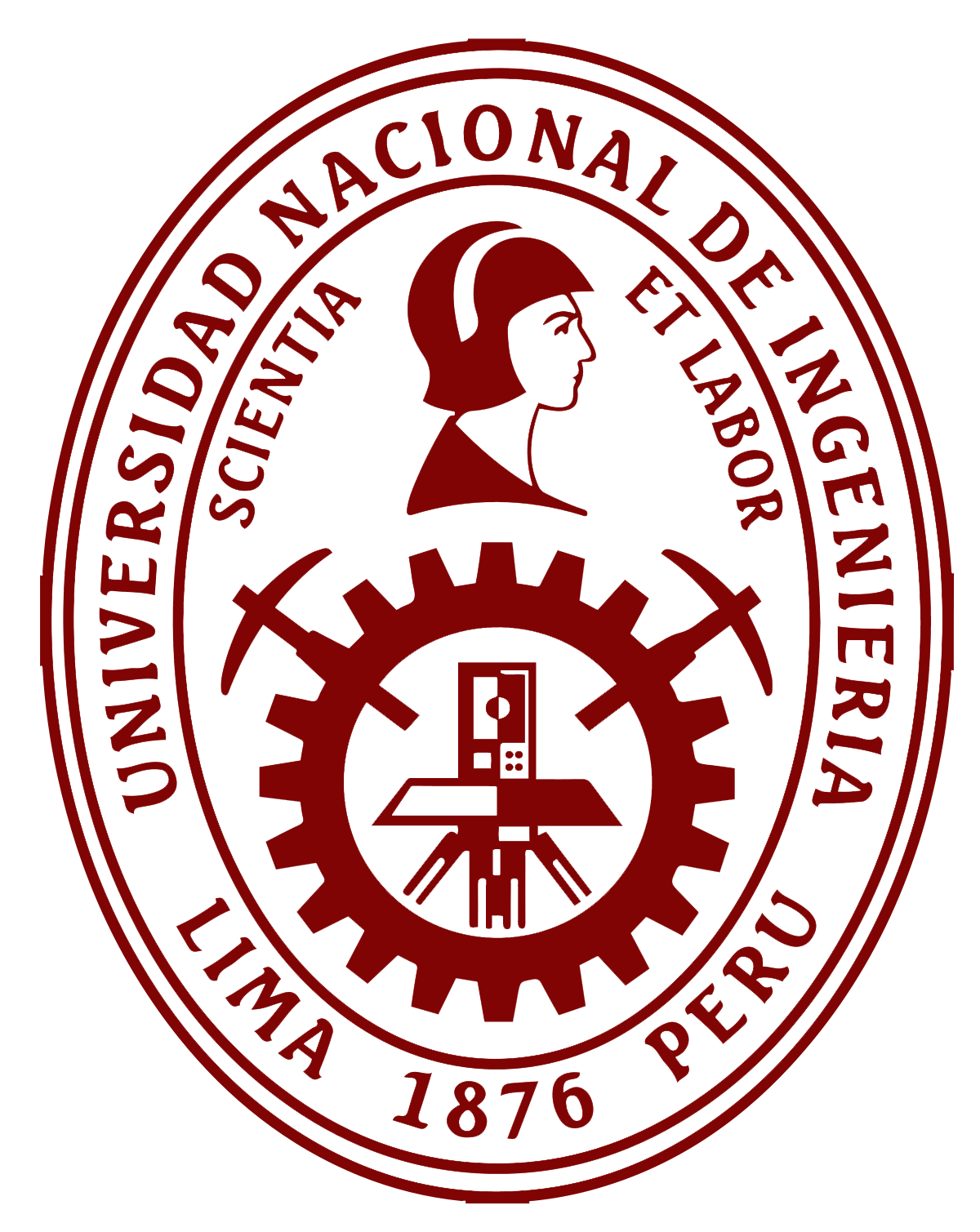
**UNIVERSIDAD NACIONAL DE INGENIERÍA**

**UNIDAD DE POSTGRADO FACULTAD DE INGENIERIA INDUSTRIAL Y DE SISTEMAS**



**PLAN DE TESIS**

Desarrollo e Implementación de un modelo de Reinforcement Learning (RL) para la optimización de dietas en acuicultura

PARA OBTENER EL GRADO ACADEMICO DE MAESTRO EN CIENCIAS CON MENCIÓN EN INTELIGENCIA ARTIFICIAL

ELABORADA POR:

BRUNO ALEJANDRO DONAYRE DONAYRE

ASESOR:

LIMA, PERU

2025

DATOS GENERALES

1) TÍTULO DEL PLAN DE TESIS:  
DESARROLLO E IMPLEMENTACIÓN DE UN MODELO DE REINFORCEMENT LEARNING (RL) PARA LA OPTIMIZACIÓN DE DIETAS EN ACUICULTURA  
  
2) NOMBRE DEL AUTOR:  
Bruno Alejandro Donayre Donayre  
  
3) NOMBRE DEL ASESOR:  
  
  
4) ÁREA INVOLUCRADA:  
Unidad de Posgrado FIIS  
  
5) LUGAR O INSTITUCIÓN DONDE SE DESARROLLA EL PROYECTO:  
Vitapro S.A., CEAs (Centros Experimentales Acuícolas) en Perú  
  
6) DURACIÓN ESTIMADA DEL PROYECTO EN MESES:  
Agosto 2025 – Julio 2026 (12 meses)

Contenido

[Capítulo 1: Planteamiento del Problema 5](#_Toc204072119)

[1.1 Diagnóstico 5](#_Toc204072120)

[1.2 Identificación y Diagnóstico del Problema de Estudio 6](#_Toc204072121)

[1.2.1 Antecedentes bibliográficos 6](#_Toc204072122)

[1.2.2 Formulación del Problema 8](#_Toc204072123)

[1.2.2.1 Formulación del Problema General 8](#_Toc204072124)

[1.2.2.2 Formulación de los Problemas Específicos 8](#_Toc204072125)

[1.2.3 Justificación y Alcances 9](#_Toc204072126)

[1.2.3.1 Justificación 9](#_Toc204072127)

[1.2.3.2 Alcances 10](#_Toc204072128)

[Capítulo 2: Objetivos 11](#_Toc204072129)

[2.1 Objetivo General 11](#_Toc204072130)

[2.2 Objetivos Específicos 11](#_Toc204072131)

[Capítulo 3: Marco Teórico y Estado del Arte 12](#_Toc204072132)

[3.1 Bases Teóricas 12](#_Toc204072133)

[3.1.1. Características biológicas del camarón blanco *(Litopeneaeus vannamei)* 12](#_Toc204072134)

[3.1.2. Fundamentos de Aprendizaje por Reforzamiento 12](#_Toc204072135)

[3.1.2.1. Procesos de Decisión de Markov (MDP) 12](#_Toc204072136)

[3.1.2.2. DDPG (Deep Deterministic Policy Gradient) 13](#_Toc204072137)

[3.1.2.3. Q-learning 13](#_Toc204072138)

[3.1.2.4. PPO (Proximal Policy Optimization) 13](#_Toc204072139)

[3.2. Definición de términos 15](#_Toc204072140)

[3.3. Estado del Arte 17](#_Toc204072141)

[3.4. Bases Teóricas 19](#_Toc204072142)

[3.4.1. Características biológicas e importancia económica del camarón blanco del Pacífico (Litopeneaeus vannamei) 19](#_Toc204072143)

[3.4.2. Aplicaciones del Reinforcement Learning (RL) en la acuicultura 19](#_Toc204072144)

[3.4.3. Aplicaciones del Reinforcement Learning (RL) en el cultivo de *Litopeneaeus vannamei* 20](#_Toc204072145)

[Capítulo 4: Metodología de Investigación 21](#_Toc204072146)

[4.1 Enfoque Metodológico 21](#_Toc204072147)

[4.2 Diseño Experimental 23](#_Toc204072148)

[4.3 Interacción con Stakeholders 24](#_Toc204072149)

[Capítulo 5: Administración del plan de tesis 25](#_Toc204072150)

[5.1 Cronograma 25](#_Toc204072151)

[5.2 Presupuesto 26](#_Toc204072152)

[5.3 Financiamiento 26](#_Toc204072153)

[Capítulo 6: Referencias Bibliográficas 27](#_Toc204072154)

[Capítulo 7: Anexos 28](#_Toc204072155)

[7.1. Esquemas de trabajo 28](#_Toc204072156)

[7.1.1. Esquema híbrido de desarrollo metodológico: CRISP-DM + Agile+AI 28](#_Toc204072157)

[7.1.1. Pipeline de Diseño experimental 28](#_Toc204072158)

[7.2. Matriz de consistencia 30](#_Toc204072159)

# Capítulo 1: Planteamiento del Problema

## 1.1 Diagnóstico

La acuicultura intensiva, especialmente en el cultivo de *Litopeneaeus vannamei*, enfrenta limitaciones significativas en la formulación y gestión eficiente de dietas. Los esquemas tradicionales de alimentación se apoyan en tablas estáticas o protocolos generalizados que no consideran la variabilidad ambiental ni la heterogeneidad entre los organismos cultivados (FAO, 2022). Estos enfoques presuponen condiciones estables de temperatura, oxígeno disuelto, salinidad y pH, cuando en la práctica los sistemas acuícolas presentan fluctuaciones continuas y complejas.

