

Guia docent

270411 - ME - Modelització Estadística

Última modificació: 10/07/2024

Unitat responsable: Facultat d'Informàtica de Barcelona
Unitat que imparteix: 715 - EIO - Departament d'Estadística i Investigació Operativa.
Titulació: GRAU EN INTEL·LIGÈNCIA ARTIFICIAL (Pla 2021). (Assignatura obligatòria).
Curs: 2024 **Crèdits ECTS:** 6.0 **Idiomes:** Català

PROFESSORAT

Professorat responsable: JORDI CORTÉS MARTÍNEZ
Altres: Primer quadrimestre:
DANTE CONTI - 11
JORDI CORTÉS MARTÍNEZ - 11, 12

CAPACITATS PRÈVIES

Introducció a l'Estadística
Coneixements de teoria de la probabilitat
Coneixements d'inferència estadística
Coneixements de models estadístics simples
Coneixements de visualització de dades
Coneixements bàsics de programació
Coneixements bàsics d'R
Àlgebra

REQUISITS

- Pre-requisit IE-GIA

COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

Específiques:

CE01. Resoldre els problemes matemàtics que puguin plantejar-se en l'àmbit de la intel·ligència artificial. Aplicar els coneixements sobre: àlgebra, càlcul diferencial i integral i mètodes numèrics; estadística i optimització.
CE09. Concebre, dissenyar i integrar sistemes d'anàlisi intel·ligent de dades amb aplicació en entorns de producció i de serveis.
CE20. Triar i emprar Tècniques de Modelització estadística i anàlisi de dades, avaluant la qualitat dels models, validant-i interpretant.

Genèriques:

CG2. Utilitzar els coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics de el futur.
CG4. Raonar, analitzant la realitat i dissenyant algoritmes i formulacions que la modelin. Identificar problemes i construir solucions algorísmiques o matemàtiques vàlides, eventualment noves, integrant el coneixement multidisciplinari necessari, valorant diferents alternatives amb esperit crític, justificant les decisions preses, interpretant i sintetitzant els resultats en el context de l' domini d'aplicació i establint generalitzacions metodològiques a partir de aplicacions concretes.
CG8. Observar un exercici ètic de la professió en totes les seves facetes, aplicant criteris ètics en el disseny de sistemes, algoritmes, experiments, utilització de dades, d'acord amb els sistemes ètics recomanats pels organismes nacionals i internacionals, amb especial èmfasi en seguretat, robustesa, privacitat, transparència, traçabilitat, prevenció de biaixos (de raça, gènere, religió, territori, etc.) i respecte als drets humans.

Transversals:

CT3. Comunicació eficaç oral i escrita. Comunicar-se de forma oral i escrita amb altres persones sobre els resultats de l'aprenentatge, de l'elaboració del pensament i de la presa de decisions; participar en debats sobre temes de la pròpia especialitat.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

CT8. Perspectiva de gènere. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.

Bàsiques:

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB4. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.

METODOLOGIES DOCENTS

L'assignatura consta de dues hores de teoria i dues de laboratori per setmana

En les classes de teoria es practicarà sempre que sigui possible l'esquema de classe invertida.

Al web de l'assignatura hi haurà el calendari de l'assignatura i els materials a portar preparats abans de cada classe. Es recorrerà a l'esquema de classe magistral puntualment quan el professor necessiti aclarir conceptes complexos que no han quedat clars amb els materials distribuïts prèviament a la classe. La classe de teoria es dedicarà fonamentalment a la presentació de casos i al desenvolupament d'activitats interactives amb els estudiants com la discussió dels casos, el desenvolupament de problemes o la realització de qüestionaris curts puntuals.

Els estudiants realitzaran per grups grans un treball pràctic amb dades que buscaran ells mateixos i que compliran certes característiques fixades pel professorat. Amb aquestes dades cada equip realitzarà les sessions de pràctiques, cada setmana aplicant les tècniques del tema treballat a la sessió de teoria. El professor farà seguiment setmanal de tots els equips de treball en les sessions de laboratori

Al final de curs els equips presentaran els seus resultats en una sessió de posta en comú on es debatrà conjuntament sobre tots els projectes

