

Guia docent

270958 - MUD - Minería de Dades No Estructurades

Última modificació: 31/01/2024

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: JORGE TURMO BORRÁS

Altres: Segon quadrimestre:
CARLOS ESCOLANO PEINADO - 12
SALVADOR MEDINA HERRERA - 11, 12
BARDIA RAFIEIAN - 11
JORGE TURMO BORRÁS - 11, 12

CAPACITATS PRÈVIES

Nivell alt de programació en python.
Coneixements d'estadística i matemàtiques a nivell de grau universitari en enginyeria o similar.
Conceptes bàsics d'aprenentatge automàtic.

COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

Específiques:

CE11. Analitzar i extreure coneixement d'informació no estructurada mitjançant tècniques de processament de llenguatge natural, minería de textos i imatges
CE12. Aplicar la ciència de dades en projectes multidisciplinaris per resoldre problemes en dominis nous o poc coneguts per la ciència de dades i que siguin econòmicament viables, socialment acceptables, i d'acord amb la legalitat vigent
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.
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

Genèriques:

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

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

METODOLOGIES DOCENTS

Classe expositiva participativa de continguts teòrics i pràctics

Classe pràctica de resolució, amb la participació dels estudiants, de casos pràctics i/o exercicis relacionats amb els continguts de la matèria

Pràctiques de laboratori - Treballen grup

Tutories.

OBJECTIUS D'APRENENTATGE DE L'ASSIGNATURA

HORES TOTALES DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores aprenentatge autònom	96,0	64.00
Hores grup petit	27,0	18.00
Hores grup gran	27,0	18.00

Dedicació total: 150 h

CONTINGUTS

(CAT) Procesamiento de lenguaje natural y su aplicación en análisis de textos.

Descripció:

(CAT)

Objectius específics:

(CAT)

Activitats vinculades:

(CAT)



(CAT) Descripción de las etapas del procesamiento de lenguaje natural.

Descripció:

(CAT)

Objectius específics:

(CAT)

Activitats vinculades:

(CAT)

(CAT) clasificación del texto y similitud de textos.

Descripció:

(CAT)

Objectius específics:

(CAT)

Activitats vinculades:

(CAT)

(CAT) Extraccion de informacion: Reconocimiento de entidades, extraccion de relaciones

Descripció:

(CAT)

Objectius específics:

(CAT)

Activitats vinculades:

(CAT)

(CAT) Técnicas de aprendizaje profundo para análisis de datos no estructurados.

Descripció:

(CAT)

Objectius específics:

(CAT)

Activitats vinculades:

(CAT)

(CAT) Principales arquitecturas de aprendizaje profundo para datos no estructurados

Descripció:

(CAT)

Objectius específics:

(CAT)

Activitats vinculades:

(CAT)

ACTIVITATS

Practica de laboratori

Objectius específics:

3

Competències relacionades:

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

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.

CE11. Analitzar i extreure coneixement d'informació no estructurada mitjançant tècniques de processament de llenguatge natural, mineria de textos i imatges

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

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.

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 autodirigida o autònoma.

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.

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.

Dedicació: 52h

Aprenentatge autònom: 52h



Examen final

Objectius específics:

1, 2

Competències relacionades:

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

CE12. Aplicar la ciència de dades en projectes multidisciplinaris per resoldre problemes en dominis nous o poc coneguts per la ciència de dades i que siguin econòmicament viables, socialment acceptables, i d'acord amb la legalitat vigent

CE11. Analitzar i extreure coneixement d'informació no estructurada mitjançant tècniques de processament de llenguatge natural, mineria de textos i imatges

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

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.

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 autodirigida o autònoma.

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.

Dedicació: 53h

Activitats dirigides: 3h

Aprenentatge autònom: 50h



NLP i les seves aplicacions

Objectius específics:

1

Competències relacionades:

CG2. Identificar i aplicar mètodes d'anàlisi, extracció de coneixement i visualització de dades recollides en formats molt diferents
CE11. Analitzar i extreure coneixement d'informació no estructurada mitjançant tècniques de processament de llenguatge natural, mineria de textos i imatges

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

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.

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.

Dedicació: 2h

Grup gran/Teoria: 2h



etapes del processament del llenguatge natural

Objectius específics:

1

Competències relacionades:

CG2. Identificar i aplicar mètodes d'anàlisi, extracció de coneixement i visualització de dades recollides en formats molt diferents
CE11. Analitzar i extreure coneixement d'informació no estructurada mitjançant tècniques de processament de llenguatge natural, mineria de textos i imatges

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

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.

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.

Dedicació: 11h 30m

Grup gran/Teoria: 8h 30m

Grup mitjà/Pràctiques: 3h



Classificació de textos, similaritat de textos.

Objectius específics:

2

Competències relacionades:

CE12. Aplicar la ciència de dades en projectes multidisciplinaris per resoldre problemes en dominis nous o poc coneguts per la ciència de dades i que siguin econòmicament viables, socialment acceptables, i d'acord amb la legalitat vigent

CE11. Analitzar i extreure coneixement d'informació no estructurada mitjançant tècniques de processament de llenguatge natural, mineria de textos i imatges

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.

