

Guia docent

270954 - PODS - Ciència de Dades Orientada a Processos

Última modificació: 23/11/2023

Unitat responsable: Facultat d'Informàtica de Barcelona
Unitat que imparteix: 723 - CS - Departament de Ciències de la Computació.

Titulació: MÀSTER UNIVERSITARI EN CIÈNCIA DE DADES (Pla 2021). (Assignatura obligatòria).

Curs: 2023 **Crèdits ECTS:** 6.0 **Idiomes:** Anglès

PROFESSORAT

Professorat responsable: CARLOS ESCOLANO PEINADO

Altres: Primer quadrimestre:
CARLOS ESCOLANO PEINADO - 11, 12

CAPACITATS PRÈVIES

Comprensió exhaustiva de la informàtica en general; bon domini de diversos llenguatges de programació; capacitat bàsica per formalitzar qüestions matemàtiques en enginyeria informàtica.

COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

Específiques:

CE13. Identificar les principals amenaces en l'àmbit de l'ètica i la privacitat de dades en un projecte de ciència de dades (tant en l'aspecte de gestió com d'anàlisi de dades) i desenvolupar i implantar mesures adequades per esmorteir aquestes amenaces.

CE5. Modelar, dissenyar i implementar sistemes complexos de dades, incloent-hi la visualització de dades

CE6. Dissenyar el procés de Ciència de Dades i aplicar metodologies científiques per a obtenir conclusions sobre poblacions i prendre decisions en conseqüència, a partir de dades estructurades o no estructurades i potencialment emmagatzemades en formats heterogenis.

CE7. Identificar les limitacions imposades per la qualitat de dades en un problema de ciència de dades i aplicar tècniques per a disminuir el seu impacte

CE9. Aplicar mètodes adequats per a l'anàlisi d'altres tipus de formats, com ara processos i grafs, dins l'àmbit de ciència de dades

Genèriques:

CG2. Identificar i aplicar mètodes d'anàlisi, extracció de coneixement i visualització de dades recollides en formats molt diferents

CG3. Definir, dissenyar i implementar sistemes complexos que cobreixin totes les fases en projectes de ciència de dades

Transversals:

CT1. EMPRENEDORIA I INNOVACIÓ: Conèixer i comprendre l'organització d'una empresa i les ciències que regeixen la seva activitat; tenir capacitat per entendre les normes laborals i les relacions entre la planificació, les estratègies industrials i comercials, la qualitat i el benefici. Conèixer i entendre els mecanismes en què es basa la recerca científica, així com els mecanismes i instruments de transferència de resultats entre els diferents agents socioeconòmics implicats en els processos d'I+D+i.

CT4. ÚS SOLVENT DELS RECURSOS D'INFORMACIÓ: Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació de l'àmbit d'especialitat, i valorar de forma crítica els resultats d'aquesta gestió.

CT5. TERCERA LLENGUA: Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i en consonància amb les necessitats que tindran els titulats i titulades.

Bàsiques:

CB10. Posseir i comprendre coneixements que aportin una base o oportunitat de ser originals en el desenvolupament i/o aplicació d'idees, sovint en un context de recerca.

CB6. Que els estudiants sàpiguen aplicar els coneixements adquirits y la seva capacitat de resolució de problemes en entorns nous o poc coneguts dins de contextos més amplis (o multidisciplinaris) relacionats amb la seva àrea d'estudi.

CB7. Que els estudiants siguin capaços d'integrar coneixements i enfrontar-se a la complexitat de formular judicis a partir d'una informació que, essent incompleta o limitada, inclogui reflexions sobre les responsabilitats socials i ètiques vinculades a l'aplicació dels seus coneixements i judicis.

CB8. Que els estudiants sàpiguen comunicar les seves conclusions i els coneixements i raons darreres que les sustenten- a públics especialitzats i no especialitzats d'una manera clara i sense ambigüitats.

CB9. Que els estudiants posseixin les habilitats d'aprenentatge que els permetin continuar estudiant d'una manera que haurà de ser en gran mesura autòdridida o autònoma.

METODOLOGIES DOCENTS

Sessions teòriques que poden incloure sessions de resolució de problemes amb o sense un component de programació, sessions pràctiques amb programari de ciència de dades orientada a processos de codi obert o comercial, desenvolupament d'un cas pràctic.