Esta dinámica ambiental afecta directamente el metabolismo, la tasa de crecimiento y el comportamiento alimenticio de los organismos. El suministro excesivo de alimento, en contextos de estrés térmico o hipoxia puede resultar en desperdicio, deterioro de la calidad del agua y aumento en la mortalidad (Aljehani et al., 2023). Por el contrario, una oferta subóptima en momentos de alta demanda energética limita el crecimiento y reduce la productividad del cultivo.

Además, los esquemas actuales tienden a ignorar la variabilidad individual y grupal en el crecimiento, aplicando estrategias uniformes que no se ajustan a las necesidades reales de los organismos (Chahid et al., 2022). Esta homogeneización conduce a decisiones subóptimas, menor eficiencia alimenticia y mayores costos operativos.

A pesar de que los centros de cultivo han empezado a recolectar datos ambientales y productivos con mayor frecuencia, la toma de decisiones sigue siendo en gran medida empírica y manual, sin aprovechar plenamente el potencial de esos datos. En este contexto, el Reinforcement Learning (RL) (Aprendizaje por Reforzamiento, AR) emerge como una técnica de inteligencia artificial prometedora. A diferencia de los métodos supervisados, el RL permite que un agente aprenda directamente de la interacción con el entorno, ajustando sus decisiones en función de recompensas acumuladas (Sutton & Barto, 2018).

Estudios recientes han demostrado que algoritmos como Q-learning, DDPG y PPO pueden aplicarse con éxito en sistemas acuícolas para optimizar tasas de alimentación y controlar parámetros ambientales (Elmessery et al., 2025; Aljehani et al., 2023). No obstante, su aplicación en el cultivo de *L. vannamei* aún es limitada, y se requieren adaptaciones metodológicas para lidiar con datos ruidosos, intermitentes y altamente variables.

En este marco, es urgente desarrollar modelos adaptativos y validados experimentalmente que permitan integrar datos históricos y en tiempo real para generar recomendaciones óptimas, personalizadas y sostenibles. La implementación de estos enfoques puede mejorar significativamente la eficiencia alimentaria, reducir la carga operativa y contribuir a una acuicultura más resiliente y basada en datos.

## 1.2 Identificación y Diagnóstico del Problema de Estudio

### 1.2.1 Antecedentes bibliográficos

La formulación de dietas en acuicultura enfrenta grandes retos debido a la alta variabilidad de los sistemas de cultivo y la poca adaptabilidad de los métodos tradicionales. Estos se basan principalmente en tablas estáticas o protocolos predefinidos que no consideran la dinámica ambiental ni la heterogeneidad entre los individuos cultivados (FAO, 2022). Las condiciones ambientales, como temperatura, oxígeno disuelto, salinidad o pH, influyen directamente en el metabolismo y consumo de alimento de especies como Litopeneaeus vannamei, haciendo ineficientes las estrategias de alimentación uniformes (Aljehani et al., 2023; Chahid et al., 2022).

El Reinforcement Learning (RL), una técnica de inteligencia artificial donde un agente aprende a partir de la interacción con el entorno mediante recompensas, ha sido identificado como una solución prometedora para enfrentar este desafío (Sutton & Barto, 2018). Modelos como Q-learning, DDPG y PPO han sido aplicados con éxito en peces como tilapia y trucha, optimizando tasas de alimentación y controlando condiciones del sistema (Chahid et al., 2022; Elmessery et al., 2025). Sin embargo, su aplicación en el cultivo de L. vannamei aún es incipiente.

Estudios como los de Elmessery et al. (2025) han demostrado que algoritmos como DDPG pueden optimizar simultáneamente tasas de alimentación y parámetros de calidad de agua en sistemas RAS. Otros trabajos, como el de Aljehani et al. (2023), revisan estrategias RL combinadas con control predictivo (MPC) para mejorar el rendimiento productivo y reducir impactos ambientales.

Además, el uso de modelos predictivos auxiliares (por ejemplo, redes neuronales para estimar crecimiento o consumo) puede mejorar la precisión del entorno simulado en el que se entrena el agente RL, acelerando el aprendizaje y mejorando la generalización (Chen et al., 2022). Esta combinación de técnicas permite desarrollar sistemas de alimentación inteligentes, adaptativos y escalables.

Aunque la mayoría de implementaciones previas se han centrado en otras especies o entornos, los resultados sugieren que el Reinforcement Learning (RL) es una tecnología con alto potencial para transformar la formulación de dietas en acuicultura, especialmente en cultivos de Litopeneaeus vannamei.

Tabla . Resumen de antecedentes bibliográficos

| **Autor (año)** | **Título del estudio** | **Metodología** | **Hallazgos clave** | **Relevancia** |
| --- | --- | --- | --- | --- |
| **Elmessery et al. (2025)** | *A deep deterministic policy gradient approach for optimizing feeding rates and water quality management in recirculating aquaculture systems* | DDPG (Reinforcement Learning (RL) continuo), simulación en RAS | Mejoras en eficiencia alimentaria y estabilidad de parámetros de calidad del agua en simulaciones RAS. | Inspira el uso de DDPG para decisiones dinámicas en sistemas acuícolas. |
| **Chahid et al. (2022)** | *Fish growth trajectory tracking using Q-learning in precision aquaculture* | Q-learning, datos de crecimiento en tilapia | Seguimiento efectivo de trayectorias de crecimiento y control de temperatura y alimentación. | Valida el uso de Q-learning para ambientes discretos en cultivo de peces. |
| **Aljehani et al. (2023)** | *Feeding control and water quality monitoring in aquaculture systems: Opportunities and challenges* | Revisión de literatura, RL + MPC | Se destacan las oportunidades del RL para alimentar, gestionar sensores, y controlar calidad del agua. | Ofrece marco conceptual para RL en acuicultura, útil para justificar el enfoque. |
| **Georgopoulou et al. (2024)** | *Real-time monitoring of feeding behavior in European sea bass using AI models and computer vision* | Visión por computador, modelos de IA | Detecta patrones de comportamiento alimentario en tiempo real. | Aporta técnicas de monitoreo que podrían integrarse con RL para decisiones. |
| **Chen et al. (2022)** | *Técnica de alimentación inteligente basada en la predicción del crecimiento del camarón* | Modelos predictivos y sensores RAS | Propone modelos de predicción de crecimiento para ajustar la alimentación. | Complementa con técnicas de predicción útiles para enriquecer las recompensas. |
| **Sutton & Barto (2018)** | *Reinforcement Learning: An Introduction* | Libro teórico | Base conceptual sobre métodos de Reinforcement Learning (RL). | Fundamento teórico del marco RL utilizado en la tesis. |
| **Mnih et al. (2015)** | *Human-level control through deep reinforcement learning* | DQN, juegos Atari | Introduce el DQN y demuestra su eficacia en entornos con retroalimentación retardada. | Soporte técnico para usar DQN como baseline en esta investigación. |

### 1.2.2 Formulación del Problema

La acuicultura enfrenta retos crecientes en la formulación de dietas, dado que los métodos tradicionales no se adaptan con suficiente rapidez a variaciones en las condiciones ambientales y en el crecimiento de los animales. Estos métodos suelen basarse en tablas de alimentación o protocolos preestablecidos que asumen condiciones ambientales estables y tasas de crecimiento promedio (FAO, 2022). Sin embargo, en la práctica, los sistemas acuícolas están sujetos a fluctuaciones constantes: cambios en la temperatura del agua, niveles de oxígeno disuelto, salinidad, pH, y la presencia de contaminantes o enfermedades pueden alterar significativamente las necesidades nutricionales de los organismos.

