UNIVERSIDAD AUTÓNOMA DE CHIHUAHUA

UNIVERSIDAD AUTÓNOMA DE CHIHUAHUA

FACULTAD DE INGENIERÍA

PROGRAMA ANALÍTICO DE LA UNIDAD DE APRENDIZAJE:

APRENDIZAJE POR REFUERZO

DES:	Ingeniería
Programa académico	Maestría en Ingeniería en Computación
Tipo de materia (Obli/Opta):	Optativa
Clave de la materia:	MICOP2311
Semestre:	2,3,4
Área en plan de estudios (G y E):	G, E
Total de horas por semana:	7
Teoría: Presencial o Virtual	2
Laboratorio o Taller:	0
Prácticas:	2
Trabajo extra-clase:	3
Créditos Totales:	7
Total de horas semestre (x 16 sem):	112
Fecha de actualización:	Agosto 2023
Prerrequisito (s):	Ninguno

DESCRIPCIÓN DEL CURSO:

En este curso el estudiante estudia y aplica métodos matemáticos y computacionales utilizados en el área de aprendizaje por refuerzo, con el objetivo de resolver tareas de control y toma de decisiones a partir de un agente que explora un entorno desconocido y determina las acciones a llevar a cabo mediante prueba y error; el agente será capaz de aprender por sí sólo obteniendo recompensas o penalizaciones, no sólo de manera inmediata sino buscando maximizar la recompensa a la larga. El estudiante desarrollará una actitud de comprensión para conceptos abstractos, pero también debe identificar escenarios de aplicación para el tema de aprendizaje por refuerzo. La habilidad que el estudiante debe desarrollar le permitirá mejorar sus fortalezas cognitivas para aplicar lo estudiado en problemas computacionales para diferentes contextos de aplicación.

COMPETENCIAS A DESARROLLAR:

Genéricas.

GESTIÓN DEL CONOCIMIENTO. Demuestra conocimientos y habilidades para la búsqueda, análisis crítico, síntesis y procesamiento de información para su transformación en conocimiento, con actitud ética.

COMUNICACIÓN CIENTÍFICA. Difunde con responsabilidad ética y social el conocimiento científico, tecnológico, artístico y/o humanístico que produce de forma objetiva.

INVESTIGACIÓN. Desarrolla investigación original, tecnología y/o innovaciones en procesos, servicios o productos que contribuyan a la solución de problemas, mejoren la convivencia, generen oportunidades para el desarrollo sustentable y propicien una mejor calidad de vida.

Específica.

INTELIGENCIA ARTIFICIAL. Aplica la Inteligencia Artificial para resolver problemas en los sectores industrial, gubernamental, académico y social bajo esquemas de colaboración ética y multidisciplinaria.

DOMINIOS	OBJETOS DE ESTUDIO	RESULTADOS DE APRENDIZAJE	METODOLOGÍA	EVIDENCIAS
Analiza y recupera información pertinente mediante diversas estrategias de búsqueda de datos científicos.	1. Introducción 1.1. El dilema de la explotación y exploración. 1.2. Algoritmo Epsilon – Greedy.	• Conoce y aplica los conceptos y fundamentos en los que se basan diferentes algoritmos computacionales	• Lectura Crítica. • Búsqueda de Información. • Implementación de Algoritmos • Computacionales.	• Tareas de Investigación. • Ejercicios extra clase resuelto • Código generado de

