

Visualització de Dades i Modelització

Codi: 43482
Crèdits: 6

Titulació	Tipus	Curs	Semestre
4313136 Modelització per a la Ciència i l'Enginyeria / Modelling for Science and Engineering	OT	0	1

La metodologia docent i l'avaluació proposades a la guia poden experimentar alguna modificació en funció de les restriccions a la presencialitat que imposin les autoritats sanitàries.

Professor/a de contacte

Nom: Pere Puig Casado

Correu electrònic: Pere.Puig@uab.cat

Utilització d'idiomes a l'assignatura

Llengua vehicular majoritària: anglès (eng)

Equip docent

Rosa Camps Camprubí

Rosario Delgado de la Torre

Juan Ramón González Ruíz

Prerequisits

Un coneixement elemental en Teoria de Probabilitats i en Inferència Estadística.

Objectius

Curs de R. Tots els exercicis pràctics es faran utilitzant el paquet estadístic R. Aquest curs introductori és bàsic pels posteriors desenvolupaments.

Visualització de grans conjunts de dades amb R GViz, Maps i Tabplot.

Simulació de dades, Bootstrap i tests permutacionals Aquestes metodologies permeten una solució ràpida per a models estadístics complexos sense un coneixement profund dels mètodes estadístics generals clàssics. Són eines indispensables entre les actuals tècniques de modelització estadística. Els estudiants completaran un programa d'entrenament bàsic, incloent el software adient, i aprendran com enfrontar-se a problemes amb dades reals.

Xarxes Bayesianes, en opinió de molts investigadors constitueixen una de les més significants contribucions de la IA en aquest segle. Es tracta d'estructures gràfiques per a representar relacions probabilístiques entre un gran nombre de variables i també per a fer inferència probabilística amb aquestes variables en un gran nombre de camps d'aplicació. Un dels objectius d'aquest curs és introduir-les, desenvolupant en els estudiants la habilitat de utilitzar-les en modelització, tant des d'un punt de vista teòric com pràctic, amb un particular èmfasi en la utilització d'un software apropiat.

Competències

- "Aplicar el pensamiento lógico/matemático: el proceso analítico a partir de principios generales para llegar a casos particulares; y el sintético, para a partir de diversos ejemplos extraer una regla general."
- Analitzar sistemes complexos de diferents camps i determinar les estructures i paràmetres bàsics del seu funcionament.
- Analitzar, sintetitzar, organitzar i planificar projectes del seu camp d'estudi.
- Aplicar la metodologia de recerca, tècniques i recursos específics per investigar en un determinat àmbit d'especialització.
- Aplicar les tècniques de resolució dels models matemàtics i els seus problemes reals d'implementació.
- Concebre i dissenyar solucions eficients, aplicant tècniques computacionals, que permetin resoldre models matemàtics de sistemes complexos.
- Extreure d'un problema complex la dificultat principal, separada d'altres qüestions d'indole menor.
- Formular, analitzar i validar models matemàtics de problemes pràctics de diferents camps.

Resultats d'aprenentatge

1. "Aplicar el pensament lògic/matemàtic: el procés analític a partir de principis generals per arribar a casos particulars; i el sintètic, para a partir de diversos exemples extreure una regla general."
2. Analitzar, sintetitzar, organitzar i planificar projectes del seu camp d'estudi.
3. Aplicar la metodologia de recerca, tècniques i recursos específics per investigar en un determinat àmbit d'especialització.
4. Aplicar tècniques de Sèries Temporals per estudiar models associats a problemes pràctics.
5. Aplicar tècniques de Sèries Temporals per predir el comportament futur de certs fenòmens.
6. Extreure d'un problema complex la dificultat principal, separada d'altres qüestions d'indole menor.
7. Identificar els paràmetres que determinen el funcionament d'un sistema
8. Implementar les solucions proposades de forma fiable i eficient.
9. Reconèixer problemes que requereixin aplicar tècniques de Sèries Temporals per construir models associats a problemes pràctics.
10. Seleccionar la millor descripció d'un sistema en funció de les seves característiques particulars
11. Usar programaris específics per a la resolució de problemes d'optimització.

Continguts

Part 1a: Introducció al R (6h)

Part 2na: Visualització de grans conjunts de dades amb R (6h)

Part 3a: Xarxes Bayesianes (14h)

- 1) Block 1: Elements bàsics.
- 2) Block 2: Xarxes causals i Inferència en Xarxes Bayesianes.
- 3) Block 3: Aprenentatge dels paràmetres de les xarxes bayesianes.

Part 4a: Simulació de dades, Bootstrap i Tests Permutacionals (12h)

- 1) Tests Permutacionals.
- 2) Jackknife.
- 3) Bootstrap paramètric.
- 4) Bootstrap no-paramètric.

Metodologia

Les classes magistrals d'aquest curs, en les que és determinant l'explicació del professor, són la base del procés d'aprenentatge. És també molt important la participació dels alumnes, combinada amb sessions pràctiques en les que l'estudiant ha d'usar el coneixement après per a solucionar problemes.

Nota: es reservaran 15 minuts d'una classe, dins del calendari establert pel centre/titulació, per a la complementació per part de l'alumnat de les enquestes d'avaluació de l'actuació del professorat i d'avaluació de l'assignatura/mòdul.

Activitats formatives

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Classes magistrals	38	1,52	4, 5, 7, 8, 9, 10
Exercicis	16	0,64	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11
Projectes+Treballs	18	0,72	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11
Tipus: Supervisades			
Sessions pràctiques	20	0,8	4, 5, 7, 8, 9, 10

Avaluació

L'avaluació del curs consisteix en una avaluació continuada.

Hi ha 4 avaluacions durant el curs, que tenen un pes de 10%, 10%, 40% i 40% respectivament.

Cada professor comptarà amb el seu tipus propi d'avaluació.

Avaluació de la part 1: Treball continuat + projecte final (anàlisi individual de dades reals amb R).

Avaluació de la part 2: Treball continuat + projecte

Avaluació de la part 3: Treball continuat + Lliurament d'alguns exercicis + examen

Avaluació de la part 4: Treball continuat + Lliurament d'alguns exercicis.

Activitats d'avaluació

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Projectes i examen	50	20	0,8	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11
Treball continuat	50	38	1,52	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11

Bibliografia

- Resampling methods: a practical guide to data Analysis. Phillip I. Good, 2006.
- The jackknife, the bootstrap and other resampling plans. Bradley Efron, 1982.

- Bootstrap methods and their application. A.C. Davison, D.V. Hinkley, 1997.
- "Learning Bayesian Networks" by R. E. Neapolitan, Prentice Hall Series in Artificial Intelligence, 2004.
- "Probabilistic Methods for Bioinformatics with an Introduction to Bayesian Networks" by R. E. Neapolitan, Elsevier, 2009.

Programari

Es farà servir el programari R.