



centro adscrito a:



UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

GUIA DOCENT DE CIÈNCIA DE DADES A L'ESPORT 2026-27

DADES GENERALS ASSIGNATURA

Nom:	Ciència de dades en l' esport
Codi:	801757
Curs:	2026-27
Titulació:	CTEF
Nº de crèdits (ECTS):	6
Ubicació en el pla d' estudis:	4t curs, 1r quadrimestre
Departament:	Sistemes d' informació i tecnologies
Responsable departament:	Dra. Cristina Cáliz
Data de l' última revisió:	Juliol 26
Professor Responsable:	Raül Pérez i Gonzalo

1. DESCRIPCIÓ GENERAL

Durant el transcurs d'un simple partit de futbol o de bàsquet es generen diversos milions de dades. Informació que, en molts casos, ni tan sols l' ull humà és capaç de percebre, però que, amb una adequada gestió i anàlisi, poden donar lloc a avantatges competitiu clars.

Un context tan complex com aquest requereix de la conjunció de diferents àrees de coneixement per donar respostes als professionals de l'esport: l'estadística –per dissenyar de forma eficient els estudis i analitzar de forma intel·ligent les dades-, la programació informàtica –per a l'organització del sistema d'informació i l'automatització dels processos-, la comunicació –per poder construir històries basades en les dades que generin impacte real en l'organització-, i, per suposat, la ciència de l' esport per poder donar-li sentit a tot aquest coneixement dins d' una àrea de negoci tan específica.

En això consisteix la ciència de dades: un nou perfil professional en enorme auge a nivell internacional, especialment en el sector esportiu, que permetrà a l' alumne aportar un gran valor afegit a l' organització.

En aquesta assignatura aprendrem, de forma eminentment pràctica, a dissenyar estudis analítics avançats, construir sistemes d'informació que els suportin, recollir, integrar, preparar i explorar les dades, generar models numèrics de classificació i regressió (aprenentatge de màquina) i, finalment, reportar els resultats de forma efectiva.

2. OBJECTIUS

L' assignatura i els seus resultats previstos responen a aquests objectius:

- Comprendre la relació entre els conceptes bàsics de l' àrea: de les matemàtiques i l' estadística a l' aprenentatge de màquina i la intel·ligència artificial.
- Dissenyar estudis d' analítica avançada de dades efectives per al sector de l' esport.

- Conèixer les tècniques més apropiades per a la preparació adequada de les dades.
- Aprendre a escollir i manejar les llibreries i algoritmes d' anàlisi avançada més adient en cada cas.
- Construir, interpretar i millorar la qualitat dels models numèrics avançats.
- Desenvolupar les habilitats de comunicació d' impacte basada en dades.
- Identificar els possibles camins en el desenvolupament professional dins d' aquesta àrea.

3. TEMARI

BLOC 1. INTRODUCCIÓ A LA CIÈNCIA DE DADES & ENGINYERIA DE DADES

Resultats específics d'aprenentatge

L'estudiant, després d'estudiar els temes i realitzar els exercicis, serà capaç de:

- Entendre la relació entre la intel·ligència artificial, l'aprenentatge de màquina, l'estadística i les matemàtiques.
- Enumerar els elements clau del disseny d'un estudi d'anàlisi avançada de dades en l'esport.
- Manejar de forma fluida estructures de dades multidimensionals i vectoritzades.
- Carregar, emmagatzemar i manipular conjunts de dades (*datasets*).
- Identificar, tractar i aplicar estratègies per solucionar problemes amb dades faltants.
- Transformar, filtrar i combinar diferents conjunts de dades per preparar-los per a la seva posterior anàlisi.

