# Guia docent
# 295601 - AB - Aprenentatge Bioestadístic

**Última modificació:** 04/06/2026

**Unitat responsable:** Escola d'Enginyeria de Barcelona Est
**Unitat que imparteix:** 749 - MAT - Departament de Matemàtiques.

**Titulació:** **Curs:** 2026 **Crèdits ECTS:** 6.0
**Idiomes:** Castellà

## PROFESSORAT

**Professorat responsable:**

**Altres:**

## CAPACITATS PRÈVIES

Durant l'etapa de formació de pregrau has vist alguns fonaments de la programació, així com certs conceptes matemàtics. En aquest curs podràs apreciar algunes utilitats, específicament:

1. Python: lús de variables, tipus de dades, declaracions, expressions, operadors i precedència.

2. Conceptes matemàtics: Al llarg del curs veureu l'aplicació d'alguns conceptes d'estadística, càlcul i àlgebra.
- Estadística: comprensió de mesures estadístiques bàsiques (mitjana, mitjana, desviació estàndard), probabilitat condicional i prova d'hipòtesis.
- Càlcul: Conceptes fonamentals de diferenciació i integració, principalment pel que fa a problemes d'optimització en ciència de dades.
- Àlgebra: Àlgebra lineal bàsica, incloses operacions matricials i espais vectorials, ja que són crucials per comprendre les estructures de dades, la complexitat algorísmica i els models daprenentatge automàtic.

Si bé una ferma comprensió d'aquests conceptes matemàtics serà útil per comprendre el material del curs, l'exercici detallat d'aquestes habilitats específiques no serà part de l'avaluació del curs.

## REQUISITS

Aquesta assignatura no pressuposa cap requisit previ. Si cal tenir un gran interès a comprendre i aplicar mètodes per a l'anàlisi de dades biomèdiques. També són essencials l'entusiasme i la voluntat per abordar els conceptes de les noves tecnologies per solucionar problemes desafiadors.

## METODOLOGIES DOCENTS

Aquest curs adopta un model daprenentatge integrat, on els conceptes teòrics i laplicació pràctica sentrellacen perfectament durant quatre hores setmanals. Cada sessió es duu a terme dins un laboratori de computació. Després de la introducció de cada concepte teòric, se n'explorarà immediatament la implementació pràctica a través de la programació a Python promovent un entorn d'aprenentatge actiu (teoria de l'aprenentatge constructivista). A l'inici de cada classe, es proporcionaran tots els materials de Python. D'aquesta manera, l'atenció se centra a aprendre i aplicar conceptes nous.

A més, el curs aprofita fins a cert punt el model d'aula invertida i assigna el 60% de la trajectòria d'aprenentatge a l'estudi autònom. S'espera que els estudiants participin en un aprenentatge autodirigit, tot fomentant la investigació independent i la consolidació de coneixements. Aquest component està dissenyat per cultivar un entorn centrat en l'alumne, permetent-los assumir la responsabilitat del seu aprenentatge i desenvolupar habilitats d'autoregulació, fonamentals en el camp de la gestió de dades. L'estratègia pedagògica distribueix les ponderacions de l'avaluació en conseqüència:

- Coneixements basats en la teoria: 20%
- Tasques i projectes de laboratori d'informàtica: 20%
- Aprenentatge autònom: 60%

Com a marc davaluació, es desenvoluparan tres treballs destinats a fomentar el compromís i laplicació continuada dels conceptes, garantint una experiència daprenentatge holística que sigui alhora rigorosa i contextualment rellevant.

## OBJECTIUS D'APRENENTATGE DE L'ASSIGNATURA

Els objectius estan dissenyats per fomentar una experiència d'aprenentatge holístic que combini coneixements tècnics amb habilitats socials essencials com el pensament crític, la col·laboració i la investigació independent:

- Obtenir una comprensió sòlida dels principis fonamentals de laprenentatge automàtic pel que fa al camp de lenginyeria biomèdica.

- Aprendre a articular i emmarcar desafiaments en diversos contextos biomèdics.

- Adquirir una comprensió integral d'un espectre dels algorismes per al processament i l'anàlisi de dades i d'imatges, inclosos els beneficis i les limitacions possibles.

- Aplicar aquests coneixements nous per abordar i resoldre problemes biomèdics de complexitat moderada, fomentant el domini de la codificació Python en contextos d'anàlisi de dades i d'imatges.

- Desenvolupar la capacitat davaluació crítica dels resultats prenent decisions metodològiques ben informades.

