

Guia docent

295626 - 295MB122 - Anàlisi de Dades i Aprenentatge Automàtic (Machine Learning)

Última modificació: 12/06/2025

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

Titulació: MÀSTER UNIVERSITARI EN TECNOLOGIES BIOMÈDIQUES AVANÇADES (Pla 2025). (Assignatura optativa).

Curs: 2025 **Crèdits ECTS:** 6.0 **Idiomes:** Català

PROFESSORAT

Professorat responsable: Pozo Montero, Francesc

Altres: Pozo Montero, Francesc

CAPACITATS PRÈVIES

Per seguir aquest curs, l'estudiantat ha de tenir coneixements previs en matemàtiques i estadística (àlgebra lineal, probabilitat, regressió), alguna experiència en programació amb Python (llibries com NumPy, Pandas, Matplotlib, Scikit-learn) i familiaritat amb anàlisi de dades biomèdiques (imatges mèdiques, senyals fisiològiques, dades clíniques). Es recomana tenir nocions bàsiques d'aprenentatge automàtic (classificació, regressió, PCA) i una bona capacitat analítica per interpretar resultats.

RESULTATS D'APRENTATGE

Coneixements:

K2. Reconèixer estructures avançades d'anàlisi de dades i modelització.

Habilitats:

S6. Interpretar dades biomèdiques mitjançant tècniques d'anàlisi de dades, aprenentatge automàtic (machine learning) i aprenentatge profund (deep learning).

Competències:

C6. Integrar els valors de la sostenibilitat, entenent la complexitat dels sistemes, a fi d'emprendre o promoure accions que estableixin i mantinguin la salut dels ecosistemes i millorin la justícia, amb la qual cosa es generaran visions per a futurs sostenibles.

C4. Usar de manera solvent els recursos d'informació, gestionant l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de la seva especialitat i valorant de manera crítica els resultats d'aquesta gestió.

C5. Utilitzar la informació científicotècnica per respondre a qualsevol demanda de modificació, innovació o millora de dispositius, productes i processos lligats a l'enginyeria biomèdica per a noves aplicacions científiques o tecnològiques.

METODOLOGIES DOCENTS

L'assignatura utilitza l'exposició de continguts teòrics (AF.1) i la resolució d'exercicis, problemes i casos (AF.2) en un 20%, les sessions en laboratoris informàtics o de simulació (AF.7) en un 20%, i la realització de treball individual i cooperatiu en un 60%.

OBJECTIUS D'APRENTATGE DE L'ASSIGNATURA

En finalitzar el curs Anàlisi de Dades i Aprenentatge Automàtic, l'alumnat haurà adquirit els coneixements i habilitats necessàries per:

1. Comprendre els fonaments de l'aprenentatge automàtic aplicat a la biomedicina, incloent-hi els tipus de dades biomèdiques, les tècniques d'anàlisi i els principals algorismes d'aprenentatge supervisat i no supervisat.
2. Aplicar tècniques de preprocessament i anàlisi exploratòria de dades biomèdiques, com la normalització, la selecció de característiques i la reducció de dimensionalitat, utilitzant eines com Scikit-learn i PyTorch.
3. Desenvolupar i avaluar models d'aprenentatge automàtic per a classificació i regressió en problemes biomèdics, implementant regressió logística, arbres de decisió, SVM i xarxes neuronals.
4. Implementar models d'aprenentatge profund (deep learning) per aplicacions biomèdiques, incloent-hi xarxes neuronals convolucionals (CNN) per al processament d'imatges mèdiques i xarxes recurrents (RNN) per a seqüències de dades.
5. Interpretar els resultats obtinguts dels models d'intel·ligència artificial i avaluar-ne la fiabilitat mitjançant mètriques adequades (precisió, sensibilitat, AUC-ROC, etc.), validació creuada i tècniques de detecció d'overfitting.
6. Aplicar metodologies d'interpretabilitat i explicabilitat per garantir l'ús ètic i segur de la intel·ligència artificial en l'àmbit biomèdic, mitjançant eines com SHAP i LIME.
7. Desenvolupar i desplegar models en entorns biomèdics reals, entenent els reptes i les consideracions pràctiques per portar un model des de la fase d'experimentació fins a la implementació en aplicacions clíniques.
8. Integrar els coneixements adquirits en un projecte final, on els estudiants dissenyaran i implementaran una solució d'aprenentatge automàtic aplicada a un problema biomèdic concret.

HORES TOTALES DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores grup petit	21,0	14.00
Hores grup gran	21,0	14.00
Hores aprenentatge autònom	108,0	72.00

Dedicació total: 150 h

CONTINGUTS

Fonaments de Programació i Aprenentatge Automàtic

Descripció:

- Introducció a Python i llibreries essencials (NumPy, Pandas, Matplotlib, Scikit-learn, PyTorch).
- Conceptes bàsics d'aprenentatge automàtic: aprenentatge supervisat vs. no supervisat, overfitting, validació creuada, mètriques d'avaluació.

Activitats vinculades:

Pràctica: Anàlisi exploratòria i preprocés de dades biomèdiques. Es proporcionarà un conjunt de dades biomèdiques (per exemple, registres clínics o mesures fisiològiques). Els estudiants hauran de fer una anàlisi exploratòria, neteja i normalització de dades fent servir Pandas, NumPy i Scikit-learn.

Dedicació: 6h

Grup gran/Teoria: 3h

Grup petit/Laboratori: 3h

Models de Regressió i Classificació

Descripció:

- Regressió lineal i polinòmica.
- Regressió logística i mètriques d'avaluació en classificació.
- Aplicacions biomèdiques: predicció de valors clínics, diagnòstic de malalties.

