETSEIB-UPC.

BIBLIOGRAFIA

Bàsica:

- Gareth James, ... [et al.]. An Introduction to statistical learning : with applications in R [en línia]. 2nd ed. New York: Springer, 2021 [Consulta: 01/10/2024]. Disponible a : <https://ebookcentral-proquest-com.recursos.biblioteca.upc.edu/lib/upcatalunya-ebooks/detail.action?pq-origsite=primo&docID=6686746>. ISBN 1071614177.
- Manly, Bryan F. J ; Navarro Alberto, Jorge A. Multivariate statistical methods : a primer. Fourth edition. Boca Raton: CRC Press, Taylor & Francis Group, [2017]. ISBN 9781498728966.

Complementària:

- Hastie, T.; Tibshirani, R.; Friedman, J. The elements of statistical learning : data mining, inference and prediction [en línia]. 2nd ed. New York [etc.]: Springer, cop. 2009 [Consulta: 25/08/2022]. Disponible a : <https://link-springer-com.recursos.biblioteca.upc.edu/book/10.1007/978-0-387-84858-7>. ISBN 0387952845.
- Peña, Daniel. Análisis de datos multivariantes [en línia]. 1a ed. Madrid: McGraw-Hill/Interamericana de España, S.L., 2013 [Consulta: 22/10/2024]. Disponible a : https://www-ingebook-com.recursos.biblioteca.upc.edu/ib/NPcd/IB_BooksVis?cod_primaria=1000187&codigo_libro=4203. ISBN 9788448191849.
- Bertsekas, Dimitri P; Tsitsiklis, John N. Introduction to probability : Dimitri P. Bertsekas and John N. Tsitsiklis. 2nd ed. Belmont, Massachussets: Athena Scientific, cop. 2008. ISBN 9781886529236.



RECURSOS

Enllaç web:

- <https://cdr-book.github.io/>. Recurs