

Entrenamiento de robot caminante por imitación y RL usando ZMP como generador de datos de referencia

Autor:

Ing. Francisco Antonio Cofré Villalón

Director:

Título y Nombre del director (pertenencia)

Índice

1. Descripción técnica-conceptual del proyecto a realizar	5
2. Identificación y análisis de los interesados	6
3. Propósito del proyecto	6
4. Alcance del proyecto	6
5. Supuestos del proyecto	7
6. Product Backlog	7
7. Criterios de aceptación de historias de usuario	8
8. Fases de CRISP-DM	9
9. Desglose del trabajo en tareas	9
10. Planificación de Sprints	11
11. Diagrama de Gantt (sprints)	13
12. Gobernanza de datos	15
13. Gestión de riesgos	15
14. Sprint Review	16
15. Sprint Retrospective	17

Registros de cambios

Revisión	Detalles de los cambios realizados	Fecha
0	Creación del documento	21 de octubre de 2025
1	Se completa hasta el punto 5 inclusive	5 de noviembre de 2025
2	Se completa hasta el punto 9 inclusive	12 de noviembre de 2025
3	Se completa hasta el punto 12 inclusive	23 de noviembre de 2025

Acta de constitución del proyecto

Buenos Aires, 21 de octubre de 2025

Por medio de la presente se acuerda con el Ing. Francisco Antonio Cofré Villalón que su Trabajo Final de la Carrera de Especialización en Inteligencia Artificial se titulará “Entrenamiento de robot caminante por imitación y RL usando ZMP como generador de datos de referencia” y consistirá en La simulación, de un robot caminante y el entrenamiento de las políticas de control de su marcha aplicando aprendizaje por refuerzo. El trabajo tendrá un presupuesto preliminar estimado de **600** horas y un costo estimado de **\$ XXX**, con fecha de inicio el 21 de octubre de 2025 y fecha de presentación pública el **15 de junio de 2026**.

Se adjunta a esta acta la planificación inicial.

Dr. Ing. Ariel Lutenberg
Director posgrado FIUBA

Nombre del cliente
Empresa del cliente

Título y Nombre del director
Director del Trabajo Final

1. Descripción técnica-conceptual del proyecto a realizar

El proyecto consiste en diseñar y entrenar en simulación, un sistema de desplazamiento para un robot caminante destinado a realizar tareas en obras de construcción. El objetivo técnico es obtener una política de control de la marcha capaz de andar de forma estable frente a perturbaciones, mientras que el objetivo de negocio es explorar tecnologías que, en una etapa posterior, permitan automatizar tareas físicamente exigentes y peligrosas de una empresa constructora.

Como consecuencia de terrenos irregulares, cambios diarios en el entorno, obstáculos imprevisibles, desniveles y riesgo de caídas, la automatización ha avanzado más lentamente en el sector de la construcción. Muchos procesos que requieren desplazamiento y manipulación en estos entornos son difíciles de automatizar con métodos tradicionales de control. Sin embargo, a mayor nivel de automatización de tareas repetitivas, mayor es el potencial de reducir costos para los clientes finales sin sacrificar calidad, además de disminuir la exposición de los operarios a condiciones peligrosas.

Este proyecto utiliza el ZMP (Zero Moment Point) como generador de datos de referencia. El ZMP es un criterio de estabilidad: si el punto donde la resultante de las fuerzas de reacción del suelo tiene un momento nulo (el ZMP) permanece dentro del polígono de apoyo, se garantiza la estabilidad en superficies planas. Proporciona una condición suficiente de estabilidad y permite el cálculo de trayectorias expertas de forma determinista, pero es inadecuado frente a perturbaciones externas y es energéticamente ineficiente.

El estado del arte combina aprendizaje profundo y aprendizaje por refuerzo (DRL) desde cero. Aquí el agente puede descubrir autónomamente estrategias de control a través de prueba y error, y adaptarse a escenarios imprevistos, pero es computacionalmente costoso.

Como se observa en la figura 2, no se utiliza ZMP como controlador final, sino como un generador de datos de referencia, o de demostraciones expertas para iniciar la política de RL. Esto puede reducir el costo de la exploración aleatoria inicial. El agente no empieza ciego, sino que ya sabe cómo caminar de forma estable, luego el RL se utiliza solo para robustecer el comportamiento ante posibles perturbaciones.

