Universitat Autònoma de Barcelona

2019/2020

Models Lineals 2

Codi: 104861
Crèdits: 6

Titulació	Tipus	Curs	Semestre
2503852 Estadística Aplicada	OB	3	1

Professor/a de contacte

Nom: Natalia Isabel Vilor Tejedor

Correu electrònic: NataliaIsabel.Vilor@uab.cat

Utilització d'idiomes a l'assignatura

Llengua vehicular majoritària: català (cat)

Grup íntegre en anglès: No

Grup íntegre en català: Sí

Grup íntegre en espanyol: No

Prerequisits

Coneixements bàsics d'estadística descriptiva i inferencial. És imprescindible un curs previ de Models Lineals.

Objectius

L'objectiu del curs és l'aprenentatge supervisat amb un enfocament en els mètodes de regressió i classificació. El pla d'estudis inclou regressió lineal i polinòmica, regressió logística i anàlisi discriminant lineal; selecció de models i mètodes de regularització (ridge i LASSO); models no lineals, splines, models additius generalitzats; mètodes basats en arbres i Random forest.

Competències

- Analitzar dades mitjançant laplicació de mètodes i tècniques estadístiques, treballant amb dades de diverses tipologies.
- Aplicar lesperit crític i el rigor per validar o refutar arguments tant propis com daltres persones.
- Avaluar de manera crítica i amb criteris de qualitat el treball realitzat.
- Dissenyar un estudi estadístic o de recerca operativa per a la resolució dun problema real.
- Formular hipòtesis estadístiques i desenvolupar estratègies per confirmar-les o refutar-les.
- Interpretar resultats, extreure conclusions i elaborar informes tècnics en el camp de lestadística.
- Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.
- Que els estudiants sàpiguen aplicar els coneixements propis a la seva feina o vocació d'una manera professional i tinguin les competències que se solen demostrar per mitjà de l'elaboració i la defensa d'arguments i la resolució de problemes dins de la seva àrea d'estudi.
- Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins de la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes destacats d'índole social, científica o ètica.
- Resumir i descobrir patrons de comportament en lexploració de les dades.
- Seleccionar els models o tècniques estadístiques per aplicar-los en estudis i problemes reals, així com conèixer-ne les eines de validació.
- Seleccionar i aplicar procediments més apropiats per a la modelització estadística i lanàlisi de dades complexes.
- Utilitzar correctament un ampli espectre del programari i llenguatges de programació estadístiques, escollint el més apropiat per a cada anàlisi i ser capaç dadaptar-lo a noves necessitats.

- Utilitzar eficaçment la bibliografia i els recursos electrònics per obtenir informació.

Resultats d'aprenentatge

1. Analitzar dades mitjançant el model de regressió lineal.
2. Analitzar dades mitjançant el model lineal generalitzat.
3. Analitzar dades mitjançant tècniques d'inferència utilitzant programari estadístic.
4. Analitzar els residus d'un model estadístic.
5. Aplicar l'esperit crític i el rigor per validar o refutar arguments, tant propis com d'altres.
6. Avaluar de manera crítica i amb criteris de qualitat la feina feta.
7. Comparar el grau d'ajustament entre diversos models estadístics.
8. Detectar i preveure interaccions entre variables explicatives.
9. Detectar i tractar la colinealitat entre variables explicatives.
10. Establir les hipòtesis experimentals de la modelització.
11. Extreure conclusions de l'adequació dels models amb la utilització i la interpretació correcta d'indicadors i gràfics.
12. Fer servir gràfics de visualització de l'ajustament i de l'adequació del model.
13. Identificar distribucions de les respostes amb l'anàlisi de residus.
14. Identificar fonts de biaix en l'obtenció de la informació.
15. Identificar les etapes en els problemes de modelització.
16. Identificar les suposicions estadístiques associades a cada procediment.
17. Identificar les variables resposta, explicatives i de control.
18. Mesurar el grau d'ajustament d'un model estadístic.
19. Modificar lleugerament el programari existent si el model estadístic proposat ho requereix.
20. Predir respostes, comparar grups (valor causal) i identificar factors rellevants.
21. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.
22. Que els estudiants sàpiguen aplicar els coneixements propis a la seva feina o vocació d'una manera professional i tinguin les competències que se solen demostrar per mitjà de l'elaboració i la defensa d'arguments i la resolució de problemes dins de la seva àrea d'estudi.
23. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins de la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes destacats d'índole social, científica o ètica.
24. Seleccionar les variables explicatives rellevants.
25. Sintetitzar i interpretar els resultats dels models lineals clàssics, generalitzats i no lineals en funció de l'objectiu de l'estudi.
26. Utilitzar eficaçment bibliografia i recursos electrònics per obtenir informació.
27. Utilitzar la regressió logística per resoldre problemes de classificació.
28. Utilitzar programari estadístic divers per ajustar i validar models lineals i les seves generalitzacions.
29. Validar els models utilitzats mitjançant tècniques d'inferència adequades.