Evalúa de manera crítica la información, considerando su calidad y pertinencia. Se comunica en forma oral y escrita con propiedad, relevancia, oportunidad y ética. Aplica los elementos fundamentales de la redacción científica. Interpreta y expresa ideas, sentimientos, teorías y corrientes de pensamiento con un enfoque ecuménico (universal). Manifiesta capacidad de innovar y creatividad al producir soluciones apropiadas para los contextos en los que se desenvuelve Comprueba los resultados obtenidos de un prototipo contra las investigaciones recientes a fin de identificar nuevas contribuciones Accede a diferentes fuentes de información (journals o revistas científicas, bases de datos, índices, etc.) de calidad. Construye prototipos innovadores utilizando herramientas de software y hardware adecuadas a la solución de problemas en cualquier ámbito de desempeño. Soluciona problemas en diversas áreas del conocimiento aplicando las ciencias computacionales	1.3. Algoritmo Optimistic Initial Values. 1.4. Algoritmo Upper Confidence Bound 1.5. Algoritmo Thompson Sampling.	para resolver problemas basados en el dilema de la exploración y explotación.	• Resolución de problemas • analíticos	ejemplos.
	2. Procesos de decisión de Markov (MDP) 2.1. Estados, acciones, recompensas y políticas. 2.2. Tipos de Procesos de decisión de Markov 2.3. Ecuaciones de Bellman y funciones valor. 2.4. Solución de la ecuación de Bellman. 2.4.1. Iteración de Valor. 2.4.2. Iteración de Política.	• Conoce e identifica las partes esenciales de las que se compone un MDP, así como los tipos de procesos que hay. • Conoce y aplica los enfoques y métodos de solución de un MDP	• Lectura Crítica. • Búsqueda de Información. • Implementación de Algoritmos • Computacionales. • Resolución de Problemas analíticos	• Tareas de Investigación. • Ejercicios extra clase resuelto • Código generado de ejemplos.
	3. Métodos para resolver MDP discretos 3.1. Programación Dinámica 3.2. Métodos Monte Carlo 3.3. Métodos de diferencia temporal. 3.3.1. Q-learning 3.3.2. SARSA. 3.4. N-step bootstrapping 3.4.1. N-step Diferencia temporal. 3.4.2. N-step SARSA	• Conoce y aplica los diferentes algoritmos computacionales para resolver problemas en varios contextos que usen MDPs en un espacio de estados discreto.	• Lectura Crítica. • Búsqueda de Información. • Implementación de Algoritmos • Computacionales. • Resolución de Problemas analíticos	• Tareas de Investigación. • Ejercicios extra clase resuelto • Código generado de ejemplos.
	4. MDP de espacio de estados continuos 4.1. State aggregation. 4.2. Algoritmo Deep-SARSA	• Conoce y aplica los diferentes algoritmos computacionales para resolver problemas en varios contextos que usen	• Lectura Crítica. • Búsqueda de Información. • Implementación de • Algoritmos	• Tareas de Investigación. • Ejercicios extra clase resuelto

	4.3. Algoritmo Deep Q-Learning. 4.4. Algoritmo Advantage Actor-Critic (A2C) 4.5. Algoritmo Asynchronous Advantage Actor-Critic (A3C) 4.6. Procesos de decisión de Markov Parcialmente Observables.	MDPs en un espacio de estados continuo donde los estados son parcialmente o completamente observables.	• Computacionales. • Resolución de Problemas analíticos	• Código generado de ejemplos.

FUENTES DE INFORMACIÓN	EVALUACIÓN DE LOS APRENDIZAJES
1. De Ponteves, H. (2019). AI Crash Course: A fun and hands-on introduction to reinforcement learning, deep learning, and artificial intelligence with Python. United Kingdom. Packt Publishing Ltd. 2. Plaat, A. (2022). Deep Reinforcement Learning. Singapure. Springer. 3. Shengbo, E. L. (2023). Reinforcement Learning for Sequential Decision and Optimal Control. Singapure. Springer	El parcial uno cubrirá los contenidos del primer y segundo objeto de estudio del curso. El parcial dos cubrirá el contenido del tercer objeto de estudio. El parcial tres cubrirá el contenido del cuarto objeto de estudio del curso. Las tareas de clase y los ejercicios extra-clase tienen una ponderación del 100% para evaluar un parcial. En caso de que un parcial se evalúe también con la implementación de algoritmos, la ponderación para la evaluación del parcial será la siguiente: - 40% para implementación de algoritmos. - 60% para tareas y ejercicios extra-clase. Para cada uno de los temas del curso se desarrollará una tarea y/o se implementará un algoritmo para la evidencia del curso.

CRONOGRAMA

Objetos de aprendizaje	Semanas															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Introducción																
2. Procesos de Decisión de Markov																
3. Métodos para resolver MDP discretos																
4. MDP de espacio de estados continuo																