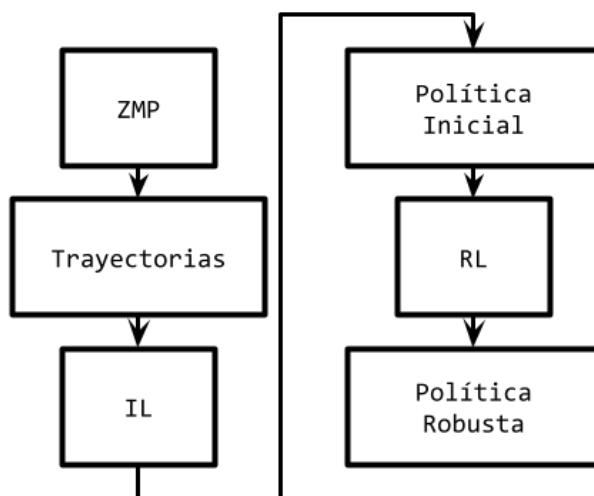


Figura 1. Diagrama en bloques del sistema.

2. Identificación y análisis de los interesados

Rol	Nombre y Apellido	Organización	Puesto
Cliente	Nombre del cliente	Empresa del cliente	-
Responsable	Ing. Francisco Antonio Cofré Villalón	FIUBA	Alumno
Colaboradores	-	-	-
Orientador	Título y Nombre del director	pertenencia	Director del Trabajo Final
Equipo	miembro1 miembro2	-	-
Opositores	-	-	-
Usuario final	-	-	-

3. Propósito del proyecto

El problema que abordará el proyecto es la automatización en entornos difíciles de predecir. A diferencia de una planta industrial, un sitio de construcción es dinámico, caracterizado por terreno irregular, escombros, obstáculos imprevistos y la necesidad de navegar en múltiples niveles. Este dominio de problema justifica el enfoque del proyecto en la locomoción bípeda. Soluciones robóticas más simples, como las plataformas con ruedas o los brazos robóticos estacionarios, son inadecuadas para la navegación y movilidad requeridas.

4. Alcance del proyecto

El proyecto incluye:

- El diseño de un modelo 3D del robot bípedo.
- La configuración de un entorno de simulación física que modele la dinámica, las colisiones y las fuerzas de reacción del suelo.
- El algoritmo de generación de trayectorias basado en el Zero Moment Point (ZMP) y el registro de las demostraciones (estados, acciones) generadas en la simulación.
- Código necesario para el entrenamiento de la política.
- Pruebas para validar la política de locomoción final.

El proyecto no incluye:

- El diseño, fabricación o costo de componentes.
- La simulación del torso, brazos y manos.
- La transferencia de la política a un robot físico.

5. Supuestos del proyecto

Para el desarrollo del presente proyecto se supone que:

- Se dispondrá de suficientes horas semanales para cumplir el cronograma estimado del proyecto y completar las aproximadamente 600 horas totales previstas.
- NVIDIA Isaac Sim, Isaac Lab, PyTorch y los frameworks de RL utilizados permanecerán disponibles durante toda la duración del proyecto, sin cambios de licencia ni compatibilidad.
- Se asume que el simulador podrá representar de manera suficientemente realista la dinámica del robot, los contactos y las fuerzas del suelo como para producir datos útiles para IL y RL.
- Se asume que el diseño del robot y su modelo 3D podrán integrarse sin errores graves al entorno de simulación.
- La disponibilidad del tiempo de cómputo será suficiente para las múltiples iteraciones requeridas por el entrenamiento de las políticas de IL y RL en el entorno de simulación física.
- Se contará con suficiente memoria para ejecutar simulaciones físicas y entrenar modelos de IL y RL sin bloqueos críticos.

6. Product Backlog

Story Points

Para obtener los Story Points (SP) de cada historia de usuario se evalúan tres factores:

Dificultad (D): cantidad de trabajo estimado.

Complejidad (C): nivel técnico requerido.