TEMA 1: LA CIÈNCIA DE DADES, LLIBRERIES I NUMPY

- 1.1 Necessitat de la ciència de dades en l'esport i en el fitness.
- 1.2 Què és un científic de dades.
- 1.3 Possibilitats de desenvolupament professional.
- 1.4 Disseny d'un projecte de ciència de dades.
- 1.5 Array multidimensionals i vectoritzats de Numpy.
- 1.6 Indexació i *slicing* (tall de dades).
- 1.7 Operadors matemàtics estàndard.
- 1.8 Guardar i carregar dades d'arrays.
- 1.9 Introducció a l'àlgebra lineal aplicada.

TEMA 2: MANEIG DE DADES: ANÀLISI I MANIPULACIÓ

- 2.1 Estructures de dades de la llibreria pandas: *Sèries* i *DataFrame*.
- 2.2 Indexació, selecció i filtratge: capacitat per treballar amb subconjunts de dades.
- 2.3 Aplicació de funcions personalitzades i mapatge (*mapping*) a les nostres dades.
- 2.4 Maneig de dades faltants (*missing data*).
- 2.5 Càrrega i emmagatzematge de dades: obrir i guardar arxius (CSV o Excel).

TEMA 3: NETEJA I PREPARACIÓ DE LES DADES

- 3.1 Consolidació de la càrrega i emmagatzematge d'arxius.
- 3.2 Estadística descriptiva aplicada a les dades.
- 3.3 Tractament avançat de dades faltants.

3.4 Transformació de dades.

3.5 Combinació i fusió (*merging*) de conjunts de dades.

BLOC 2. EXPLORACIÓ DE LES DADES I VISUALITZACIÓ

Resultats específics d'aprenentatge

L'estudiant, després d'estudiar el capítol i realitzar els exercicis, serà capaç de:

- Aplicar tècniques d'estadística descriptiva per resumir i comprendre les característiques principals de les dades.
- Identificar i calcular mesures de posició, dispersió i associació lineal entre variables.
- Seleccionar els gràfics més adequats per representar visualment variables qualitatives i quantitatives.
- Dissenyar, personalitzar i exportar visualitzacions de dades efectives utilitzant la llibreria `matplotlib`.

TEMA 4: VISUALITZACIÓ DE DADES I ESTADÍSTICA DESCRIPTIVA

4.1 Estadística descriptiva: taules i gràfics.

4.2 Gràfics de variables qualitatives i quantitatives.

4.3 Mesures de posició.

4.4 Mesures de dispersió.

4.5 Associació lineal entre variables quantitatives.

TEMA 5: LLIBRERIA DE MATPLOTLIB

5.1 Tipus de gràfics (*plots*) més comuns.

5.2 Figures i *subplots*.

5.3 Ús de colors, marcadors i estils de línia.

5.4 Exportació i guardat de gràfics en arxius.

BLOC 3. INTRODUCCIÓ A L'APRENTATGE AUTOMÀTIC

Resultats específics d'aprenentatge

L'estudiant, després d'estudiar el capítol i realitzar els exercicis, serà capaç de:

- Entendre la relació entre la intel·ligència artificial, l'aprenentatge automàtic (*Machine Learning*), l'estadística i les matemàtiques.

- Diferenciar entre aprenentatge supervisat i no supervisat, així com avaluar els desafiaments i reptes de l'entrenament de models.
- Preparar adequadament les dades de cara al modelatge numèric.
- Seleccionar, entrenar i ajustar hiperparàmetres de models de regressió i classificació.
- Avaluar el rendiment dels models predictius interpretant correctament les mètriques d'error i corbes d'aprenentatge.

TEMA 6: INTRODUCCIÓ AL MACHINE LEARNING

- 6.1 Què és el Machine Learning?
- 6.2 Supervisió durant l'entrenament.
- 6.3 Aprenentatge basat en instàncies vs. basat en models.
- 6.4 Reptes i desafiaments del Machine Learning.
- 6.5 Proves i validació.
- 6.6 Ajust d'hiperparàmetres i selecció del model.