OBJECTIUS D'APRENTATGE DE L'ASSIGNATURA

1. Dissenyar jocs de proves i d'entrenament solvents i orientats a objectius
2. Identificar quin model predictiu és adequat per un problema concret i unes dades concretes
3. Construir i interpretar models vàlids per l'evolució temporal d'una variable numèrica
4. Identificar classes en un conjunt de dades i saber-les validar i interpretar conceptualment
5. Caracteritzar les relacions multivariants en un conjunt de dades amb tècniques d'anàlisi factorial
6. Poder fer l'anàlisi bàsica no supervisada d'una base de dades textual amb tècniques bàsiques de topic modelling i anàlisi multivariant per dades textuais
7. Saber construir i validar el model adequat per una situació real nova
8. Saber integrar els continguts dels diferents temes d'aquest curs i els previs en una solució global per un problema complex
9. Saber planificar a llarg termini la modelització d'un problema real complex i resoldre'l al llarg del curs en equip

HORES TOTALES DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores aprenentatge autònom	90,0	60.00
Hores grup petit	30,0	20.00
Hores grup gran	30,0	20.00

Dedicació total: 150 h

CONTINGUTS

Models lineals generalitzats

Descripció:

Introducció als conceptes de models lineals generalitzats. Models logístics

Disseny d'experiments

Descripció:

Dissenys 2k complets i fraccionals. Anàlisi de sensibilitat i explicabilitat dels models. Identificació d'efectes principals i interaccions. Disseny de conjunts de dades d'entrenament per aprenentatge automàtic. Disseny de jocs de test per validació de models de dades

Sèries temporals

Descripció:

Introducció als processos estocàstics. Sèrie cronològica vs Sèrie temporal

Metodologia Box-Jenkins

Principals models de sèries temporals: MA, AR , ARIMA, SARIMA (concepte i estudi de casos)

Clustering

Descripció:

Introducció. Principal models de classificació. Distàncies.

Profiling

Descripció:

Descripció de les classificacions a partir de l'estudi de significativitat de variables

Anàlisi factorial

Descripció:

Mètodes de reducció de dimensionalitat

ACTIVITATS

Treball en equip

Descripció:

Els alumnes s'organitzen en grups i busquen unes dades reals que compleixen certs requisits marcats pel professor. Les utilitzen per anar aplicant les tècniques i metodologies que es vegin al llarg del curs. Al final presenten un informe amb els resultats i fan una presentació oral amb els resultats més rellevants de l'estudi.

Objectius específics:

1, 2, 3, 4, 5, 6, 7, 8, 9

Competències relacionades:

CB4. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CE01. Resoldre els problemes matemàtics que puguin plantejar-se en l'àmbit de la intel·ligència artificial. Aplicar els coneixements sobre: àlgebra, càlcul diferencial i integral i mètodes numèrics; estadística i optimització.

CE20. Triar i emprar Tècniques de Modelització estadística i anàlisi de dades, avaluant la qualitat dels models, validant-i interpretant.

CE09. Concebre, dissenyar i integrar sistemes d'anàlisi intel·ligent de dades amb aplicació en entorns de producció i de serveis.

CG2. Utilitzar els coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics de el futur.

CG4. Raonar, analitzant la realitat i dissenyant algorismes i formulacions que la modelin. Identificar problemes i construir solucions algorísmiques o matemàtiques vàlides, eventualment noves, integrant el coneixement multidisciplinari necessari, valorant diferents alternatives amb esperit crític, justificant les decisions preses, interpretant i sintetitzant els resultats en el context de l'àmbit d'aplicació i establint generalitzacions metodològiques a partir de aplicacions concretes.

CG8. Observar un exercici ètic de la professió en totes les seves facetes, aplicant criteris ètics en el disseny de sistemes, algorismes, experiments, utilització de dades, d'acord amb els sistemes ètics recomanats pels organismes nacionals i internacionals, amb especial èmfasi en seguretat, robustesa, privacitat, transparència, traçabilitat, prevenció de biaixos (de raça, gènere, religió, territori, etc.) i respecte als drets humans.

CT8. Perspectiva de gènere. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.

CT3. Comunicació eficaç oral i escrita. Comunicar-se de forma oral i escrita amb altres persones sobre els resultats de l'aprenentatge, de l'elaboració del pensament i de la presa de decisions; participar en debats sobre temes de la pròpia especialitat.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

Dedicació: 68h

Aprenentatge autònom: 41h

Grup petit/Laboratori: 27h

Quiz 1

Descripció:

Durant el curs es realitzaran proves de resposta curta per fixar peces d'aprenentatge. Es farà al final de certes classes de laboratori

Objectius específics:

2

Competències relacionades:

CE01. Resoldre els problemes matemàtics que puguin plantejar-se en l'àmbit de la intel·ligència artificial. Aplicar els coneixements sobre: àlgebra, càlcul diferencial i integral i mètodes numèrics; estadística i optimització.

