

Guia docent

270220 - CAI - Cerca i Anàlisi de la Informació

Última modificació: 19/07/2023

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

Titulació: GRAU EN CIÈNCIA I ENGINYERIA DE DADES (Pla 2017). (Assignatura obligatòria).

Curs: 2023 **Crèdits ECTS:** 6.0 **Idiomes:** Català, Castellà, Anglès

PROFESSORAT

Professorat responsable: MARTA ARIAS VICENTE

Altres: Primer quadrimestre:
MARTA ARIAS VICENTE - 11, 12, 13

CAPACITATS PRÈVIES

Les donades per les assignatures dels Quadrimestres 1 a 4 del grau. L'assignatura és en bona part "comprehensiva" de molt el fet anteriorment, i especialment conceptes de matemàtica (discreta, àlgebra, una mica de càlcul), probabilitat i estadística, algorísmia general (en particular, de grafs), aprenentatge automàtic i anàlisi de dades, bases de dades, i sistemes de computadors distribuïts i paral·lels. Els laboratoris es fan en el llenguatge de programació python.

COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

Específiques:

CE1. Utilitzar amb destresa els conceptes i mètodes matemàtics subjacents els problemes de la ciència i l'enginyeria de les dades.
CE4. Utilitzar els sistemes de computació actuals, inclosos els sistemes d'alt rendiment, per al procés de grans volums de dades des del coneixement de la seva estructura, funcionament i particularitats.
CE6. Construir o utilitzar sistemes de processat i comprensió del llenguatge escrit, integrant-lo en altres sistemes dirigits per les dades. Dissenyar sistemes de recerca d'informació textual o hipertextual i d'anàlisi de xarxes socials.
CE7. Demostrar coneixement i capacitat d'aplicació de les eines necessàries per a l'emmagatzematge, el processament i l'accés a les dades.

Genèriques:

CG2. Elegir i aplicar els mètodes i tècniques més adequats a un problema definit per dades que representin un repte pel seu volum, velocitat, varietat o heterogeneïtat, inclosos mètodes informàtics, matemàtics, estadístics i de processament del senyal.
CG3. Treballar en equips i projectes multidisciplinaris relacionats amb el processat i explotació de dades complexes, interactuant fluidament amb enginyers i professionals d'altres disciplines.
CG4. Identificar oportunitats per a aplicacions innovadores orientades a dades en entorns tecnològics en contínua evolució.
CG5. Poder recórrer a coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics del futur.

Transversals:

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.
CT5. Ús solvent dels recursos d'informació. Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de l'especialitat i valorar de forma crítica els resultats d'aquesta gestió.
CT6. Aprenentatge autònom. Detectar deficiències en el propi coneixement i superar-les mitjançant la reflexió crítica i l'elecció de la millor actuació per ampliar aquest coneixement.
CT7. Tercera llengua. Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i d'acord amb les necessitats que tindran els titulats i titulades.

Bàsiques:

CB2. Que els estudiants sàpiguen aplicar els seus coneixements al seu treball o vocació d'una manera professional i posseixin les competències que solen demostrar-se mitjançant l'elaboració i defensa d'arguments i la resolució de problemes dins la seva àrea d'estudi.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloquin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB4. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.

METODOLOGIES DOCENTS

Classes "de teoria" expositives per part del professor. Es proposaran un cert nombre d'exercicis a resoldre fora de classe per a la propera sessió.

Classes "de teoria" dedicades a la resolució. Es comentaran en comú les solucions dels exercicis proposats a la/les sessions precedents. S'esperarà que els estudiants hagin intentat resoldre'ls.

Classes "de laboratori": A partir d'un guió que rebran al principi de la sessió, els estudiants duran a terme alguna tasca amb ordinador per consolidar els conceptes vistos a les classes de "teoria". Típicament serà l'implementació i experimentació amb algun algorisme, o l'anàlisi d'algun conjunt de dades.

