

Guia docent

230329 - MLR - Aprenentatge Automàtic per Reforç

Última modificació: 29/04/2020

Unitat responsable: Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona
Unitat que imparteix: 739 - TSC - Departament de Teoria del Senyal i Comunicacions.

Titulació: GRAU EN CIÈNCIES I TECNOLOGIES DE TELECOMUNICACIÓ (Pla 2010). (Assignatura optativa).
GRAU EN ENGINYERIA DE SISTEMES AUDIOVISUALS (Pla 2009). (Assignatura optativa).
GRAU EN ENGINYERIA DE SISTEMES ELECTRÒNICS (Pla 2009). (Assignatura optativa).
GRAU EN ENGINYERIA DE SISTEMES DE TELECOMUNICACIÓ (Pla 2010). (Assignatura optativa).
GRAU EN ENGINYERIA TELEMÀTICA (Pla 2010). (Assignatura optativa).
GRAU EN ENGINYERIA DE TECNOLOGIES I SERVEIS DE TELECOMUNICACIÓ (Pla 2015). (Assignatura optativa).

Curs: 2020 **Crèdits ECTS:** 2.0 **Idiomes:** Anglès, Català

PROFESSORAT

Professorat responsable: Vidal Manzano, Jose
Altres: Cabrera Bean, Margarita Asuncion
Giró Nieto, Xavier

CAPACITATS PRÈVIES

Àlgebra, Probabilitat i processos, Senyals i sistemes

REQUISITS

Àlgebra, Probabilitat i processos, Senyals i sistemes

COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

Genèriques:

08 CRPE. CAPACITAT PER IDENTIFICAR, FORMULAR I RESOLDRE PROBLEMES D'ENGINYERIA. Capacitat per plantejar i resoldre problemes d'enginyeria en l'àmbit TIC amb iniciativa, presa de decisions i creativitat. Desenvolupar un mètode d'anàlisi i solució de problemes sistemàtic i creatiu.

Transversals:

04 COE N2. COMUNICACIÓ EFICAC ORAL I ESCRITA - Nivell 2: Utilitzar estratègies per preparar i dur a terme les presentacions orals i redactar textos i documents amb un contingut coherent, una estructura i un estil adequats i un bon nivell ortogràfic i gramatical.

06 URI N1. ÚS SOLVENT DELS RECURSOS D'INFORMACIÓ - Nivell 1: Identificar les pròpies necessitats d'informació i utilitzar les col·leccions, els espais i els serveis disponibles per dissenyar i executar cerques simples adequades a l'àmbit temàtic.

06 URI N2. ÚS SOLVENT DELS RECURSOS D'INFORMACIÓ - Nivell 2: Després d'identificar les diferents parts d'un document acadèmic i d'organitzar-ne les referències bibliogràfiques, dissenyar-ne i executar-ne una bona estratègia de cerca avançada amb recursos d'informació especialitzats, seleccionant-hi la informació pertinent tenint en compte criteris de rellevància i qualitat.

06 URI N3. ÚS SOLVENT DELS RECURSOS D'INFORMACIÓ - Nivell 3: Planificar i utilitzar la informació necessària per a un treball acadèmic (per exemple, per al treball de fi de grau) a partir d'una reflexió crítica sobre els recursos d'informació utilitzats.

METODOLOGIES DOCENTS

Classes teòriques i laboratori



OBJECTIUS D'APRENTATGE DE L'ASSIGNATURA

Dominar els principis de l'aprenentatge per reforç com a eina d'intel·ligència artificial basada en la interacció de la màquina amb el seu entorn i que està en la base de sistemes com els vehicles autònoms, els softwares que juguen a escacs o go, o l'organització de sistemes complexos de comunicacions. Es treballarà en la seva implementació pràctica i l'evaluació en casos concrets.

HORES TOTALS DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores grup gran	20,0	40.00
Hores aprenentatge autònom	30,0	60.00

Dedicació total: 50 h

CONTINGUTS

1. Introducció a l'aprenentatge per reforç

Descripció:

Descriure amb exemples els conceptes fonamentals i els problemes que podrem resoldre.

Dedicació: 2h 30m

Grup gran/Teoria: 2h 30m

2. Processos de decisió de Markov

Descripció:

- L'interfaç agent-entorn
- Objectius i recompenses
- Processos de decisió de Markov
- Funcions de valor y optimalitat: equació de Bellman

Dedicació: 2h 30m

Grup gran/Teoria: 2h 30m

3. Programació dinàmica

Descripció:

- Avaluació de polítiques, millores i iteracions
- Programació dinàmica basada en els MDP

Dedicació: 2h 30m

Grup gran/Teoria: 2h 30m



4. Mètodes de Monte-Carlo

Descripció:

- Aprenentatge sense models.
- Mètodes de Montecarlo de primera visita
- Mètodes de Montecarlo de cada visita
- Exploració i explotació
- Mètodes on-policy i off-policy

Dedicació: 2h 30m

Grup gran/Teoria: 2h 30m

5. Aprenentatge per diferencia temporal i funcions Q

Descripció:

- Aprenentatge sense models fent servir diferències temporals
- Q-learning i accions discretes
- Jocs

Dedicació: 2h 30m

Grup gran/Teoria: 2h 30m

6. Mètodes basats en polítiques de gradient

Descripció:

- Polítiques de gradient

Dedicació: 2h 30m

Grup gran/Teoria: 2h 30m

7. Aprenentatge profund per reforç

Descripció:

Modelatge de funcions Q amb xarxes neuronals profundes

Dedicació: 2h 30m

Grup gran/Teoria: 2h 30m

8. Laboratori i aplicacions

Descripció:

Laboratoris en Matlab i/o Python que s'aniran distribuïnt al llarg dels temes

- Assignació dinàmica de recursos
- Blackjack
- Programador de tasques de taller

Dedicació: 2h 30m

Grup gran/Teoria: 2h 30m



SISTEMA DE QUALIFICACIÓ

Examen i avaluació de les pràctiques

BIBLIOGRAFIA

Bàsica:

- Sutton, R.S.; Barto, A.G. Reinforcement learning: an introduction [en línia]. 2nd ed. Cambridge, Mass.: MIT Press, 2018 [Consulta: 14/09/2020]. Disponible a: <https://ebookcentral.proquest.com/lib/upcatalunya-ebooks/detail.action?docID=6260249>. ISBN 9780262039246.