CE20. Triar i emprar Tècniques de Modelització estadística i anàlisi de dades, avaluant la qualitat dels models, validant-i interpretant.

CE09. Concebre, dissenyar i integrar sistemes d'anàlisi intel·ligent de dades amb aplicació en entorns de producció i de serveis.

CG4. Raonar, analitzant la realitat i dissenyant algoritmes i formulacions que la modelin. Identificar problemes i construir solucions algorísmiques o matemàtiques vàlides, eventualment noves, integrant el coneixement multidisciplinari necessari, valorant diferents alternatives amb esperit crític, justificant les decisions preses, interpretant i sintetitzant els resultats en el context de l'àmbit d'aplicació i establint generalitzacions metodològiques a partir de aplicacions concretes.

Quiz 2

Descripció:

Durant el curs es realitzaran proves de resposta curta per fixar peces d'aprenentatge. Es farà al final de certes classes de laboratori

Objectius específics:

2, 3

Competències relacionades:

CE01. Resoldre els problemes matemàtics que puguin plantejar-se en l'àmbit de la intel·ligència artificial. Aplicar els coneixements sobre: àlgebra, càlcul diferencial i integral i mètodes numèrics; estadística i optimització.

CE20. Triar i emprar Tècniques de Modelització estadística i anàlisi de dades, avaluant la qualitat dels models, validant-i interpretant.

CE09. Concebre, dissenyar i integrar sistemes d'anàlisi intel·ligent de dades amb aplicació en entorns de producció i de serveis.

CG4. Raonar, analitzant la realitat i dissenyant algoritmes i formulacions que la modelin. Identificar problemes i construir solucions algorísmiques o matemàtiques vàlides, eventualment noves, integrant el coneixement multidisciplinari necessari, valorant diferents alternatives amb esperit crític, justificant les decisions preses, interpretant i sintetitzant els resultats en el context de l'àmbit d'aplicació i establint generalitzacions metodològiques a partir de aplicacions concretes.

CT3. Comunicació eficaç oral i escrita. Comunicar-se de forma oral i escrita amb altres persones sobre els resultats de l'aprenentatge, de l'elaboració del pensament i de la presa de decisions; participar en debats sobre temes de la pròpia especialitat.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

Quiz 3

Descripció:

Durant el curs es realitzaran proves de resposta curta per fixar peces d'aprenentatge. Es farà al final de certes classes de laboratori

Objectius específics:

2, 3

Competències relacionades:

CE01. Resoldre els problemes matemàtics que puguin plantejar-se en l'àmbit de la intel·ligència artificial. Aplicar els coneixements sobre: àlgebra, càlcul diferencial i integral i mètodes numèrics; estadística i optimització.

CE20. Triar i emprar Tècniques de Modelització estadística i anàlisi de dades, avaluant la qualitat dels models, validant-i interpretant.

CE09. Concebre, dissenyar i integrar sistemes d'anàlisi intel·ligent de dades amb aplicació en entorns de producció i de serveis.

CG4. Raonar, analitzant la realitat i dissenyant algorismes i formulacions que la modelin. Identificar problemes i construir solucions algorísmiques o matemàtiques vàlides, eventualment noves, integrant el coneixement multidisciplinari necessari, valorant diferents alternatives amb esperit crític, justificant les decisions preses, interpretant i sintetitzant els resultats en el context de l' domini d'aplicació i establint generalitzacions metodològiques a partir de aplicacions concretes.

CT3. Comunicació eficaç oral i escrita. Comunicar-se de forma oral i escrita amb altres persones sobre els resultats de l'aprenentatge, de l'elaboració del pensament i de la presa de decisions; participar en debats sobre temes de la pròpia especialitat.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

Presentació inicial de la pràctica

Descripció:

Presentació inicial de la pràctica

Objectius específics:

1, 2, 3, 4, 5, 6, 7, 8, 9

Competències relacionades:

CB4. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CE01. Resoldre els problemes matemàtics que puguin plantejar-se en l'àmbit de la intel·ligència artificial. Aplicar els coneixements sobre: àlgebra, càlcul diferencial i integral i mètodes numèrics; estadística i optimització.

CE20. Triar i emprar Tècniques de Modelització estadística i anàlisi de dades, avaluant la qualitat dels models, validant-i interpretant.

CE09. Concebre, dissenyar i integrar sistemes d'anàlisi intel·ligent de dades amb aplicació en entorns de producció i de serveis.

CG2. Utilitzar els coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics de el futur.

