



UNIVERSIDAD
POLITÉCNICA
DE MADRID

PROCESO DE
COORDINACIÓN DE LAS
ENSEÑANZAS PR/CL/001



E.T.S. de Ingenieros de
Telecomunicacion

ANX-PR/CL/001-01

GUÍA DE APRENDIZAJE

ASIGNATURA

93000948 - Aprendizaje De Refuerzo

PLAN DE ESTUDIOS

09AT - Master Universitario En Teoria De La Señal Y Comunicaciones

CURSO ACADÉMICO Y SEMESTRE

2024/25 - Segundo semestre

Índice

Guía de Aprendizaje

1. Datos descriptivos.....	1
2. Profesorado.....	1
3. Conocimientos previos recomendados.....	2
4. Competencias y resultados de aprendizaje.....	2
5. Descripción de la asignatura y temario.....	4
6. Cronograma.....	6
7. Actividades y criterios de evaluación.....	8
8. Recursos didácticos.....	11
9. Otra información.....	12

1. Datos descriptivos

1.1. Datos de la asignatura

Nombre de la asignatura	93000948 - Aprendizaje de Refuerzo
No de créditos	3 ECTS
Carácter	Optativa
Curso	Primer curso
Semestre	Segundo semestre
Período de impartición	Febrero-Junio
Idioma de impartición	Inglés/Castellano
Titulación	09AT - Master Universitario en Teoría de la Señal y Comunicaciones
Centro responsable de la titulación	09 - Escuela Tecnica Superior De Ingenieros De Telecomunicacion
Curso académico	2024-25

2. Profesorado

2.1. Profesorado implicado en la docencia

Nombre	Despacho	Correo electrónico	Horario de tutorías *
Julian Cabrera Quesada	C-320	julian.cabrera@upm.es	Sin horario. Arrange the meeting by email
Santiago Zazo Bello	C-326	santiago.zazo@upm.es	Sin horario. Arrange the meeting by email

Juan Parras Moral (Coordinador/a)	C-303	j.parras@upm.es	Sin horario. Arrange the meeting by email
Mateo Jose Camara Largo	C-303	mateo.camara@upm.es	Sin horario. Arrange the meeting by email

* Las horas de tutoría son orientativas y pueden sufrir modificaciones. Se deberá confirmar los horarios de tutorías con el profesorado.

3. Conocimientos previos recomendados

3.1. Asignaturas previas que se recomienda haber cursado

- Fundamentos De Optimización
- Técnicas De Optimización Para Análisis De Datos Masivos

3.2. Otros conocimientos previos recomendados para cursar la asignatura

- Statistical Signal Processing

4. Competencias y resultados de aprendizaje

4.1. Competencias

CB06 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación

CB07 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio

CB08 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

CB09 - 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

CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo

CE02 - Evaluar y sintetizar los resultados de un trabajo en equipo en proyectos relacionados con la teoría de la señal y las comunicaciones, en un entorno internacional.

CE03 - Valorar y contrastar la utilización de las diferentes técnicas disponibles para la resolución de problemas reales dentro del área de teoría de la señal y comunicaciones.

CT01 - Capacidad para comprender los contenidos de clases magistrales, conferencias y seminarios en lengua inglesa

CT03 - Capacidad para adoptar soluciones creativas que satisfagan adecuadamente las diferentes necesidades planteadas

CT04 - Capacidad para trabajar de forma efectiva como individuo, organizando y planificando su propio trabajo, de forma independiente o como miembro de un equipo

4.2. Resultados del aprendizaje

RA1 - Capacidad para desarrollar técnicas de tratamiento de señal específicas para datos masivos y diseñar aplicaciones sobre señales como: imágenes, señales de video, voz, audio y las procedentes de sensores de diversa naturaleza

RA12 - Capability to construct probabilistic models from experimental data using inference tools.

RA14 - Capability to model real phenomena using probability theory.

RA15 - Capability to relate the foundations of statistical inference with standard machine learning schemes.

RA4 - Formular problemas relacionados con la ingeniería como problemas de optimización en forma estándar

RA2 - Capacidad para planificar, diseñar y realizar aplicaciones que integren técnicas de tratamiento de señal, análisis estadístico y aprendizaje automático sobre datos masivos.

RA25 - Handle with ease the bases of linear algebra and calculus necessary to formulate problems optimization.