OBJECTIUS D'APRENTATGE DE L'ASSIGNATURA

1.Descriure diferents models per avaluar similitud entre textos, i com s'apliquen a la cerca textual. Decidir quin dels models és el més adequat a un escenari concret que involucra cerca de textos. Implementar els models des de zero (en un sistema molt bàsic) o damunt un sistema d'indexació de textos que sigui altament escalable.

2.Descriure els avantatges, de cara a efectuar cerques efectives, d'utilitzar la informació donada pels enllaços en estructures hiperenllaçades, com ara la web, les xarxes socials digitals, i la web semàntica. Descriure els principals paràmetres que s'usen per caracteritzar aquestes estructures enllaçades. Reproduir els algorismes més utilitzats per establir importància en aquestes estructures (p.ex. pagerank), per descobrir estructura en elles (p.ex., descobriment de comunitats) i per millorar els resultats de cerques proposades per un usuari. Implementar aquests algorismes des de zero en un sistema molt bàsic, o damunt de sistemes de processament de dades massives de manera que puguin escalar.

3.Avaluar l'efectivitat de sistemes de cerca en sistemes complexos, descrivint-la en termes de mesures dures com ara "recall" i "precision" però també en termes de mesures toves com ara la satisfacció dels usuaris, la novetat i la finalització de tasques. Adaptar el funcionament i la presentació de sistemes de cerca d'informació amb el feedback de l'experiència d'usuaris metòdicament recollida.

4.Definir el problema de la recomanació i les diferències amb altres problemes relacionats amb informació prèviament emmagatzemada (cerca, aprenentatge, ...). Descriure els principals abordatges al problema de la recomanació d'items i els avantatges i inconvenients de cadascun. Descriure els principals algorismes de cadascun dels abordatges. Ser capaç d'implementar-ne versions bàsiques des de zero, o versions avançades damunt de sistemes de processament massius de dades. Avaluar l'efectivitat de sistemes de recomanació, tant en termes de mesures dures com de mesures toves com ara la satisfacció de l'usuari. Decidir les formes de recomanació més adequades a escenaris reals simples, incloent la caracterització dels usuaris potencials.

5.Recórrer de manera a paradigmes algorísmics coneguts per tractar amb problemes de dades caracteritzats per alt volum i alta velocitat. S'inclouen: algorismes de "streaming" que tracten fluxos de dades amb poc temps per element, i poca memòria. Algorismes per respondre preguntes de proximitat, en particular amb informació geolocalitzada. Algorismes que utilitzen mostreig per extreure conclusions fiables sobre grans volums de dades. Integració de les tècniques vistes a la resta del curs amb tècniques algorísmiques d'altres assignatures, com ara "machine learning", "clustering" i "pattern mining". Tècniques per tractar amb dades sensibles, com ara anonimització i "privacy-preserving machine learning". "Caching" consistent i distribuït.

6.Integrar les tècniques descrites als anteriors objectius en un projecte petit però realista. Tenir la capacitat per dissenyar l'arquitectura d'un sistema complex i triar les tècniques, i les tecnologies vistes amb anterioritat durant el curs que s'hi aplicarien. L'objectiu no és finalitzar la implementació del sistema, sinó arribar a un nivell de detall del disseny que permetés encarregar a un equip de programació la seva finalització.

7.Avaluar de manera elemental les implicacions dels sistemes que s'aprenen a construir en l'assignatura pel que fa a privacitat, seguretat, ètica i drets de les persones. S'enten per "de manera elemental" ser capaç de detectar que aquestes implicacions són prou significatives com per buscar l'opinió d'un expert en la matèria, en particular en relació al RGPD i la necessitat de dur a terme anàlisis de risc i impacte.



HORES TOTALES DE DEDICACIÓ DE L'ESTUDIANTAT

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

Dedicació total: 150 h

CONTINGUTS

Cerca i anàlisi d'informació textual

Descripció:

Models booleà i vectorial. Cerca basada en paraules clau. Preprocés dels textos. Indexació. Avaluació d'estratègies de cerca. Formació de grups i classificació de textos. Models generatius (LSI, LDA).

