ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

93000948 - Reinforcement learning

DEGREE PROGRAMME

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

ACADEMIC YEAR & SEMESTER

2018/19 - Semester 2

Index

Learning guide

1. Description

1.1. Subject details

Name of the subject	93000948 - Reinforcement learning
No of credits	3 ECTS
Type	Optional
Academic year ot the programme	First year
Semester of tuition	Semester 2
Tuition period	February-June
Tuition languages	English
Degree programme	09AT - Master universitario en teoria de la señal y comunicaciones
Centre	09 - Escuela Tecnica Superior de Ingenieros de Telecomunicacion
Academic year	2018-19

2. Faculty

2.1. Faculty members with subject teaching role

Name and surname	Office/Room	Email	Tutoring hours *
Santiago Zazo Bello (Subject coordinator)	C-326	santiago.zazo@upm.es	Sin horario. Arrange the meeting by email
Julian Cabrera Quesada	C-320	julian.cabrera@upm.es	Sin horario. Arrange the meeting by email

* The tutoring schedule is indicative and subject to possible changes. Please check tutoring times with the faculty member in charge.

3. Prior knowledge recommended to take the subject

3.1. Recommended (passed) subjects

- Fundamentos de optimización

- Técnicas de optimización para análisis de datos masivos

3.2. Other recommended learning outcomes

- Statistical Signal Processing

4. Skills and learning outcomes *

4.1. Skills to be learned

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. Learning outcomes

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 diversanaturaleza

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

* The Learning Guides should reflect the Skills and Learning Outcomes in the same way as indicated in the Degree Verification Memory. For this reason, they have not been translated into English and appear in Spanish.

5. Brief description of the subject and syllabus

5.1. Brief description of the subject

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 scale problems

5.2. Syllabus

1. Introduction

2. Multi-Armed Bandits

3. Markov Decision Processes

4. Dynamic Programming

5. Model-Free methods. Prediction and Control

6. Linear Approximation

7. Non-Linear Approximation

6. Schedule

6.1. Subject schedule*

Week	Face-to-face classroom activities	Face-to-face laboratory activities	Other face-to-face activities	Assessment activities
1				
2				
3				
4				
5				
6				
7				
8	**Introduction to Reinforcement Learning Introduction and bandits** Duration: 02:00 Lecture	**Case study related to chapter 2** Duration: 01:00 Laboratory assignments		**Personal work programming a problem related to chapter 2** Problem-solving test Continuous assessment Duration: 01:00
9	**Introduction to Reinforcement Learning. Markov Decision Processes** Duration: 02:00 Lecture	**Case study related to chapter 3** Duration: 01:00 Laboratory assignments		**Personal work programming a problem related to chapter 3** Problem-solving test Continuous assessment Duration: 01:00
10	**Planning and learning in small scale problems. Dynamic programming.** Duration: 02:00 Lecture	**Case study related to chapter 4** Duration: 01:00 Laboratory assignments		**Personal work programming a problem related to chapter 4** Problem-solving test Continuous assessment Duration: 01:00
11	**Planning and learning in small scale problems. Model free methods.** Duration: 02:00 Lecture	**Case study related to chapter 5** Duration: 01:00 Laboratory assignments		**Personal work programming a problem related to chapter 5** Problem-solving test Continuous assessment Duration: 01:00
12	**Learning in large scale problems. Base of functions. LInear approximation** Duration: 01:00 Lecture	**Case study related to chapter 6** Duration: 01:00 Laboratory assignments		**Personal work programming a problem related to chapter 6** Problem-solving test Continuous assessment Duration: 01:00 **Mid term exam corresponding to chapters 1-5. The student will have to solve an RL problem equivalent to the laboratory activities** Problem-solving test Continuous assessment Duration: 01:00
13	**Learning in large scale problems. Linear approximation** Duration: 02:00 Lecture	**Case study related to chapter 6** Duration: 01:00 Laboratory assignments		**Personal work programming a problem related to chapter 6** Problem-solving test Continuous assessment Duration: 01:00

14	**Learning in large scale problems. Non linear approximation** Duration: 02:00 Lecture	**Personal work related to chapter 7** Duration: 01:00 Laboratory assignments		**Personal work programming a problem related to chapter 7** Problem-solving test Continuous assessment Duration: 01:00
15				
16				
17				**Final exam. The student will have to solve an RL problem equivalent to the laboratory activities It will cover all the chapters** Problem-solving test Final examination Duration: 02:00 **Second exam. The student will have to solve an RL problem equivalent to the laboratory activities** Problem-solving test Continuous assessment Duration: 01:00

The independent study hours are training activities during which students should spend time on individual study or individual assignments.