- Fomentar el desenvolupament d'habilitats d'aprenentatge autònom, capacitant els estudiants per navegar i dominar diferents camps de manera independent.

- Conrear un entorn d'aprenentatge col·laboratiu on es valori i es promogui l'intercanvi d'idees i la participació constructiva del grup.

- Inculcar una mentalitat impulsada per la investigació, impulsant els estudiants a explorar més enllà dels materials proporcionats i sintetitzar informació de diverses fonts.

- Dotar els estudiants de la capacitat de comparar críticament diverses tècniques i recomanar les més adequades per a determinades qüestions biomèdiques.

## HORES TOTALS DE DEDICACIÓ DE L'ESTUDIANTAT

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

**Dedicació total:** 150 h

## CONTINGUTS

### 1. Introducció

**Descripció:**
Les sessions estan dissenyades per proporcionar una visió general del desenvolupament històric dels processos industrials i establir les bases per comprendre el paper del processament automatitzat a l'anàlisi contemporani de dades i imatges dins de les aplicacions biomèdiques. S'examinarà la importància de l'adquisició de dades, el control de processos i el monitoratge, juntament amb diversos models utilitzats per a l'anàlisi i la interpretació de dades.

**Objectius específics:**
Familiaritzar-se amb els fonaments conceptuals de l'aprenentatge automàtic i la seva aplicació a l'enginyeria biomèdica. Això inclou comprendre la importància de les dades per impulsar millores en els processos, els conceptes bàsics de l'adquisició i la instrumentació de dades i una introducció als tipus d'algorismes que es trobaran al llarg del curs.

**Activitats vinculades:**
Classes teòriques: Presentació del curs i Introducció.
Sessions de laboratori: Introducció a Python, Matplotlib, NumPy, Pandes i exploració de dades.

**Dedicació:** 10h
Grup gran/Teoria: 2h
Grup petit/Laboratori: 2h
Aprenentatge autònom: 6h

### 2. Aprenentatge supervisat

**Descripció:**
L'aprenentatge supervisat és un tipus d'aprenentatge automàtic on un algorisme aprèn a partir de dades d'entrenament etiquetades i les utilitza per fer prediccions o decisions sense intervenció humana. Es presentaran els fonaments bàsics de regressió i classificació. S'estudiaran les principals tècniques emprades per al model predictiu dins el domini de la ciència de dades.
La regressió abraça la metodologia estadística per modelar la relació entre una variable dependent quantitativa i una o més variables independents. Es posarà èmfasi en la comprensió de la teoria, els supòsits subjacents i la implementació pràctica de models de regressió en el context de l'anàlisi de dades biomèdiques.
A les sessions de classificació s'ensenyaran algorismes destinats a la categorització de dades a classes predeterminades. S'hi examinaran una varietat de tècniques de classificació. Els exercicis pràctics facilitaran la comprensió de laplicació daquests mètodes en la classificació de conjunts de dades biomèdiques.

**Objectius específics:**
Dotar les habilitats necessàries per a construir, interpretar i avaluar models predictius utilitzant regressió i algoritmes de classificació. Se centrarà en l'aplicació d'aquests models a conjunts de dades biomèdiques, la interpretació dels resultats amb ull crític i la comprensió de les implicacions dels models en escenaris biomèdics del món real. Aquest coneixement servirà com a trampolí per a temes més avançats i la presa de decisions.

**Activitats vinculades:**
Classes teòriques i sessions de laboratori: Regressió Lineal, Bayes-Naives, Regressió Logística, Anàlisi discriminant, K veïns propers (KNN), Màquines de Suport Vectorial (SVM). Índexs per a l'avaluació de models i la matriu de confusió.
Treball 1.

**Dedicació:** 40h
Grup gran/Teoria: 8h
Grup petit/Laboratori: 8h
Aprenentatge autònom: 24h

## 3. Remostreig i aprenentatge en conjunt

**Descripció:**
Aquesta part del curs presentarà metodologies avançades essencials per millorar el rendiment del model predictiu. S'exploren estratègies de remostreig (resampling), com la validació creuada (cross-validation), bootstrapping i els conjunts de validació, pel seu paper fonamental en l'avaluació, la selecció i l'estimació de la precisió dels models. Els mètodes d'aprenentatge conjunt (Ensemble Learning) abasten enfocaments integradors com ara embutxacat (bagging), impuls (boosting) i apilament (stacking), que combinen múltiples models per superar la precisió i l'estabilitat de models predictius singulars.