Estas variaciones ambientales impactan directamente en el metabolismo, el comportamiento alimentario y la salud de especies cultivadas como *Litopeneaeus vannamei*, lo que genera desequilibrios entre la oferta alimenticia y la demanda real (Aljehani et al., 2023). Por ejemplo, un exceso de alimento en condiciones de estrés térmico o hipoxia puede no solo desperdiciar recursos, sino también deteriorar la calidad del agua y aumentar los riesgos de mortalidad. Por otro lado, una oferta insuficiente en momentos de alta demanda puede limitar el crecimiento y la productividad.

Además, las variaciones en el crecimiento individual y grupal son inherentes a cualquier cultivo acuícola. Los modelos tradicionales no consideran esta heterogeneidad y aplican estrategias uniformes que no responden a las necesidades específicas de cada grupo o individuo (Chahid et al., 2022). Esto se traduce en decisiones subóptimas, desperdicio de recursos y costos elevados para los productores.

Por estas razones, se hace evidente la necesidad de un enfoque dinámico y adaptativo que permita ajustar las estrategias de alimentación en tiempo real, optimizando así el rendimiento productivo, reduciendo el impacto ambiental y mejorando la sostenibilidad del cultivo (Sutton y Barto, 2018). El Reinforcement Learning (RL) y modelos predictivos avanzados ofrecen soluciones prometedoras, al aprender directamente de la interacción con el sistema y adaptarse a las condiciones cambiantes sin depender de reglas fijas (Elmessery et al., 2025). La capacidad de estos enfoques para adaptarse a la dinámica ambiental y fisiológica representa un avance clave hacia una acuicultura más eficiente y resiliente.

#### 1.2.2.1 Formulación del Problema General

¿Cómo aplicar un modelo de Reinforcement Learning (RL) profundo para optimizar la rentabilidad de decisiones productivas en CEAs usando datos históricos de 2020 a 2025?

#### 1.2.2.2 Formulación de los Problemas Específicos

* ¿Qué variables explican mejor la eficiencia productiva en los CEAs?
* ¿Qué arquitectura de DRL (DQN, DDPG, PPO) es más efectiva en este entorno?
* ¿Qué políticas de exploración permiten mejores decisiones en escenarios con datos limitados y alta variabilidad?
* ¿Cómo validar experimentalmente las recomendaciones del agente DRL?

### 1.2.3 Justificación y Alcances

#### 1.2.3.1 Justificación

El presente proyecto propone una solución innovadora para la formulación de dietas en acuicultura, integrando técnicas de Reinforcement Learning (RL) y modelos predictivos que permiten decisiones personalizadas y adaptativas basadas en datos históricos, condiciones ambientales en tiempo real y el crecimiento de los organismos (Deldjoo et al., 2023). Esta estrategia tiene el potencial de reducir significativamente los costos operativos mediante la optimización del uso de insumos alimenticios, al mismo tiempo que minimiza el desperdicio y sus efectos negativos sobre la calidad del agua y el medio ambiente.

Asimismo, al incorporar algoritmos que se ajustan dinámicamente a las fluctuaciones de las condiciones del cultivo, se promueve una mayor resiliencia frente a escenarios cambiantes, lo cual es esencial para garantizar la sostenibilidad a largo plazo del sector acuícola. La propuesta se concibe como una herramienta escalable y adaptable, capaz de ser implementada en diferentes tipos de sistemas acuícolas y especies, incluyendo el cultivo intensivo de Litopeneaeus vannamei, contribuyendo así a la mejora de la rentabilidad, sostenibilidad y competitividad del sector.

#### 1.2.3.2 Alcances

El presente proyecto se centra en el desarrollo y validación de un modelo de Reinforcement Learning (RL) para optimizar la formulación de dietas en el cultivo intensivo de *Litopeneaeus vannamei*, permitiendo decisiones adaptativas basadas en datos históricos y actuales. Los alcances del proyecto se definen en los siguientes términos:

* **Análisis de datos históricos y actuales**: Se utilizarán registros continuos de parámetros ambientales (temperatura, oxígeno disuelto, salinidad, pH) y registros intermitentes (cada 7 días) de peso real y consumo de alimento, como insumo para construir modelos predictivos auxiliares y caracterizar patrones de crecimiento y consumo (FAO, 2022; Elmessery et al., 2025).
* **Desarrollo y comparación de modelos de RL**: Se diseñarán, implementarán y compararán algoritmos como **Q-learning**, **DDPG** y **PPO**, seleccionando el enfoque más adecuado en función de su capacidad de adaptarse a las condiciones reales del cultivo, considerando la intermitencia de los registros y la variabilidad de los datos (Mnih et al., 2015; Chahid et al., 2022; Aljehani et al., 2023).
* **Simulación y validación preliminar**: La validación del sistema se realizará a través de simulaciones basadas en datos reales, con la participación de expertos en acuicultura para evaluar el impacto del modelo propuesto en la eficiencia alimenticia, el crecimiento, la sostenibilidad y la adaptabilidad frente a cambios ambientales y de crecimiento (Deldjoo et al., 2023).
* **Limitaciones del alcance**: Este proyecto se desarrollará en un entorno de simulación utilizando datos históricos y no contempla la implementación piloto a gran escala en campo real. Sin embargo, se proporcionarán recomendaciones y lineamientos para futuras aplicaciones prácticas y validaciones en entornos comerciales (Georgopoulou et al., 2024).

El proyecto representa un avance hacia la adopción de soluciones inteligentes y adaptativas en la acuicultura, estableciendo las bases para su futura escalabilidad y aplicación a otras especies y sistemas de cultivo (Elmessery et al., 2025).

# Capítulo 2: Objetivos

## 2.1 Objetivo General

Desarrollar un modelo de Reinforcement Learning (RL) que optimice la formulación de dietas en acuicultura, ajustando recomendaciones de alimentación de forma dinámica según condiciones ambientales y productivas.

## 2.2 Objetivos Específicos

* Analizar datos históricos y actuales para construir modelos predictivos auxiliares de crecimiento y consumo.
* Desarrollar un modelo de RL basado en Q-learning, DDPG o PPO, eligiendo el más adecuado en función de los resultados experimentales y la adaptabilidad a los datos reales del cultivo
* Validar el sistema con expertos y métricas productivas, evaluando su impacto en la eficiencia alimenticia y sostenibilidad.

# Capítulo 3: Marco Teórico y Estado del Arte

## 3.1 Bases Teóricas

### 3.1.1. Características biológicas del camarón blanco *(Litopeneaeus vannamei)*

Camarón blanco del Pacífico (*Litopeneaeus vannamei*) es una de las especies de crustáceos más cultivadas a nivel mundial debido a su rápido crecimiento, alta tasa de supervivencia y adaptabilidad a diversas condiciones ambientales. Originario de las costas del Pacífico oriental, desde México hasta Perú, su cultivo se ha expandido significativamente en América Latina y Ásia (FAO, 2007).

Esta especie se caracteriza por su capacidad de adaptación a diferentes salinidades, permitiendo su cultivo tanto en aguas marinas como en sistemas de agua dulce. Además, presenta una madurez sexual temprana y una alta fecundidad, lo que facilita su reproducción en cautiverio (Cobo & Pérez, 2018).

Los camarones peneidos pertenecen a la familia Penaeidae, la cual incluye numerosas especies. Entre estas, destaca *Litopeneaus vannamei*, conocido como el camarón blanco del Pacífico (Pérez-Farfante & Kensley, 1997). Este crustáceo desempeña un papel crucial en la industria pesquera y acuícola, debido a su gran valor comercial (Han et al., 2021). Su sabor suave, versatilidad culinaria y rápida adaptación a diferentes condiciones lo han convertido en una de las especies más demandadas en el mercado global de mariscos (Dayal et al., 2013).

Su capacidad y elevada tasa de reproducción rápida, su adaptación a diversas condiciones en el cultivo y su resistencia a enfermedades lo han convertido en el camarón blanco del Pacífico en una especie de especial interés para la acuicultura, sobre todo en regiones tropicales y subtropicales del mundo (FAO, 2022).