Cerca i anàlisi en estructures enllaçades

Descripció:

La web: Algorismes d'avaluació en estructures hiperenllaçades. "Crawling" i "scraping". Xarxes socials: Mesures de centralitat. Comunitats. Influència. Web semàntica.

Recomanació

Descripció:

Sistemes recomanadors. Recomanació basada en contingut i recomanació basada en la comunitat ("collaborative filtering"). Consideracions pràctiques.

Algorismes per a dades massives

Descripció:

Resums (sketches) i fluxos de dades (streaming). Mostratge (sampling). Preguntes de proximitat. Dades geolocalitzades. "Caching" consistent i distribuït. Tractament de dades sensibles: anonimització, "end-to-end encryption" i "privacy-preserving machine learning"

ACTIVITATS

Activitat sobre el contingut "Cerca i anàlisi d'informació textual"

Descripció:

A teoria, el professor presenta les motivacions i principals conceptes, i en acabar professor i estudiants resolen conjuntament 2-3 problemes de consolidació.

A laboratori, els estudiants resolen un cas relacionat amb el contingut.

Objectius específics:

1, 3, 6, 7

Competències relacionades:

CG3. Treballar en equips i projectes multidisciplinaris relacionats amb el processat i explotació de dades complexes, interactuant fluidament amb enginyers i professionals d'altres disciplines.

CG4. Identificar oportunitats per a aplicacions innovadores orientades a dades en entorns tecnològics en contínua evolució.

CG2. Elegir i aplicar els mètodes i tècniques més adequats a un problema definit per dades que representin un repte pel seu volum, velocitat, varietat o heterogeneïtat, inclosos mètodes informàtics, matemàtics, estadístics i de processament del senyal.

CG5. Poder recórrer a coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics del futur.

CE7. Demostrar coneixement i capacitat d'aplicació de les eines necessàries per a l'emmagatzematge, el processament i l'accés a les dades.

CE6. Construir o utilitzar sistemes de processat i comprensió del llenguatge escrit, integrant-lo en altres sistemes dirigits per les dades. Dissenyar sistemes de recerca d'informació textual o hipertextual i d'anàlisi de xarxes socials.

CE1. Utilitzar amb destresa els conceptes i mètodes matemàtics subjacents als problemes de la ciència i l'enginyeria de les dades.

CE4. Utilitzar els sistemes de computació actuals, inclosos els sistemes d'alt rendiment, per al procés de grans volums de dades des del coneixement de la seva estructura, funcionament i particularitats.

CT5. Ús solvent dels recursos d'informació. Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de l'especialitat i valorar de forma crítica els resultats d'aquesta gestió.

CT6. Aprenentatge autònom. Detectar deficiències en el propi coneixement i superar-les mitjançant la reflexió crítica i l'elecció de la millor actuació per ampliar aquest coneixement.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

CT7. Tercera llengua. Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i d'acord amb les necessitats que tindran els titulats i titulades.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB4. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.

CB2. Que els estudiants sàpiguen aplicar els seus coneixements al seu treball o vocació d'una manera professional i posseeixin les competències que solen demostrar-se mitjançant l'elaboració i defensa d'arguments i la resolució de problemes dins la seva àrea d'estudi.

Dedicació: 24h

Grup gran/Teoria: 6h

Grup petit/Laboratori: 6h

Aprenentatge autònom: 12h

Activitat sobre el contingut "Cerca i anàlisi en estructures enllaçades"

Descripció:

A teoria, el professor presenta les motivacions i principals conceptes, i en acabar professor i estudiants resolen conjuntament 2-3 problemes de consolidació.

A laboratori, els estudiants resolen un cas relacionat amb el contingut.

