

<p style="text-align: center;">UNIVERSIDAD AUTÓNOMA DE CHIHUAHUA</p>  <p style="text-align: center;">FACULTAD DE INGENIERÍA</p> <p style="text-align: center;">PROGRAMA ANALÍTICO DE LA UNIDAD DE APRENDIZAJE:</p> <p style="text-align: center;">APRENDIZAJE POR REFUERZO</p>	DES:	Ingeniería
	Programa académico	Doctorado en Ingeniería
	Tipo de materia (Obli/Opta):	Optativa
	Clave de la materia:	DI24OP24
	Semestre:	1, 2, 3
	Área en plan de estudios (B, P y E):	G, E
	Total de horas por semana:	4
	<i>Teoría: Presencial o Virtual</i>	2
	<i>Laboratorio o Taller:</i>	0
	<i>Prácticas:</i>	2
	<i>Trabajo extra-clase:</i>	6
	<i>Créditos Totales:</i>	10
	Total de horas semestre (x 16 sem):	160
	Fecha de actualización:	Marzo 2024
<i>Prerrequisito (s):</i>	Ninguno	
<p>DESCRIPCIÓN DEL CURSO</p> <p>En este curso, los estudiantes aprenden y aplican métodos matemáticos y computacionales en el área del aprendizaje por refuerzo. El objetivo es resolver tareas de control y toma de decisiones, donde un agente interactúa con un entorno desconocido, aprendiendo de la experiencia para maximizar las recompensas a largo plazo. Se fomenta la comprensión de conceptos abstractos y la identificación de escenarios de aplicación del aprendizaje por refuerzo. Los estudiantes desarrollan habilidades para aplicar estos conocimientos en problemas computacionales de diversos contextos, preparándolos para enfrentar desafíos en áreas como la robótica, la inteligencia artificial y la toma de decisiones automatizada.</p>		
<p>COMPETENCIAS A DESARROLLAR</p> <p>GESTIÓN DE PROYECTOS Coordina y administra de forma responsable, proyectos que atiendan criterios de sustentabilidad y que contribuyan a mejorar la calidad de vida.</p> <p>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.</p> <p>COMUNICACIÓN CIENTÍFICA Difunde con responsabilidad ética y social el conocimiento científico, tecnológico, artístico y/o humanístico que produce de forma objetiva para aportar ideas y hallazgos científicos.</p> <p>DISEÑO Y GESTIÓN DE INFRAESTRUCTURAS SOSTENIBLES PARA EL DESARROLLO Diseña y gestiona infraestructuras seguras, eficientes y sostenibles que promueven el desarrollo socioeconómico y ambiental, integrando conocimientos de áreas como infraestructura para el transporte, estructura y materiales, computación e hidrología. Este diseño y gestión considera la sostenibilidad en todos sus aspectos y se rige por altos estándares éticos y profesionales.</p>		