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 autodirigida o autònoma.

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.

Dedicació: 2h

Grup gran/Teoria: 1h 30m

Grup mitjà/Pràctiques: 0h 30m



Extracció d'informació: Reconeixement d'entitats. Extracció de relacions.

Objectius específics:

1, 2

Competències relacionades:

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

CE12. Aplicar la ciència de dades en projectes multidisciplinaris per resoldre problemes en dominis nous o poc coneguts per la ciència de dades i que siguin econòmicament viables, socialment acceptables, i d'acord amb la legalitat vigent

CE11. Analitzar i extreure coneixement d'informació no estructurada mitjançant tècniques de processament de llenguatge natural, mineria de textos i imatges

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

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.

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 autòdrida o autònoma.

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.

Dedicació: 2h

Grup gran/Teoria: 1h 30m

Grup mitjà/Pràctiques: 0h 30m



Tècniques de deep learning per a l'anàlisi de dades no estructurades.

Objectius específics:

3

Competències relacionades:

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

CE11. Analitzar i extreure coneixement d'informació no estructurada mitjançant tècniques de processament de llenguatge natural, mineria de textos i imatges

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

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 LENGUA: 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.

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 autodirigida o autònoma.

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.

Dedicació: 7h

Grup gran/Teoria: 5h

Grup mitjà/Pràctiques: 2h



Principals arquitectures de deep learning per a dades no estructurades

Objectius específics:

3

Competències relacionades:

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

CE11. Analitzar i extreure coneixement d'informació no estructurada mitjançant tècniques de processament de llenguatge natural, mineria de textos i imatges

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

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.

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 autòdrida o autònoma.

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.

Dedicació: 5h 30m

Grup gran/Teoria: 4h

Grup mitjà/Pràctiques: 1h 30m



pràctica de laboratori en reconeixement d'entitats

Objectius específics:

1, 2

Competències relacionades:

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

CE12. Aplicar la ciència de dades en projectes multidisciplinaris per resoldre problemes en dominis nous o poc coneguts per la ciència de dades i que siguin econòmicament viables, socialment acceptables, i d'acord amb la legalitat vigent

CE11. Analitzar i extreure coneixement d'informació no estructurada mitjançant tècniques de processament de llenguatge natural, mineria de textos i imatges

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

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.

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 autòdrida o autònoma.

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.

Dedicació: 5h

Grup petit/Laboratori: 5h



pràctica de laboratori en extracció de relacions

Objectius específics:

1, 2

Competències relacionades:

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

CE12. Aplicar la ciència de dades en projectes multidisciplinaris per resoldre problemes en dominis nous o poc coneguts per la ciència de dades i que siguin econòmicament viables, socialment acceptables, i d'acord amb la legalitat vigent

CE11. Analitzar i extreure coneixement d'informació no estructurada mitjançant tècniques de processament de llenguatge natural, mineria de textos i imatges

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

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.

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 autodirigida o autònoma.

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.

Dedicació: 5h

Grup petit/Laboratori: 5h

pràctica de laboratori sobre xarxes neuronals

Objectius específics:

3

Competències relacionades:

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

CE11. Analitzar i extreure coneixement d'informació no estructurada mitjançant tècniques de processament de llenguatge natural, mineria de textos i imatges

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

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.

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 autodirigida o autònoma.

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.

Dedicació: 5h

Grup petit/Laboratori: 5h

SISTEMA DE QUALIFICACIÓ

pràctica de laboratori 50% + examen final 50%

BIBLIOGRAFIA

Bàsica:

- Jurafsky, Dan; Martin, James H. Speech and language processing : an introduction to natural language processing, computational linguistics, and speech recognition. 2nd ed. Upper Saddle River: Prentice Hall, 2008. ISBN 9789332518414.
- Indurkha, Nitin ; Damerau, Fred J. Handbook of natural language processing. 2nd ed. Boca Raton: Taylor–Francis, 2010. ISBN 9780429149207.
- Manning, Christopher D; Schütze, Hinrich. Foundations of statistical natural language processing. Cambridge, Mass: MIT Press, 1999. ISBN 0262133601.
- Mitkov, Ruslan. The Oxford handbook of computational linguistics. Oxford: Oxford University Press, 2003. ISBN 0198238827.
- Indurkha, Nitin; Damerau, Frederick J. Handbook of natural language processing. Boca Raton: Chapman and Hall/CRC, 2010. ISBN 9781420085938.
- Clark, Alexander; Fox, Chris; Lappin, Shalom. The Handbook of computational linguistics and natural language processing. Oxford: Wiley-Blackwell, 2010. ISBN 9781444324044.
- Manning, C.; See, A. Natural language processing with deep learning. Stanford University, 2022.
- Collins, M. Natural language processing. Columbia University,