**Objectius específics:**
Impartir una comprensió integral de les tècniques de remostreig per a una validació sòlida i mètodes daprenentatge en conjunt per augmentar lefectivitat del model. L'aplicació d'aquestes metodologies es demostrarà utilitzant Python en conjunts de dades biomèdiques, proveint amb les habilitats necessàries per mitigar el sobreajustament, la variació i el biaix en el desenvolupament de models. Dominar aquestes tècniques és crucial per desenvolupar models avançats i generalitzables a l'anàlisi de dades.

**Activitats vinculades:**
Classes teòriques i sessions de laboratori: validació creuada, Bootstrapping, Ensemble Learning, bagging, boosting, stacking, arbres de decisió.

**Dedicació:** 20h
Grup gran/Teoria: 4h
Grup petit/Laboratori: 4h
Aprenentatge autònom: 12h

## 4. Aprenentatge no supervisat

**Descripció:**
En aquestes sessions s'aprofundirà l'exploració d'estructures de dades sense resultats preetiquetats. Es prestarà una atenció especial a les tècniques d'agrupació, com les k-mitges i l'agrupació jeràrquica, que agrupen punts de dades en funció de mètriques de similitud. S'examinarà la utilitat dels mètodes de reducció de dimensionalitat, com és l'anàlisi de components principals (PCA) per simplificar conjunts de dades complexes i així revelar patrons intrínsecs.

**Objectius específics:**
Aplicar tècniques d'aprenentatge no supervisades de manera efectiva per destriar estructures subjacents en dades biomèdiques. Això inclou identificar agrupacions naturals dins de conjunts de dades, reduir la complexitat de les dades d'alta dimensió per a una visualització i interpretació millor i preparar les dades per a anàlisis posteriors. A través d'això, s'adquirirà competència per revelar patrons i tendències.

**Activitats vinculades:**
Classes teòriques i sessions de laboratori: k-means, Density-Based Spatial Clustering of Applications with noise (DBSCAN), hierarchical clusterin, Principal Component Analysis (PCA).

Treball 2.

**Dedicació:** 20h
Grup gran/Teoria: 4h
Grup petit/Laboratori: 4h
Aprenentatge autònom: 12h

## 5. Introducció a l'aprenentatge profund

**Descripció:**
Aquesta secció del curs abordarà l'aprenentatge profund, posant especial èmfasi a les xarxes neuronals i les seves diverses arquitectures. S'introduiran conceptes fonamentals com les xarxes neuronals convolucionals (CNN) i les xarxes neuronals recurrents (RNN), a més de tècniques essencials com ara l'ajust de paràmetres i la regularització. S'explorarà l'ús pràctic de les xarxes profundes a l'enginyeria biomèdica, particularment a la classificació d'imatges i de documents, ia la predicció de sèries temporals. L'estudiantat aprendrà a implementar i entrenar models daprenentatge profund usant Python i biblioteques populars daprenentatge profund.

**Objectius específics:**
Impartir una comprensió sòlida de l'estructura i el funcionament de les xarxes neuronals artificials, permetent la implementació pràctica d'aquests models per resoldre problemes d'enginyeria biomèdica del món real. Els participants adquiriran les habilitats per dissenyar i entrenar arquitectures de xarxes neuronals, comprendran els seus avantatges i limitacions i aprofitaran aquest coneixement per desenvolupar solucions efectives per a desafiaments complexos danàlisi de dades biomèdiques. Aquesta fundació té com a objectiu aplanar el camí per a estudis avançats en aprenentatge automàtic i aplicacions innovadores dins del camp biomèdic.

**Activitats vinculades:**
Classes teòriques i sessions de laboratori: Estructura de xarxes neuronals (neurones, capes i funcions dactivació). Algorismes amb xarxes Feedforward and backpropagation. Diferents tipus d'arquitectures de xarxes neuronals com ara Perceptrons Multicapa (MLP), xarxes neuronals convolucionals (CNN) i xarxes neuronals recurrents (RNN). Estratègies per prevenir el sobreajustament, com a mètodes de regularització i abandonament (dropout).