DOMINIOS	OBJETOS DE ESTUDIO	RESULTADOS DE APRENDIZAJE	METODOLOGÍA	EVIDENCIAS
<p>Aplica procesos metodológicos para el desarrollo de investigación o intervención, en congruencia con el planteamiento y objetivos del proyecto a abordar.</p> <p>Muestra habilidad para la observación del fenómeno u objeto de estudio en su campo atencional.</p>	<p>1. Introducción</p> <p>1.1 El dilema de la explotación y exploración.</p> <p>1.2 Algoritmo Epsilon – Greedy.</p> <p>1.3 Algoritmo Optimistic Initial Values.</p> <p>1.4 Algoritmo Upper Confidence Bound</p> <p>1.5 Algoritmo Thompson Sampling.</p>	<p>Los estudiantes serán capaces de aplicar algoritmos de aprendizaje por refuerzo, como Epsilon-Greedy, Optimistic Initial Values, Upper Confidence Bound y Thompson Sampling, para resolver problemas de exploración y explotación en entornos dinámicos.</p> <p>Los estudiantes podrán comprender y aplicar conceptos clave de procesos de decisión de Markov (MDP), como estados, acciones, recompensas y políticas.</p> <p>Los estudiantes serán capaces de utilizar las ecuaciones de Bellman y las funciones de valor para modelar y resolver problemas de MDP.</p> <p>Los estudiantes podrán implementar métodos para resolver MDP discretos, incluyendo Programación Dinámica, Métodos Monte Carlo y Métodos de Diferencia Temporal, como Q-learning y SARSA.</p> <p>Los estudiantes serán capaces de aplicar técnicas de aprendizaje por</p>	<p>Estrategias: Clases magistrales para introducir los conceptos fundamentales. Ejercicios prácticos para implementar algoritmos de aprendizaje por refuerzo. Trabajo en grupos pequeños para discutir y resolver problemas.</p> <p>Secuencia: Introducción a los conceptos básicos de aprendizaje por refuerzo. Explicación de los algoritmos Epsilon-Greedy, Optimistic Initial Values, Upper Confidence Bound y Thompson Sampling. Implementación práctica de los algoritmos utilizando lenguajes de programación como Python. Resolución de problemas y discusión en grupos pequeños.</p> <p>Recursos Didácticos: Presentaciones multimedia. Código fuente y ejemplos de implementaciones.</p> <p>Lecturas complementarias y artículos científicos sobre aprendizaje por refuerzo.</p>	<p>Implementación de algoritmos de aprendizaje por refuerzo en un entorno de simulación y presentación de informes sobre su desempeño en la resolución de problemas específicos.</p> <p>Participación en discusiones y análisis de casos prácticos relacionados con la aplicación de algoritmos como Epsilon-Greedy, Upper Confidence Bound, entre otros.</p> <p>Entrega de trabajos escritos que expliquen en detalle la aplicación de los algoritmos de aprendizaje por refuerzo en un contexto específico, junto con su justificación y análisis de resultados.</p>

		refuerzo en entornos con espacios de estados continuos, utilizando métodos como State Aggregation, Deep SARSA, Deep Q-Learning y Advantage Actor-Critic.		
<p>Desarrolla el pensamiento científico y humanista con base en los fundamentos epistemológicos de la investigación.</p> <p>Asume una actitud ética al procesar la información derivada de los resultados de investigación.</p>	<p>2. Procesos de decisión de Markov (MDP)</p> <p>2.1 Estados, acciones, recompensas y políticas.</p> <p>2.2 Tipos de Procesos de decisión de Markov</p> <p>2.3 Ecuaciones de Bellman y funciones valor.</p> <p>2.4 Solución de la ecuación de Bellman.</p> <p>2.4.1 Iteración de Valor.</p> <p>2.4.2 Iteración de Política.</p>	<p>Los estudiantes serán capaces de identificar y comprender los componentes clave de los procesos de decisión de Markov (MDP), incluyendo estados, acciones, recompensas y políticas.</p> <p>Los estudiantes podrán aplicar las ecuaciones de Bellman y las funciones de valor para modelar y resolver problemas de MDP.</p> <p>Los estudiantes serán capaces de utilizar métodos para resolver MDP discretos, como Programación Dinámica, Métodos Monte Carlo y Métodos de Diferencia Temporal, como Q-learning y SARSA.</p>	<p>Estrategias: Clases teóricas para comprender los conceptos de MDP. Estudios de casos para aplicar los conceptos en situaciones reales. Sesiones prácticas de resolución de problemas utilizando herramientas computacionales.</p> <p>Secuencia: Introducción a los procesos de decisión de Markov. Estudio de las ecuaciones de Bellman y las funciones de valor. Aplicación de métodos para resolver MDP discretos. Resolución de casos prácticos y discusión en grupo.</p> <p>Recursos Didácticos: Material de lectura sobre MDP. Ejemplos de casos prácticos.</p> <p>Herramientas de software para resolver problemas de MDP.</p>	<p>Resolución de problemas y ejercicios prácticos basados en la aplicación de las ecuaciones de Bellman y funciones de valor en diferentes escenarios de procesos de decisión de Markov.</p> <p>Desarrollo de modelos y simulaciones de MDP para analizar y predecir el comportamiento de sistemas en condiciones diversas.</p> <p>Presentación de informes de investigación que incluyan la aplicación de métodos para resolver MDP discretos, con análisis crítico de los resultados obtenidos.</p>
<p>Propone áreas de oportunidad sobre los procesos y logros del proyecto que</p>	<p>3. Métodos para resolver MDP discretos</p>	<p>Los estudiantes serán capaces de aplicar técnicas de aprendizaje por</p>	<p>Estrategias: Talleres prácticos para aplicar diferentes</p>	<p>Desarrollo y presentación de proyectos de investigación</p>