Incertidumbre (I): grado de riesgo o novedad.

Cada dimensión se puntúa en una escala de uno a diez, se suman y luego se redondea hacia el número superior más próximo de la serie de Fibonacci:

1, 2, 3, 5, 8, 13, 21, 34...

■ Épica 1: Modelo 3D y entorno de simulación

- Como responsable de simulación, quiero un modelo 3D del robot bípedo para poder ajustar dimensiones, masas, calcular trayectorias y entrenar agentes.
- D: 3, C: 3, I: 3, SP: 13.
- Como responsable de simulación, quiero un entorno de simulación para calcular trayectorias y entrenar agentes.
- D: 3, C: 3, I: 3, SP: 13.

■ **Épica 2: Generación de trayectorias expertas con ZMP**

- Como responsable de control, quiero que el sistema genere trayectorias de marcha estables usando ZMP para poder utilizarlas como demostraciones expertas en el entrenamiento por imitación.
- D: 3, C: 3, I: 3, SP: 13.
- Como responsable de control, quiero poder ver indicados el ZMP, el centro de masa y los contactos con el suelo durante la marcha para confirmar que las trayectorias cumplen las condiciones de estabilidad.
- D: 3, C: 3, I: 3, SP: 13.

■ **Épica 3: Entrenamiento por imitación y RL de la marcha bípeda**

- Como responsable de ML, quiero entrenar una política inicial a partir de las demostraciones generadas con ZMP para obtener un caminante estable en pocas iteraciones.
- D: 3, C: 3, I: 3, SP: 13.
- Como responsable de ML, quiero mejorar la rapidez y estabilidad frente a perturbaciones y variaciones del terreno usando RL.
- D: 3, C: 3, I: 3, SP: 13.

■ **Épica 4: Experimentación y análisis de resultados**

- Como responsable de ML, quiero evaluar el progreso de la política para medir la mejora del agente durante el entrenamiento.
- D: 3, C: 3, I: 3, SP: 13.
- Como responsable de ML, quiero ver gráficos que resuman la estabilidad, rapidez y que comparan ZMP e IL+RL.
- D: 3, C: 3, I: 3, SP: 13.

7. Criterios de aceptación de historias de usuario

■ **Épica 1: Modelo 3D y entorno de simulación**

- URDF/USD con articulaciones (cadera, rodilla, tobillo) y masa superior. El modelo se exporta e integra sin errores en el entorno de simulación seleccionado.
- El mundo simula contacto, fricción y colisión. Se pueden salvar y reproducir series de acciones.

■ **Épica 2: Generación de trayectorias expertas con ZMP**

- Dataset de N pasos sin caídas exportado.
- Se muestra polígono de soporte y ZMP en tiempo real.

■ **Épica 3: Entrenamiento por imitación y RL de la marcha bípeda**

- Entrena política inicial. Logra caminar X metros sin caídas en entorno base. Secuencia es reproducible.
- Supera la política inicial en caídas, velocidad y energía/distancia.

■ **Épica 4: Experimentación y análisis de resultados**

- Pruebas de tasa de caídas, velocidad, trayectorias CoM/ZMP, energía/distancia.
Pruebas con perturbaciones, rugosidad, obstáculos.
- Metodología, registros, videos, gráficos.

8. Fases de CRISP-DM

1. **Comprensión del negocio:** Movilidad de robots en obras de construcción. Éxito:
Conseguir costo inferior a comprar el sistema en el mercado.
2. **Comprensión de los datos:** Simulador, trayectorias ZMP. Calidad:
3. **Preparación de los datos:** características clave, transformaciones necesarias.
4. **Modelado:** tipo de problema, algoritmos posibles.
5. **Evaluación del modelo:** Distancia sin caída, velocidad y consumo de energía cercanos
al estado del arte, estabilidad ante empujes y rugosidad, comparación ZMP vs. IL+RL.