Objectius específics:

2, 6, 7

Competències relacionades:

CG3. Treballar en equips i projectes multidisciplinaris relacionats amb el processat i explotació de dades complexes, interactuant fluidament amb enginyers i professionals d'altres disciplines.

CG4. Identificar oportunitats per a aplicacions innovadores orientades a dades en entorns tecnològics en contínua evolució.

CG2. Elegir i aplicar els mètodes i tècniques més adequats a un problema definit per dades que representin un repte pel seu volum, velocitat, varietat o heterogeneïtat, inclosos mètodes informàtics, matemàtics, estadístics i de processament del senyal.

CG5. Poder recórrer a coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics del futur.

CE7. Demostrar coneixement i capacitat d'aplicació de les eines necessàries per a l'emmagatzematge, el processament i l'accés a les dades.

CE6. Construir o utilitzar sistemes de processat i comprensió del llenguatge escrit, integrant-lo en altres sistemes dirigits per les dades. Dissenyar sistemes de recerca d'informació textual o hipertextual i d'anàlisi de xarxes socials.

CE1. Utilitzar amb destresa els conceptes i mètodes matemàtics subjacents als problemes de la ciència i l'enginyeria de les dades.

CE4. Utilitzar els sistemes de computació actuals, inclosos els sistemes d'alt rendiment, per al procés de grans volums de dades des del coneixement de la seva estructura, funcionament i particularitats.

CT5. Ús solvent dels recursos d'informació. Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de l'especialitat i valorar de forma crítica els resultats d'aquesta gestió.

CT6. Aprenentatge autònom. Detectar deficiències en el propi coneixement i superar-les mitjançant la reflexió crítica i l'elecció de la millor actuació per ampliar aquest coneixement.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

CT7. Tercera llengua. Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i d'acord amb les necessitats que tindran els titulats i titulades.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB4. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.

CB2. Que els estudiants sàpiguen aplicar els seus coneixements al seu treball o vocació d'una manera professional i posseeixin les competències que solen demostrar-se mitjançant l'elaboració i defensa d'arguments i la resolució de problemes dins la seva àrea d'estudi.

Dedicació: 24h

Grup gran/Teoria: 6h

Grup petit/Laboratori: 6h

Aprenentatge autònom: 12h



Activitat sobre el tema "Recomanació"

Descripció:

A teoria, el professor presenta les motivacions i principals conceptes, i en acabar professor i estudiants resolen conjuntament 2-3 problemes de consolidació.

A laboratori, els estudiants resolen un cas relacionat amb el contingut.

Objectius específics:

4, 6, 7

Competències relacionades:

CG3. Treballar en equips i projectes multidisciplinaris relacionats amb el processat i explotació de dades complexes, interactuant fluidament amb enginyers i professionals d'altres disciplines.

CG4. Identificar oportunitats per a aplicacions innovadores orientades a dades en entorns tecnològics en contínua evolució.

CG2. Elegir i aplicar els mètodes i tècniques més adequats a un problema definit per dades que representin un repte pel seu volum, velocitat, varietat o heterogeneïtat, inclosos mètodes informàtics, matemàtics, estadístics i de processament del senyal.

CG5. Poder recórrer a coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics del futur.

CE7. Demostrar coneixement i capacitat d'aplicació de les eines necessàries per a l'emmagatzematge, el processament i l'accés a les dades.

CE6. Construir o utilitzar sistemes de processat i comprensió del llenguatge escrit, integrant-lo en altres sistemes dirigits per les dades. Dissenyar sistemes de recerca d'informació textual o hipertextual i d'anàlisi de xarxes socials.

CE1. Utilitzar amb destresa els conceptes i mètodes matemàtics subjacents als problemes de la ciència i l'enginyeria de les dades.

CE4. Utilitzar els sistemes de computació actuals, inclosos els sistemes d'alt rendiment, per al procés de grans volums de dades des del coneixement de la seva estructura, funcionament i particularitats.

CT5. Ús solvent dels recursos d'informació. Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de l'especialitat i valorar de forma crítica els resultats d'aquesta gestió.

