

Guia docent

270420 - APRNS - Aprenentatge per Reforç i No Supervisat

Última modificació: 10/07/2024

Unitat responsable: Facultat d'Informàtica de Barcelona
Unitat que imparteix: 723 - CS - Departament de Ciències de la Computació.
Titulació: GRAU EN INTEL·LIGÈNCIA ARTIFICIAL (Pla 2021). (Assignatura obligatòria).
Curs: 2024 **Crèdits ECTS:** 6.0 **Idiomes:** Català, Castellà

PROFESSORAT

Professorat responsable: JAVIER BÉJAR ALONSO - MARIO MARTÍN MUÑOZ

Altres: Primer quadrimestre:
JAVIER BÉJAR ALONSO - 11, 12
MARIO MARTÍN MUÑOZ - 11, 12

CAPACITATS PRÈVIES

Coneixements bàsics de Deep Learning i de Machine Learning.

COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

Específiques:

CE18. Adquirir i desenvolupar tècniques d'aprenentatge computacional i dissenyar i implementar aplicacions i sistemes que les utilitzin, incloent les dedicades a extracció automàtica d'informació i coneixement a partir de grans volums de dades.

Genèriques:

CG2. Utilitzar els coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics de el futur.

CG4. Raonar, analitzant la realitat i dissenyant algorismes i formulacions que la modelin. Identificar problemes i construir solucions algorísmiques o matemàtiques vàlides, eventualment noves, integrant el coneixement multidisciplinari necessari, valorant diferents alternatives amb esperit crític, justificant les decisions preses, interpretant i sintetitzant els resultats en el context de l'domini d'aplicació i establint generalitzacions metodològiques a partir de aplicacions concretes.

Transversals:

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.

Bàsiques:

CB5. Que els estudiants hagin desenvolupat aquelles habilitats d'aprenentatge necessàries per emprendre estudis posteriors amb un alt grau d'autonomia

METODOLOGIES DOCENTS

Les classes estan dividides en sessions de teoria, problemes i laboratori.

A les sessions de teoria es desenvoluparan els coneixements de l'assignatura, intercalant l'exposició de nou material teòric amb exemples i la interacció amb els alumnes per tal de discutir els conceptes.

A les classes de laboratori es desenvoluparan petites pràctiques utilitzant eines i fent servir llibreries específiques que permetran practicar i reforçar els coneixements de les classes de teoria.

OBJECTIUS D'APRENTATGE DE L'ASSIGNATURA

1. Conèixer quins tipus de problemes es poden modelitzar com un problema d'aprenentatge per reforç i identificar les tècniques que es poden aplicar per resoldre-les
2. Entendre la necessitat, fonaments i particularitats de l'aprenentatge de comportaments i les seves diferències respecte a l'aprenentatge automàtic de classificació i no-supervisat.
3. Conèixer els algorismes més importants i estat de l'art en l'àrea d'aprenentatge per reforçament
4. Saber formalitzar computacionalment un problema del món real com a aprenentatge per reforçament i saber implementar en els entorns més actuals els algorismes d'aprenentatge que els resoln
5. Coneixer els problemes que es poden modelitzar amb algorismes no supervisats profunds
6. Entendre les particularitats dels algorismes no supervisats profunds
7. Coneixer els algorismes més importants i l'estat de l'art del aprenentatge no supervisat profund
8. Saber implementar i aplicar a un problema algorismes d'aprenentatge profund fent servir el entorn més actuals

HORES TOTALES DE DEDICACIÓ DE L'ESTUDIANTAT

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

Dedicació total: 150 h

CONTINGUTS

Introducció: L'aprenentatge de comportaments en agents i descripció dels principals elements en l'aprenentatge de reforç

Descripció:

Intuïció, motivació i definició del marc de l'aprenentatge per reforçament (RL). Elements clau en RL.

Trobar polítiques òptimes mitjançant la programació dinàmica

Descripció:

Com aprendre un comportament amb coneixement complet del model del món: solució algebraica, avaluació iterada de polítiques i avaluació iterada de valors.

Introducció als enfocaments sense models del món.

Descripció:

Algorismes bàsics per a l'aprenentatge de reforç: Montecarlo, Q-learning, Sarsa, TD (λ). La necessitat d'exploració. Diferències entre els mètodes On-policy i Off-policy.

Aproximació de funcions en aprenentatge per reforçament

Descripció:

La necessitat de l'aproximació de funcions i mètodes incrementals en RL. L'aproximació del descens del gradient. RL amb aproximació de funció lineal. La triada mortal per a l'aproximació de funcions en RL. Mètodes per lots i xarxes neuronals per a l'aproximació de funcions.

Aprentatge per reforç profund (DRL)

Descripció:

Introducció de DL a RL. Com tractar la mortal tríada amb l'algorisme DQN. Aplicació de DQN al cas dels jocs Atari. Evolucions de l'algorisme DQN: Double DQN, Prioritized Experience Replay, aprenentatge en múltiples passos i funcions de valor distribuïdes. Rainbow: l'algoritme d'última generació per un espai d'acció discret.

