

**FITXA IDENTIFICATIVA****Dades de l'Assignatura**

Codi	44650
Nom	Estadística i optimització
Cicle	Màster
Crèdits ECTS	6.0
Curs acadèmic	2022 - 2023

Titulació/titulacions

Titulació	Centre	Curs	Període
2221 - M.U. en Ciència de Dades	Escola Tècnica Superior d'Enginyeria	1	Primer quadrimestre

Matèries

Titulació	Matèria	Caràcter
2221 - M.U. en Ciència de Dades	2 - Estadística i optimització	Obligatòria

Coordinació

Nom	Departament
MARTINEZ GAVARA, ANNA	130 - Estadística i Investigació Operativa

RESUM

Aquesta assignatura es divideix en dos blocs temàtics importants. En la part d'Estadística, es presenten els principals conceptes, tècniques estadístiques i sèries temporals, els quals són d'utilitat per a l'anàlisi de dades i *data science*. Coneguda la Teoria de Probabilitat, en aquesta part s'estudien els principals mètodes i conceptes per a l'estimació i la formulació de hipòtesis. Finalment, en aquest bloc s'introduiran les sèries temporals, estudiant entre elles els models ARIMA.

En la part d'Optimització s'introdueixen els conceptes bàsics d'optimització irrestringida i restringida, i a partir d'ells es desenvolupen els mètodes d'optimització (com ara l'algorisme EM), mètodes de recerca local (gradients i derivats) i mètodes de recerca global (genètics).



CONEIXEMENTS PREVIS

Relació amb altres assignatures de la mateixa titulació

No heu especificat les restriccions de matrícula amb altres assignatures del pla d'estudis.

Altres tipus de requisits

COMPETÈNCIES

2221 - M.U. en Ciència de Dades

- Que els estudiants posseïsquen les habilitats d'aprenentatge que els permeten continuar estudiant d'una forma que haurà de ser en gran manera autodirigida o autònoma.
- Ser capaços de valorar la necessitat de completar la seua formació tècnica, científica, en llengües, en informàtica, en literatura, en ètica, social i humana en general, i d'organitzar el seu propi autoaprenentatge amb un alt grau d'autonomia
- Ser capaços d'accedir a ferramentes d'informació (bibliogràfiques i d'ocupació) i utilitzar-les apropiadament.
- Ser capaços d'assumir la responsabilitat del seu propi desenvolupament professional i de la seua especialització en un o més camps d'estudi, aplicant els coneixements adquirits en la identificació d'eixides professionals i jaciments d'ocupació.
- Extraure coneixement de conjunts de dades en diferents formats.
- Entendre la utilitat de la ciència de dades i els seus elements associats, així com la seua aplicació en la resolució de problemes, triant les tècniques més adequades a cada problema, aplicant de forma correcta les tècniques d'avaluació i, finalment, interpretant els models i resultats.

RESULTATS DE L'APRENTATGE

Conèixer les bases matemàtiques que donen suport a les diferents tècniques d'anàlisi de dades.

Proporcionar a l'estudiant una base perquè puga desenvolupar nous algorismes/tècniques d'anàlisi de dades.

Processar dades en els quals existeix una dependència temporal (sèries temporals).

Conèixer l'ús dels contrastes d'hipòtesis en l'anàlisi de dades.

Aplicar mètodes de recerca local (gradients i derivats) i mètodes de recerca global (genètics).

DESCRIPCIÓ DE CONTINGUTS



1. Descripció de dades i sèries temporals

Definicions. Sèries temporals: Classificació, anàlisi clàssic i models d'autocorrelació

2. Inferència Estadística

Probabilitat. Mostra aleatòria. Funció de Versemblança.

Estimació: puntual i per intervals. Mètode de Màxima Versemblança.

