



Inteligencia Artificial con Open AI Gym y Ray RLLib para el Aprendizaje Interactivo de la Planificación de Requerimiento de Materiales

Artificial Intelligence with Open AI Gym and Ray RLLib for Interactive Learning of Material Requirements Planning

Julio C. Serrano-Ruiz^a, David Peidro^{a,b}, Josefa Mula^{a,b}, Raúl Poler^{a,b}

^aCentro de Investigación en Gestión e Ingeniería de la Producción (CIGIP)

Universitat Politècnica de València (ESPAÑA)

^bDepartamento de Organización de Empresas, Escuela Politécnica Superior de Alcoy

Universitat Politècnica de València (ESPAÑA)

How to cite: Julio C. Serrano-Ruiz, David Peidro, Josefa Mula y Raúl Poler. 2022. Inteligencia Artificial con Open AI Gym y Ray RLLib para el Aprendizaje Interactivo de la Planificación de Requerimiento de Materiales. En libro de actas: *VIII Congreso de Innovación Educativa y Docencia en Red*. Valencia, 6-8 de julio de 2022. <https://doi.org/10.4995/INRED2022.2022.15877>

Abstract

Material requirements planning (MRP) is a process whose purpose is to guarantee the flow of materials into production, ensuring that each of the necessary materials is received in the required quantity and on the required date. As a problem, MRP involves products, bills of materials and components, inventory, purchase orders, and production orders, among other input variables. All this abundant dataset intervening in the solution configures a combinatorial optimisation problem of great complexity. Indeed, the MRP belongs to the group of NP-hard problems, since the time required to calculate the optimal solution, in terms of computability, is of non-polynomial order so that it increases exponentially as the volume of data increases. In real-world environments, the problem reaches such a magnitude that it usually becomes intractable for exact approximation methods. This paper proposes the use of Open AI Gym y RLLib, two well-known frameworks for deep reinforcement learning (DRL), to carry out MRP simulation practices based on the project-based learning (PBL) teaching technique, in the educational context of the Master's Degree in Industrial Organisation and Logistics (MUIOL) currently taught at the Alcoy Campus of the Universitat Politècnica de València (UPV). The contribution of this study is twofold: i) it brings artificial intelligence closer to the teaching context, and ii) it provides a reference for developing teaching materials for the study of MRP.

Keywords: *Material requirements planning, Open AI Gym, deep reinforcement learning, project-based learning, Master's Degree in Industrial Organisation and Logistics.*

Resumen

La planificación de requerimiento de materiales, actividad conocida como MRP por sus siglas en inglés, es un proceso cuyo propósito es garantizar el flujo de materiales en la producción, asegurando que cada uno de los materiales necesarios sea recibido en la

cantidad y fecha requeridas. Como problema, el MRP involucra productos, listas de materiales y componentes, inventario, pedidos de compra, y ordenes de producción, entre otras variables de entrada. Todo este abundante conjunto de datos interviniendo en la solución configura un problema de optimización combinatoria de gran complejidad. En efecto, el MRP pertenece al grupo de los problemas NP-hard, ya que el tiempo requerido para calcular la solución óptima, en términos de computabilidad, es de orden no polinomial, de modo que aumenta de forma exponencial frente a incrementos en el volumen de datos. En entornos reales el problema alcanza tal dimensión que, por lo general, se vuelve intratable para los métodos exactos de aproximación. Este artículo propone el empleo de Open AI Gym y Ray RLLib, dos conocidos marcos de trabajo de aprendizaje por refuerzo profundo (ARP), para la realización de prácticas de simulación del MRP sobre la base de la técnica docente del aprendizaje basado en proyectos (ABP), en el contexto educativo del Máster Universitario en Ingeniería de Organización y Logística (MUIOL) que, actualmente, se imparte en el Campus de Alcoy de la Universitat Politècnica de València (UPV). La contribución de este estudio es doble: i) aproxima la inteligencia artificial al contexto de la enseñanza; y ii) proporciona una referencia para desarrollar materiales didácticos para el estudio del MRP.

Palabras clave: *planificación de requerimiento de materiales, Open AI Gym, aprendizaje por refuerzo profundo, aprendizaje basado en proyectos, Master Universitario en Ingeniería de Organización y Logística.*

Introducción

Es prácticamente imposible encontrar algún sector de nuestra sociedad y economía que, de algún modo u otro, no se vea afectado por la transformación digital a la que venimos asistiendo en años recientes (Serrano-Ruiz et al., 2021). No obstante, su impacto no es uniforme, y existen ciertos contextos en los que la presión transformadora es considerable. Uno de ellos es el de la cadena de suministro industrial. Una prueba de ello es la génesis, desde hace una década, de paradigmas de producción como la Industria 4.0, y derivadas tales como la Cadena de Suministro 4.0 o Logistics 4.0, así como la generación de algunos paradigmas equivalentes, entre los que destacan *Smart Manufacturing* y *Digital Manufacturing*, y también versiones de última generación, como la iniciativa europea de la Industria 5.0 (Breque et al., 2021). Todas estas iniciativas han impulsado una corriente de cambio que algunos creen una nueva revolución industrial y que, aunque inicialmente inspiró a los académicos, posteriormente, empezó también a extenderse y arraigar entre los cuadros técnicos y gerenciales de las empresas industriales.