Activitats vinculades:

Pràctica: Predicció de valors clínics amb regressió. Els estudiants treballaran amb un conjunt de dades de pacients per predir una variable clínica (per exemple, nivells de glucosa en sang). Implementaran regressió lineal i Ridge/Lasso i compararan el rendiment amb mètriques com RMSE i R^2 .

Dedicació: 6h

Grup gran/Teoria: 3h

Grup petit/Laboratori: 3h

Reducció de Dimensionalitat i Agrupament

Descripció:

- PCA i tècniques avançades de reducció de dimensions.
- Clustering: K-means, DBSCAN i clustering jeràrquic.
- Interpretabilitat i aplicacions en dades biomèdiques.

Activitats vinculades:

Pràctica: Visualització de dades biomèdiques en 2D amb PCA i t-SNE. S'utilitzarà un conjunt de dades amb moltes variables (per exemple, imatges transformades en vectors o dades genòmiques) per reduir la dimensionalitat amb PCA i t-SNE i visualitzar els grups latents.

Dedicació: 6h

Grup mitjà/Pràctiques: 3h

Grup petit/Laboratori: 3h

Models Avançats: Arbres de Decisió i Mètodes Ensemble

Descripció:

- Arbres de decisió, Random Forest i Gradient Boosting.
- Selecció i extracció de característiques.
- Aplicacions en biomarcadors i diagnòstic assistit.

Activitats vinculades:

Resolució de problemes: Comparació de models per a la detecció d'anomalies clíniques. Es donaran diversos models (Àrbre de decisió, Random Forest, Gradient Boosting) i els estudiants hauran d'analitzar quin és millor en termes de precisió, recall i importància de característiques en un cas de detecció d'anomalies en ECGs.

Dedicació: 6h

Grup gran/Teoria: 3h

Grup petit/Laboratori: 3h

Xarxes Neuronals i Aprenentatge Profund

Descripció:

- Introducció a les xarxes neuronals profundes.
- Xarxes totalment connectades (Fully Connected Neural Networks).
- Regularització i optimització en xarxes neuronals.

Activitats vinculades:

Pràctica: Construcció d'una xarxa neuronal per a diagnòstic mèdic. Es farà servir PyTorch o TensorFlow per construir una MLP per classificar dades mèdiques (per exemple, detecció de diabetis amb la base de dades Pima Indians Diabetes). S'experimentarà amb optimització, dropout i regularització L2.

Dedicació: 9h

Grup gran/Teoria: 4h 30m

Grup petit/Laboratori: 4h 30m

Xarxes Convolucionals i Aplicacions en Imatges Biomèdiques

Descripció:

- Fonaments de CNNs: convolucions, pooling i arquitectures clàssiques.
- Aplicacions biomèdiques: classificació d'imatges mèdiques, segmentació, diagnòstic automàtic.
- Avaluació de models i consideracions ètiques.

Activitats vinculades:

Pràctica: Classificació d'imatges mèdiques amb CNNs. Es treballarà amb un conjunt de dades d'imatges mèdiques (per exemple, raigs X pulmonars per detectar pneumònia). Els estudiants implementaran una CNN bàsica amb Keras/PyTorch, analitzaran matrius de confusió, interpretabilitat (Grad-CAM) i biaixos ètics.

Dedicació: 9h

Grup gran/Teoria: 4h 30m

Grup petit/Laboratori: 4h 30m

SISTEMA DE QUALIFICACIÓ

Primer examen parcial = 35 %

Segon examen parcial = 45%

Treballs, projectes, activitats dirigides = 20%

L'avaluació es durà a terme mitjançant la valoració per part del professorat. L'estudiantat pot superar l'assignatura mitjançant l'avaluació continuada a partir de dos exàmens parcials (primer parcial a mitjan curs, i últim parcial en el període habilitat per a la realització d'aquestes proves) i la realització de treballs, projectes i activitats dirigides. L'estructura de cada examen té un bloc teòric (questions) i un bloc pràctic (problemes), de dues hores de durada.

No hi haurà examen de reavaluació en aquesta assignatura.

BIBLIOGRAFIA

Bàsica:

- Deprez, Maria; Robinson, Emma C.. Machine Learning for Biomedical Applications. Academic Press, 2024. ISBN 9780128229040.

Complementària:

- Kose, Utku; Deperlioglu, Omer; Hemanth, Jude. Deep Learning for Biomedical Applications. 1st Edition . CRC Press, 2024. ISBN 9781032033235.

- Shaikh, T.A.; Hakak, S.; Rasool, T.; Wasid, M.. Machine Learning and Artificial Intelligence in Healthcare Systems: Tools and Techniques. CRC Press, ISBN 9781032549217.



- Balas, V.E. et al.. Handbook of Deep Learning in Biomedical Engineering. Academic Press, 2020. ISBN 9780128230145.
- Mohanty, S. N.; Nalinipriya, G.; Jena, O. P.; Sarkar, A. . Machine Learning for Healthcare Applications. Wiley, 2021. ISBN 9781119791812.
- Kumar Rana, A.; Sharma, S.; Rana, S.; Chaudhary, R.. Evolution of Machine Learning and Internet of Things Applications in Biomedical Engineering. Routledge, 2023. ISBN 9781032759234.
- Karthik, C.; Rajalakshmi, M.; Mohanty, S.N.; Chowdhury, S.. Machine Learning for Healthcare Systems. River Publishers, 2023. ISBN 978-8770228114.
- Singh, B.K.; Sinha, G.R.. Machine Learning in Healthcare. Fundamentals and Recent Applications. CRC Press, 2024. ISBN 9780367564438.

RECURSOS

Enllaç web:

- Machine Learning (MIT OCW). Machine Learning (MIT OCW)

Altres recursos:

<https://ocw.mit.edu/courses/6-867-machine-learning-fall-2006/>