Mètodes del gradient en la política

Descripció:

Què fer en espais d'acció continus. Com les polítiques probabilístiques permeten aplicar el mètode de gradient directament a la xarxa de polítiques. L'algoritme REINFORCE. Els algoritmes Actor-Critic. Algoritmes d'última generació en espais d'acció continus: DDPG, TD3 i SAC.

Temes avançats: Com tractar el problema del reforç espars

Descripció:

El problema de la recompensa esparsa. Introducció a tècniques avançades d'exploració: curiositat i empoderament en RL. Introducció a l'aprenentatge curricular per facilitar l'aprenentatge de l'objectiu. RL jeràrquic per aprendre tasques complexes. L'aprenentatge de les funcions de valor universals i Hindsight Experience Replay (HER).

Temes avançats: Aprentatge de reforç en el marc multi-agent

Descripció:

Aprentatge de conductes en un entorn on actuen diversos agents. Aprentatge de conductes cooperatives, Aprentatge de conductes competitives i casos mixtos. Algoritmes d'última generació. El cas especial dels jocs: el cas AlfaGo i l'extensió a Alfa-Zero.

Introducció: L'aprenentatge no supervisat profund

Descripció:

Introducció a la necessitat del aprenentatge no supervisat profund i les seves aplicacions

Models autoregressius

Descripció:

Introducció a l'aprenentatge de distribucions de probabilitat definides como distribucions autoregresives i principals models

Fluxes normalitzants

Descripció:

Introducció als fluxes normalitzants per a l'aprenentatge de distribucions de probabilitat

Models de variables latents

Descripció:

Introducció als models basats en variables latents i als autocodificadors variacionals

Xarxes adversàries generatives

Descripció:

Introducció a les xarxes adversàries generatives, generació condicionada i sense condicionar, separació d'atributs

Xarxes de difusió

Descripció:

Introducció a models basats en difusió de soroll, xarxes per eliminació de soroll, condicionament, generació multimodal

Aprenentatge per autosupervisió

Descripció:

Introducció al aprenentatge per autosupervisió pel entrenament de xarxes generadores de característiques, mètodes contrastius i no contrastius, enmascarament

ACTIVITATS

Introducció: L'aprenentatge de comportaments en agents i descripció dels principals elements en l'aprenentatge de reforç

Dedicació: 10h

Aprenentatge autònom: 6h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Trobar polítiques òptimes mitjançant la programació dinàmica

Descripció:

Com aprendre un comportament amb coneixement complet del model del món: solució algebraica, avaluació iterada de polítiques i avaluació iterada de valors.

Dedicació: 10h

Aprenentatge autònom: 6h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Introducció als enfocaments sense models. Monte-Carlo, Q-learning, Sarsa, TD (λ)

Descripció:

Desenvolupament del tema corresponent de l'assignatura

Dedicació: 10h

Aprenentatge autònom: 6h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Aproximació de funciones en RL

Dedicació: 10h

Aprenentatge autònom: 6h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Aprenentatge per reforç profund

Dedicació: 10h

Aprenentatge autònom: 6h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Mètodes del gradient en la política

Descripció:

Què fer en espais d'acció continus. Com les polítiques probabilístiques permeten aplicar el mètode de gradient directament a la xarxa de polítiques. L'algoritme REINFORCE. Els algorismes Actor-Critic. Algorismes d'última generació en espais d'acció continus: DDPG, TD3 i SAC.

Dedicació: 10h

Aprenentatge autònom: 6h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Temes avançats: Com tractar el problema del reforç espars

Descripció:

El problema de la recompensa esparsa. Introducció a tècniques avançades d'exploració: curiositat i empoderament en RL. Introducció a l'aprenentatge curricular per facilitar l'aprenentatge de l'objectiu. RL jeràrquic per aprendre tasques complexes. L'aprenentatge de les funcions de valor universals i Hindsight Experience Replay (HER).

Dedicació: 10h

Aprenentatge autònom: 6h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Temes avançats: Aprenentatge de reforç en el marc multi-agent

Descripció:

Aprenentatge de conductes en un entorn on actuen diversos agents. Aprenentatge de conductes cooperatives, Aprenentatge de conductes competitives i casos mixtos. Algorismes d'última generació. El cas especial dels jocs: el cas AlfaGo i l'extensió a Alfa-Zero.

Dedicació: 13h

Aprenentatge autònom: 9h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Control de la part d'aprenentatge per reforçament

Objectius específics:

1, 2, 3, 4

Competències relacionades:

CE18. Adquirir i desenvolupar tècniques d'aprenentatge computacional i dissenyar i implementar aplicacions i sistemes que les utilitzin, incloent les dedicades a extracció automàtica d'informació i coneixement a partir de grans volums de dades.

CG2. Utilitzar els coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics de el futur.

