Course guides
230817 - ARAP - Reinforcement Learning and Deep Learning

Last modified: 30/06/2020

Unit in charge: Barcelona School of Telecommunications Engineering
Teaching unit: 739 - TSC - Department of Signal Theory and Communications.

Degree: BACHELOR'S DEGREE IN TELECOMMUNICATIONS TECHNOLOGIES AND SERVICES ENGINEERING (Syllabus 2015). (Optional subject).
BACHELOR'S DEGREE IN DATA SCIENCE AND ENGINEERING (Syllabus 2017). (Optional subject).

Academic year: 2020 **ECTS Credits:** 6.0 **Languages:** Catalan, English

LECTURER

Coordinating lecturer: Giró Nieto, Xavier

Others: Cabrera Bean, Margarita Asuncion

Vidal Manzano, Jose

PRIOR SKILLS

Algebra, Probability and stochastic processes, Signals and systems, Machine and Deep Learning.

REQUIREMENTS

For GRETST students, Introduction to deep learning (IDL 230325). For GCED students, Machine Leaning 2 (AA2 270222).

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Generical:
08 CRPE. ABILITY TO IDENTIFY, FORMULATE AND SOLVE ENGINEERING PROBLEMS. To plan and solve engineering problems in the ICT with initiative, making decisions and with creativity. To develop a method of analysis and problem solving in a systematic and creative way.

Transversal:
04 COE N2. EFFICIENT ORAL AND WRITTEN COMMUNICATION - Level 2. Using strategies for preparing and giving oral presentations. Writing texts and documents whose content is coherent, well structured and free of spelling and grammatical errors.
06 URI N1. EFFECTIVE USE OF INFORMATI0N RESOURCES - Level 1. Identifying information needs. Using collections, premises and services that are available for designing and executing simple searches that are suited to the topic.
06 URI N2. EFFECTIVE USE OF INFORMATI0N RESOURCES - Level 2. Designing and executing a good strategy for advanced searches using specialized information resources, once the various parts of an academic document have been identified and bibliographical references provided. Choosing suitable information based on its relevance and quality.
06 URI N3. EFFECTIVE USE OF INFORMATI0N RESOURCES - Level 3. Planning and using the information necessary for an academic assignment (a final thesis, for example) based on a critical appraisal of the information resources used.

TEACHING METHODOLOGY

Classroom lectures and labs

LEARNING OBJECTIVES OF THE SUBJECT

Master the principles of reinforcement learning as an artificial intelligence tool based on the interaction of the machine with its environment, with applications to control tasks (eg. robotics, autonomous driving) o decision making (eg. resource optimization in wireless communication networks).
Design and train deep neural networks capable of learning with little or no supervision, both for discriminative and generative tasks, with special attention on multimedia applications (vision, language and speech).

STUDY LOAD

Type	Hours	Percentage
Self study	98,0	65.33
Hours small group	26,0	17.33
Hours large group	26,0	17.33

Total learning time: 150 h

CONTENTS

Introduction to reinforcement learning

Description:
Describe with examples the fundamental concepts and the problems that can be solved.

Full-or-part-time: 2h
Theory classes: 2h

The exploration-exploitation trade-off

Description:
- The exploration-exploitation trade-off
- Study case: Multi-armed bandits.

Full-or-part-time: 1h
Theory classes: 1h

Markov decision processes (MDP)

Description:
- The agent-environment interface
- Goals and rewards
- Markov Decision Processes
- Value functions and optimality: Bellman equation

Full-or-part-time: 2h
Theory classes: 2h

Dynamic programming

Description:
- Policy evaluation, improvement and iteration
- Dynamic programming based on MDP

Full-or-part-time: 1h
Theory classes: 1h

Monte-Carlo methods

Description:
- First-visit Monte-Carlo methods
- Every-visit Monte-Carlo methods
- Exploration and exploitation
- On-policy and off-policy methods

Full-or-part-time: 2h 30m
Theory classes: 2h 30m

Temporal-Difference (TD) Learning

Description:
- Model-free learning using time diferences
- SARSA and Q-learning with discrete actions
- Games

Full-or-part-time: 2h 30m
Theory classes: 2h 30m

Policy gradient methods

Description:
- Policy gradient
- Value function approximation
- Actor-Critic methods
- Baseline functions

Full-or-part-time: 2h 30m
Theory classes: 2h 30m

Deep reinforcement learning

Description:
Modeling of q-value functions and policies with deep neural networks

Full-or-part-time: 2h 30m
Theory classes: 2h 30m

Advanced deep learning

Description:
- Generative models: GANs, VAEs and Flows.
- Advanced Recurrent Neural Networks.
- Attention mechanism and Transformers.
- Graph Neural Networks.
- Self-supervised learning and meta-learning.
- Supercomputation for deep learning.

Full-or-part-time: 10h
Theory classes: 10h

Reinforcement learning labs

Description:
Labs in Matlab and/or Python distributed throught the lectures
- Dynamic channel allocation
- Blackjack
- Job-shop scheduling
- Tabular Q-Learning

Full-or-part-time: 13h
Laboratory classes: 13h

Deep Learning Labs

Description:
Labs in PyTorch about:
- Q-learning with neural networks.
- Vanilla policy gradients (REINFORCE).
- Optimizers for deep neural networks
- Attention models
- Transfer learning
- Generative Adversarial Networks (GANs)
- Variational Autoencoders (VAEs)

Full-or-part-time: 13h
Laboratory classes: 13h

GRADING SYSTEM

Exams and evaluation of labs

BIBLIOGRAPHY

Basic:
- Sutton, R.S.; Barto, A.G. Reinforcement learning: an introduction [on line]. 2nd ed. Cambridge, Mass.: MIT Press, 2018 [Consultation: 14/09/2020]. Available on: https://ebookcentral.proquest.com/lib/upcatalunya-ebooks/detail.action?docID=6260249. ISBN 9780262039246.
- Goodfellow, I.; Bengio, Y.; Courville, A. Deep learning [on line]. Cambridge, Massachusetts: The MIT Press, [2016] [Consultation: 02/07/2020]. Available on: http://www.deeplearningbook.org/. ISBN 978-0262035613.

RESOURCES

Other resources:
Course website: https://telecombcn-dl.github.io/drl-2020/