OBJECTIUS D'APRENENTATGE DE L'ASSIGNATURA

1. Conèixer el conjunt teòric i pràctic de problemes que constitueixen la ciència de dades orientada a processos i comprendre els principals algorismes per abordar-la: tant a nivell conceptual com a nivell d'aplicació a través d'algunes eines i llibreries actuals.

2. Adquirir i demostrar la capacitat de posar en pràctica els coneixements obtinguts durant el curs i relacionar-los amb les perspectives organitzatives i d'equip com a projecte de ciència de dades orientat al procés que s'executa en una organització real.

HORES TOTALES DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores activitats dirigides	6,0	4.00
Hores grup petit	24,0	16.00
Hores aprenentatge autònom	96,0	64.00
Hores grup gran	24,0	16.00

Dedicació total: 150 h

CONTINGUTS

Models de processos i dades d'events

Descubrimient automàtic de models de processos

Verificació de conformitat de models de processos i dades d'events

Millora de processos basada en l'evidència de les dades

Selecció d'aplicacions i tècniques avançades



ACTIVITATS

Models de processos i dades d'eventms

Descripció:

En aquesta activitat s'introduiran els models de processos per especificar processos en les organitzacions, i les dades que parlen d'events que s'originen en l'execució dels processos.

Objectius específics:

1

Competències relacionades:

CG3. Definir, dissenyar i implementar sistemes complexos que cobreixin totes les fases en projectes de ciència de dades
CG2. Identificar i aplicar mètodes d'anàlisi, extracció de coneixement i visualització de dades recollides en formats molt diferents
CE6. Dissenyar el procés de Ciència de Dades i aplicar metodologies científiques per a obtenir conclusions sobre poblacions i prendre decisions en conseqüència, a partir de dades estructurades o no estructurades i potencialment emmagatzemades en formats heterogenis.

CE7. Identificar les limitacions imposades per la qualitat de dades en un problema de ciència de dades i aplicar tècniques per a disminuir el seu impacte

CE13. Identificar les principals amenaces en l'àmbit de l'ètica i la privacitat de dades en un projecte de ciència de dades (tant en l'aspecte de gestió com d'anàlisi de dades) i desenvolupar i implantar mesures adequades per esmorteir aquestes amenaces.

CE5. Modelar, dissenyar i implementar sistemes complexos de dades, incloent-hi la visualització de dades

CE9. Aplicar mètodes adequats per a l'anàlisi d'altres tipus de formats, com ara processos i grafs, dins l'àmbit de ciència de dades

CT4. ÚS SOLVENT DELS RECURSOS D'INFORMACIÓ: Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació de l'àmbit d'especialitat, i valorar de forma crítica els resultats d'aquesta gestió.

CT5. TERCERA LLENGUA: Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i en consonància amb les necessitats que tindran els titulats i titulades.

CB10. Posseir i comprendre coneixements que aportin una base o oportunitat de ser originals en el desenvolupament i/o aplicació d'idees, sovint en un context de recerca.

CB7. Que els estudiants siguin capaços d'integrar coneixements i enfrontar-se a la complexitat de formular judicis a partir d'una informació que, essent incompleta o limitada, inclogui reflexions sobre les responsabilitats socials i ètiques vinculades a l'aplicació dels seus coneixements i judicis.

CB9. Que els estudiants posseeixin les habilitats d'aprenentatge que els permetin continuar estudiant d'una manera que haurà de ser en gran mesura autòdrida o autònoma.

Dedicació: 24h 54m

Grup gran/Teoria: 5h

Grup petit/Laboratori: 4h

Aprenentatge autònom: 15h 54m



Descubrimient automàtic de models de processos

Descripció:

En aquesta activitat s'introduiran diverses tècniques que extreuen models de processos en diversos formalismes a partir de dades d'events.

Objectius específics:

1

Competències relacionades:

CG3. Definir, dissenyar i implementar sistemes complexos que cobreixin totes les fases en projectes de ciència de dades

CG2. Identificar i aplicar mètodes d'anàlisi, extracció de coneixement i visualització de dades recollides en formats molt diferents

CE6. Dissenyar el procés de Ciència de Dades i aplicar metodologies científiques per a obtenir conclusions sobre poblacions i prendre decisions en conseqüència, a partir de dades estructurades o no estructurades i potencialment emmagatzemades en formats heterogenis.