<p>contribuyan a implementar mejores prácticas en futuros proyectos.</p>	<p>3.1 Programación Dinámica</p> <p>3.2 Métodos Monte Carlo</p> <p>3.3 Métodos de diferencia temporal.</p> <p>3.3.1 Q-learning</p> <p>3.3.2 SARSA.</p> <p>3.4 N-step bootstrapping</p> <p>3.4.1 N-step Diferencia temporal.</p> <p>3.4.2 N-step SARSA.</p>	<p>refuerzo en entornos con espacios de estados continuos, utilizando métodos como State Aggregation, Deep SARSA, Deep Q-Learning y Advantage Actor-Critic.</p> <p>Los estudiantes podrán comprender y aplicar técnicas de resolución de MDP que sean adecuadas para espacios de estados continuos, como State Aggregation y algoritmos de aprendizaje profundo.</p> <p>Los estudiantes serán capaces de diseñar y desarrollar algoritmos de aprendizaje por refuerzo para problemas complejos en ingeniería con un enfoque en el desarrollo sostenible.</p>	<p>métodos de resolución de MDP. Proyectos de investigación donde los estudiantes diseñan y desarrollan algoritmos de aprendizaje por refuerzo para problemas específicos. Seminarios para discutir avances en la investigación en el campo del aprendizaje por refuerzo.</p> <p>Secuencia: Revisión de métodos tradicionales para resolver MDP discretos. Talleres prácticos de implementación de algoritmos como Q-learning y SARSA. Desarrollo de proyectos de investigación. Presentación y discusión de los resultados de los proyectos.</p> <p>Recursos Didácticos: Conjunto de datos para proyectos de investigación. Software especializado en aprendizaje por refuerzo.</p> <p>Guías de proyectos y rúbricas de evaluación.</p>	<p>aplicada donde se implementen y evalúen algoritmos de aprendizaje por refuerzo para resolver problemas específicos en ingeniería.</p> <p>Elaboración de informes técnicos detallados que documenten el proceso de diseño, implementación y resultados obtenidos en la aplicación de métodos como Q-learning, SARSA, entre otros.</p> <p>Participación en seminarios y discusiones grupales donde se presenten y analicen los avances en la investigación sobre métodos para resolver MDP discretos.</p>
<p>Modelado Multidisciplinario para la Sostenibilidad.</p> <p>Uso Eficiente de Tecnologías y</p>	<p>4. MDP de espacio de estados continuos</p> <p>4.1 State aggregation.</p>	<p>Los estudiantes serán capaces de aplicar modelos multidisciplinarios para representar las condiciones actuales y futuras de las</p>	<p>Estrategias: Conferencias magistrales sobre modelado multidisciplinario y uso eficiente de</p>	<p>Investigación y experimentación para caracterizar materiales y desarrollar nuevos materiales sostenibles</p>