CT6. Aprenentatge autònom. Detectar deficiències en el propi coneixement i superar-les mitjançant la reflexió crítica i l'elecció de la millor actuació per ampliar aquest coneixement.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

CT7. Tercera llengua. Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i d'acord amb les necessitats que tindran els titulats i titulades.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloquin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB4. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.

CB2. Que els estudiants sàpiguen aplicar els seus coneixements al seu treball o vocació d'una manera professional i posseeixin les competències que solen demostrar-se mitjançant l'elaboració i defensa d'arguments i la resolució de problemes dins la seva àrea d'estudi.

Dedicació: 16h

Grup gran/Teoria: 4h

Grup petit/Laboratori: 4h

Aprenentatge autònom: 8h



Activitat sobre el contingut "Algorismes per a dades massives"

Descripció:

A teoria, el professor presenta les motivacions i principals conceptes, i en acabar professor i estudiants resolen conjuntament 2-3 problemes de consolidació.

A laboratori, els estudiants resolen un cas relacionat amb el contingut.

Objectius específics:

5, 6, 7

Competències relacionades:

CG3. Treballar en equips i projectes multidisciplinaris relacionats amb el processat i explotació de dades complexes, interactuant fluidament amb enginyers i professionals d'altres disciplines.

CG4. Identificar oportunitats per a aplicacions innovadores orientades a dades en entorns tecnològics en contínua evolució.

CG2. Elegir i aplicar els mètodes i tècniques més adequats a un problema definit per dades que representin un repte pel seu volum, velocitat, varietat o heterogeneïtat, inclosos mètodes informàtics, matemàtics, estadístics i de processament del senyal.

CG5. Poder recórrer a coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics del futur.

CE7. Demostrar coneixement i capacitat d'aplicació de les eines necessàries per a l'emmagatzematge, el processament i l'accés a les dades.

CE6. Construir o utilitzar sistemes de processat i comprensió del llenguatge escrit, integrant-lo en altres sistemes dirigits per les dades. Dissenyar sistemes de recerca d'informació textual o hipertextual i d'anàlisi de xarxes socials.

CE1. Utilitzar amb destresa els conceptes i mètodes matemàtics subjacents als problemes de la ciència i l'enginyeria de les dades.

CE4. Utilitzar els sistemes de computació actuals, inclosos els sistemes d'alt rendiment, per al procés de grans volums de dades des del coneixement de la seva estructura, funcionament i particularitats.

CT5. Ús solvent dels recursos d'informació. Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de l'especialitat i valorar de forma crítica els resultats d'aquesta gestió.

CT6. Aprenentatge autònom. Detectar deficiències en el propi coneixement i superar-les mitjançant la reflexió crítica i l'elecció de la millor actuació per ampliar aquest coneixement.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

CT7. Tercera llengua. Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i d'acord amb les necessitats que tindran els titulats i titulades.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB4. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.

CB2. Que els estudiants sàpiguen aplicar els seus coneixements al seu treball o vocació d'una manera professional i posseeixin les competències que solen demostrar-se mitjançant l'elaboració i defensa d'arguments i la resolució de problemes dins la seva àrea d'estudi.

Dedicació: 34h

Grup gran/Teoria: 8h

Grup petit/Laboratori: 8h

Aprenentatge autònom: 18h

Integració. Construcció de sistemes reals. Implicacions en privacitat, seguretat i drets de les persones.

Descripció:

A teoria, el professor presenta les motivacions i principals conceptes, i en acabar professor i estudiants resolen conjuntament 2-3 problemes de consolidació.

A laboratori, els estudiants resolen un cas relacionat amb el contingut.

Objectius específics:

6, 7

Competències relacionades:

CG3. Treballar en equips i projectes multidisciplinaris relacionats amb el processat i explotació de dades complexes, interactuant fluidament amb enginyers i professionals d'altres disciplines.