Este cambio en la cadena de suministro industrial precisa gestión. Los objetivos de los promotores de la transformación digital en cada organización pueden diferir en los niveles de competitividad, eficiencia, calidad, o respeto medioambiental, entre otras capacidades perseguidas, pero, generalmente, coinciden en lo primordial: ser más resilientes y sostenibles, i.e., adquirir la cualidad de adaptarse a las amenazas y adversidades para pervivir en el tiempo sin comprometer el futuro de la organización y de su entorno. La gestión que implica la consecución y perfeccionamiento de estos objetivos supone todo un reto dentro de la organización tanto para sus promotores como para el personal a cargo de su implementación, y afrontar exitosamente este desafío requiere el dominio de ciertas competencias profesionales, en cuya adquisición

las universidades deben jugar un rol protagonista, especialmente, a través de sus titulaciones en ingeniería industrial y afines.

La educación recibida en la universidad por el estudiante de ingeniería en organización industrial ha de prepararlo para su futuro profesional en este contexto de transformación digital. Los que se empleen en áreas de planificación y control de la producción precisarán haber adquirido un conocimiento teórico y práctico sobre los problemas que enfrenta esta área, pero dicho aprendizaje debe efectuarse desde esta nueva perspectiva marcada por los principios de diseño de la Industria 4.0 y sus tecnologías habilitadoras para que, consecuentemente, la adquisición de conocimiento derive en las competencias que requiere este nuevo contexto profesional.

En el nivel táctico de decisión del área de planificación y control de la producción se encuentra uno de sus problemas más conocidos: la planificación de requerimiento de materiales (MRP), que hace posible al planificador conocer las fechas en las que será necesario aprovisionar cada material y componente, y en qué cantidad, para dar así cumplimiento al plan maestro de producción sin perturbar o interrumpir el flujo planificado de producción (Whybark and Williams, 1976; Mula et al. 2003). Es un problema complejo, dependiente de un número considerable de variables. En un entorno real de producción, dinámico y estocástico, es difícil obtener la solución óptima, incluso en el caso de los problemas de menor dimensión, volviéndose intratable cuando el volumen de datos a procesar crece un poco más (Guillaume et al., 2017). Tradicionalmente, las dificultades planteadas por el problema han sido abordadas admitiendo las soluciones sub-óptimas obtenidas por procedimientos no analíticos, e.g. la simulación o los métodos heurísticos. Pero la expansión de la inteligencia artificial en el ámbito industrial, fundamentalmente en la forma de aprendizaje automático, está cambiando el panorama abriendo un espacio de nuevas oportunidades de desarrollo en este campo (Serrano-Ruiz et al., 2022). El aprendizaje por refuerzo es un método de aprendizaje automático inspirado en el procedimiento que, según la psicología conductista, emplea el aprendizaje humano, y que consiste en que el agente decisor aprende interaccionando con su entorno, de modo que escoge cada acción a efectuar sobre él tras observar su estado y evaluar el beneficio obtenido de sus acciones, que se obtiene en la forma de premio o recompensa. El empleo de redes neuronales como soporte al aprendizaje por refuerzo (Park et al., 2020), conocido como aprendizaje por refuerzo profundo (ARP), eleva las capacidades del método constituyendo una prometedora herramienta para todos aquellos problemas modelables como procesos de decisión de Markov, tal y como ocurre con MRP. La aparición de marcos de trabajo open-source especializados en ARP (Kuhnle et al., 2019), como es el caso de Open AI Gym y Ray RLLib, aproximan y facilitan a los ingenieros el acceso a la inteligencia artificial, sean éstos investigadores, profesionales de las organizaciones o estudiantes.

El presente artículo propone una metodología que combina la fórmula docente del aprendizaje basado en proyectos (ABP), y las herramientas Open AI Gym y Ray RLLib, para su utilización conjunta como instrumentos de apoyo en el aprendizaje práctico del problema de MRP desde la perspectiva de la inteligencia artificial. Esta propuesta está orientada a los alumnos del Máster Universitario en Ingeniería de Organización y Logística (MUIOL) impartido en el Campus de Alcoy de la Universitat Politècnica de València (UPV). Las contribuciones principales de este trabajo residen, esencialmente, en: (i) potenciar el proceso educativo a través de procedimientos prácticos respaldados por tecnologías de simulación e inteligencia artificial, y (ii) erigir referencias que fundamenten el desarrollo de nuevos contenidos con los que avanzar en la docencia práctica del entorno definido por MRP.

1. Descripción del marco docente

Se consideran cuatro elementos que configuran el marco docente de la presente propuesta: (i) el título de grado, (ii) las asignaturas impartidas, y (iii) el problema de MRP en este contexto.

1.1. Máster Universitario en Ingeniería de Organización Industrial y Logística (MUIOL)

El MUIOL impartido por la UPV es un título universitario homologado que pertenece al área de Ingeniería y Arquitectura. Se halla incluido desde el año 2.014 en el Registro de Universidades, Centros y Títulos del Ministerio de Universidades. Para recibir el título es preciso cursar 2 años y obtener 90 créditos del Sistema Europeo de Transferencia y Acumulación de Créditos (ECTS), los cuales están distribuidos en 42 ECTS para asignaturas obligatorias, 36 ECTS para asignaturas optativas y 12 ECTS para la preparación y defensa del Trabajo Fin de Máster (Ministerio de Universidades de España, 2022).