### 3.1.2. Fundamentos de Aprendizaje por Reforzamiento

El Reinforcement Learning (RL) (RL, por sus siglas en inglés) es un paradigma del aprendizaje automático en el que un agente aprende a tomar decisiones secuenciales mediante la interacción con un entorno, con el objetivo de maximizar una señal de recompensa acumulativa (Sutton & Barto, 2018). A continuación, se presentan los fundamentos más relevantes:

#### 3.1.2.1. Procesos de Decisión de Markov (MDP)

Los procesos de decisión de Markov son la base formal para definir problemas de RL. Un MDP se describe mediante una tupla (S,A,P,R,γ), donde:

* S: espacio de estados;
* A: espacio de acciones;
* P (*s′|s, a*): función de transición de estado;
* R (*s, a*): función de recompensa;
* γ: factor de descuento.

Estos modelos asumen que las transiciones dependen solo del estado y acción actuales, cumpliendo la **propiedad de Markov** (Sutton & Barto, 2018).

#### 3.1.2.2. DDPG (Deep Deterministic Policy Gradient)

DDPG es un algoritmo basado en actores-críticos, adecuado para espacios de acción **continuos**. Utiliza dos redes neuronales:

* Una red **actor**, que aprende una política determinista μ(s)
* Una red **crítica**, que estima la función Q (*s, a*) (Lillicrap et al., 2016).

Este enfoque permite generar recomendaciones de acciones con alta resolución, como gramajes de alimento personalizados por estado del cultivo.

#### 3.1.2.3. Q-learning

Q-learning es un algoritmo libre de modelo (model-free) que busca aprender una función de acción-valor 𝑄(𝑠,𝑎) que representa la utilidad esperada de realizar la acción a en el estado 𝑠 y seguir una política óptima. Su regla de actualización es:

**Q(s,a)←Q(s,a)+α[r+γmax​Q(s′,a′)−Q(s,a)]**

(Sutton & Barto, 2018). Es adecuado para entornos discretos, como decisiones basadas en rangos de peso o categorías de alimentación.

#### 3.1.2.4. PPO (Proximal Policy Optimization)

PPO es un algoritmo robusto basado en políticas, que busca actualizar de forma estable usando técnicas de **clipping** para evitar desviaciones grandes entre políticas sucesivas. La optimización se basa en la función:

**LCLIP(θ)=Et​[min(rt​(θ)At​, clip(rt​(θ),1−ϵ,1+ϵ) At​)]**

(Schulman et al., 2017). Se ha usado exitosamente en entornos complejos, como robótica y simulaciones biológicas.

* + 1. **Metodologías de desarrollo como Agile+AI, CRISP-DM y MLOps**

El desarrollo de sistemas inteligentes en acuicultura, especialmente aquellos basados en Reinforcement Learning (RL), requiere el uso de metodologías que guíen el ciclo de vida del proyecto desde la concepción hasta la evaluación. A continuación, se describen tres enfoques relevantes: CRISP-DM, Agile+AI y MLOps (Tabla 2).

Tabla . Descripción de metodologías aplicadas en el desarrollo de proyectos de inteligencia artificial

|  |  |
| --- | --- |
| Metodología | Descripción general |
| CRISP-DM | Está compuesta por seis fases: comprensión del negocio, comprensión de datos, preparación de los datos, modelado, evaluación y despliegue (Chapman et al. 2000) |
| Agile+AI | Adaptación de principios ágiles al desarrollo de soluciones con inteligencia artificial. En lugar de fases secuenciasles, se privilegia la entrega iterative y continua de valor, que se alinea con la naturaleza ecperimental de modelos RL (Amershi et al. 2019). |
| MLOps | Conjunto de prácticas orientadas a la automatización del ciclo complete de vida de modelos de ML. Incluye cinco pasos: control de versions, integración y entrega continua, monitoreo en producción y reentrenamiento automático (Kreuzberger et al., 2022). |

* + 1. **Integración con modelos predictivos auxiliares.**

La integración de modelos predictivos auxiliares dentro del entorno simulado de aprendizaje por refuerzo (RL) constituye una estrategia clave para mejorar la calidad del entrenamiento del agente. Estos modelos permiten estimar variables críticas como el crecimiento real del camarón, el consumo de alimento y la eficiencia alimenticia (FCA), a partir de datos históricos y parámetros ambientales.

Los modelos predictivos funcionan como componentes del entorno, aportando respuestas simuladas cuando el agente realiza una acción. Esto resulta particularmente útil en entornos donde los datos son intermitentes o costosos de recolectar en tiempo real, como es el caso de los sistemas acuícolas. Al utilizar predicciones realistas, el entorno simulado adquiere mayor fidelidad, lo que mejora la calidad del aprendizaje del agente y acelera su convergencia.

Entre las técnicas más empleadas para estos modelos auxiliares se encuentran las redes neuronales artificiales (ANN), XGBoost y regresión múltiple. Estas herramientas permiten predecir, por ejemplo, el peso del camarón al día N+7, en función del alimento suministrado, condiciones ambientales y peso inicial. A su vez, estas predicciones pueden alimentar la función de recompensa, vinculando las decisiones del agente con resultados proyectados más cercanos al comportamiento productivo real.

De esta manera, la combinación de RL con modelos supervisados proporciona un marco híbrido robusto, en el cual el agente aprende no solo de recompensas acumuladas, sino también de las proyecciones a corto y mediano plazo, mejorando así la anticipación y la calidad de sus decisiones.

## Definición de términos

| **Término** | **Definición** |
| --- | --- |
| RL (Reinforcement Learning) | Rama del aprendizaje automático en la que un agente aprende mediante recompensas a tomar decisiones en un entorno secuencial. |
| Q-Learning | Algoritmo de RL basado en una tabla que evalúa el valor de una acción dada en un estado específico sin requerir modelo del entorno. |
| DDPG (Deep Deterministic Policy Gradient) | Algoritmo de RL continuo basado en actores-críticos, útil para espacios de acción continuos como gramajes de alimento. |
| PPO (Proximal Policy Optimization) | Algoritmo de RL basado en políticas que estabiliza el entrenamiento mediante técnicas de recorte (clipping). Ideal para entornos complejos. |
| GAN (Generative Adversarial Network) | Arquitectura de redes neuronales donde un generador crea datos sintéticos y un discriminador evalúa su realismo, permitiendo simular datos no observados. |
| Boruta | Algoritmo de selección de variables basado en árboles aleatorios, que identifica las características más relevantes para el modelo. |
| FCA (Factor de Conversión Alimenticia) | Relación entre el alimento suministrado y el incremento de biomasa. Se usa como indicador de eficiencia alimenticia. |
| ROI (Return on Investment) | Indicador económico que mide el retorno generado por la inversión en el sistema productivo. |
| MAE (Mean Absolute Error) | Promedio de las diferencias absolutas entre valores reales y predichos. Usado para validar modelos de regresión. |
| MSE (Mean Squared Error) | Media del cuadrado de los errores. Penaliza fuertemente grandes desviaciones. |
| Hold-out | Técnica de validación donde se reserva una porción fija del dataset como conjunto de prueba. |
| K-Fold Cross Validation | Método de validación donde se divide el dataset en K partes y se entrena/valida K veces rotando el conjunto de prueba. |
| Bootstrapping | Método de validación basado en la generación de múltiples subconjuntos con reemplazo para estimar la variabilidad de una métrica. |
| A/B Testing | Técnica experimental que compara dos versiones de un modelo o política bajo condiciones similares para identificar cuál rinde mejor. |
| Replay Buffer | Mecanismo en RL que almacena experiencias pasadas para reutilizarlas durante el entrenamiento y estabilizar el aprendizaje. |
| Exploration vs. Exploitation | Dilema fundamental en RL que balancea entre explorar nuevas acciones o explotar el conocimiento actual para maximizar la recompensa. |
| Entorno simulado | Representación computacional del sistema de cultivo que permite entrenar el agente RL sin riesgo para la producción real. |
| Métricas UX | Indicadores cualitativos que miden la experiencia del usuario al interactuar con el sistema, tales como usabilidad, satisfacción, interpretabilidad y confianza. |

## Estado del Arte

La revisión de la literatura evidencia un crecimiento sostenido en la aplicación de inteligencia artificial en acuicultura, destacando el uso de modelos de aprendizaje por refuerzo (RL) y técnicas auxiliares como visión por computador, optimización bayesiana y aprendizaje supervisado. Trabajos como los de **Elmessery et al. (2025) y Chahid et al. (2022)** validan el uso de algoritmos como DDPG y Q-learning en ambientes acuícolas, mientras que otros estudios incorporan enfoques híbridos (RL + MPC) para la mejora del control ambiental y alimenticio (**Aljehani, 2023).**

Asimismo, investigaciones recientes exploran el uso de redes neuronales convolucionales, árboles de decisión y técnicas de procesamiento multisensorial para el monitoreo automatizado de peces y **crustáceos (Omankwu, 2024; Hossam, 2024; Georgopoulou, 2024).** Estas herramientas permiten una caracterización más precisa del entorno productivo, habilitando la generación de entornos simulados más realistas.