CG4. Identificar oportunitats per a aplicacions innovadores orientades a dades en entorns tecnològics en contínua evolució.

CG2. Elegir i aplicar els mètodes i tècniques més adequats a un problema definit per dades que representin un repte pel seu volum, velocitat, varietat o heterogeneïtat, inclosos mètodes informàtics, matemàtics, estadístics i de processament del senyal.

CG5. Poder recórrer a coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics del futur.

CE7. Demostrar coneixement i capacitat d'aplicació de les eines necessàries per a l'emmagatzematge, el processament i l'accés a les dades.

CE6. Construir o utilitzar sistemes de processat i comprensió del llenguatge escrit, integrant-lo en altres sistemes dirigits per les dades. Dissenyar sistemes de recerca d'informació textual o hipertextual i d'anàlisi de xarxes socials.

CE1. Utilitzar amb destresa els conceptes i mètodes matemàtics subjacents als problemes de la ciència i l'enginyeria de les dades.

CE4. Utilitzar els sistemes de computació actuals, inclosos els sistemes d'alt rendiment, per al procés de grans volums de dades des del coneixement de la seva estructura, funcionament i particularitats.

CT5. Ús solvent dels recursos d'informació. Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de l'especialitat i valorar de forma crítica els resultats d'aquesta gestió.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

CT7. Tercera llengua. Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i d'acord amb les necessitats que tindran els titulats i titulades.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB4. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.

CB2. Que els estudiants sàpiguen aplicar els seus coneixements al seu treball o vocació d'una manera professional i posseeixin les competències que solen demostrar-se mitjançant l'elaboració i defensa d'arguments i la resolució de problemes dins la seva àrea d'estudi.

Dedicació: 16h

Grup gran/Teoria: 4h

Grup petit/Laboratori: 4h

Aprenentatge autònom: 8h



Resolució de l'examen parcial

Descripció:

Resolució de l'examen final fora de classe. L'estudiant tindrà 1 setmana per lliurar la seva solució des que es publica l'enunciat.

Objectius específics:

1, 2, 3

Competències relacionades:

CG3. Treballar en equips i projectes multidisciplinaris relacionats amb el processat i explotació de dades complexes, interactuant fluidament amb enginyers i professionals d'altres disciplines.

CG4. Identificar oportunitats per a aplicacions innovadores orientades a dades en entorns tecnològics en contínua evolució.

CG2. Elegir i aplicar els mètodes i tècniques més adequats a un problema definit per dades que representin un repte pel seu volum, velocitat, varietat o heterogeneïtat, inclosos mètodes informàtics, matemàtics, estadístics i de processament del senyal.

CG5. Poder recórrer a coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics del futur.

CE7. Demostrar coneixement i capacitat d'aplicació de les eines necessàries per a l'emmagatzematge, el processament i l'accés a les dades.

CE6. Construir o utilitzar sistemes de processat i comprensió del llenguatge escrit, integrant-lo en altres sistemes dirigits per les dades. Dissenyar sistemes de recerca d'informació textual o hipertextual i d'anàlisi de xarxes socials.

CE1. Utilitzar amb destresa els conceptes i mètodes matemàtics subjacents als problemes de la ciència i l'enginyeria de les dades.

CE4. Utilitzar els sistemes de computació actuals, inclosos els sistemes d'alt rendiment, per al procés de grans volums de dades des del coneixement de la seva estructura, funcionament i particularitats.

CT5. Ús solvent dels recursos d'informació. Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de l'especialitat i valorar de forma crítica els resultats d'aquesta gestió.

CT6. Aprenentatge autònom. Detectar deficiències en el propi coneixement i superar-les mitjançant la reflexió crítica i l'elecció de la millor actuació per ampliar aquest coneixement.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

CT7. Tercera llengua. Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i d'acord amb les necessitats que tindran els titulats i titulades.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB4. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.