RA32 - Capability for planning, design and implement applications, incorporating signal processing, statistical analysis and machine learning

RA34 - Capability to develop and evaluate machine-learning techniques and to design big data learning systems

RA17 - Capacidad para aplicar conocimientos de modelado estadístico, técnicas de optimización y modelos de series temporales en el análisis de datos y como base para el desarrollo de algoritmos de aprendizaje automático

RA26 - Ability of oral and written communication

RA18 - Knowledge of tools for description, analysis and modeling of discrete-time random processes

RA7 - Capacidad para desarrollar y evaluar técnicas de aprendizaje automático y diseñar sistemas de aprendizaje para datos masivos

RA5 - Saber resolver problemas de optimización básicos como los de programación lineal o cuadrática

5. Descripción de la asignatura y temario

5.1. Descripción de la asignatura

This subject describes the problem where an agent has to make optimum decisions just by interacting with the environment that provides a reward. Starting from MDP (Markov Decision Processes) as a proper model of the problem we will emphasize the case where this model is unknown and has to be inferred. Finally, different solutions suitable for large scale problems are developed where value and policy functions are approximated by linear and non linear architectures. We will distinguish three main blocks

1. Fundamentals of Reinforcement Learning
2. Planning and learning in small scale problems
3. Learning in large and continuous spaces

5.2. Temario de la asignatura

1. Introduction
2. Multi-Armed Bandits
3. Markov Decision Processes
4. Planning by Dynamic Programming
5. Model-Free methods.
6. Linear Approximation
7. Non-Linear Approximation. Deep RL

6. Cronograma

6.1. Cronograma de la asignatura *

Sem	Actividad tipo 1	Actividad tipo 2	Tele-enseñanza	Actividades de evaluación
1				
2				
3				
4				
5				
6				
7				
8	1. Introduction 2. Multiarmed bandits Duración: 03:00 LM: Actividad del tipo Lección Magistral	Personal work related to chapter 2 Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio		
9	3. Markov decision processes Duración: 03:00 LM: Actividad del tipo Lección Magistral	Personal work related to chapter 3 Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio		
10	4. Planning by dynamic programming Duración: 03:00 LM: Actividad del tipo Lección Magistral	Personal work related to chapter 4 Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio		
11	5. Model free methods Duración: 03:00 LM: Actividad del tipo Lección Magistral	Personal work related to chapter 5 Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio		
12	6. Linear approximation Duración: 02:00 LM: Actividad del tipo Lección Magistral Mid term exam corresponding to chapters 1-4 Duración: 02:00 OT: Otras actividades formativas / Evaluación			Mid term exam corresponding to chapters 1-4. EP: Técnica del tipo Examen de Prácticas Evaluación Progresiva Presencial Duración: 02:00
13	7. Non-linear approximation. Deep RL Duración: 02:00 LM: Actividad del tipo Lección Magistral	Personal work related to chapter 6 Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
14	7. Non-linear approximation. Deep RL Duración: 02:00 LM: Actividad del tipo Lección Magistral	Personal work related to chapter 7 Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		

15				
16				
17				<p>First exam. EP: Técnica del tipo Examen de Prácticas Evaluación Global Presencial Duración: 02:00</p> <p>Second exam. EP: Técnica del tipo Examen de Prácticas Evaluación Progresiva y Global Presencial Duración: 02:00</p> <p>Deliver the final report with solved exercises. TI: Técnica del tipo Trabajo Individual Evaluación Progresiva y Global No presencial Duración: 00:00</p>

Para el cálculo de los valores totales, se estima que por cada crédito ECTS el alumno dedicará dependiendo del plan de estudios, entre 26 y 27 horas de trabajo presencial y no presencial.