El MUIOL capacita a sus titulados para la gestión y dirección de empresas de producción y servicios en áreas como: compras, logística, producción, procesos, productos, costes, recursos humanos o sistemas de información. El objetivo de este máster universitario es formar profesionales para la empresa que sean capaces de analizar, modelar, diseñar, implantar y mejorar sistemas complejos compuestos por personas, materiales, dinero, información, máquinas, tecnología y energía, con el fin de ofrecer productos y servicios en el menor plazo y con la mayor productividad, calidad, fiabilidad y eficiencia posible. Sus titulados son capacitados para el diseño de sistemas productivos y logísticos sujetos a restricciones técnicas y de recursos, la evaluación del rendimiento de dichos sistemas detectando y priorizando áreas de mejora, y la gestión de estos sistemas en funcionamiento. El máster está dirigido a graduados en ingeniería y titulados en ingenierías técnicas, procedentes de universidades españolas y extranjeras, que estén interesados en complementar su formación en el ámbito de la ingeniería de organización y la logística (Universitat Politècnica de València [B], 2022).

1.2. Asignaturas objeto de la propuesta metodológica

Las asignaturas objetivo de la presente propuesta son Métodos Cuantitativos de Organización Industrial (MCOI) y Organización de la Producción (OP):

1.2.1 Asignatura de MCOI

Se trata de una asignatura obligatoria impartida en el primer curso del MUIOL con una carga lectiva de 6 ECTS. La asignatura tiene un enfoque, eminentemente, práctico y pretende dotar a los alumnos de conocimientos sobre determinados modelos y herramientas cuantitativas (programación lineal, entera y no lineal; teoría de inventarios; teoría de grafos; teoría de la decisión; decisión multicriterio; teoría de colas; simulación) que les ayuden a abordar con éxito problemas de organización industrial y logística.

Tabla 1. Distribución horaria de la asignatura de MCOI.

| Unidad didáctica | Teoría Aula | Teoría Seminario | Práctica Aula | Práctica Laboratorio |
|---|----------------|---------------------|------------------|-------------------------|
| 1. Introducción a los métodos cuantitativos | 1 | --- | --- | --- |
| 2. Programación lineal, entera y no lineal | 6 | 1 | 4 | 5 |
| 3. Teoría de inventarios | 3 | 1 | 2 | --- |
| 4. Teoría de grafos | 4 | 1 | 2 | 2 |
| 5. Teoría de la decisión | 4 | 1 | 4 | 4 |

| Unidad didáctica | Teoría Aula | Teoría Seminario | Práctica Aula | Práctica Laboratorio |
|--------------------|----------------|---------------------|------------------|-------------------------|
| 6. Teoría de colas | 3 | 1 | 2 | --- |
| 7. Simulación | 4 | --- | 1 | 4 |
| Total horas | 25 | 5 | 15 | 15 |

1.2.2 Asignatura de OP

Se trata de una asignatura obligatoria impartida en el primer curso del MUIOL con una carga lectiva de 6 ECTS. La asignatura pretende proporcionar modelos, métodos y herramientas para resolver los diferentes problemas de planificación, programación y control de la producción: (i) la asignación de recursos globales (capacidades de producción, necesidades de recursos humanos, almacenes) mediante la planificación agregada de la producción; (ii) la desagregación de los planes agregados mediante la programación maestra de la producción; (iii) el cálculo de las necesidades de materiales y recursos mediante la planificación de necesidades de materiales y recursos; (iv) la programación de la producción mediante reglas de prioridad; y (v) la aportación de elementos y procedimientos para la producción ajustada.

Tabla 2. Distribución horaria de la asignatura de OP.

| Unidad didáctica | Teoría Aula | Teoría Seminario | Práctica Aula | Práctica Laboratorio |
|---|----------------|---------------------|------------------|-------------------------|
| 1. Introducción al Sistema de Planificación y Control de Producción | 2 | 1 | --- | 1 |
| 2. Planificación de la Producción | 5 | 1 | 3 | 4 |
| 3. Planificación de Requerimientos de Materiales y Recursos | 6 | 1 | 5 | 4 |
| 4. Programación y Control de Producción | 6 | 1 | 5 | 4 |
| 5. Sistemas de Producción Ajustada | 6 | 1 | 2 | 2 |
| Total horas | 25 | 5 | 15 | 15 |

1.3. La planificación de requerimiento de materiales

Se trata de un procedimiento de planificación de la producción cuyo objetivo es aprovisionar, en la fecha óptima para su entrada en producción, los materiales y componentes requeridos (Ptak et al., 2013). Se trata de un problema con décadas de investigación, que empezó a aplicarse en la industria militar de los Estados Unidos de Norte América a finales de los años 50 del pasado siglo, para controlar los procesos complejos de fabricación en los que se hallaban implicados productos con un número significativo de materiales, piezas, componentes, accesorios, etc. provenientes de múltiples proveedores, resultando de gran utilidad para dimensionar la capacidad de producción e inventario, equilibrar la producción, ajustar la mano de obra, estimar los plazos de fabricación y, en definitiva, organizar la producción.

El problema es estudiado en la tercera unidad didáctica de la asignatura de OP durante el segundo semestre del primer curso, después de haber cursado la asignatura de MCOI, que dota de conocimientos y herramientas al alumno para su aprendizaje práctico, principalmente, mediante programación matemática, teoría de inventarios y simulación.

