

Guia docent (versió completa)

230817 - ARAP - Aprenentatge per Reforç i Aprenentatge Profund

Última modificació: 17/05/2022

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 ENGINYERIA DE TECNOLOGIES I SERVEIS DE TELECOMUNICACIÓ (Pla 2015). (Assignatura optativa).

GRAU EN CIÈNCIA I ENGINYERIA DE DADES (Pla 2017). (Assignatura optativa).

GRAU EN ENGINYERIA ELECTRÒNICA DE TELECOMUNICACIÓ (Pla 2018). (Assignatura optativa).

Curs: 2022

Crèdits ECTS: 6.0

Idiomes: Català, Anglès

PROFESSORAT

Professorat responsable: Consultar aquí / See here:
<https://telecos.upc.edu/ca/estudis/curs-actual/professorat-responsables-coordinadors/responsables-assignatura>

Altres: Consultar aquí / See here:
<https://telecos.upc.edu/ca/estudis/curs-actual/professorat-responsables-coordinadors/professorat-assignat-idioma>

CAPACITATS PRÈVIES

Àlgebra, Probabilitat i processos, Senyals i sistemes, Aprenentatge automàtic i profund.

REQUISITS

Pels estudiants de GRETST i GREELEC, Introducció a l'aprenentatge profund (IDL 230325). Pels estudiants del GCED, Aprenentatge Automàtic 2 (AA2 270222).

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 EFICAÇ 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, amb aplicacions en tasques de control (ex. robòtica, vehicles autònoms), preses de decisions (ex. optimització de recursos en xarxes de comunicacions sense fils).

Dissenyar i entrenar xarxes neuronals profundes que siguin capaces d'aprendre amb poca o nula supervisió, tant per a tasques discriminatives com generatives, amb especial èmfasi en les aplicacions multimèdia (visió, llenguatge i àudio).

HORES TOTALES DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores grup gran	26,0	17.33
Hores grup petit	26,0	17.33
Hores aprenentatge autònom	98,0	65.33

Dedicació total: 150 h

CONTINGUTS

Introducció a l'aprenentatge per reforç

Descripció:

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

Dedicació: 2h

Grup gran/Teoria: 2h

El compromís exploració-explotació

Descripció:

- El compromís exploració-explotació
- Cas d'estudi: la sala de màquines escurabutxaques

Dedicació: 1h

Grup gran/Teoria: 1h

Processos de decisió de Markov (MDP)

Descripció:

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

Dedicació: 2h

Grup gran/Teoria: 2h

Programació dinàmica

Descripció:

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

Dedicació: 1h

Grup gran/Teoria: 1h

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

Aprenentatge per diferencia temporal (TD)

Descripció:

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

Dedicació: 2h 30m

Grup gran/Teoria: 2h 30m

Gradients sobre les polítiques

Descripció:

- Gradients sobre les polítiques
- Aproximació del valor de funcions
- Mètodes Actor-Crític
- Funcions baseline

Dedicació: 2h 30m

Grup gran/Teoria: 2h 30m

Aprenentatge profund per reforç

Descripció:

Modelatge de funcions Q i de les polítiques amb xarxes neuronals profundes

Dedicació: 2h 30m

Grup gran/Teoria: 2h 30m

Aprentatge profund avançat

Descripció:

- Models generatius: GANs, VAEs and Flows.
- Xarxes Neuronals Recurrents (RNNs) avançades.
- Mecanismes d'atenció i Transformers.
- Xarxes Neuronals amb Grafs.
- Aprentatge auto-supervisat i meta-aprenentatge.
- Supercomputació per a aprenentatge profund.

Dedicació: 10h

Grup gran/Teoria: 10h

Laboratoris d'aprenentatge per reforç

Descripció:

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

- Assignació dinàmica de recursos
- Blackjack
- Programador de tasques de taller
- Aprentatge-Q amb taules

Dedicació: 13h

Grup petit/Laboratori: 13h

Laboratoris d'aprenentatge profund

Descripció:

Laboratoris en PyTorch sobre:

- Aprentatge-Q amb xarxes neuronals
- Polítiques de gradient amb xarxes neuronals (REINFORCE)
- Optimitzadors per a xarxes neuronals profundes
- Models d'atenció
- Transferència del coneixement
- Xarxes Adversaries Generatives (GAN)
- Auto-codificadors variacionals (VAE)

Dedicació: 13h

Grup petit/Laboratori: 13h

SISTEMA DE QUALIFICACIÓ

Exàmens 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.

- Goodfellow, I.; Bengio, Y.; Courville, A. Deep learning [en línia]. Cambridge, Massachusetts: The MIT Press, [2016] [Consulta: 02/07/2020]. Disponible a: <http://www.deeplearningbook.org/>. ISBN 978-0262035613.



RECURSOS

Altres recursos:

Web del curs: <https://telecombcn-dl.github.io/drl-2020/>