Sin embargo, se identifican **vacíos importantes en la implementación práctica de modelos RL en cultivos de camarón,** en especial en ***Litopenaeus vannamei***, donde las condiciones ambientales variables, la intermitencia de datos y la complejidad del sistema dificultan la adopción de decisiones automatizadas. La mayoría de aplicaciones se concentra en peces como tilapia o especies europeas, y pocas investigaciones validan modelos RL en cultivo de crustáceos en Latinoamérica.

En cuanto a las fortalezas observadas, destaca el uso de **entornos simulados multiescala y** el diseño de recompensas adaptativas que integran múltiples variables (crecimiento, consumo, calidad del agua). No obstante, persisten desafíos relacionados con la interpretabilidad del sistema, la robustez frente al ruido, y la falta de integración de variables económicas como el **retorno sobre la inversión (ROI)** o el **factor de conversión alimenticia (FCA)** como métricas explícitas de optimización.

En este contexto, el presente proyecto se posiciona como una contribución original al proponer un modelo de aprendizaje por refuerzo profundo, validado mediante simulación multivariable y retroalimentación experta, específicamente aplicado al cultivo intensivo *de L. vannamei* en el Perú. Esta aproximación busca reducir la brecha entre el desarrollo teórico y su aplicabilidad real, incorporando además variables económicas, productivas y técnicas en la formulación de dietas inteligentes.

Tabla . Taxonomía de métodos en IA aplicada.

| **#** | **Paper** | **Metodología Principal** | **Modelos/Algoritmos Clave** |
| --- | --- | --- | --- |
| 1 | Saad2024 | Integración de datos de sensores, video y manuales, catalogación en base grafo, y predicción de alimentación con ML. | Gradient Boosting Regressor, Elastic Net Regression, SVR. |
| 2 | Cyrex2024 | (No detallado, se centra en revisión) | Revisión de modelos ML para acuicultura. |
| 3 | Omankwu2024 | Recolección multisensorial, procesamiento en pipeline, modelos ML para monitoreo agrícola, IA para predicción y control. | CNN, SVM, Random Forest, RL. |
| 4 | Hossam2024 | Visión por computadora con YOLOv8, detección de puntos clave, monitoreo IoT, estimación de peso, control remoto. | YOLOv8, Pixel-depth conversion, Keypoint detection. |
| 5 | Georgopoulou2024 | Cámaras y AI en jaulas marinas, análisis de velocidad/densidad, cálculo del FBI, seguimiento con modelos de tracking. | YOLOv5, DEEPSORT, Polynomial fitting. |
| 6 | Elmessery2025 | Control RAS con RL (DDPG), definición de estados, acciones y recompensas, validación multiescala, análisis de rendimiento. | Deep Deterministic Policy Gradient (DDPG), Reward shaping. |
| 7 | Aljehani2023 | Revisión de control por modelos y RL, discusión sobre retos de incertidumbre, propuesta de enfoques híbridos. | MPC, RL (Q-learning, Actor-Critic). |
| 8 | Chahid2022 | Q-learning en control de crecimiento de tilapia, MDP en simulaciones, políticas óptimas, control de alimento y temperatura. | Q-learning, RL, MDP, Trajectory tracking. |
| 9 | Terayama2024 | Optimización de dietas con Bayesian Optimization (BO) y análisis ómicos, ajuste iterativo y validación en laboratorio. | Bayesian Optimization, ML para modelado de crecimiento, Ómicas. |

## Bases Teóricas

### Características biológicas e importancia económica del camarón blanco del Pacífico (Litopeneaeus vannamei)

Camarón blanco del Pacífico (Litopeneaeus vannamei) es una de las especies de crustáceos más cultivadas a nivel mundial debido a su rápido crecimiento, alta tasa de supervivencia y adaptabilidad a diversas condiciones ambientales. Originario de las costas del Pacífico oriental, desde México hasta Perú, su cultivo se ha expandido significativamente en América Latina y Ásia (FAO, 2007).

Esta especie se caracteriza por su capacidad de adaptación a diferentes salinidades, permitiendo su cultivo tanto en aguas marinas como en sistemas de agua dulce. Además, presenta una madurez sexual temprana y una alta fecundidad, lo que facilita su reproducción en cautiverio (Cobo & Pérez, 2018).

Los camarones peneidos pertenecen a la familia Penaeidae, la cual incluye numerosas especies. Entre estas, destaca Litopeneaus vannamei, conocido como el camarón blanco del Pacífico (Pérez-Farfante & Kensley, 1997). Este crustáceo desempeña un papel crucial en la industria pesquera y acuícola, debido a su gran valor comercial (Han et al., 2021). Su sabor suave, versatilidad culinaria y rápida adaptación a diferentes condiciones lo han convertido en una de las especies más demandadas en el mercado global de mariscos (Dayal et al., 2013).

Su capacidad y elevada tasa de reproducción rápida, su adaptación a diversas condiciones en el cultivo y su resistencia a enfermedades lo han convertido en el camarón blanco del Pacífico en una especie de especial interés para la acuicultura, sobre todo en regiones tropicales y subtropicales del mundo (FAO, 2022).

### Aplicaciones del Reinforcement Learning (RL) en la acuicultura

El Reinforcement Learning (RL) es una técnica de IA donde un agente aprende a través de interacciones con un entorno, recibiendo retroalimentación en forma de recompensas (Sutton y Barto, 2018). Esta metodología es particularmente adecuada para entornos acuícolas, caracterizados por condiciones dinámicas, variabilidad ambiental y relaciones no lineales entre las variables de control y los resultados productivos (Aljehani et al., 2023).

Uno de los algoritmos más relevantes para la implementación del RL en acuicultura es el Proximal Policy Optimization (PPO), considerado estable y eficiente para espacios de acción continuos, como aquellos que describen el suministro progresivo de alimento o la regulación de parámetros ambientales en sistemas acuícolas (Mnih et al., 2015). PPO permite ajustar las dosis de alimento de manera dinámica, considerando factores como temperatura, oxígeno disuelto, calidad del agua y tasa de crecimiento proyectada. Esto permite un aprendizaje continuo y adaptativo, superando las limitaciones de los métodos tradicionales, que suelen depender de reglas estáticas o de la interpretación subjetiva del personal técnico (Aljehani et al., 2023).

Además, el RL puede integrarse con modelos predictivos basados en redes neuronales, capaces de anticipar el crecimiento, el consumo de alimento y la eficiencia alimenticia. Estas redes pueden utilizarse como parte del entorno simulado del RL, proporcionando datos sintéticos y complejos sobre el comportamiento del sistema, lo cual acelera el aprendizaje y mejora la generalización del agente (Chen et al., 2022). De esta forma, se facilita la adopción de sistemas inteligentes capaces de realizar ajustes finos en la alimentación, optimizando no solo la tasa de crecimiento, sino también la conversión alimenticia y la sostenibilidad del cultivo.

En resumen, la aplicación del RL en acuicultura representa un avance hacia sistemas de control autónomos y adaptativos, capaces de mejorar el rendimiento productivo y reducir costos operativos. Aunque la mayoría de estas implementaciones se han desarrollado y probado en el cultivo de peces (como tilapia y trucha arcoíris), la extensión de estas técnicas al cultivo de Litopeneaeus vannamei es una oportunidad prometedora para futuras investigaciones y desarrollos (Aljehani et al., 2023).