2. Recursos docentes

Son tres, básicamente, los instrumentos que requiere la implementación de la metodología propuesta: (i) un entorno de simulación y prueba de algoritmos de ARP, que en este caso es proporcionado por el software Open AI Gym; (ii) una librería de algoritmos de ARP compatible con el anterior entorno como Ray RLlib; y (iii) un método docente, ABP, con el que llevar a cabo la propuesta.

2.1. Open AI Gym

Open AI Gym es un *toolkit* basado en un interfaz de código abierto que permite modelar entornos caracterizables como procesos de decisión de Markov sujetos a la actuación de agentes basados en algoritmos de aprendizaje por refuerzo. Está, especialmente, diseñado para llevar a cabo la prueba y experimentación de la acción de estos agentes inteligentes interactuando con un entorno elegido entre los incluidos en su librería, o uno personalizado desarrollado y registrado con la aplicación. El, comunmente, denominado Gym es un instrumento popular entre los desarrolladores de ARP, pues posee varias características destacables que, conjuntamente, le proporcionan un gran potencial: (i) está basado en Python, un estándar académico y profesional en cuanto a lenguajes de programación; (ii) es compatible con cualquier librería de computación numérica para aprendizaje profundo, i.e. redes neuronales, como TensorFlow o Theano; (iii) ofrece una librería propia de entornos con los que entrenar y evaluar agentes de ARP, principalmente, basados en varios de los conocidos juegos Atari; y (iv) permite la generación de entornos propios personalizados en un contexto basado en los procesos de decisión de Markov y sus principales elementos: estado, acción y recompensa (Open AI Gym, 2022).

2.2. Ray RLlib

Ray RLlib es una biblioteca de código abierto de algoritmo de aprendizaje por refuerzo basada en Python, que ofrece soporte para cargas de trabajo altamente distribuidas, al tiempo que mantiene interfaces de programación de aplicaciones unificadas y sencillas para una gran variedad de aplicaciones industriales. Su diseño le hace ampliamente flexible: (i) permite trabajar con agentes únicos o configuraciones multiagente; (ii) puede entrenar agentes con datos nuevos o históricos, y sintéticos o reales; (iii) permite la acción del agente tanto en línea como fuera de la misma; y (iv) aunque no incorpora su propio marco de aprendizaje profundo, es compatible con TensorFlow y PyTorch. Ray RLlib ofrece soluciones sencillas adaptables a cada necesidad de toma de decisiones. No es necesario ser un experto en aprendizaje por refuerzo para utilizar la aplicación. RLlib es utilizado en entornos de control del clima, producción y logística, finanzas, juegos, automóvil, robótica, y diseño de barcos, entre otros (Ray RLlib, 2022).

2.3. Aprendizaje basado en proyectos

El planteamiento de la metodología ABP se basa en la hipótesis de que los problemas que el alumno encuentra en el mundo real despiertan su interés, a la par que le facilitan la adquisición de nuevos conocimientos en un entorno de resolución de problemas que le invita a su aplicación (Efstratia, 2014). El método se desarrolla teóricamente en 7 etapas: (i) enunciación de objetivos de aprendizaje; (ii) conocimiento de los materiales; (iii) formación de capacidades, (iv) selección del tema; (v) propuesta de proyecto; (vi) desarrollo del proyecto; y (vii) presentación del proyecto (Jalinus et al., 2017), si bien admite ciertas variaciones. Con esta metodología se estimula a los estudiantes a aprender a aprender, si bien cabe admitir que esta metodología experiencial requiere un esfuerzo adicional por parte del alumno.

3. Propuesta de proyecto docente

En esta sección se propone un proyecto docente para el aprendizaje de MRP en un contexto experiencial de ABP basado en el desarrollo de un entorno Markoviano programado en Open AI Gym que simula el proceso, permitiendo la toma de decisiones en ese entorno de agentes de ARP de la biblioteca Ray RLlib, para: (i) aproximarse al conocimiento de las herramientas de inteligencia artificial basadas en aprendizaje automático; y (ii) emplear este conocimiento para captar experiencialmente la dinámica que impulsa el proceso de MRP. El proyecto está aplicado a las asignaturas de MCOI y OP del MUIOL que son impartidas en el Campus de Alcoy de la UPV.

En primer lugar, se expone el sistema de experimentación que configura el marco técnico en el cual desarrollar la experimentación sobre MRP y, a continuación, los detalles del proyecto docente.