9. Desglose del trabajo en tareas

Historia de usuario	Tarea técnica	Estimación	Prioridad
HU1	Determinar y especificar componentes	6 h	Alta
HU1	Modelo 3D sensores actuadores vínculos (URDF/USD)	8 h	Alta
HU2	Configurar una escena base en simulador	5 h	Media
HU2	Cargar y validar robot en simulador	6 h	Alta
HU3	Calcular ZMP	6 h	Alta
HU3	Implementar el generador de trayectorias	8 h	Alta
HU4	Visualizar de ZMP, polígono y centro de masa	5 h	Media
HU4	Tarea 2 HU2	6 h	Alta
HU5	Entrenar la política por imitación con el dataset de demostraciones	6 h	Alta
HU5	Integrar la política entrenada en el simulador	8 h	Alta
HU6	Configurar y ejecutar el algoritmo de RL elegido	5 h	Media
HU6	irregularidades, cambios de velocidad objetivo	6 h	Alta
HU7	Registrar evolución de métricas	6 h	Alta
HU7	Automatizar registro y almacenamiento de métricas y logs	8 h	Alta
HU8	notebooks que generen gráficos	5 h	Media
HU8	Tarea 2 HU2	6 h	Alta

10. Planificación de Sprints

Cuadro 1. Formato sugerido

Sprint	HU o fase	Tarea	SP	Responsable	% Completado
Sprint 0	Planificación	Definir alcance y cronograma	10 h	Alumno	X %
Sprint 0	Planificación	Reunión con el tutor/cliente	5 h	Alumno	X %
Sprint 0	Planificación	Ajuste de los entregables	6 h	Alumno	X %
Sprint 1	HU1	Determinar y especificar componentes, parámetros físicos y sensores	6 h / 3 SP	Alumno	0 %
Sprint 1	HU1	Modelo 3D con sensores, actuadores y vínculos (URDF/USD)	10 h / 5 SP	Alumno	0 %
Sprint 2	HU2	Configurar una escena base en el simulador (mundo, gravedad, suelo)	7 h / 5 SP	Alumno	0 %
Sprint 2	HU2	Cargar y validar el robot bípedo en el simulador	10 h / 5 SP	Alumno	0 %
Sprint 3	HU3	Implementar cálculo del ZMP a partir de fuerzas de contacto	7 h / 5 SP	Alumno	0 %
Sprint 3	HU3	Implementar el generador de trayectorias de marcha basadas en ZMP	10 h / 5 SP	Alumno	0 %
Sprint 4	HU4	Visualizar ZMP, polígono de soporte y centro de masa en tiempo real	7 h / 5 SP	Alumno	0 %

Cuadro 2. Formato sugerido

Sprint	HU o fase	Tarea	SP	Responsable	% Completado
Sprint 4	HU4	Integrar visualización en la escena de simulación ya configurada	10 h / 5 SP	Alumno	0 %
Sprint 5	HU5	Entrenar la política inicial por imitación (behavior cloning)	7 h / 5 SP	Alumno	0 %
Sprint 5	HU5	Integrar la política entrenada en el simulador y validar marcha base	10 h / 5 SP	Alumno	0 %
Sprint 6	HU6	Configurar y ejecutar el algoritmo de RL elegido	7 h / 5 SP	Alumno	0 %
Sprint 6	HU6	Entrenar robustez a irregularidades y cambios de velocidad objetivo	10 h / 5 SP	Alumno	0 %
Sprint 7	HU7	Diseñar y registrar métricas (caídas, velocidad, energía/distancia)	7 h / 5 SP	Alumno	0 %
Sprint 7	HU7	Automatizar registro y almacenamiento de métricas y logs	10 h / 5 SP	Alumno	0 %
Sprint 8	HU8	Desarrollar notebooks para generación de gráficos y tablas	7 h / 5 SP	Alumno	0 %
Sprint 8	HU8	Integrar dashboards/notebooks en el flujo de experimentación	10 h / 5 SP	Alumno	0 %
Sprint 9	Escritura	Redacción memoria	50 h / 34 SP	Alumno	0 %
Sprint 10	Defensa	Preparación de la exposición	20 h / 13 SP	Alumno	0 %

11. Diagrama de Gantt (sprints)

Visualizar en un diagrama de Gantt la planificación temporal del proyecto, tomando como base los sprints definidos en la sección anterior. Debe contemplar todas las horas del proyecto.