Tabla . Comparación entre métodos en aprendizaje por reforzamiento encontrados en acuicultura

| **Característica** | **Q-learning** | **DDPG** | **PPO** |
| --- | --- | --- | --- |
| Naturaleza del algoritmo | Modelo libre (off-policy), discreto | Actor-Critic continuo (off-policy) | Actor-Critic continuo (on-policy) |
| Tipo de acción | Discreta (tablas Q) | Continua (redes neuronales) | Continua o discreta (redes neuronales) |
| Adaptabilidad a variabilidad | Limitada por discretización y escalabilidad | Alta adaptabilidad, ajusta políticas a entornos complejos | Alta adaptabilidad con políticas suaves y regulación (clip) para mejorar estabilidad |
| Tolerancia a intermitencia | Baja (depende de buena exploración y cobertura de espacio de estados) | Media-Alta (puede lidiar con datos intermitentes si bien entrenado y con buffer adecuado) | Alta (robustez al ruido y estabilidad en actualizaciones de política) |
| Eficiencia computacional | Baja para entornos grandes o continuos | Media (exige recursos por doble red Actor-Critic) | Alta (actualiza políticas más rápido y eficiente que otros algoritmos on-policy) |
| Escalabilidad a cultivo real | Limitada (no ideal para espacios de estado y acción grandes) | Alta (diseñado para problemas de control continuo, ideal para adaptarse a variabilidad) | Alta (buen balance entre estabilidad y adaptabilidad, ideal para entornos con variabilidad) |
| Robustez y estabilidad | Baja, propenso a exploración ineficiente | Media-Alta, requiere buen diseño de red y ajuste | Alta, con regulación mediante clipping para evitar cambios abruptos |
| Referencias clave | Chahid et al., 2022 | Elmessery et al., 2025; Aljehani et al., 2023 | Mnih et al., 2015 |

### Aplicaciones del Reinforcement Learning (RL) en el cultivo de *Litopeneaeus vannamei*

El avance de las tecnologías de la información ha permitido la incorporación de técnicas de aprendizaje automático en la acuicultura, especialmente en el cultivo de L. vannamei. Estas herramientas han sido utilizadas para optimizar la alimentación, mejorar la predicción del crecimiento y monitorear la calidad de agua.

Chen et al. (2022) desarrollaron un sistema de alimentación inteligente basado en la predicción del crecimiento del camarón en sistemas de recirculación acuícola, logrando una mejora en la eficiencia alimentaria y reducción de desperdicios.

Asimismo, Arfiati et al. (2025) realizaron un análisis comparativo de modelos de regresión de aprendizaje automático para predecir el crecimiento del camarón blanco en sistemas de acuicultura ecológica, destacando la eficacia de estas técnicas en la gestión del cultivo.

Además, Rahman et al. (2023) implementaron modelos de aprendizaje automático para la estimación de parámetros de calidad del agua en sistemas de cultivo de camarón en interiores, utilizando técnicas de "soft sensing" para predecir variables difíciles de medir directamente.

# Capítulo 4: Metodología de Investigación

4.1 Enfoque Metodológico

El presente estudio se enmarca en una **investigación tecnológica aplicada**, orientada al diseño, desarrollo y validación de un sistema inteligente para la optimización de dietas en acuicultura, mediante técnicas de Reinforcement Learning (RL). Este enfoque permite generar conocimiento nuevo a través de la aplicación práctica de herramientas avanzadas de inteligencia artificial en un contexto productivo real.

La metodología adoptada combina el marco estructurado de CRISP-DM (Cross Industry Standard Process for Data Mining) con elementos iterativos propios de Agile+AI, permitiendo un desarrollo flexible pero controlado del sistema propuesto. CRISP-DM aporta una guía clara en cuanto a las etapas de comprensión del negocio, preparación de los datos, modelado y evaluación, mientras que Agile+AI se integra en la fase de modelado mediante ciclos cortos (sprints) que facilitan el ajuste progresivo de la función de recompensa, la arquitectura del agente y el entorno simulado, en coordinación con expertos del dominio.

No se contempla el uso completo de MLOps, ya que el alcance de esta tesis es de carácter experimental. Sin embargo, sus principios serán considerados como guía futura para una eventual implementación operativa del modelo en un entorno productivo real.

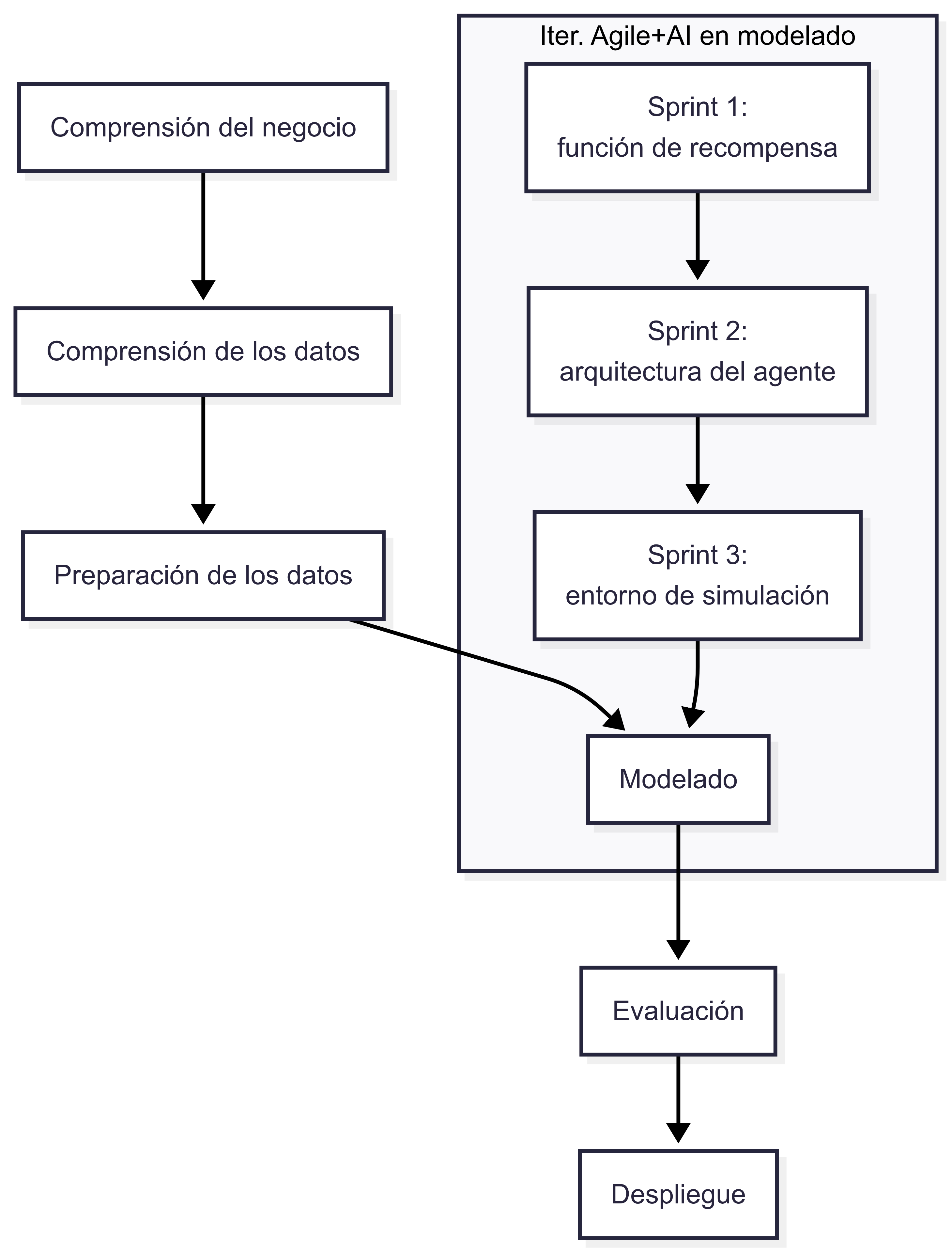


Ilustración Esquema híbrido CRISP-DM + Agile+AI. *El proceso de modelado incluye iteraciones ágiles para el ajuste progresivo del agente de Reinforcement Learning (RL).*

El ciclo de investigación sigue cuatro etapas principales:

* Diseño: incluye la formulación del problema, definición del entorno RL, variables clave, y arquitectura inicial del agente.
* Implementación: abarca la construcción del entorno simulado, desarrollo del agente RL (Q-learning, DDPG o PPO), e integración con modelos predictivos auxiliares.
* Validación: implica pruebas experimentales en simulación, análisis de métricas (ROI, FCA, tasa de crecimiento) y revisión por expertos del sector acuícola.
* Comunicación: comprende la documentación del proceso, análisis de resultados y elaboración del informe final de tesis.

