

<b>Título de proyecto</b>	Optimización Industrial mediante Aprendizaje por Refuerzo: Aplicaciones en Navegación y Gestión de Inventarios
<b>Código UTN</b>	SFCCEC544
<b>Director/a</b>	Yuan, Rebeca
<b>Dirección de correo</b>	rebecayuan@gmail.com
<b>Codirector/a</b>	Redolfi, Javier Andrés
<b>Dirección de correo</b>	javierredolfi@gmail.com
<b>Desde</b>	01/04/2025
<b>Hasta</b>	31/03/2027
<b>Resumen técnico del PID</b>	<p>En la vanguardia de la inteligencia artificial, el aprendizaje por refuerzo emerge como un área de investigación de gran relevancia y promesa. Este tipo de aprendizaje se fundamenta en la programación dinámica (Bellman, 1967); esta técnica de optimización utilizada en ciencias de la computación y matemáticas resuelve problemas complejos descomponiéndolos en subproblemas más simples. Se centra en cómo los agentes inteligentes (Russell y Norvig, 2005) pueden aprender a tomar decisiones óptimas a través de la interacción con su entorno. La secuencia de sus acciones lo aproximan al objetivo buscado. Dentro del aprendizaje automático, el aprendizaje por refuerzo no requiere de datos etiquetados, sino que permite que los agentes aprendan a través de la experiencia y la retroalimentación recibida de su entorno (Rufo Paris, J.,2022). El proceso de aprendizaje se basa en la información que el agente adquiere mediante la interacción con el entorno. Cada vez que el agente cambia de estado por una acción realizada, recibe información del entorno. Esta información, que se presenta como un premio o penalización, le permite al agente aprender a tomar la mejor decisión para la elección de la acción que lo aproxime al resultado deseado (Díaz-Iza, Henry, 2015). Este tipo de aprendizaje se puede aplicar a problemas cuya solución implica la toma de decisiones secuenciales y aparecen en una amplia variedad de campos como control automático, robótica, investigación operativa, economía, medicina entre otras (Wiering, et. al,2012). El objetivo de los algoritmos de Aprendizaje por Refuerzo es encontrar una política que maximice la recompensa obtenida por el agente a lo largo del tiempo, cantidad conocida como retorno. Estos conceptos forman parte de lo que se llama modelo de optimalidad (Pina Navarro, 2022). La navegación de robots y la optimización de inventarios son áreas críticas en la industria y la logística. La navegación de robot es utilizada en distintos entornos, ya sea industriales o sociales, acelerando la entrega de productos o herramientas; en las grandes urbes ya podemos ver la oferta de taxis autónomos (La Nación, 2022). Los sistemas de inventario gestionan grandes bases de datos, registrando entradas, salidas y movimientos de productos. Actualmente muchos sistemas de inventarios se comunican con dispositivos electrónicos para actualizar la</p>

información en tiempo real, estas soluciones son llamadas inventarios digitales (De Giovanni P.,2019), los mismo incluyen información como la cantidad de cada artículo, su ubicación y costo. Un inventario digital es la versión electrónica de tu inventario físico. Los algoritmos de aprendizaje por refuerzo pueden acoplarse y mejorar estas soluciones, en ambos campos, la gestión de inventario y la navegación de robots. Buscamos aplicar los modelos del aprendizaje por refuerzo en el ámbito industrial, para mejorar la eficiencia y la seguridad en la navegación de robots en entornos complejos de producción. A través de sus sensores, los robots pueden obtener información del ambiente que sirva de entrada, no sólo para que el robot aprenda a mejorar su trayectoria, sino también para esa información se integre a otros sistemas, en este trabajo, con el sistema de inventario. Al referirnos a la optimización de inventarios, exploraremos cómo el mismo tipo de aprendizaje puede aplicarse para gestionar de manera más eficiente los niveles de inventario en almacenes, esperando reducir los costos de almacenamiento y mejorar la disponibilidad de productos para los clientes, optimizando la tarea de compras. Ambos ambientes, navegación de robot y automatización en la gestión de inventarios, se complementan en una solución integral (Lin et al., 2024). La presente investigación busca abstraer los modelos de aprendizaje por refuerzo para poder reutilizarlos en distintos dominios, navegación y sistemas de inventario. Se espera validar la generalización y adaptabilidad de los modelos. Si un mismo algoritmo puede aplicarse eficazmente en múltiples dominios, se optimizan los recursos en términos de tiempo y costo (Silver, D., et al. ,2016). Estos dominios permitirán verificar y evaluar la adaptabilidad y robustez de los modelos del aprendizaje por refuerzo. La navegación autónoma de robots y la gestión de inventarios se presentan como áreas complementarias que brindan una solución integral, el inventario digital. Las características interdisciplinarias de nuestro grupo de investigación (Ingeniería Electrónica e Ingeniería en Sistemas de Información), permitirá llevar a cabo un desarrollo experimental que aproveche los conocimientos de ambas áreas (Manual de Frascati, 2015). Se realizará el diseño y construcción de un robot (de menos escala) que permitirá experimentar el modelo a través de la navegación. Esto generará un producto útil para aplicar el aprendizaje del modelo y también para facilitar la transferencia de conocimiento, a través de demostraciones y uso. En cuanto al modelo aplicado a inventarios, se abordará analizando sistemas de inventario, evaluando el diseño de los mismos, la comunicación de entrada de datos y sus restricciones; la visualización de los ajustes para el aprendizaje. La integración de los dominios, abre la oportunidad de la extensión del trabajo a futuro. El conocimiento adquirido, permitirá diseñar la arquitectura de un inventario digital; donde el aprendizaje por refuerzo aporte la optimización de la respuesta, donde los datos se obtengan a través redes de comunicación eficientes que permitan la transmisión de datos entre dispositivos y sistemas de inventario; permitiendo esto el control de productos en forma eficiente y la toma de decisiones en tiempo real.