

Guía docente

270420 - APRNS - Aprendizaje por Refuerzo y No Supervisado

Última modificación: 10/07/2024

Unidad responsable: Facultad de Informática de Barcelona
Unidad que imparte: 723 - CS - Departamento de Ciencias de la Computación.
Titulación: GRADO EN INTELIGENCIA ARTIFICIAL (Plan 2021). (Asignatura obligatoria).
Curso: 2024 **Créditos ECTS:** 6.0 **Idiomas:** Catalán, Castellano

PROFESORADO

Profesorado responsable: JAVIER BÉJAR ALONSO - MARIO MARTÍN MUÑOZ

Otros: Primer quadrimestre:
JAVIER BÉJAR ALONSO - 11, 12
MARIO MARTÍN MUÑOZ - 11, 12

CAPACIDADES PREVIAS

Conocimientos básicos de Deep Learning y de Machine Learning.

COMPETENCIAS DE LA TITULACIÓN A LAS QUE CONTRIBUYE LA ASIGNATURA

Específicas:

CE18. Adquirir y desarrollar técnicas de aprendizaje computacional y diseñar e implementar aplicaciones y sistemas que las utilicen, incluyendo las dedicadas a extracción automática de información y conocimiento a partir de grandes volúmenes de datos.

Genéricas:

CG2. Utilizar los conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.

CG4. Razonar, analizando la realidad y diseñando algoritmos y formulaciones que la modelen. Identificar problemas y construir soluciones algorítmicas o matemáticas válidas, eventualmente nuevas, integrando el conocimiento multidisciplinar necesario, valorando distintas alternativas con espíritu crítico, justificando las decisiones tomadas, interpretando y sintetizando los resultados en el contexto del dominio de aplicación y estableciendo generalizaciones metodológicas a partir de aplicaciones concretas.

Transversales:

CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.

Básicas:

CB5. Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía

METODOLOGÍAS DOCENTES

Las clases están divididas en sesiones de teoría, problemas y laboratorio.

En las sesiones de teoría se desarrollarán los conocimientos de la asignatura, intercalando la exposición de nuevo material teórico con ejemplos y la interacción con los alumnos para discutir los conceptos.

En las clases de laboratorio se desarrollarán pequeñas prácticas utilizando herramientas y utilizando librerías específicas que permitirán practicar y reforzar los conocimientos de las clases de teoría.

OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA

1. Conocer qué tipo de problemas se pueden modelizar como un problema de aprendizaje por refuerzo e identificar las técnicas que se pueden aplicar para resolverlas
2. Comprender la necesidad, los fundamentos y las particularidades del aprendizaje conductual y las diferencias que tiene con el aprendizaje automático supervisado y no supervisado.
3. Comprender los algoritmos más importantes y el estado del arte en el área del aprendizaje por refuerzo.
4. Saber formalizar computacionalmente un problema del mundo real como aprendizaje para reforzamiento y saber implementar en los entornos más actuales los algoritmos de aprendizaje que los resuelve
5. Conocer los problemas que se pueden modelizar con algoritmos no supervisados profundos
6. Entender las particularidades de los algoritmos no supervisados profundos
7. Conocer los algoritmos más importantes y el estado del arte del aprendizaje no supervisado profundo
8. Saber implementar y aplicar a un problema algoritmos de aprendizaje profundo utilizando el entorno más actual

HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTE

Tipo	Horas	Porcentaje
Horas grupo grande	30,0	20.00
Horas grupo pequeño	30,0	20.00
Horas aprendizaje autónomo	90,0	60.00

Dedicación total: 150 h

CONTENIDOS

Introducción: Aprendizaje del comportamiento en agentes y descripción de los principales elementos en el aprendizaje de refuerzo

Descripción:

Intuición, motivación y definición del marco del aprendizaje por refuerzo (RL). Elementos clave en RL.

Encontrando políticas óptimas mediante la programación dinámica

Descripción:

Como aprender un comportamiento con conocimiento completo del modelo del mundo: solución algebraica, evaluación iterada de políticas y evaluación iterada de valores.

Introducción a los enfoques sin modelos del mundo.

Descripción:

Algoritmos básicos para el aprendizaje por refuerzo: Monte-Carlo, Q-learning, Sarsa, TD (λ). La necesidad de exploración. Diferencias entre los métodos On-policy y Off-policy

Aproximación de funciones en el aprendizaje por refuerzo

Descripción:

La necesidad de la aproximación de funciones y métodos incrementales en RL. El enfoque de Gradient Descent. RL con aproximación de función lineal. La tríada mortal para la aproximación de funciones en RL. Métodos por lotes y redes neuronales para la aproximación de funciones.

Aprendizaje por refuerzo profundo (DRL)

Descripción:

Introducción de DL en RL. Como tratar la mortal tríada con el algoritmo DQN. Aplicación de DQN al caso de los juegos Atari. Evoluciones del algoritmo DQN: Double DQN, Prioritized Experience Replay, aprendizaje en múltiples pasos y funciones de valor distribuidas. Rainbow: el algoritmo de última generación para un espacio de acción discreto.