Consigna:

- Elaborar un diagrama de Gantt que muestre la secuencia temporal de los sprints.
- Cada fila debe representar un sprint (con su número o nombre), y el eje horizontal debe indicar el tiempo (en semanas o fechas concretas).
- Las tareas técnicas derivadas de HU deben diferenciarse visualmente (por ejemplo, con un color distinto) de las tareas no técnicas (planificación, redacción, defensa).
- Incluir todas las tareas estimadas en cada sprint.

Recomendaciones para el Gantt:

- Podés usar herramientas gratuitas como TeamGantt, ClickUp, GanttProject, [Google Sheets], [Trello + Planyway], entre otras.
- Ordená los sprints de forma cronológica, comenzando con Sprint 0 (planificación) y finalizando con el sprint de defensa.
- Asegurate de reflejar la duración realista de cada sprint según tu disponibilidad y el cronograma general del posgrado.
- Incluí hitos importantes: reuniones, entregas parciales, defensa.

Incluir una imagen legible del diagrama de Gantt. Si es muy ancho, presentar primero la tabla y luego el gráfico de barras.

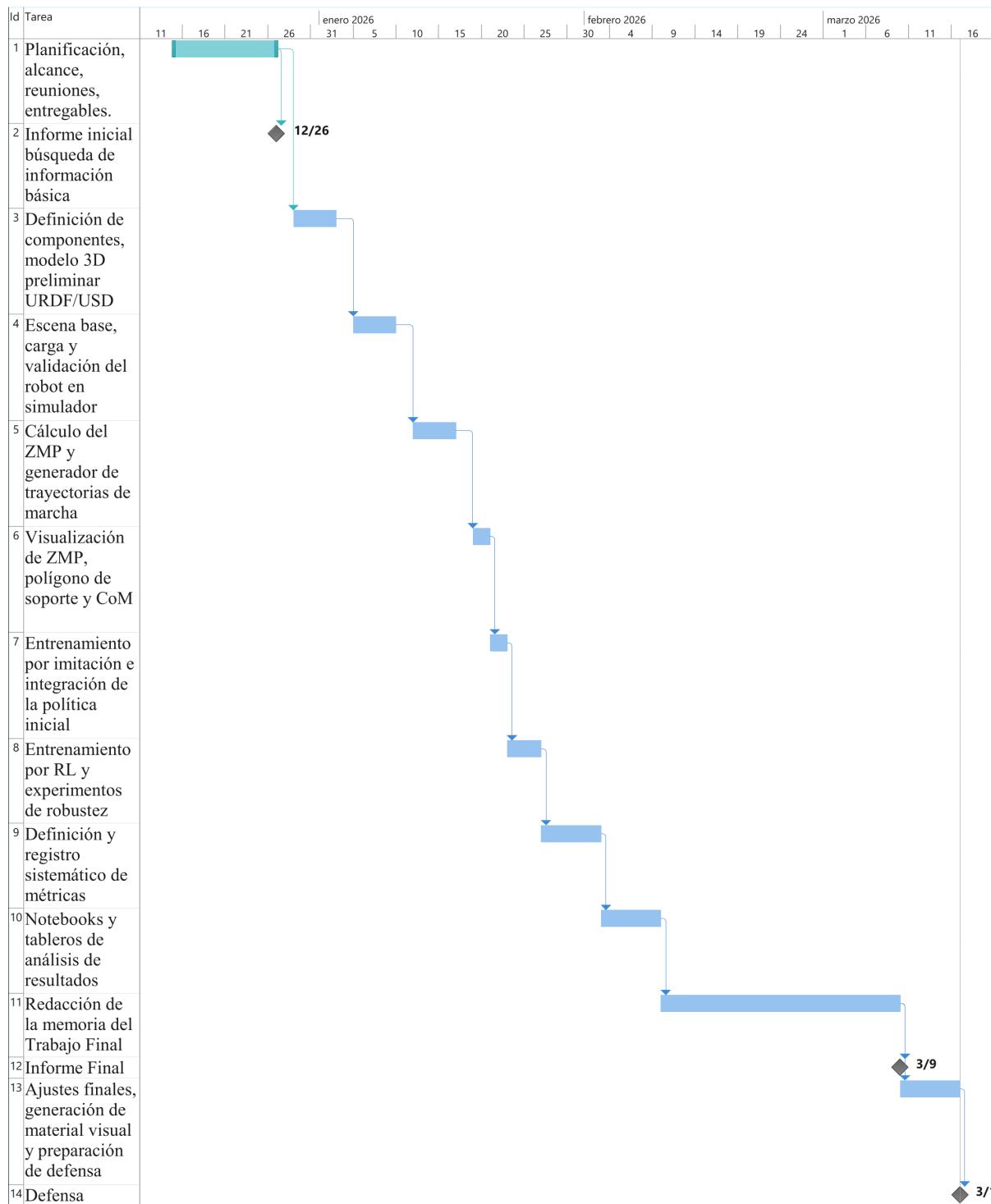


Figura 2. Diagrama en bloques del sistema.

12. Gobernanza de datos

No se capturan ni procesan datos personales de individuos ni información sensible; todo el comportamiento del agente se entrena sobre entornos sintéticos. El único material que se prevé publicar de forma abierta son el código fuente y las políticas entrenadas, junto con scripts de evaluación y documentación del experimento.

12.1. Cumplimiento normativo

El simulador y los entornos asociados (NVIDIA Isaac Sim e Isaac Lab) se utilizarán bajo sus licencias de desarrollador, exclusivamente con fines académicos y no productivos. No se redistribuirán escenas ni assets propietarios.

Las librerías de aprendizaje automático como PyTorch, frameworks de RL y arquitecturas tipo ResNet-18 se emplearán bajo sus licencias open source (BSD-3, MIT, Apache-2.0). En el repositorio del proyecto se incluirá un apartado con licencias y avisos correspondientes.

12.2. Ética en el uso de inteligencia artificial