**Dedicació:** 40h
Grup gran/Teoria: 8h
Grup petit/Laboratori: 24h
Aprenentatge autònom: 8h

**6. Introducció als Transformers**

**Descripció:**
Aquest mòdul final del curs introdueix l'arquitectura de xarxa neuronal profunda Transformers, iniciant pel mòdul d'atenció i estudiant els fonaments de l'arquitectura Transformer original de codificador i descodificador (encoder-decoder). S'exploraran les arquitectures bàsiques derivades del codificador com a BERT, i de la part descodificador com a GPT. També introduirà una modificació del Transformer per aplicar-lo en imatges biomèdiques, anomenada Transformadors de Visió (ViT). S'exploraran aquestes arquitectures avançades i es comprendran els principis fonamentals que n'impulsen l'acompliment i les formes innovadores en què s'apliquen a la resolució de problemes complexos. El mòdul també proporcionarà una exposició pràctica per implementar aquests algoritmes utilitzant marcs moderns daprenentatge profund i demostrar la seva aplicació.

**Objectius específics:**
Proveir una comprensió profunda de les arquitectures daprenentatge profund avançades i les seves aplicacions dotant de les habilitats necessàries no només per comprendre els fonaments teòrics daquests models sofisticats. El domini d'aquestes tècniques avançades és crucial per contribuir significativament al camp de l'anàlisi de dades biomèdiques, impulsant la innovació i ampliant els límits del que és possible amb la IA a l'atenció mèdica i la medicina.

**Activitats vinculades:**
Classes teòriques i pràctiques de laboratori per entendre i aplicar l'arquitectura dels Transformers, començant per l'estudi del mecanisme d'atenció i com s'utilitza a l'estructura encoder-decoder. S'analitzaran i posaran en pràctica arquitectures derivades del Transformer com a BERT per al processament del llenguatge natural i GPT per a generació de text, així com l'adaptació de Transformers al camp de la visió per ordinador amb els ViT. Aprendrem sobre la importància de l'atenció i com ha revolucionat l'anàlisi de dades seqüencials. Els estudiants tindran l'oportunitat d'entrenar models de Transformer en tasques de classificació d'imatges i comprensió del llenguatge natural utilitzant frameworks d'aprenentatge profund actuals. S'encoratjarà els estudiants a experimentar amb aquests models avançats per desenvolupar solucions innovadores a problemes d'anàlisi de dades biomèdiques.

Treball 3

**Dedicació:** 20h
Grup gran/Teoria: 4h
Grup petit/Laboratori: 4h
Aprenentatge autònom: 12h

## SISTEMA DE QUALIFICACIÓ

Qualsevol acte de frau acadèmic, plagi, ús o tinença a l'abast de mitjans no autoritzats en qualsevol activitat d'avaluació comportarà la qualificació de zero (0) en la prova o lliurament afectat. A més, d'acord amb la normativa de la Universitat, la possible derivació dels fets per a l'obertura d'un expedient disciplinari implicarà que l'assignatura quedi en l'estat provisional de "pendent d'avaluació" fins a la resolució de l'expedient. La gestió d'aquestes incidències es desplega d'acord amb el **Marc d'actuació per a la integritat acadèmica en l'avaluació de la UPC**.

Lavaluació es realitzarà de forma contínua durant el desenvolupament de cada mòdul mitjançant treballs pràctics.

## NORMES PER A LA REALITZACIÓ DE LES PROVES.

.

## BIBLIOGRAFIA

**Bàsica:**
- James, G.; Witten, D.; Hastie, T.; Tibshirani, R. An introduction to statistical learning with applications in R. Springer, 2013. ISBN 9781461471370.
- Bishop, Christopher M. Pattern recognition and machine learning. Springer, 2006. ISBN 9780387310732.
- Géron, Aurélien. Hands-on machine learning with scikit-learn & tensorflow : concepts, tools, and techniques to build intelligent systems [en línia]. Sebastopol, CA: O'Reilly Media, Inc, 2017 [Consulta: 15/04/2020]. Disponible a: https://ebookcentral.proquest.com/lib/upcatalunya-ebooks/detail.action?docID=4822582. ISBN 9781491962268.
- Raschka, Sebastian. Python machine learning : machine learning and deep learning with Python, scikit-learn, and TensorFlow [en línia]. 2nd ed. Birmingham, UK: Packt Publishing Ltd, 2017 [Consulta: 14/04/2020]. Disponible a: https://ebookcentral.proquest.com/lib/upcatalunya-ebooks/detail.action?docID=5050960. ISBN 9781787126022|.

**Complementària:**
- Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome. The Elements of statistical learning : data mining, inference, and prediction [en línia]. 2nd ed. New York, NY: Springer Series in Statistics, 2001 [Consulta: 27/08/2018]. Disponible a: http://dx.doi.org/10.1007/978-0-387-84858-7. ISBN 9780387848587.

## RECURSOS

**Altres recursos:**
Material disponible a ATENEA per part dels responsables del curs.