7. Actividades y criterios de evaluación

7.1. Actividades de evaluación de la asignatura

7.1.1. Evaluación (progresiva)

Sem.	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
12	Mid term exam corresponding to chapters 1-4.	EP: Técnica del tipo Examen de Prácticas	Presencial	02:00	30%	3.5 / 10	CB06 CB07 CB08 CB09 CB10 CT01 CT03 CT04 CE02 CE03
17	Second exam.	EP: Técnica del tipo Examen de Prácticas	Presencial	02:00	30%	3.5 / 10	CB06 CB07 CB08 CB09 CB10 CT01 CT03 CT04 CE02 CE03
17	Deliver the final report with solved exercises.	TI: Técnica del tipo Trabajo Individual	No Presencial	00:00	40%	3 / 10	CB06 CB07 CB08 CB09 CB10 CT01 CT03 CT04 CE02 CE03

7.1.2. Prueba evaluación global

Sem	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
-----	-------------	-----------	------	----------	-----------------	-------------	------------------------

17	First exam.	EP: Técnica del tipo Examen de Prácticas	Presencial	02:00	30%	3.5 / 10	CB06 CB07 CB08 CB09 CB10 CT01 CT03 CT04 CE02 CE03
17	Second exam.	EP: Técnica del tipo Examen de Prácticas	Presencial	02:00	30%	3.5 / 10	CB06 CB07 CB08 CB09 CB10 CT01 CT03 CT04 CE02 CE03
17	Deliver the final report with solved exercises.	TI: Técnica del tipo Trabajo Individual	No Presencial	00:00	40%	3 / 10	CB06 CB07 CB08 CB09 CB10 CT01 CT03 CT04 CE02 CE03

7.1.3. Evaluación convocatoria extraordinaria

Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
Final exam.	EP: Técnica del tipo Examen de Prácticas	Presencial	03:00	60%	3.5 / 10	CB06 CB07 CB08 CB09 CB10 CT01 CT03 CT04 CE02

						CE03
Deliver the final report with solved exercises	TI: Técnica del tipo Trabajo Individual	Presencial	00:00	40%	3 / 10	CB06 CB07 CB08 CB09 CB10 CT01 CT03 CT04 CE02 CE03

7.2. Criterios de evaluación

The assessment procedure for the progressive evaluation will be as follows:

- One midterm exam including the first 4 chapters counting 30% of the final mark has to be completed. The student will have to solve a certain number of theoretical / practical issues similar to those contents of the practices and lectures. A minimum mark (3.5) is required.
- One second term exam including the last chapters from 5 to 7 counting 30% of the final mark has to be completed. The student will have to solve a certain number of theoretical / practical issues similar to those contents of the practices and lectures. A minimum mark (3.5) is required.
- A report including all requested exercises, counting 40% of the final mark.

The assessment procedure for the global evaluation, as well as the extraordinary examination, will be as follows:

- A final exam counting 60% of the final mark has to be completed. The student will have to solve a certain number of theoretical / practical issues similar to those contents of the practices and lectures.
- A report including all requested exercises, counting 40% of the final mark.

Any assessment or report may require a complementary oral assessment by the professor in order to validate that the task has been done by the student without help. According to the current assessment norms at UPM, if academic fraud is detected on any assessment, the student(s) will receive a grade of zero in the final grade of the examination to which the assessment belonged (ordinary or extraordinary).

8. Recursos didácticos

8.1. Recursos didácticos de la asignatura

Nombre	Tipo	Observaciones
Reinforcement learning	Bibliografía	Notes describing all the contents of the course
Matlab code of case studies	Otros	We provide a Matlab code solving all the case studies proposed in the course
Course slides	Bibliografía	Slides to be presented by the instructor to support the explanations
Reinforcement learning. An introduction. R. Sutton, A. Barto. The MIT Press. Draft second edition, 2015	Bibliografía	Main reference of chapters 1 - 5
Dynamic Programming and Optimal Control. D. Bertsekas. Third edition. Vol. 2. Athena Scientific Pub.	Bibliografía	Important reference to chapters 3 and 4
Reinforcement learning and dynamic programming using function approximators. L. Busoniu et al. CRC Press 2010	Bibliografía	Important reference chapters 6 and 7

9. Otra información

9.1. Otra información sobre la asignatura

This subject shows the fundamental ideas of reinforcement learning that could be used to model different ODS objectives in dynamic forms, as biological models (ODS 3), climate changing (ODS 13) or ecosystems (ODSs 14 y 15). It could also be applied to the efficient use of resources as water (ODS 6) or energy (ODS 7).

In more general terms, we teach applied mathematics used exhaustively in engineering, in particular will affect telecommunications infrastructures (ODS 9).

This course also contributes to subobjetives 4.4: to increase the number of persons with professional competences and techniques to access to employment and entrepreneurship and 4.7, to guarantee that all students acquire solid practical and theoretical knowledge required to promote sustainable development