El objetivo del proyecto es mejorar la autonomía de robots caminantes para tareas de construcción, un ámbito donde una caída o comportamiento errático podría implicar riesgos para personas y bienes. Este trabajo se desarrolla exclusivamente en simulación, no se recomienda el uso de las políticas resultantes en plataformas físicas sin medidas de seguridad. Desde una perspectiva social, el proyecto busca automatizar tareas físicamente demandantes que son realizadas en entornos riesgosos. Al disminuir la exposición directa de los operarios a estas condiciones peligrosas, se contribuye significativamente a mejorar su seguridad.

13. Gestión de riesgos

a) Identificación de los riesgos (al menos cinco) y estimación de sus consecuencias:

Riesgo 1: detallar el riesgo (riesgo es algo que si ocurre altera los planes previstos de forma negativa)

- Severidad (S): mientras más severo, más alto es el número (usar números del 1 al 10). Justificar el motivo por el cual se asigna determinado número de severidad (S).
- Probabilidad de ocurrencia (O): mientras más probable, más alto es el número (usar del 1 al 10). Justificar el motivo por el cual se asigna determinado número de (O).

Riesgo 2:

- Severidad (S): X.
Justificación...
- Ocurrencia (O): Y.
Justificación...

Riesgo 3:

- Severidad (S): X.
Justificación...
- Ocurrencia (O): Y.
Justificación...

b) Tabla de gestión de riesgos: (El RPN se calcula como $RPN=SxO$)

Riesgo	S	O	RPN	S*	O*	RPN*

Criterio adoptado:

Se tomarán medidas de mitigación en los riesgos cuyos números de RPN sean mayores a...

Nota: los valores marcados con (*) en la tabla corresponden luego de haber aplicado la mitigación.

c) Plan de mitigación de los riesgos que originalmente excedían el RPN máximo establecido:

Riesgo 1: plan de mitigación (si por el RPN fuera necesario elaborar un plan de mitigación). Nueva asignación de S y O, con su respectiva justificación:

- Severidad (S*): mientras más severo, más alto es el número (usar números del 1 al 10). Justificar el motivo por el cual se asigna determinado número de severidad (S).
- Probabilidad de ocurrencia (O*): mientras más probable, más alto es el número (usar del 1 al 10). Justificar el motivo por el cual se asigna determinado número de (O).

Riesgo 2: plan de mitigación (si por el RPN fuera necesario elaborar un plan de mitigación).