Continguts

1. Regressió lineal
Regressió lineal simple
Regressió lineal múltiple
Ampliació dels models lineals.

2. Mètodes de Classificació
Introducció dels mètodes de classificació.
Regressió logística. El model logístic. Estimació dels coeficients de regressió. Prediccions.
Regressió logística múltiple
Anàlisi discriminant lineal.

3. Selecció de models i Regularització.
Selecció de subconjunts: Selecció per passos, i selecció de model òptim.
Mètodes de regularització: Ridge i regressió LASSO. Selecció del paràmetre d'afinació dels mètodes de

regularització.
Mètodes de reducció de dimensionalitat: Anàlisi de components principals (PCA) i mínims quadrats parcials (PLS).

4. Més enllà de la linealitat
Regressió polinomica
Models additius generalitzats (GAM)

5. Mètodes basats en arbres
Arbres de decisió: arbres de regressió i arbres de classificació
Random Forest

Metodologia

L'assignatura té, al llarg del quadrimestre, dues hores setmanals de teoria i problemes més dues hores setmanals de pràctiques. El material del curs (apunts de teoria, llistes de problemes i enunciats de pràctiques) estarà disponible al campus virtual, de manera progressiva al llarg del curs.

Activitats formatives

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Pràctiques amb ordinador	50	2	1, 2, 3, 5, 6, 10, 12, 15, 16, 19, 23, 25, 26, 27
Teoria	50	2	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29
Tipus: Supervisades			
problemes/exercicis per resoldre	16	0,64	3, 5, 6, 11, 23, 25, 26, 28
Tipus: Autònomes			
Preparació Examens	10	0,4	3, 5, 6, 16, 21, 25

Avaluació

PR: Lliurament de les pràctiques que s'indiquin al llarg del curs. Valoració de PR: 2 punts sobre 10. Aquesta part NO és recuperable.

P1: Prova parcial 1 (teoria, problemes i pràctiques). Valoració de P1: 3 punts sobre 10.

P2: Prova parcial 2 (teoria, problemes i pràctiques). Valoració de P2: 5 punts sobre 10.

Per a que un examen es tingui en compte a l'hora de calcular la nota final és necessari obtenir una nota mínima de 4 punts. La Nota final serà: Nota Final = PR + P1 + P2. Al gener es té un dia assignat per fer una prova final, PF, que permet només la recuperació de P1 i P2 (8 punts sobre 10) pels alumnes que no hagin aprovat per curs. Cas que es faci la prova de síntesi, la nota final serà: Nota Final = PR + min(PF, 5).

Atenció: "Sense perjudici d'altres mesures disciplinàries que s'estimin oportunes, i d'acord amb la normativa acadèmica vigent, es qualificaran amb un zero les irregularitats comeses per l'estudiant que puguin conduir a

una variació de la qualificació d'un acte d'avaluació. Per tant, plagiar, copiar o deixar copiar una pràctica o qualsevol altra activitat d'avaluació implicarà suspendre-la amb un zero i no es podrà recuperar en el mateix curs acadèmic. Si aquesta activitat té una nota mínima associada, aleshores l'assignatura quedarà suspesa."

Activitats d'avaluació

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Parcial 1	30%	4	0,16	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29
Parcial 2	50%	4	0,16	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29
Pràctiques	20%	16	0,64	3, 5, 6, 10, 12, 15, 16, 19, 21, 22, 23, 25, 26, 28

Bibliografia

Montgomery, D. Peck, A. Vining, G.; *Introduction to Linear Regression Analysis*. Wiley. 2001.

Christopher Hay-Jahans; *An R Companion to Linear Statistical Models*. Chapman and Hall, 2012.

John Fox and Sandord Weisberg; *An R Companion to Applied Regression*, 2nd edition, Sage Publications, 2011.

Daniel Peña; *Regresión y diseño de Experimentos*, Alianza Editorial (Manuales de Ciencias Sociales), 2002.

Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani; An Introduction to Statistical Learning, Springer texts in Statistics, 2013.