Este enfoque metodológico permite no solo evaluar el desempeño técnico del modelo propuesto, sino también su aplicabilidad práctica en entornos reales, abriendo la posibilidad de futuras implementaciones piloto en campo.

4.2 Diseño Experimental

El diseño experimental propuesto sigue una secuencia estructurada de etapas para garantizar un entrenamiento robusto del agente de aprendizaje por refuerzo. A continuación, se describen los pasos principales:

1. **Ingesta de datos**: Conexión a la base de datos en Google BigQuery para la extracción de registros históricos de variables ambientales y productivas (2020–2025).
2. **Limpieza**: Eliminación de registros incompletos, imputación de valores faltantes y verificación de consistencia temporal.
3. **Transformaciones y discretización**: Aplicación de técnicas como KBinsDiscretizerpara estandarizar el espacio de acciones en modelos como Q-learning.
4. **Generación de datos sintéticos**: Implementación de redes generativas adversariales (GANs) para ampliar la base de entrenamiento con escenarios sintéticos basados en los patrones históricos reales.
5. **Inyección de ruido controlado**: Introducción de ruido gaussiano para simular variabilidad ambiental no observada, fortaleciendo la generalización del modelo.
6. **Selección de variables**: Aplicación del algoritmo Boruta para identificar las variables con mayor poder explicativo.
7. **Entrenamiento del agente RL:** Desarrollo e implementación de tres variantes (Q-learning, PPO, DDPG) en entornos simulados, comparando desempeño según métricas técnicas (MAE, MSE, ROI).
8. **Validación del modelo**: Evaluación mediante técnicas de hold-out, K-Fold y bootstrapping. Complementariamente, se considera A/B Testing para simular escenarios de adopción.
9. **Evaluación de desempeño**:
   1. **Para el agente RL:** se evaluará la recompensa acumulada, la estabilidad del entrenamiento, la tasa de convergencia, el ROI simulado, el IC y el FCA (eficiencia alimenticia).
   2. **Para los modelos auxiliares:** se utilizarán métricas de regresión como MAE y MSE para validar las predicciones de crecimiento y consumo.
   3. **Evaluación cualitativa (expertos):** se incluirá la retroalimentación de expertos mediante encuestas estructuradas, entrevistas y análisis de métricas UX tales como usabilidad percibida, confianza en las recomendaciones, interpretabilidad de las decisiones, carga cognitiva y satisfacción general.

Tabla . Principales métricas UX

| **Métrica UX** | **Descripción breve** |
| --- | --- |
| **Usabilidad percibida** | **¿El sistema es fácil de entender y utilizar por el usuario final?** |
| **Confianza en las decisiones** | **¿Los usuarios confían en las recomendaciones del agente RL?** |
| **Interpretabilidad** | **¿Se entienden los motivos detrás de una recomendación específica?** |
| **Tiempo de adopción** | **¿Cuánto tiempo toma a un experto usar eficazmente el sistema?** |
| **Carga cognitiva** | **¿Qué tanto esfuerzo mental requiere usar el sistema? (evaluable vía NASA-TLX)** |
| **Satisfacción del usuario** | **¿Qué tan satisfechos están los usuarios con la herramienta?** |
| **Aceptación de uso futuro** | **¿Considerarían integrar el modelo RL en sus flujos de decisión reales?** |

Diagrama

El contenido generado por IA puede ser incorrecto.

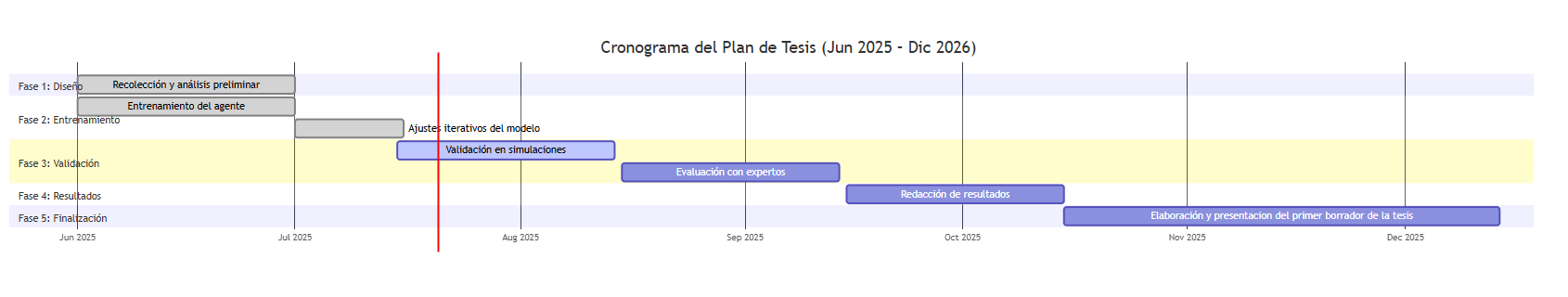
Ilustración Esquema del flujo metodológico para el desarrollo, entrenamiento y evaluación del modelo de aprendizaje por refuerzo en acuicultura

4.3 Interacción con Stakeholders

Se establecerá un plan de interacción continua con los stakeholders del proyecto, que incluye sesiones programadas de consulta y retroalimentación. Estas reuniones permitirán validar el enfoque del modelo, ajustar parámetros en función de criterios técnicos y operativos, y asegurar la alineación del sistema con los objetivos estratégicos de la organización. Además, se documentará cada iteración relevante como parte de una estrategia de desarrollo ágil e incremental.

# Capítulo 5: Administración del plan de tesis

5.1 Cronograma



| **Mes** | **Actividad principal** |
| --- | --- |
| junio 2025 | Recolección y análisis preliminar de datos históricos; diseño del modelo de Reinforcement Learning (RL) y del entorno simulado. |
| junio - agosto 2025 | Entrenamiento del agente RL (Q-learning, DDPG o PPO); ajustes iterativos en la función de recompensa y arquitectura del modelo. |
| agosto - septiembre 2025 | Validación del modelo en escenarios simulados con datos reales; análisis de métricas productivas. |
| septiembre - octubre 2026 | Revisión del sistema con expertos del sector; aplicación de mejoras según retroalimentación técnica. |
| noviembre 2026 | Redacción del informe de resultados; presentación de hallazgos y discusión de impactos potenciales. |
| diciembre 2026 | Elaboración final del documento de tesis; revisión por parte del asesor; proceso de sustentación académica. |

5.2 Presupuesto

Para el desarrollo, entrenamiento, despliegue y validación de los modelos de inteligencia artificial propuestos en esta tesis, se estima el uso de diversas herramientas de Google Cloud Platform (GCP), considerando los requerimientos técnicos asociados al procesamiento intensivo, la automatización de tareas y la ejecución de servicios en contenedores.

A continuación, se detalla el presupuesto mensual y anual estimado. Este presupuesto se sugiere a la organización como parte de los costos operativos anuales a considerar para la integración sostenible de soluciones basadas en inteligencia artificial.

Tabla . Herramientas institucionales

| **Herramienta** | **Funcionalidad** | **Costo Estimado** | **Costo Mensual (USD)** | **Costo Anual (USD)** |
| --- | --- | --- | --- | --- |
| **Cloud Scheduler** | Automatiza tareas como actualizaciones de tablas, limpieza de datos o envío de reportes. | $0.10/job/mes × 5 jobs | $0.50 | $6.00 |
| **Cloud Composer** | Orquesta flujos complejos tipo ETL: cargar datos → procesar → calcular → guardar. | $300/mes por entorno | $300.00 | $3,600.00 |
| **Cloud Workflows** | Coordina ejecución de múltiples servicios o scripts en secuencia. | ~$0.025 por ejecución × estimaciones | $5.00 | $60.00 |
| **Compute Engine (VM)** | Ejecuta notebooks pesados, simulaciones o pruebas de modelos sin GPU. | ~$25–80/mes (sin GPU) | $80.00 | $960.00 |
| **Colab Enterprise** | Uso colaborativo de notebooks con mayor potencia y conexión directa a GCP. | ~$50–100/mes por usuario activo | $75.00 | $900.00 |
| **Vertex AI (GPU)** | Entrena modelos complejos (ej. RL o deep learning) estimando uso diario de 3 horas. | $1–8/hora × 3h × 30 días | $720.00 | $8,640.00 |
| **Cloud Run (API)** | Despliega scripts o apps con contenedores Docker, útil para exponer APIs del modelo. | $3–$20/mes según uso | $20.00 | $240.00 |
|  | **Total estimado** |  | **$1,200.50** | **$14,406.00** |

5.3 Financiamiento  
  
Las herramientas tecnológicas requeridas para el desarrollo del presente proyecto, así como el tiempo de dedicación del tesista, serán cubiertos por la institución beneficiada directamente, en el marco de sus actividades de investigación e innovación. No se contempla un financiamiento externo adicional al momento de la elaboración del presente plan.