Depending on the programme study plan, total values will be calculated according to the ECTS credit unit as 26/27 hours of student face-to-face contact and independent study time.

* The subject schedule is based on a previous theorical planning of the subject plan and might go to through experience some unexpected changes along throughout the academic year.

7. Activities and assessment criteria

7.1. Assessment activities

7.1.1. Continuous assessment

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
8	Personal work programming a problem related to chapter 2	Problem-solving test	Face-to-face	01:00	5%	0 / 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CB10
9	Personal work programming a problem related to chapter 3	Problem-solving test	Face-to-face	01:00	5%	0 / 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CB10
10	Personal work programming a problem related to chapter 4	Problem-solving test	Face-to-face	01:00	5%	0 / 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CB10
11	Personal work programming a problem related to chapter 5	Problem-solving test	Face-to-face	01:00	10%	0 / 10	CB08 CB09 CT01 CT03 CB06 CE02 CT04 CE03 CB10

12	Personal work programming a problem related to chapter 6	Problem-solving test	Face-to-face	01:00	5%	0 / 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CB10
12	Mid term exam corresponding to chapters 1-5. The student will have to solve an RL problem equivalent to the laboratory activities	Problem-solving test	Face-to-face	01:00	30%	0 / 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CB10
13	Personal work programming a problem related to chapter 6	Problem-solving test	Face-to-face	01:00	5%	0 / 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CB10
14	Personal work programming a problem related to chapter 7	Problem-solving test	Face-to-face	01:00	5%	0 / 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CB10
17	Second exam. The student will have to solve an RL problem equivalent to the laboratory activities	Problem-solving test	Face-to-face	01:00	30%	0 / 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CB10

7.1.2. Final examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
17	Final exam. The student will have to solve an RL problem equivalent to the laboratory activities It will cover all the chapters	Problem-solving test	Face-to-face	02:00	100%	0 / 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CB10

7.1.3. Referred (re-sit) examination

Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
One exam with theoretical questions and also some aspects related to the practical works	Written test	Face-to-face	03:00	100%	0 / 10	CB08 CB09 CT01 CB07 CT03 CB06 CE02 CT04 CE03 CB10

7.2. Assessment criteria

Students will be qualified through continuous evaluation by default. According to the Normativa de Evaluación del Aprendizaje de la Universidad Politécnica de Madrid, students willing to renounce to continuous evaluation must complete the Moodle task entitled "Renounce to continuous evaluation" before the first intermediate exam (deadline will be announced in Moodle).
Evaluation will assess if students have acquired all the competences of the subject. Thus, evaluation through final assessment will be carried out considering all the evaluation techniques used in continuous evaluation (EX, ET, TG, etc.), and will be celebrated in the exam period approved by Junta de Escuela for the current academic semester and year. Evaluation activities that assess learning outcomes that cannot be evaluated through a single exam can be carried out along the semester.
Extraordinary examination will be carried out exclusively by the final assessment method.

After each chapter, the student will have to program an RL problem in Matlab similar to the case study described by the instructor. After completing the program and running some simulations, a short report has to be written and sent to the instructor for evaluation. Each report will be scored with 5-10% of the final mark

One mid term exam including the first 5 chapters counting 30% of the final mark has to be completed. The student will have to solve a certain RL problem and answer some theoretical aspects.

A second mid term exam including the chapters 6 and 7 counting 30% of the final mark has to be completed. The student will have to solve a certain RL problem and answer some theoretical aspects.

The evaluation procedure for the final and re-sit examination will be as follows:

A short report including the code and derivations for every exercise in the course has to be written and sent to the instructor for evaluation. All reports will be scored with 40 % of the final mark.
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 as those contents of the practices and lectures.

8. Teaching resources

8.1. Teaching resources for the subject

Name	Type	Notes
Reinforcement learning	Bibliography	Notes describing all the contents of the course
Matlab code of case studies	Others	We provide a Matlab code solving all the case studies proposed in the course
Course slides	Bibliography	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	Bibliography	Main reference of chapters 1 - 5
Dynamic Programming and Optimal Control. D. Bertsekas.Third edition. Vol. 2. Athena Scientific Pub.	Bibliography	Important reference to chapters 3 and 4
Reinforcement learning and dynamic programming using function approximators. L. Busoniu et al. CRC Press 2010	Bibliography	Important reference chapters 6 and 7