



1. **Código:** 35679 **Nombre:** Reinforcement learning
2. **Créditos:** 6,00 **--Teoría:** 3,00 **--Prácticas:** 3,00 **Carácter:** Optativo
- Titulación:** 2314-Máster Universitario en Ingeniería de Telecomunicación
- Módulo:** 3-Módulo de Optativas **Materia:** 5-Formación Optativa
- Centro:** E.T.S.I. DE TELECOMUNICACIÓN

3. **Coordinador:** Igual García, Jorge
- Departamento:** COMUNICACIONES

4. Bibliografía

Reinforcement learning : an introduction Sutton, Richard.; Barto, Andrew G.
Deep reinforcement learning with Python : with PyTorch, TensorFlow and Sanghi, Nimish
OpenAI Gym
Deep learning with PyTorch Stevens, Eli

5. Descripción general de la asignatura

Objetivos de la asignatura

The main objective of the course is to provide students with an understanding of the fundamentals and applications of this branch of machine learning: reinforcement learning (RL). Theory classes will introduce the essential aspects of RL and a variety of learning techniques designed to address various types of real-world problems that will be applied in practical classes, allowing students to understand its internal logic and practical applications.

Hands-on practice, both classroom and laboratory, will be an integral part of the course. In them, students will apply the concepts and algorithms learned in theory through application examples of limited complexity, which will allow them to develop practical skills and consolidate their theoretical understanding.

To emphasize this practical orientation, students will learn Pytorch, one of the most advanced programming environments for deep learning that will be useful for other deep learning applications.

Contextualización de la asignatura

This course studies another of the fast growing branches of artificial intelligence: reinforcement learning, or ¿ reinforcement learning¿ (RL).

RL techniques are used to teach an agent to make decisions in an environment whose operation may be unknown. The agent is faced with making sequential decisions based on the current state of the environment and what it has learned up to that point. As a result of the decision, the agent receives a reward. The agent's decisions also have an impact on the evolution of the state of the environment. Each decision the agent makes results in a reward, and also has an impact on the evolution of the state of the environment. The main goal of RL techniques is for the agent to learn what is the best decision in each state, so that the total reward that the agent accumulates in the long run is maximized.

The knowledge and application of RL has experienced tremendous growth in recent years. These techniques are used in a wide variety of fields, such as robotics, gaming, finance, medicine and telecommunications. Some examples of application in the telecommunications field could be the following: dynamic management of network resources; activation and deactivation of network elements for energy saving; automatic selection of networks, routes or radio channels; intrusion or fault detection; and application of damage recovery or mitigation measures.

6. Conocimientos recomendados

- (35475) Inteligencia artificial para sistemas de comunicaciones y multimedia
(35677) Advanced methods of artificial vision
(35678) Signal and natural language processing with deep learning

7. Resultados

Resultados fundamentales

BA4(GE) Que los estudiantes sepan comunicar sus conclusiones ¿y los conocimientos y razones últimas que las sustentan¿ a públicos especializados y no especializados de un modo claro y sin ambigüedades;

Competencias transversales

- (4) Comunicación efectiva
- Actividades desarrolladas relacionadas con la adquisición de la competencia

Document signat electrònicament per Documento firmado electrónicamente por Electronically signed document by	UNIVERSITAT POLITÈCNICA DE VALÈNCIA	Data/Fecha/Date 06/06/2025	1 / 2	
Autenticitat verificable mitjançant Codi Segur Verificació Autenticidad verificable mediante Código Seguro Verificación Original document can be verified by Secure Verification Code	ALUX2SLI6EI https://sede.upv.es/e/Verificador			



7. Resultados

Competencias transversales

- Development of a pytorch project from start to finish on an artificial intelligence problem.
- Criterios de evaluación
- Presentation of the work done, including various support media (slides, programming code, examples, etc.)
- Resultados de Aprendizaje Específicos
- RA4.4 - Demostrar destreza en la comunicación digital utilizando medios de apoyo variados y adaptados a la situación y a la audiencia.

8. Unidades didácticas

1. Introduction to reinforcement learning
2. Model free methods
3. Function approximation and policy gradient methods
4. Deep reinforcement learning
5. Pytorch: deep learning framework

9. Método de enseñanza-aprendizaje

<u>UD</u>	<u>TA</u>	<u>SE</u>	<u>PA</u>	<u>PL</u>	<u>PC</u>	<u>PI</u>	<u>EVA</u>	<u>TP</u>	<u>TNP</u>	<u>TOTAL HORAS</u>
1	6,00	--	2,00	--	--	0,00	--	8,00	5,00	13,00
2	8,00	--	3,00	--	--	1,00	--	12,00	20,00	32,00
3	6,00	--	3,00	--	--	1,00	--	10,00	20,00	30,00
4	5,00	--	3,00	--	--	1,00	--	9,00	22,00	31,00
5	5,00	--	7,00	--	--	9,00	--	21,00	30,00	51,00
TOTAL HORAS	30,00	--	18,00	--	--	12,00	--	60,00	97,00	157,00

UD: Unidad Didáctica. TA: Teoría de Aula. SE: Seminario. PA: Práctica de Aula. PL: Práctica de Laboratorio. PC: Práctica de Campo. PI: Práctica de Informática. EVA: Actividades de Evaluación. TP: Trabajo Presencial. TNP: Trabajo No Presencial.

10. Evaluación

Descripción

- (14) Prueba escrita
(09) Proyecto

<u>Nº Actos</u>	<u>Peso (%)</u>
2	60
1	40

The evaluation consists of two parts. The first one consists on two tests at the end of some thematic units (2 acts of evaluation, 30% each of the final grade). The second one consists on the development of a project in group related to Reinforcement Learning / Pytorch (1 act of evaluation, 40%). The first part can be recovered with a global test. The second part can be recovered improving the work until it passes the evaluation.

In the case of students with dispensation of attendance obligation, the evaluation will be carried out by means of the same type of evaluation acts as the face-to-face students.

If a student has lost the right to be evaluated in an evaluation act by application of the Academic Integrity Regulations (NIA), he/she will not be able to take advantage of the continuous evaluation and must take a final test corresponding to 100% of the grade of the course.

11. Porcentaje máximo de ausencia

<u>Actividad</u>	<u>Porcentaje</u>	<u>Observaciones</u>
Teoría Aula	40	
Teoría Seminario	40	
Práctica Aula	40	
Práctica Laboratorio	40	
Práctica Informática	40	
Práctica Campo	40	