TEMA 7: MACHINE LEARNING - REGRESSIÓ

- 7.1 Homologació i visualització de les dades enfocada al model.
- 7.2 Neteja de les dades.
- 7.3 Maneig d'atributs categòrics.
- 7.4 Escalat de característiques (*Feature Scaling*) i transformació.
- 7.5 Selecció i entrenament d'un model.
- 7.6 Ajust fi del model (*Fine-tune*).

TEMA 8: MACHINE LEARNING - CLASSIFICACIÓ

- 8.1 Entrenament d'un classificador binari.
- 8.2 Mesures de rendiment.
- 8.3 Classificació multiclasse.
- 8.4 Anàlisi d'errors.
- 8.5 Classificació multietiqueta (*Multilabel*).
- 8.6 Classificació multisalida (*Multioutput*).

TEMA 9: ENTRENAMENT DE MODELS

- 9.1 Regressió Lineal.
- 9.2 Equació Normal.

9.3 Descens de Gradient.

9.4 Corbes d'aprenentatge.

9.5 Regularització.

9.6 Aprenentatge no supervisat: Algoritme K-Means.

4. METODOLOGIA D'ENSENYAMENT I APRENTATGE

Es basa en classes pràctiques desenvolupades amb els equips i eines informàtiques de l'aula (o de l'alumne) i complementades amb l'exposició dels continguts teòrics fonamentals.

5. SISTEMA D' AVALUACIÓ

D'acord amb el Pla Bolonya, el model premia l'esforç constant i continuat de l'estudiantat. Un 40% de la nota s'obté de l'avaluació contínua de les activitats dirigides i el 60% percentatge restant, de l'examen final presencial. L'examen final té dues convocatòries.

La nota final de l'assignatura (NF) es calcularà a partir de la següent fórmula:

- **NF = Nota Examen Final x 60% + Nota Avaluació Continuada x 40%**
- Nota mínima de l'examen final per calcular la NF serà de 40 punts sobre 100.
- L'assignatura queda aprovada amb una NF igual o superior a 50 punts sobre 100.

Activitats d'avaluació contínua:

Tipus d' activitat	Descripció	% Avaluació contínua	
Lliuraments:			40 %
Treball	Objectiu general del projecte, descripció de l' estudi, dades i codi Python per a la seva preparació automàtica. Així com, Homologació de dades, model final i mètriques d' avaluació.	75%	
Presentació	Presentació oral del treball i estudi realitzat.	25%	
Examen final:			60 %
	Examen final	100%	

6. BIBLIOGRAFIA

6.1. BIBLIOGRAFIA BÀSICA

- Haslwanter, T. (2016). *An introduction to statistics with Python: With applications in the life sciences (Statistics and computing)*. Springer.
- Géron, A. (2022). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (3rd ed.). O'Reilly Media.
- McKinney, W. (2022). *Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter* (3rd ed.). O'Reilly Media.
- Miller, T. W. (2015). *Sports analytics and data science: Winning the game with methods and models*. FT Press.
- Raschka, S., & Mirjalili, V. (2019). *Python machine learning: With applications in the life sciences (Statistics and computing)*. Springer.
- Spiegelhalter, D. (2019). *The art of statistics: Learning from data*. Pelican.

6.2 BIBLIOGRAFIA COMPLEMENTÀRIA

- Barça Innovation Hub. (n.d.). Què es necessita aprendre per treballar en football analytics? Retrieved from <https://barcainnovationhub.com/es/que-se-necesita-aprender-para-trabajar-en-football-analytics/>
- Data Science Central. (n.d.). Retrieved from <https://www.datasciencecentral.com/>
- Ferrero, R. (n.d.). [Twitter profile]. Retrieved from <https://x.com/RosanaFerrero>
- KDnuggets. (n.d.). The data science process. Retrieved from <https://www.kdnuggets.com/>
- Ortega, E., Ortiz, E., & Artés, M. (2009). Manual d' estadística aplicada a les ciències de l' activitat física i l' esport. Diego Marín.
- Towards Data Science. (n.d.). The data science process. Retrieved from <https://towardsdatascience.com/the-data-science-process-a19eb7ebc41b>