Contrast d'hipòtesi: hipòtesi estadística, regió crítica i regió d'acceptació, estadístic de contrast, regla de decisió, error de tipus I i tipus II, nivell de significativitat, funció de potència del test, p-valor

3. Anàlisi estadístic d'una o diverses poblacions

Anàlisi estadístic d'una població. Estudi sobre una proporció. Estudi sobre una mitjana poblacional.

Comparació de diverses poblacions (dades contínues). Mostres emparellades i independents. Alternatives no paramètriques.

Comparació de diverses poblacions (dades categòriques). Taules de contingència.

4. Optimització irrestringida

Conceptes bàsics d'optimització sense restriccions. Mètodes de recerca local. Algorisme EM

5. Optimització amb restriccions

Conceptes bàsics d'optimització amb restriccions. Programació Lineal i Entera. Mètodes de recerca local

6. Mètodes de recerca global. Algorismes metaheurístics

Mètodes de recerca global. Algorismes metaheurístics. Algorismes genètics

**VOLUM DE TREBALL**

ACTIVITAT	Hores	% Presencial
Classes teoricopràctiques	60,00	100
Elaboració de treballs individuals	20,00	0
Estudi i treball autònom	12,00	0
Lectures de material complementari	3,00	0
Preparació d'activitats d'avaluació	12,00	0
Preparació de classes de teoria	20,00	0
Preparació de classes pràctiques i de problemes	13,00	0
Resolució de casos pràctics	10,00	0
TOTAL	150,00	

METODOLOGIA DOCENT

Les classes combinaran la part teòrica amb la pràctica, sense distinció de les sessions dedicades a teoria i pràctica. Totes les sessions s'impartiran en aula d'informàtica.

En la part teòrica s'introduiran els conceptes i mètodes de l'assignatura, acompanyant-los d'exemples i exercicis proposats. Les sessions pràctiques estaran sincronitzades amb la part teòrica i en aquestes l'aprenentatge es realitzarà mitjançant la resolució de problemes, exercicis i casos d'estudi a través dels quals s'adquireixen competències sobre els diferents aspectes de la matèria.

AVALUACIÓ

L'avaluació de l'aprenentatge dels coneixements i competències obtinguts pels estudiants es realitzarà de forma continuada al llarg del curs, i constarà dels següents blocs d'avaluació:

1. Exercicis i treballs entregats durant el curs i/o exàmens parcials: 60% de la nota final.
2. Prova o proves individuals: 40% de la nota final.

Les qualificacions obtingudes en l'apartat 1 es conservaran en les dues convocatòries del curs acadèmic en què s'hagen realitzat, donat que la seua avaluació sols és possible en el període de docència.

REFERÈNCIES



Bàsiques

- Barber, D., (2012). Bayesian Reasoning and Machine Learning. Ed. Cambridge University Press
- Peña, D. , (2001). Fundamentos de Estadística. Ed. Alianza Editorial.
- Shumway, R.H., Stoffer, D. S., (2011). Time Series Analysis and Its Applications, Ed. Springer.
- Walpole, R. E., Myers, R.H. y Myers, S.L. (1999). Probabilidad y Estadística para Ingenieros. Ed. Prentice Hall.
- Lange, K. (2004) Optimization. Ed. Springer.
- Li, L. (2015) Selected Applications of Convex Optimization. Ed. Springer
- Nocedal, J; Wright, S. (2006) Numerical Optimization. Ed. Springer, 2nd edition
- Snyman, J.A. (2005) Practical Mathematical Optimization. Ed. Springer
- James G.; Witten D.; Hastie T.; Tibshirani R. (2017). An introduction to Statistical Learning with Applications in R. Springer.
- Bruce, P.; Bruce, A. (2017). Practical Statistics for Data Scientists. O'Reilly Media, Inc.

Complementàries

- John Braun Duncan, W., Murdoch, J. (2007). A first course in statistical programming with R. Cambridge University Press.
- Venables, V.N. (2013). An Introduction to R. <http://cran.r-project.org>.
- Bertsekas, D.P. (2016) Nonlinear Programming. Ed. Athena Scientific, 3rd Edition