CE7. Identificar les limitacions imposades per la qualitat de dades en un problema de ciència de dades i aplicar tècniques per a disminuir el seu impacte

CE13. Identificar les principals amenaces en l'àmbit de l'ètica i la privacitat de dades en un projecte de ciència de dades (tant en l'aspecte de gestió com d'anàlisi de dades) i desenvolupar i implantar mesures adequades per esmorteir aquestes amenaces.

CE5. Modelar, dissenyar i implementar sistemes complexos de dades, incloent-hi la visualització de dades

CE9. Aplicar mètodes adequats per a l'anàlisi d'altres tipus de formats, com ara processos i grafs, dins l'àmbit de ciència de dades

CT4. ÚS SOLVENT DELS RECURSOS D'INFORMACIÓ: Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació de l'àmbit d'especialitat, i valorar de forma crítica els resultats d'aquesta gestió.

CT5. TERCERA LLENGUA: Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i en consonància amb les necessitats que tindran els titulats i titulades.

CB10. Posseir i comprendre coneixements que aportin una base o oportunitat de ser originals en el desenvolupament i/o aplicació d'idees, sovint en un context de recerca.

CB7. Que els estudiants siguin capaços d'integrar coneixements i enfrontar-se a la complexitat de formular judicis a partir d'una informació que, essent incompleta o limitada, inclogui reflexions sobre les responsabilitats socials i ètiques vinculades a l'aplicació dels seus coneixements i judicis.

CB9. Que els estudiants posseeixin les habilitats d'aprenentatge que els permetin continuar estudiant d'una manera que haurà de ser en gran mesura autodirigida o autònoma.

Dedicació: 28h

Grup gran/Teoria: 6h

Grup petit/Laboratori: 6h

Aprenentatge autònom: 16h



Verificació de conformitat de models de processos i dades d'events

Descripció:

En aquesta activitat s'introduiran algorismes per establir la relació entre el comportament modelat i observat d'un procés.

Objectius específics:

1

Competències relacionades:

CG3. Definir, dissenyar i implementar sistemes complexos que cobreixin totes les fases en projectes de ciència de dades

CG2. Identificar i aplicar mètodes d'anàlisi, extracció de coneixement i visualització de dades recollides en formats molt diferents

CE6. Dissenyar el procés de Ciència de Dades i aplicar metodologies científiques per a obtenir conclusions sobre poblacions i prendre decisions en conseqüència, a partir de dades estructurades o no estructurades i potencialment emmagatzemades en formats heterogenis.

CE7. Identificar les limitacions imposades per la qualitat de dades en un problema de ciència de dades i aplicar tècniques per a disminuir el seu impacte

CE13. Identificar les principals amenaces en l'àmbit de l'ètica i la privacitat de dades en un projecte de ciència de dades (tant en l'aspecte de gestió com d'anàlisi de dades) i desenvolupar i implantar mesures adequades per esmorteir aquestes amenaces.

CE5. Modelar, dissenyar i implementar sistemes complexos de dades, incloent-hi la visualització de dades

CE9. Aplicar mètodes adequats per a l'anàlisi d'altres tipus de formats, com ara processos i grafs, dins l'àmbit de ciència de dades

CT4. ÚS SOLVENT DELS RECURSOS D'INFORMACIÓ: Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació de l'àmbit d'especialitat, i valorar de forma crítica els resultats d'aquesta gestió.

CT5. TERCERA LLENGUA: Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i en consonància amb les necessitats que tindran els titulats i titulades.

CB10. Posseir i comprendre coneixements que aportin una base o oportunitat de ser originals en el desenvolupament i/o aplicació d'idees, sovint en un context de recerca.

CB7. Que els estudiants siguin capaços d'integrar coneixements i enfrontar-se a la complexitat de formular judicis a partir d'una informació que, essent incompleta o limitada, inclogui reflexions sobre les responsabilitats socials i ètiques vinculades a l'aplicació dels seus coneixements i judicis.

CB9. Que els estudiants posseeixin les habilitats d'aprenentatge que els permetin continuar estudiant d'una manera que haurà de ser en gran mesura autodirigida o autònoma.

Dedicació: 28h

Grup gran/Teoria: 6h

Grup petit/Laboratori: 6h

Aprenentatge autònom: 16h



Millora de processos basada en l'evidència de les dades

Descripció:

En aquesta activitat es presentaran tècniques per utilitzar dades d'esdeveniments per projectar i millorar models de processos i registres d'esdeveniments.