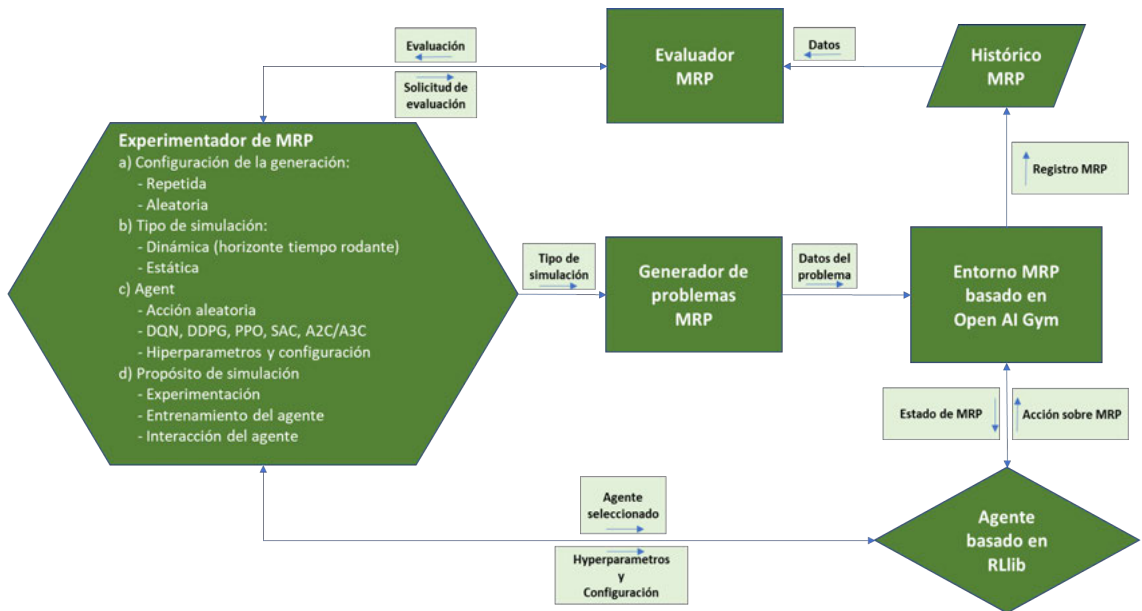


Fig. 1 Elementos constituyentes y configuración del experimentador de MRP

3.1. Experimentador de MRP con Open AI Gym

Esta propuesta se basa en el desarrollo *ad hoc* y el empleo de un entorno de simulación de un sistema de MRP a partir del software Open AI Gym (Figura 1), diseñado como simulador de tiempo discreto. Las características dimensionales del problema son configuradas, previamente, de forma aleatoria o manualmente por el usuario. Un generador de problemas genera un escenario de planificación sintético basado en las anteriores características dimensionales que, en cada paso de simulación correspondiente a un periodo de planificación, solicita la acción del agente seleccionado. Este sistema de interacción entorno-agente está controlado por el experimentador, a través del cual se establece lo siguiente: (i) la configuración del generador de problemas, estableciendo si tras cada episodio completo se repite el mismo entorno, o si se genera uno nuevo; (ii) el tipo de simulación, i.e. dinámica, en la que la planificación progresa de forma continua, sin restricciones temporales, o estática, en la que la planificación se realiza durante un periodo de tiempo limitado y predefinido, transcurrido el cual, ésta finaliza; (iii) el tipo de agente, bien un decisor de base aleatoria, bien uno seleccionado entre los algoritmos de ARP disponibles en la librería Ray RLlib, e.g. el algoritmo *deep Q-network* (DQN), el *deep deterministic policy gradient* (DDPG), el *proximal policy*

optimisation (PPO), el *soft actor-critic* (SAC), o el *advantage actor-critic* (A2C, A3C); más la configuración de la red neuronal y los hiperparámetros del algoritmo; (iv) la finalidad de la simulación, que será la simple experimentación, el entrenamiento del agente, o su interacción con el entorno para la realización de evaluaciones de rendimiento. El experimentador descrito se concibe como un sistema impulsado por datos y un registro histórico con el objetivo de permitir posteriores procesamientos de información o nuevas evaluaciones de los datos.

3.2. Proyecto docente basado en ABP y Open AI Gym para el aprendizaje práctico de MRP

El proyecto docente se centra en el aprendizaje experiencial y, por ello, está focalizado en sesiones prácticas: las prácticas de aula y de laboratorio.

Las sesiones prácticas de aula se basan, principalmente, en conferencias de índole participativa y en la resolución de los problemas propuestos durante las sesiones de teoría de aula o seminario. Estas sesiones prácticas se centran, principalmente, en el análisis y la resolución en grupo de los problemas incluidos en los contenidos cubiertos por las asignaturas, haciendo hincapié en el análisis y la modelización de problemas, lo que a posteriori será de utilidad para llevar a cabo variaciones y extensiones de los problemas estudiados.

Respecto a las sesiones prácticas de laboratorio, éstas persiguen mejorar la comprensión e integración de los conocimientos teóricos y prácticos tratados en las anteriores sesiones mediante su ejercicio práctico, y promover que el alumno practique competencias transversales relacionadas con el trabajo en equipo, lo que se fomenta redactando informes, presentando oralmente su trabajo en público, y debatiendo los resultados obtenidos. Además, se emplea equipamiento informático para la resolución de los problemas.

El modelo propuesto se basa en un proceso temporal que abarca los dos semestres del primer curso del MUIOL y se estructura en cuatro etapas: (i) fase de desarrollo e implementación del experimentador de MRP con Open AI Gym, etapa que no forma parte de la propuesta docente, pero que debe ser tenida en consideración desde el punto de vista de su planificación temporal, obtención de recursos, programación del entorno y depuración; (ii) fase de familiarización del alumno con el lenguaje de programación Python, que es una cuestión fundamental para el desarrollo de las prácticas de aula y laboratorio, y que debe ser fomentada desde el inicio del curso y desde todos los ámbitos del máster pero, especialmente, desde la primera clase de la asignatura de MCOI; (iii) fase de familiarización del alumno con el entorno de desarrollo formado por Open AI Gym y Ray RLLib en la asignatura de MCOI; y (iv) fase de familiarización con el experimentador MRP y aprendizaje de la dinámica del problema de MRP mediante la experimentación, en el marco de la asignatura de OP.