CG4. Raonar, analitzant la realitat i dissenyant algoritmes i formulacions que la modelin. Identificar problemes i construir solucions algorísmiques o matemàtiques vàlides, eventualment noves, integrant el coneixement multidisciplinari necessari, valorant diferents alternatives amb esperit crític, justificant les decisions preses, interpretant i sintetitzant els resultats en el context de l'àmbit d'aplicació i establint generalitzacions metodològiques a partir de aplicacions concretes.

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.

Dedicació: 2h

Activitats dirigides: 2h

Introducció: L'aprenentatge no supervisat profund

Descripció:

Introducció a la necessitat del aprenentatge no supervisat profund i les seves aplicacions

Dedicació: 10h

Aprenentatge autònom: 6h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Models autoregressius

Descripció:

Introducció a l'aprenentatge de distribucions de probabilitat definides como distribucions autoregressives i principals models

Dedicació: 10h

Aprenentatge autònom: 6h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Fluxes normalitzants

Descripció:

Introducció als fluxes normalitzants per a l'aprenentatge de distribucions de probabilitat

Dedicació: 10h

Aprenentatge autònom: 6h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Models de variables latents

Descripció:

Introducció als models basats en variables latents i als autocodificadors variacionals

Dedicació: 10h

Aprenentatge autònom: 6h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Xarxes adversàries generatives

Descripció:

Introducció a les xarxes adversàries generatives, generació condicionada i sense condicionar, separació d'atributs

Dedicació: 10h

Aprenentatge autònom: 6h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Xarxes de difusió i Aprenentatge per autosupervisió

Descripció:

Introducció a models basats en difusió de soroll, xarxes per eliminació de soroll, condicionament, generació multimodal

Dedicació: 15h

Aprenentatge autònom: 9h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 4h

Control del temari d'aprenentatge no supervisat

Objectius específics:

5, 6, 7, 8

Competències relacionades:

CB5. Que els estudiants hagin desenvolupat aquelles habilitats d'aprenentatge necessàries per emprendre estudis posteriors amb un alt grau d'autonomia

CE18. Adquirir i desenvolupar tècniques d'aprenentatge computacional i dissenyar i implementar aplicacions i sistemes que les utilitzin, incloent les dedicades a extracció automàtica d'informació i coneixement a partir de grans volums de dades.

CG2. Utilitzar els coneixements fonamentals i metodologies de treball sòlides adquirits durant els estudis per adaptar-se als nous escenaris tecnològics de el futur.

CG4. Raonar, analitzant la realitat i dissenyant algorismes i formulacions que la modelin. Identificar problemes i construir solucions algorísmiques o matemàtiques vàlides, eventualment noves, integrant el coneixement multidisciplinari necessari, valorant diferents alternatives amb esperit crític, justificant les decisions preses, interpretant i sintetitzant els resultats en el context de l'àmbit d'aplicació i establint generalitzacions metodològiques a partir de aplicacions concretes.

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.

SISTEMA DE QUALIFICACIÓ

L'assignatura comprendrà els següents actes avaluatoris:

- Informes de les activitats de laboratori, que caldrà haver lliurat dins un termini indicat per a cada sessió (orientativament, 2 setmanes). A partir d'una mitjana ponderada de les notes d'aquests informes es calcularà una nota de laboratori, L.
- Un primer examen parcial, fet cap a meitat del curs, de la matèria vista fins llavors. Sigui P1 la nota obtinguda en aquest examen.
- En el dia designat dins del període d'exàmens, un segon examen parcial de la matèria no coberta pel primer parcial. Sigui P2 la nota obtinguda en aquest examen.

Les tres notes L, P1, P2 són entre 0 i 10. La nota final de l'assignatura serà:

$$0.4*L + 0.3*P1 + 0.3*P2$$

BIBLIOGRAFIA

Bàsica:

- Sutton, Richard S; Barto, Andrew G. Reinforcement learning : an introduction. Second edition. Cambridge, Massachusetts: The MIT Press, [2020]. ISBN 9780262039246.
- Morales, Miguel. Grokking deep reinforcement learning. Shelter Island: Manning Publications, 2020. ISBN 9781617295454.
- Foster, D. Generative deep learning: teaching machines to paint, write, compose, and play. 2nd ed. Sebastopol: O'Reilly Media, Incorporated, 2023. ISBN 9781098134143.
- Cheong, S.Y. Hands-on image generation with TensorFlow: a practical guide to generating images and videos using deep learning. Birmingham: Packt Publishing, 2020. ISBN 9781838821104.

Complementària:

- Zai, Alexander; Brown, Brandon. Deep reinforcement learning in action. Shelter Island, NY: Manning Publications Co, 2020. ISBN 9781617295430.
- Babcock, J.; Bali, R. Generative AI with Python and TensorFlow 2: harness the power of generative models to create images, text, and music. Birmingham, England ; Mumbai: Packt Publishing, 2021. ISBN 9781800208506.