CB2. Que els estudiants sàpiguen aplicar els seus coneixements al seu treball o vocació d'una manera professional i posseeixin les competències que solen demostrar-se mitjançant l'elaboració i defensa d'arguments i la resolució de problemes dins la seva àrea d'estudi.

Dedicació: 13h

Activitats dirigides: 3h

Aprenentatge autònom: 10h



Examen final

Descripció:

Resolució de l'examen final, un cop acabat el període de docència, en una aula i un temps determinat.

Objectius específics:

1, 2, 3, 4, 5, 7

Competències relacionades:

CG3. Treballar en equips i projectes multidisciplinaris relacionats amb el processat i explotació de dades complexes, interactuant fluidament amb enginyers i professionals d'altres disciplines.

CG4. Identificar oportunitats per a aplicacions innovadores orientades a dades en entorns tecnològics en contínua evolució.

CG2. Elegir i aplicar els mètodes i tècniques més adequats a un problema definit per dades que representin un repte pel seu volum, velocitat, varietat o heterogeneïtat, inclosos mètodes informàtics, matemàtics, estadístics i de processament del senyal.

CG5. Poder recórrer a coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics del futur.

CE7. Demostrar coneixement i capacitat d'aplicació de les eines necessàries per a l'emmagatzematge, el processament i l'accés a les dades.

CE6. Construir o utilitzar sistemes de processat i comprensió del llenguatge escrit, integrant-lo en altres sistemes dirigits per les dades. Dissenyar sistemes de recerca d'informació textual o hipertextual i d'anàlisi de xarxes socials.

CE1. Utilitzar amb destresa els conceptes i mètodes matemàtics subjacents als problemes de la ciència i l'enginyeria de les dades.

CE4. Utilitzar els sistemes de computació actuals, inclosos els sistemes d'alt rendiment, per al procés de grans volums de dades des del coneixement de la seva estructura, funcionament i particularitats.

CT5. Ús solvent dels recursos d'informació. Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de l'especialitat i valorar de forma crítica els resultats d'aquesta gestió.

CT6. Aprenentatge autònom. Detectar deficiències en el propi coneixement i superar-les mitjançant la reflexió crítica i l'elecció de la millor actuació per ampliar aquest coneixement.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

CT7. Tercera llengua. Conèixer una tercera llengua, preferentment l'anglès, amb un nivell adequat oral i escrit i d'acord amb les necessitats que tindran els titulats i titulades.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB4. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.

CB2. Que els estudiants sàpiguen aplicar els seus coneixements al seu treball o vocació d'una manera professional i posseeixin les competències que solen demostrar-se mitjançant l'elaboració i defensa d'arguments i la resolució de problemes dins la seva àrea d'estudi.

Dedicació: 15h

Activitats dirigides: 3h

Aprenentatge autònom: 12h

SISTEMA DE QUALIFICACIÓ

P = nota d'un examen parcial take-home a mig curs.

F = nota d'un examen final fet a final de curs.

L = nota dels informes a lliurar sobre les sessions de laboratori.

La nota final es calcularà com $20\% P + 40\% F + 40\% L$.

La nota de la competència transversal CT6 "aprenentatge autònom" es calcularà en base a un tema proposat durant el curs que els i les estudiants hauran d'estudiar pel seu compte. S'avaluarà si s'han assimilats els conceptes d'aquest tema a partir de preguntes als examens i/o a informació plasmada als informes de les pràctiques.



BIBLIOGRAFIA

Bàsica:

- Leskovec, J.; Rajaraman, A.; Ullman, J.D. Mining of massive datasets. 3rd ed. Cambridge: Cambridge University Press, 2020. ISBN 9781108476348.
- Stephens-Davidowitz, S. Everybody lies : what the internet can tell us about who we really are. Bloomsbury Publishing, 2018. ISBN 9781408894736.
- Baeza-Yates, R.; Ribeiro-Neto, B. Modern information retrieval: the concepts and technology behind search. 2nd ed. Addison-Wesley / Pearson, 2011. ISBN 9780321416919.