CG4. Raonar, analitzant la realitat i dissenyant algorismes i formulacions que la modelin. Identificar problemes i construir solucions algorísmiques o matemàtiques vàlides, eventualment noves, integrant el coneixement multidisciplinari necessari, valorant diferents alternatives amb esperit crític, justificant les decisions preses, interpretant i sintetitzant els resultats en el context de l'àmbit d'aplicació i establint generalitzacions metodològiques a partir de aplicacions concretes.

CG8. Observar un exercici ètic de la professió en totes les seves facetes, aplicant criteris ètics en el disseny de sistemes, algorismes, experiments, utilització de dades, d'acord amb els sistemes ètics recomanats pels organismes nacionals i internacionals, amb especial èmfasi en seguretat, robustesa, privacitat, transparència, traçabilitat, prevenció de biaixos (de raça, gènere, religió, territori, etc.) i respecte als drets humans.

CT8. Perspectiva de gènere. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.

CT3. Comunicació eficaç oral i escrita. Comunicar-se de forma oral i escrita amb altres persones sobre els resultats de l'aprenentatge, de l'elaboració del pensament i de la presa de decisions; participar en debats sobre temes de la pròpia especialitat.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

Dedicació: 12h

Aprenentatge autònom: 10h

Grup petit/Laboratori: 2h

Presentació final pràctica

Descripció:

Presentació final pràctica

Dedicació: 10h

Aprenentatge autònom: 10h

Classes de teoria del temari de l'assignatura

Descripció:

classes de teoria del temari de l'assignatura

Objectius específics:

2, 3, 4, 5, 6, 7

Competències relacionades:

CE01. Resoldre els problemes matemàtics que puguin plantejar-se en l'àmbit de la intel·ligència artificial. Aplicar els coneixements sobre: àlgebra, càlcul diferencial i integral i mètodes numèrics; estadística i optimització.

CE20. Triar i emprar Tècniques de Modelització estadística i anàlisi de dades, avaluant la qualitat dels models, validant-i interpretant.

CE09. Concebre, dissenyar i integrar sistemes d'anàlisi intel·ligent de dades amb aplicació en entorns de producció i de serveis.

CG2. Utilitzar els coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics de el futur.

CG4. Raonar, analitzant la realitat i dissenyant algorismes i formulacions que la modelin. Identificar problemes i construir solucions algorísmiques o matemàtiques vàlides, eventualment noves, integrant el coneixement multidisciplinari necessari, valorant diferents alternatives amb esperit crític, justificant les decisions preses, interpretant i sintetitzant els resultats en el context de l'domini d'aplicació i establint generalitzacions metodològiques a partir de aplicacions concretes.

CT3. Comunicació eficaç oral i escrita. Comunicar-se de forma oral i escrita amb altres persones sobre els resultats de l'aprenentatge, de l'elaboració del pensament i de la presa de decisions; participar en debats sobre temes de la pròpia especialitat.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

Dedicació: 60h

Aprenentatge autònom: 30h

Grup gran/Teoria: 30h

SISTEMA DE QUALIFICACIÓ

(T) Treball en equip realitzat al llarg del curs 20%

(O) Prova oral de control de coneixements 10% (discussió amb el professorat en la presentació oral dels treballs en equip)

(WT) Qualitat i rendiment de l'equip de treball (TG). 10%

(C) Comunicació oral i escrita 10%

(E) Ètica de l'equip de treball i del treball pròpiament dit 10%

(G) Perspectiva de gènere de l'equip i del treball 10%

(A) Assistència i participació en classes i laboratoris (AP). 10%

(Q) 3 Quiz al llarg del curs 20%

$$N=0,2T+0,1*O+0,1*WT+0,1*C+0,1*E+0,1*G+0,1*A+0,2*Q$$

$$Q=(Q1+Q2+Q3)/3$$

BIBLIOGRAFIA

Bàsica:

- Bruce, Peter; Bruce, Andrew; Gedeck, Peter. Practical statistics for data scientists: 50+ essential concepts using R and Python. 2nd ed. O'Reilly, [2020]. ISBN 9781492072942.
- Maindonald, J. H; Braun, John. Data analysis and graphics using R : an example-based approach. 3rd ed. Cambridge University, 2010. ISBN 9780521762939.
- Hyndman, R.J.; Athanasopoulos, G. Forecasting: principles and practice. 3rd print edition. [Melbourne]: O Texts, 2021. ISBN 9780987507136.
- Nielsen, Aileen. Practical time series analysis: prediction with statistics and machine learning. Sebastopol: O'Reilly Media, Inc, 2019. ISBN 9781492041658.