Procede indicar que es recomendable que la fase de familiarización con el lenguaje Python se fomente en el ámbito privado del alumno y fuera de la distribución horaria planificada para las asignaturas de MCOI y OP, si bien el profesor prestará soporte en horario de tutoría, solucionando las dudas que puedan surgir a los alumnos noveles en la materia.

A continuación, se detallan los objetivos de la innovación docente, la organización propuesta para el proyecto en ambas asignaturas, y la metodología propuesta de evaluación de los resultados de la innovación.

3.2.1 Objetivos de la innovación docente

En primer y más importante lugar, el modelo propuesto pretende obtener un impacto positivo en el rendimiento del alumnado. Dicho impacto se materializará abordando tres de las cinco dimensiones clave

identificadas por Mauri et al., 2017 en la evaluación de la calidad de los proyectos de innovación docente universitaria:

- Potenciar la autonomía y la autoregulación del aprendizaje del alumno.
- Fomentar el trabajo cooperativo.
- Impulsar las tecnologías digitales como instrumento de los procesos de enseñanza-aprendizaje.

3.2.2 Desarrollo del proyecto en la asignatura de MCOI

Se propone emplear las cinco horas de prácticas de aula y laboratorio de las programadas para la séptima unidad didáctica, Simulación, para abordar la adquisición de conocimiento práctico por parte de los alumnos en la herramienta Open AI Gym y la simulación de tiempo discreto, así como en la interacción de los agentes de ARP que integran la librería Ray RLlib con el entorno simulado por Open AI Gym. Esta tarea se realizará en tres sesiones distintas: (i) La única hora destinada a las prácticas de aula se empleará, primeramente, para la organización de los grupos. A continuación, se iniciará a los alumnos en la interfaz de programación basada en Python del Open AI Gym, desde su procedimiento de instalación hasta el aprendizaje de sus procedimientos `init`, `reset`, `step`, y `render`, pasando por el registro de entornos personalizados, y por la integración de la acción del agente en el entorno. Adicionalmente, se explicarán a los alumnos las normas para la elección y preparación del problema, el cual debe ser seleccionado por los alumnos en grupo. El profesor sugerirá emplear un problema sencillo de entre los abordados durante la tercera unidad didáctica, Teoría de Inventarios. (ii) Las primeras 2 horas de prácticas de laboratorio servirán para entrenar y enseñar a los alumnos a dar sus primeros pasos con la herramienta, interaccionando, inicialmente, con ella mediante el empleo, por ejemplo, de alguno de los entornos preprogramados de Atari, o bien programando un ejemplo básico y estático de inventario. (iii) Las últimas 2 horas de prácticas de laboratorio se emplearán en la presentación oral en público del informe del trabajo efectuado por cada grupo, con su correspondiente turno de preguntas y respuestas.

Los grupos estarán formados, preferentemente, por tres alumnos, con un máximo de cuatro, para fomentar tanto la organización de tareas como el análisis crítico y la discusión constructiva en el seno del grupo. El trabajo se llevará a cabo en las siguientes nueve fases: (i) selección del problema, para lo que el profesor sugerirá algunas posibilidades; por ejemplo, del ámbito industrial, como el control de inventario de algún componente sencillo de automoción, o más general, como el control de las existencias de alimentos en un restaurante, pero los alumnos también podrán salirse de estas sugerencias si el problema elegido responde a las especificaciones docentes planificadas; (ii) modelización del problema como un proceso de decisión de Markov, identificando el espacio de estados, el espacio de acción, y las recompensas locales y/o globales; (iii) programación del entorno de inventario, definiendo las funciones `init`, `reset`, `step` y `render`, así como el fichero instanciador; (iv) registro del entorno en Open AI Gym, para su activación; (v) prueba con un agente de acción aleatoria, para la evaluación del funcionamiento del entorno de inventario programado; (vi) prueba con dos de los algoritmos ARP proporcionados por la librería Ray RLlib, uno de tipo *value-based*, como el DQN, y otro del tipo *actor-critic*, a elegir entre PPO, SAC, A2C o A3C; (vii) experimentación para evaluar el rendimiento de cada agente en su interacción con el entorno; (viii) redacción del informe, justificando los resultados; y (ix) presentación oral en público, para difundir y discutir los resultados durante la sesión final (Figura 2).

3.2.3 Desarrollo del proyecto en la asignatura de OP

Respecto al caso de esta asignatura, se destinarán al proyecto cinco de las nueve horas de las sesiones prácticas que están programadas para la tercera unidad didáctica, denominada Planificación de

Requerimientos de Materiales y Recursos, empleando para ello una de las horas de prácticas de aula, y las cuatro horas de prácticas de laboratorio. El proyecto consiste en abordar el conocimiento práctico del problema de MRP y conocer su dinámica desde la perspectiva de la inteligencia artificial, usando para ello la herramienta Open AI Gym. El trabajo se realizará en tres sesiones diferentes distribuidas como sigue: (i) en las horas de prácticas de aula se organizarán, en primer lugar, los grupos, y después se iniciará a los alumnos en el entorno MRP y sus principales componentes: experimentador, generador de problemas, entorno MRP propiamente dicho, agente ARP, base de datos con registro histórico, y evaluador; al final de la hora se explicarán las normas de selección del problema, que deberá basarse en un caso real; (ii) en las dos primeras horas de las prácticas de laboratorio los alumnos introducirán las instancias de su problema en el sistema y experimentarán con él, variando los hiperparámetros de la red neuronal, el tipo de agentes ARP y su configuración, registrando y evaluando los resultados; y (iii) en las dos últimas horas de las prácticas de laboratorio cada grupo presentará los resultados de su trabajo y responderá a las preguntas de los otros alumnos y del profesor.