Riesgo 3: plan de mitigación (si por el RPN fuera necesario elaborar un plan de mitigación).

14. Sprint Review

La revisión de sprint (*Sprint Review*) es una práctica fundamental en metodologías ágiles. Consiste en revisar y evaluar lo que se ha completado al finalizar un sprint. En esta instancia, se presentan los avances y se verifica si las funcionalidades cumplen con los criterios de aceptación establecidos. También se identifican entregables parciales y se consideran ajustes si es necesario.

Aunque el proyecto aún se encuentre en etapa de planificación, esta sección permite proyectar cómo se evaluarán las funcionalidades más importantes del backlog. Esta mirada anticipada favorece la planificación enfocada en valor y permite reflexionar sobre posibles obstáculos.

Objetivo: anticipar cómo se evaluará el avance del proyecto a medida que se desarrolle las funcionalidades, utilizando como base al menos cuatro historias de usuario del *Product Backlog*.

Seleccionar al menos 4 HU del Product Backlog. Para cada una, completar la siguiente tabla de revisión proyectada:

Formato sugerido:

HU seleccionada	Tareas asociadas	Entregable esperado	¿Cómo sabrás que está cumplida?	Observaciones o riesgos
HU1	Tarea 1	Módulo funcional	Cumple criterios de aceptación definidos	Falta validar con el tutor
	Tarea 2			
HU3	Tarea 1	Reporte generado	Exportación disponible y clara	Requiere datos reales
	Tarea 2			
HU5	Tarea 1	Panel de gestión	Roles diferenciados operativos	Riesgo en integración
	Tarea 2			
HU7	Tarea 1	Informe trimestral	PDF con gráficos y evolución	Puede faltar tiempo para ajustes
	Tarea 2			

15. Sprint Retrospective

La retrospectiva de sprint es una práctica orientada a la mejora continua. Al finalizar un sprint, el equipo (o el alumno, si trabaja de forma individual) reflexiona sobre lo que funcionó bien, lo que puede mejorarse y qué acciones concretas pueden implementarse para trabajar mejor en el futuro.

Durante la cursada se propuso el uso de la **Estrella de la Retrospectiva**, que organiza la reflexión en torno a cinco ejes:

- ¿Qué hacer más?
- ¿Qué hacer menos?
- ¿Qué mantener?
- ¿Qué empezar a hacer?
- ¿Qué dejar de hacer?

Aun en una etapa temprana, esta herramienta permite que el alumno planifique su forma de trabajar, identifique anticipadamente posibles dificultades y diseñe estrategias de organización personal.

Objetivo: reflexionar sobre las condiciones iniciales del proyecto, identificando fortalezas, posibles dificultades y estrategias de mejora, incluso antes del inicio del desarrollo.

Completar la siguiente tabla tomando como referencia los cinco ejes de la Estrella de la Retrospectiva (*Starfish* o estrella de mar). Esta instancia te ayudará a definir buenas prácticas

desde el inicio y prepararte para enfrentar el trabajo de forma organizada y flexible. Se deberá completar la tabla al menos para 3 sprints técnicos y 1 no técnico.

Formato sugerido:

Sprint tipo y N°	¿Qué hacer más?	¿Qué hacer menos?	¿Qué mantener?	¿Qué empezar a hacer?	¿Qué dejar de hacer?
Sprint técnico - 1	Validaciones continuas con el alumno	Cambios sin versión registrada	Pruebas con datos simulados	Documentar cambios propuestos	Ajustes sin análisis de impacto
Sprint técnico - 2	Verificar configuraciones en múltiples escenarios	Modificar parámetros sin guardar historial	Perfiles reutilizables	Usar logs para configuración	Repetir pruebas manuales innecesarias
Sprint técnico - 8	Comparar correlaciones con casos previos	Cambiar parámetros sin justificar	Revisión cruzada de métricas	Anotar configuraciones usadas	Trabajar sin respaldo de datos
Sprint no técnico - 12 (por ej.: “Defensa”)	Ensayos orales con feedback	Cambiar contenidos en la memoria	Material visual claro	Dividir la presentación por bloques	Agregar gráficos difíciles de explicar