Métodos del gradiente en la política

Descripción:

Qué hacer en espacios de acción continuos. Cómo las políticas probabilísticas permiten aplicar el método de gradiente directamente en la red de políticas. El algoritmo REINFORCE. Los algoritmos Actor-Critic. Algoritmos de última generación en espacios de acción continuos: DDPG, TD3 y SAC.

Temas avanzados: Como tratar el problema del refuerzo esparso

Descripción:

El problema de la recompensa esparsa. Introducción a técnicas avanzadas de exploración: curiosidad y empoderamiento en RL. Introducción al aprendizaje curricular para facilitar el aprendizaje del objetivo. RL jerárquico para aprender tareas complejas. El aprendizaje de las funciones de valor universales y Hindsight Experience Replay (HER).

Aprendizaje por refuerzo en el marco de múltiples agentes

Descripción:

Aprendizaje de comportamientos en un entorno donde actúan varios agentes. Aprendizaje de conductas cooperativas, Aprendizaje de conductas competitivas y casos mixtos. Algoritmos de última generación. El caso especial de los juegos: el caso AlphaGo y la extensión a Alpha-Zero.

Introducción: El aprendizaje no supervisado profundo

Descripción:

Introducción a la necesidad del aprendizaje no supervisado profundo y sus aplicaciones

Modelos autoregresivos

Descripción:

Introducción al aprendizaje de distribuciones de probabilidad definidas como distribuciones autorregresivas y principales modelos

Flujos normalizantes

Descripción:

Introducción a los flujos normalizados para el aprendizaje de distribuciones de probabilidad

Modelos de variables latentes

Descripción:

Introducción a los modelos basados en variables latentes y a los autocodificadores variacionales



Redes adversarias Generativas

Descripción:

Introducción a las redes adversarias generativas, generación condicionada y sin condicionar, separación de atributos

Redes de difusión

Descripción:

Introducción a modelos basados en difusión de ruido, redes para eliminación de ruido, condicionamiento, generación multimodal

Aprendizaje por auto supervisión

Descripción:

Introducción al aprendizaje por autosupervisión para el entrenamiento de redes generadoras de características, metodos contrastivos y no contrastivos, enmascaramiento

ACTIVIDADES

Introducción: Aprendizaje del comportamiento en agentes y descripción de los principales elementos en el aprendizaje de refuerzo

Dedicación: 10h

Aprendizaje autónomo: 6h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Encontrando políticas óptimas mediante la programación dinámica

Descripción:

Como aprender un comportamiento con conocimiento completo del modelo del mundo: solución algebraica, evaluación iterada de políticas y evaluación iterada de valores.

Dedicación: 10h

Aprendizaje autónomo: 6h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Introducción a los enfoques sin modelos. Monte-Carlo, Q-learning, Sarsa, TD (λ)

Descripción:

Desarrollo del tema correspondiente de la asignatura

Dedicación: 10h

Aprendizaje autónomo: 6h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aproximación de funciones en RL

Dedicación: 10h

Aprendizaje autónomo: 6h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aprendizaje por refuerzo profundo

Dedicación: 10h

Aprendizaje autónomo: 6h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Métodos del gradiente en la política

Descripción:

Qué hacer en espacios de acción continuos. Cómo las políticas probabilísticas permiten aplicar el método de gradiente directamente en la red de políticas. El algoritmo REINFORCE. Los algoritmos Actor-Critic. Algoritmos de última generación en espacios de acción continuos: DDPG, TD3 y SAC.

Dedicación: 10h

Aprendizaje autónomo: 6h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Temas avanzados: Como tratar el problema del refuerzo esparso

Descripción:

El problema de la recompensa esparsa. Introducción a técnicas avanzadas de exploración: curiosidad y empoderamiento en RL. Introducción al aprendizaje curricular para facilitar el aprendizaje del objetivo. RL jerárquico para aprender tareas complejas. El aprendizaje de las funciones de valor universales y Hindsight Experience Replay (HER).

Dedicación: 10h

Aprendizaje autónomo: 6h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Aprendizaje por refuerzo en el marco de múltiples agentes

Descripción:

Aprendizaje de comportamientos en un entorno donde actúan varios agentes. Aprendizaje de conductas cooperativas, Aprendizaje de conductas competitivas y casos mixtos. Algoritmos de última generación. El caso especial de los juegos: el caso AlfaGo y la extensión a Alfa-Zero.

Dedicación: 13h

Aprendizaje autónomo: 9h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Control de la parte de aprendizaje por reforzamiento

Objetivos específicos:

1, 2, 3, 4

Competencias relacionadas:

CE18. Adquirir y desarrollar técnicas de aprendizaje computacional y diseñar e implementar aplicaciones y sistemas que las utilicen, incluyendo las dedicadas a extracción automática de información y conocimiento a partir de grandes volúmenes de datos.
CG2. Utilizar los conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.