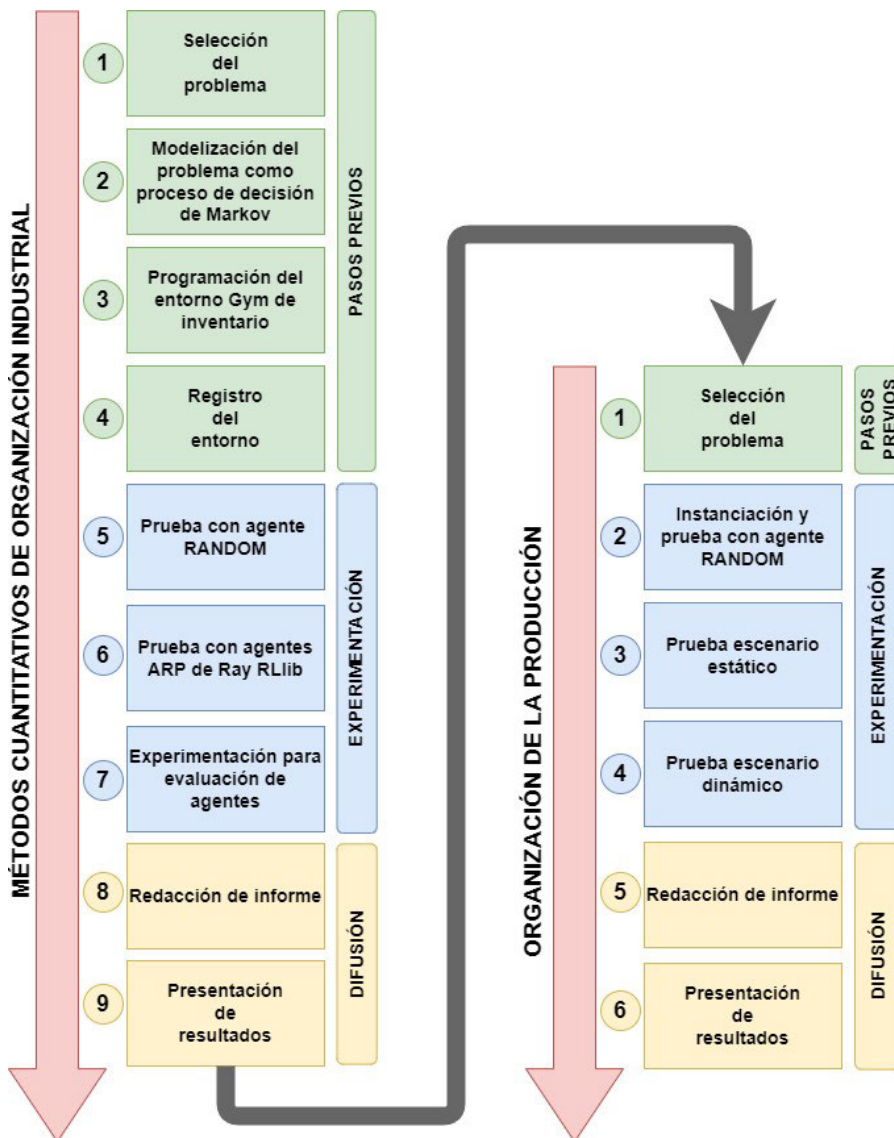


Fig. 2 Secuencia de tareas del proyecto en ambas asignaturas

Los grupos estarán formados, preferentemente, por dos alumnos, con un máximo de tres, quienes efectuarán el trabajo en las siguientes seis etapas: (i) selección del problema real, para lo que el profesor sugerirá algunas posibilidades dentro del contexto industrial del ámbito provincial, aunque los alumnos podrán utilizar casos de estudio fuera de ese contexto, siempre y cuando el problema elegido responda a las especificaciones docentes planificadas; (ii) instanciación del problema y prueba con un agente de acción aleatoria en dos escenarios, el estático y el dinámico, verificando que la instanciación haya sido efectuada con éxito; (iii) prueba comparativa en el escenario estático con tres de los algoritmos ARP proporcionados por la librería Ray RLlib, uno de tipo value-based, como el DQN, y otros dos del tipo actor-critic, a elegir entre PPO, SAC, A2C o A3C; (iv) prueba comparativa en el escenario dinámico con los mismos algoritmos ARP seleccionados en la etapa anterior; (v) redacción del informe justificando los resultados; y (vi) presentación oral en público, para difundir y discutir los resultados durante la sesión final (Figura 2).

3.2.4 Evaluación de los resultados

El grado de consecución de los objetivos de la propuesta docente conllevará la medición previa y posterior de las siguientes variables:

- Tasa de rendimiento académico, tomando como indicadores las calificaciones parciales obtenidas en las sesiones prácticas, así como las finales.
- Grado de satisfacción del alumnado, tomando como indicador el resultado de las encuestas de evaluación del profesorado.