Objectius específics:

1

Competències relacionades:

CG3. Definir, dissenyar i implementar sistemes complexos que cobreixin totes les fases en projectes de ciència de dades

CG2. Identificar i aplicar mètodes d'anàlisi, extracció de coneixement i visualització de dades recollides en formats molt diferents

CE6. Dissenyar el procés de Ciència de Dades i aplicar metodologies científiques per a obtenir conclusions sobre poblacions i prendre decisions en conseqüència, a partir de dades estructurades o no estructurades i potencialment emmagatzemades en formats heterogenis.

CE7. Identificar les limitacions imposades per la qualitat de dades en un problema de ciència de dades i aplicar tècniques per a disminuir el seu impacte

CE13. Identificar les principals amenaces en l'àmbit de l'ètica i la privacitat de dades en un projecte de ciència de dades (tant en l'aspecte de gestió com d'anàlisi de dades) i desenvolupar i implantar mesures adequades per esmorteir aquestes amenaces.

CE5. Modelar, dissenyar i implementar sistemes complexos de dades, incloent-hi la visualització de dades

CE9. Aplicar mètodes adequats per a l'anàlisi d'altres tipus de formats, com ara processos i grafs, dins l'àmbit de ciència de dades

CT4. ÚS SOLVENT DELS RECURSOS D'INFORMACIÓ: Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació de l'àmbit d'especialitat, i valorar de forma crítica els resultats d'aquesta gestió.

CT5. TERCERA LLENGUA: Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i en consonància amb les necessitats que tindran els titulats i titulades.

CB10. Posseir i comprendre coneixements que aportin una base o oportunitat de ser originals en el desenvolupament i/o aplicació d'idees, sovint en un context de recerca.

CB7. Que els estudiants siguin capaços d'integrar coneixements i enfrontar-se a la complexitat de formular judicis a partir d'una informació que, essent incompleta o limitada, inclogui reflexions sobre les responsabilitats socials i ètiques vinculades a l'aplicació dels seus coneixements i judicis.

CB9. Que els estudiants posseeixin les habilitats d'aprenentatge que els permetin continuar estudiant d'una manera que haurà de ser en gran mesura autodirigida o autònoma.

Dedicació: 24h

Grup gran/Teoria: 4h

Grup petit/Laboratori: 4h

Aprenentatge autònom: 16h



Selecció d'aplicacions i tècniques avançades

Descripció:

Diverses tècniques per resoldre problemes de la vida real de ciències de dades orientada a processos

Objectius específics:

1, 2

Competències relacionades:

CG3. Definir, dissenyar i implementar sistemes complexos que cobreixin totes les fases en projectes de ciència de dades

CG2. Identificar i aplicar mètodes d'anàlisi, extracció de coneixement i visualització de dades recollides en formats molt diferents

CE6. Dissenyar el procés de Ciència de Dades i aplicar metodologies científiques per a obtenir conclusions sobre poblacions i prendre decisions en conseqüència, a partir de dades estructurades o no estructurades i potencialment emmagatzemades en formats heterogenis.

CE7. Identificar les limitacions imposades per la qualitat de dades en un problema de ciència de dades i aplicar tècniques per a disminuir el seu impacte

CE13. Identificar les principals amenaces en l'àmbit de l'ètica i la privacitat de dades en un projecte de ciència de dades (tant en l'aspecte de gestió com d'anàlisi de dades) i desenvolupar i implantar mesures adequades per esmorteir aquestes amenaces.

CE5. Modelar, dissenyar i implementar sistemes complexos de dades, incloent-hi la visualització de dades

CE9. Aplicar mètodes adequats per a l'anàlisi d'altres tipus de formats, com ara processos i grafs, dins l'àmbit de ciència de dades

CT1. EMPRENEDORIA I INNOVACIÓ: Conèixer i comprendre l'organització d'una empresa i les ciències que regeixen la seva activitat; tenir capacitat per entendre les normes laborals i les relacions entre la planificació, les estratègies industrials i comercials, la qualitat i el benefici. Conèixer i entendre els mecanismes en què es basa la recerca científica, així com els mecanismes i instruments de transferència de resultats entre els diferents agents socioeconòmics implicats en els processos d'I+D+i.

CT4. ÚS SOLVENT DELS RECURSOS D'INFORMACIÓ: Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació de l'àmbit d'especialitat, i valorar de forma crítica els resultats d'aquesta gestió.

CT5. TERCERA LLENGUA: Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i en consonància amb les necessitats que tindran els titulats i titulades.