CG4. Razonar, analizando la realidad y diseñando algoritmos y formulaciones que la modelen. Identificar problemas y construir soluciones algorítmicas o matemáticas válidas, eventualmente nuevas, integrando el conocimiento multidisciplinar necesario, valorando distintas alternativas con espíritu crítico, justificando las decisiones tomadas, interpretando y sintetizando los resultados en el contexto del dominio de aplicación y estableciendo generalizaciones metodológicas a partir de aplicaciones concretas.

CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.

Dedicación: 2h

Actividades dirigidas: 2h

Introducción: El aprendizaje no supervisado profundo

Descripción:

Introducción a la necesidad del aprendizaje no supervisado profundo y sus aplicaciones

Dedicación: 10h

Aprendizaje autónomo: 6h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Modelos autoregresivos

Descripción:

Introducción al aprendizaje de distribuciones de probabilidad definidas como distribuciones autorregresivas y principales modelos

Dedicación: 10h

Aprendizaje autónomo: 6h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Flujos normalizantes

Descripción:

Introducción a los flujos normalizados para el aprendizaje de distribuciones de probabilidad

Dedicación: 10h

Aprendizaje autónomo: 6h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Modelos de variables latentes

Descripción:

Introducción a los modelos basados en variables latentes y a los autocodificadores variacionales

Dedicación: 10h

Aprendizaje autónomo: 6h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Redes adversarias Generativas

Descripción:

Introducción a las redes adversarias generativas, generación condicionada y sin condicionar, separación de atributos

Dedicación: 10h

Aprendizaje autónomo: 6h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 2h

Redes de difusión y Aprendizaje por auto supervisión

Descripción:

Introducción a modelos basados en difusión de ruido, redes para eliminación de ruido, condicionamiento, generación multimodal

Dedicación: 15h

Aprendizaje autónomo: 9h

Grupo grande/Teoría: 2h

Grupo pequeño/Laboratorio: 4h

Control del temario de aprendizaje no supervisado

Objetivos específicos:

5, 6, 7, 8

Competencias relacionadas:

CB5. Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía

CE18. Adquirir y desarrollar técnicas de aprendizaje computacional y diseñar e implementar aplicaciones y sistemas que las utilicen, incluyendo las dedicadas a extracción automática de información y conocimiento a partir de grandes volúmenes de datos.

CG2. Utilizar los conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.

CG4. Razonar, analizando la realidad y diseñando algoritmos y formulaciones que la modelen. Identificar problemas y construir soluciones algorítmicas o matemáticas válidas, eventualmente nuevas, integrando el conocimiento multidisciplinar necesario, valorando distintas alternativas con espíritu crítico, justificando las decisiones tomadas, interpretando y sintetizando los resultados en el contexto del dominio de aplicación y estableciendo generalizaciones metodológicas a partir de aplicaciones concretas.

CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.

SISTEMA DE CALIFICACIÓN

La asignatura comprenderá los siguientes actos evaluatorios:

- Informes de las actividades de laboratorio, que será necesario haber entregado dentro de un plazo indicado para cada sesión (orientativamente, 2 semanas). A partir de una media ponderada de las notas de estos informes se calculará una nota de laboratorio, L.
- Un primer examen parcial, realizado a mitad del curso, de la materia vista hasta entonces. Sea P1 la nota obtenida en este examen.
- En el día designado dentro del período de exámenes, un segundo examen parcial de la materia no cubierta por el primer parcial. Sea P2 la nota obtenida en este examen.

Las tres notas L, P1, P2 son entre 0 y 10.

La nota final de la asignatura será: $0.4 * L + 0.3 * P1 + 0.3 * P2$

BIBLIOGRAFÍA

Básica:

- Sutton, Richard S; Barto, Andrew G. Reinforcement learning : an introduction. Second edition. Cambridge, Massachusetts: The MIT Press, [2020]. ISBN 9780262039246.
- Morales, Miguel. Grokking deep reinforcement learning. Shelter Island: Manning Publications, 2020. ISBN 9781617295454.
- Foster, D. Generative deep learning: teaching machines to paint, write, compose, and play. 2nd ed. Sebastopol: O'Reilly Media, Incorporated, 2023. ISBN 9781098134143.
- Cheong, S.Y. Hands-on image generation with TensorFlow: a practical guide to generating images and videos using deep learning. Birmingham: Packt Publishing, 2020. ISBN 9781838821104.

Complementaria:

- Zai, Alexander; Brown, Brandon. Deep reinforcement learning in action. Shelter Island, NY: Manning Publications Co, 2020. ISBN 9781617295430.
- Babcock, J.; Bali, R. Generative AI with Python and TensorFlow 2: harness the power of generative models to create images, text, and music. Birmingham, England ; Mumbai: Packt Publishing, 2021. ISBN 9781800208506.