4. Conclusiones

Este artículo ha propuesto la utilización de las herramientas Open AI Gym y Ray RLlib para el aprendizaje experiencial e interactivo del problema de MRP en un contexto de ABP, en el ámbito docente de dos de las asignaturas del MUIOL que se imparten en el Campus de Alcoy de la UPV: MCOI y OP. Para ello, se ha definido el contexto educativo, integrado por el MUIOL como titulación, las dos asignaturas involucradas, y el problema de MRP. A continuación, se han introducido los recursos docentes necesarios para el desarrollo del proyecto: las herramientas de software open-source basadas en Python denominadas Open AI Gym y Ray RLlib, por un lado, y la metodología del ABP, por otro. Finalmente, se han pormenorizado las actividades a desarrollar en el proyecto, definiendo los objetivos de la innovación docente, detallando su distribución temporal y la secuencia de realización, y los criterios e indicadores de evaluación del resultado de la metodología. El modelo de ABP presentado recurre a las técnicas de simulación y aprendizaje automático, permitiendo al alumno no solo adquirir un conocimiento profundo del problema de MRP y su dinámica, sino hacerlo, además, mediante la inmersión en un contexto de modelización basado en la inteligencia artificial, lo que enriquece, significativamente, el aprendizaje. El proyecto docente presentado podría facilitar la adquisición de conocimientos en planificación de la producción, procesos de decisión de Markov, entornos de simulación de tiempo discreto, aprendizaje por refuerzo, y redes neuronales, ayudando con ello a que el alumno dé un salto cualitativo en el desarrollo de sus competencias como futuro ejerciente de la ingeniería de organización dentro de entornos de producción basados en la Industria 4.0. Finalmente, la metodología expuesta podría proporcionar una base referencial para el desarrollo futuro de nuevos materiales docentes para el aprendizaje de MRP.

5. Referencias

- EFSTRATIA, D. (2014). “Experiential Education through Project Based Learning”. *Procedia - Social and Behavioral Sciences*. Elsevier BV; Oct;152:1256–60
- BREQUE, M., DE NUL, L., PETRIDIS, A. (2021). “*Industry 5.0: towards a sustainable, human-centric and resilient European industry*”, European Commission, Directorate-General for Research and Innovation. Publications Office, <https://data.europa.eu/doi/10.2777/308407>
- GUILLAUME, R., THIERRY, C., ZIELIŃSKI, P. (2017). “Robust material requirement planning with cumulative demand under uncertainty”. *International Journal of Production Research*. Taylor and Francis Ltd.; Jul 19;55(22):6824–45
- JALINUS, N., NABAWI, R.A., MARDIN, A. (2017). “The Seven Steps of Project Based Learning Model to Enhance Productive Competences of Vocational Students”. *Proceedings of the International Conference on Technology and Vocational Teachers (ICTVT 2017)* [Internet]. Atlantis Press; 2017. p. 251–6. Available from: <https://doi.org/10.2991/ictvt-17.2017.43>
- KUHNLE, A., SCHÄFER, L., STRICKER, N., LANZA, G. (2019). “Design, Implementation and Evaluation of Reinforcement Learning for an Adaptive Order Dispatching in Job Shop Manufacturing Systems”. *Procedia CIRP*. Elsevier; 2019 Jan 1;81:234–9
- MINISTERIO DE UNIVERSIDADES (ESPAÑA). *Registro de Universidades, Centros y Títulos (RUCT)*. <<https://www.educacion.gob.es/ruct/estudio.action?codigoCiclo=SC&codigoTipo=G&CodigoEstudio=4314071&actual=estudios>> [Consulta: 06/03/2022]
- OPEN AI GYM. *Documentación*. <<https://gym.openai.com/>> [Consulta: 06/03/2022]
- PARK, I.B., HUH, J., KIM, J. (2020). “A Reinforcement Learning Approach to Robust Scheduling of Semiconductor Manufacturing Facilities”; *IEEE Transactions on Automation Science and Engineering*; 17(3)
- POLER, R., RODRÍGUEZ VILLALOBOS, A., MULA JOSEFA (2003). *Programación lineal continua. Planificación de requerimientos de materiales (MRP)*. Valencia: Editorial UPV
- PTAK, CA., SMITH C., ORLICKY, J. (2013). *Orlicky's material requirements planning*. New York: McGraw-Hill. <<http://www.accessengineeringlibrary.com/browse/orlickys-material-requirements-planning-third-edition>>
- RAY RLlib. *Documentación*. <<https://docs.ray.io/en/latest/rllib/index.html>> [Consulta: 06/03/2022]
- SERRANO-RUIZ, J.C., MULA, J., POLER, R. (2021). “Smart Master Production Schedule for the Supply Chain: A Conceptual Framework”. *Computers* [Internet]. Nov 23;10(12):156
- SERRANO-RUIZ, J.C., MULA, J., POLER, R. (2022). “Development of a multidimensional conceptual model for job shop smart manufacturing scheduling from the Industry 4.0 perspective”. *Journal of Manufacturing Systems* (in press)
- UNIVERSIDAD POLITÈCNICA DE VALÈNCIA [A]. Home. <<https://www.upv.es/>> [Consulta: 06/03/2022]
- UNIVERSIDAD POLITÈCNICA DE VALÈNCIA [B]. *Máster Universitario en Ingeniería de Organización y Logística (MUIOL)*. <<https://www.upv.es/titulaciones/MUIOL/indexc.html>> [Consulta: 06/03/2022]
- WHYBARK, D.C., WILLIAMS, J.G. (1976). “Material requirements planning under uncertainty”. *Decision sciences*. Wiley Online Library; 7(4):595–606