<p>Herramientas Especializadas.</p> <p>Investigación y Desarrollo de Materiales Sostenibles.</p>	<p>4.2 Algoritmo Deep-SARSA</p> <p>4.3 Algoritmo Deep Q-Learning.</p> <p>4.4 Algoritmo Advantage Actor-Critic (A2C)</p> <p>4.5 Algoritmo Asynchronous Advantage Actor-Critic (A3C)</p> <p>4.6 Procesos de decisión de Markov Parcialmente Observables.</p>	<p>infraestructuras, considerando aspectos ambientales, sociales y económicos.</p> <p>Los estudiantes podrán utilizar software especializado para resolver problemas de ingeniería con un enfoque en el desarrollo sostenible, reconociendo y abordando sus limitaciones de manera responsable y eficiente.</p> <p>Los estudiantes serán capaces de realizar investigación y experimentación para caracterizar materiales y desarrollar nuevos materiales sostenibles aplicables en la construcción de infraestructuras.</p>	<p>tecnologías especializadas.</p> <p>Laboratorios prácticos para experimentar con software especializado en la resolución de problemas de ingeniería.</p> <p>Proyectos de investigación aplicada para desarrollar nuevos materiales sostenibles.</p> <p>Secuencia: Introducción al modelado multidisciplinario para la sostenibilidad. Laboratorios prácticos utilizando software especializado. Desarrollo de proyectos de investigación aplicada. Presentación y discusión de los resultados de los proyectos.</p> <p>Recursos Didácticos: Software de modelado multidisciplinario. Materiales de laboratorio.</p> <p>Bibliografía sobre materiales sostenibles y desarrollo de infraestructuras.</p>	<p>aplicables en la construcción de infraestructuras, con presentación de informes técnicos y resultados de laboratorio.</p> <p>Desarrollo de proyectos de diseño de infraestructuras sostenibles donde se apliquen modelos multidisciplinarios y tecnologías especializadas, con entrega de documentación técnica y presentación de resultados.</p> <p>Participación en foros de discusión y presentaciones públicas donde se compartan los hallazgos y avances en la investigación sobre materiales sostenibles y desarrollo de infraestructuras.</p>
--	--	--	--	---

FUENTES DE INFORMACIÓN	EVALUACIÓN DE LOS APRENDIZAJES
<ul style="list-style-type: none"> Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction. MIT Press. [Disponible en: http://incompleteideas.net/book/the-book-2nd.html] 	<p>Criterios de Evaluación: Comprensión de conceptos teóricos y fundamentales en ingeniería, desarrollo sostenible y aprendizaje por refuerzo.</p>

<ul style="list-style-type: none"> • Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. <i>Journal of Artificial Intelligence Research</i>, 4, 237-285. [Disponible en: https://www.jair.org/index.php/jair/article/view/10182] • Szepesvári, C. (2010). <i>Algorithms for Reinforcement Learning</i>. Morgan & Claypool Publishers. [Disponible en: https://sites.ualberta.ca/~szepesva/RLBook.html] • Levine, S., Finn, C., Darrell, T., & Abbeel, P. (2016). End-to-end training of deep visuomotor policies. <i>Journal of Machine Learning Research</i>, 17(39), 1-40. [Disponible en: http://jmlr.org/papers/v17/15-522.html] • Goodfellow, I., Bengio, Y., & Courville, A. (2016). <i>Deep Learning</i>. MIT Press. [Disponible en: http://www.deeplearningbook.org/] • Silver, D., Schrittwieser, J., Simonyan, K., et al. (2017). Mastering the game of Go without human knowledge. <i>Nature</i>, 550(7676), 354-359. [Disponible en: https://www.nature.com/articles/nature24270] 	<p>Aplicación de métodos y técnicas para resolver problemas de ingeniería y tomar decisiones sostenibles utilizando el aprendizaje por refuerzo. Capacidad para realizar investigación original y contribuir al avance del conocimiento en el campo de estudio. Habilidad para comunicar eficazmente ideas, resultados y hallazgos de manera oral y escrita. Actitud ética y profesional en la realización de trabajos académicos y proyectos de investigación.</p> <p>Ponderación: Comprensión de conceptos teóricos y fundamentales: 20% Aplicación de métodos y técnicas: 30% Investigación original: 30% Comunicación eficaz: 15% Actitud ética y profesional: 5%</p> <p>Instrumentos de Evaluación: Exámenes escritos: Evaluarán la comprensión de conceptos teóricos y fundamentales en ingeniería y aprendizaje por refuerzo. Proyectos de investigación: Permitirán evaluar la capacidad de los estudiantes para aplicar métodos y técnicas de aprendizaje por refuerzo en la resolución de problemas de ingeniería y desarrollo sostenible. Presentaciones orales: Evaluarán la habilidad de los estudiantes para comunicar resultados de investigación y hallazgos de manera clara y efectiva. Informes técnicos y trabajos escritos: Permitirán evaluar la capacidad de los estudiantes para comunicar de manera escrita ideas, resultados y hallazgos de investigación.</p> <p>Observación y seguimiento del desempeño ético y profesional durante todo el programa.</p>
--	--

Cronograma del avance programático

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																