CB6. Que els estudiants sàpiguen aplicar els coneixements adquirits y la seva capacitat de resolució de problemes en entorns nous o poc coneguts dins de contextos més amplis (o multidisciplinaris) relacionats amb la seva àrea d'estudi.

CB10. Posseir i comprendre coneixements que aportin una base o oportunitat de ser originals en el desenvolupament i/o aplicació d'idees, sovint en un context de recerca.

CB7. Que els estudiants siguin capaços d'integrar coneixements i enfrontar-se a la complexitat de formular judicis a partir d'una informació que, essent incompleta o limitada, inclogui reflexions sobre les responsabilitats socials i ètiques vinculades a l'aplicació dels seus coneixements i judicis.

CB8. Que els estudiants sàpiguen comunicar les seves conclusions i els coneixements i raons darreres que les sustenten- a públics especialitzats i no especialitzats d'una manera clara i sense ambigüitats.

CB9. Que els estudiants posseeixin les habilitats d'aprenentatge que els permetin continuar estudiant d'una manera que haurà de ser en gran mesura autòdirigida o autònoma.

Dedicació: 24h

Grup gran/Teoria: 4h

Grup petit/Laboratori: 4h

Aprenentatge autònom: 16h



Metodologia per projectes de ciència de dades orientats a processos

Descripció:

Descripció general de com gestionar un projecte PODS

Objectius específics:

2

Competències relacionades:

CT1. EMPRENEDORIA I INNOVACIÓ: Conèixer i comprendre l'organització d'una empresa i les ciències que regeixen la seva activitat; tenir capacitat per entendre les normes laborals i les relacions entre la planificació, les estratègies industrials i comercials, la qualitat i el benefici. Conèixer i entendre els mecanismes en què es basa la recerca científica, així com els mecanismes i instruments de transferència de resultats entre els diferents agents socioeconòmics implicats en els processos d'I+D+i.

CB6. Que els estudiants sàpiguen aplicar els coneixements adquirits y la seva capacitat de resolució de problemes en entorns nous o poc coneguts dins de contextos més amplis (o multidisciplinaris) relacionats amb la seva àrea d'estudi.

CB8. Que els estudiants sàpiguen comunicar les seves conclusions i els coneixements i raons darreres que les sustenten- a públics especialitzats i no especialitzats d'una manera clara i sense ambigüitats.

Dedicació: 21h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 3h

Aprenentatge autònom: 16h

SISTEMA DE QUALIFICACIÓ

L'avaluació de l'assignatura consisteix en dos elements: exàmen final (60%), pràctiques de curs (40%).

L'exàmen final contindrà preguntes i problemes sobre els continguts teòrics que s'expliquen a les classes de teoria.

Les pràctiques de curs seràn pràctiques guiades que es realitzaran durant el curs sobre diverses eines i plataformes de mineria de procesos. Les pràctiques podran fer-se per parelles o individualment.

BIBLIOGRAFIA

Bàsica:

- Aalst, Wil van der. Process mining : data science in action [en línia]. 2nd ed. Berlin: Springer, 2016 [Consulta: 10/01/2024]. Disponible a : <https://ebookcentral-proquest-com.recursos.biblioteca.upc.edu/lib/upcatalunya-ebooks/detail.action?pq-origsite=primo&docID=4505537>. ISBN 9783662498514.

- Carmona Vargas, Josep; Van Dongen, Boudewijn; Solti, Andreas; Weidlich, Matthias. Conformance checking : relating processes and models. Cham: Springer International Publishing, 2018. ISBN 9783319994130.

Complementària:

- Dumas, M.; La Rosa, M.; Mendling, J.; Reijers, H.A. Fundamentals of business process management. 2nd ed. Springer, 2018. ISBN 9783662565087.

- Reinkemeyer, Lars. Process mining in action : principles, use cases and outlook [en línia]. Cham: Springer, 2020 [Consulta: 10/01/2024]. Disponible a : <https://ebookcentral-proquest-com.recursos.biblioteca.upc.edu/lib/upcatalunya-ebooks/detail.action?pq-origsite=primo&docID=6134217>. ISBN 9783030401726.

- Weske, Mathias. Business process management: concepts, languages, architectures. 3rd ed. Berlin: Springer, 2019. ISBN 9783662594346.

- Wolfgang Reisig. Understanding petri nets. Berlin: Springer, 2013. ISBN 9783642332777.



RECURSOS

Enllaç web:

- <https://www.tf-pm.org/>