# Capítulo 6: Referencias Bibliográficas

1. Amershi, S., Begel, A., Bird, C., DeLine, R., Gall, H., Kamar, E., ... & Zimmermann, T. (2019). Software engineering for machine learning: A case study. Proceedings of the IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP), 291–300. https://doi.org/10.1109/ICSE-SEIP.2019.00030
2. Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step data mining guide. SPSS Inc.

Disponible en: https://www.the-modeling-agency.com/crisp-dm.pdf

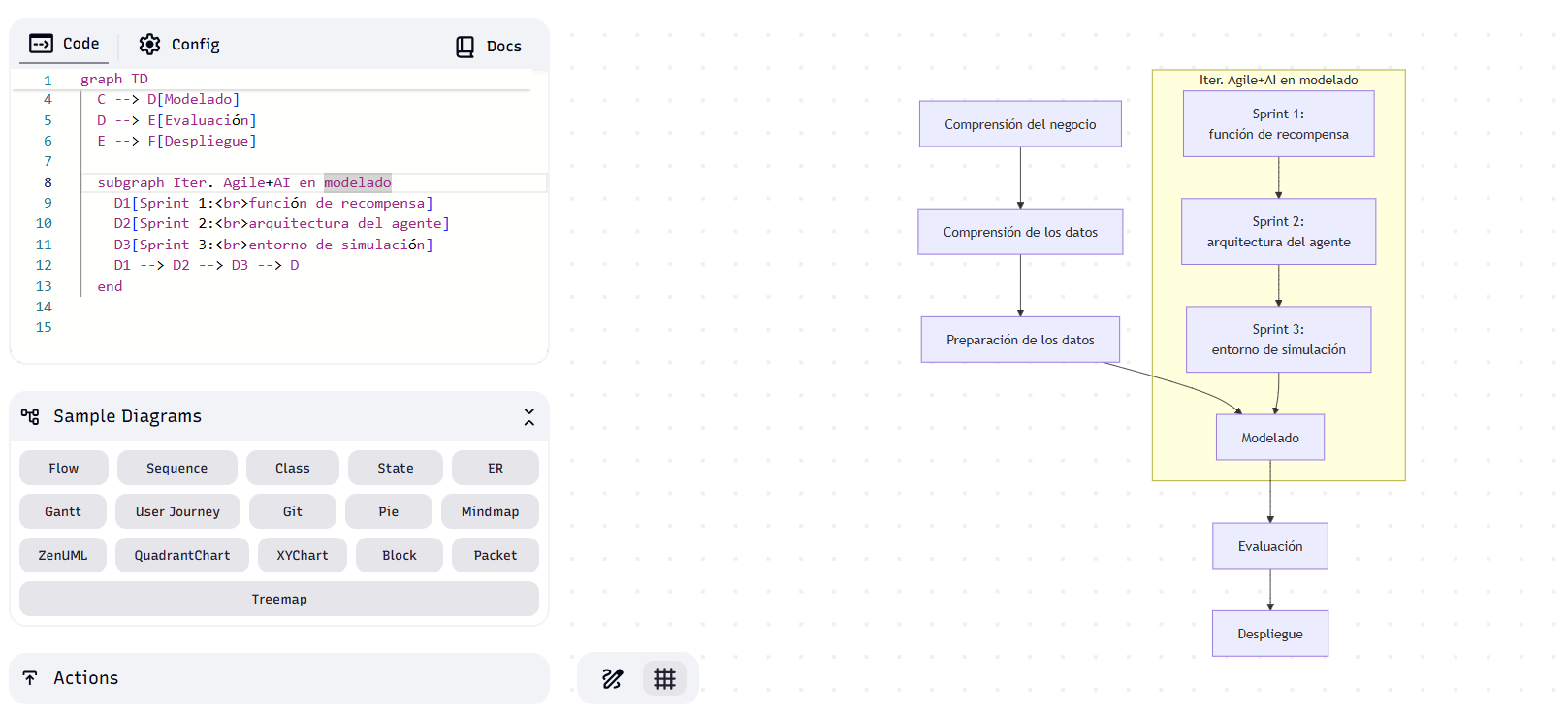
1. Kreuzberger, D., Kühl, N., & Hirschl, S. (2022). Machine Learning Operations (MLOps): Overview, definition, and architecture. arXiv preprint arXiv:2205.02302. https://arxiv.org/abs/2205.02302
2. Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction (2nd ed.). MIT Press. https://www.andrew.cmu.edu/course/10-703/textbook/BartoSutton.pdf
3. Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., ... & Wierstra, D. (2016). Continuous control with deep reinforcement learning. arXiv preprint arXiv:1509.02971. https://arxiv.org/abs/1509.02971
4. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal Policy Optimization Algorithms. arXiv preprint arXiv:1707.06347. https://arxiv.org/abs/1707.06347FAO (2022). Aquaculture Feed and Feeding Practices. Food and Agriculture Organization of the United Nations.
5. Salfner, F. et al. (20YY). A Survey of Online Failure Prediction Methods. ACM Journal.
6. Deldjoo, Y. et al. (2023). A Review of Modern Fashion Recommender Systems. ACM Computing Surveys.
7. Aljehani, F., N’Doye, I., & Laleg-Kirati, T. M. (2023). Feeding control and water quality monitoring in aquaculture systems: Opportunities and challenges. Elsevier Preprint. https://arxiv.org/abs/2306.09920
8. Chen, Y., et al. (2022). Técnica de alimentación inteligente basada en la predicción del crecimiento del camarón en un sistema de recirculación acuícola. Panorama Acuícola.
9. Mnih, V., et al. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529-533.
10. Chahid, A., et al. (2022). *Fish growth trajectory tracking using Q-learning in precision aquaculture*. *Aquaculture*, 550, 737838.
11. Elmessery, W. M., et al. (2025). *A deep deterministic policy gradient approach for optimizing feeding rates and water quality management in recirculating aquaculture systems*. *Aquaculture International*.
12. Georgopoulou, D., et al. (2024). *Real-time monitoring of feeding behavior in European sea bass using AI models and computer vision*. *IEEE IROS Workshop*.

# Capítulo 7: Anexos

## 7.1. Esquemas de trabajo

### 7.1.1. Esquema híbrido de desarrollo metodológico: CRISP-DM + Agile+AI

Este esquema ilustra la integración de prácticas de desarrollo ágil (Agile+AI) dentro de la fase de modelado del enfoque CRISP-DM. Las iteraciones representadas como "sprints" permiten ajustar progresivamente componentes clave del agente de Reinforcement Learning (RL), como la función de recompensa, la arquitectura del agente y el entorno de simulación. El resto del flujo sigue la estructura tradicional de CRISP-DM, con etapas bien definidas desde la comprensión del negocio hasta el despliegue del modelo.



### 7.1.2. Pipeline de Diseño experimental

graph TD

    A[Ingesta<br>BigQuery: datos ambientales y productivos] --> B[Limpieza y transformación<br>Imputación, discretización, normalización]

    B --> C[Generación de datos sintéticos<br>Redes GAN entrenadas con datos reales]

    C --> D[Inyección de ruido<br>Ruido gaussiano para simular incertidumbre]

    D --> E[Selección de variables<br>Algoritmo Boruta]

    E --> F[Modelado de entorno simulado<br>Estados, acciones, recompensas, transiciones]

    F --> G1[Entrenamiento con Q-Learning]

    F --> G2[Entrenamiento con PPO]

    F --> G3[Entrenamiento con DDPG]

    G1 --> H[Validación del modelo<br>Hold-out, K-Fold, bootstrapping, A/B testing]

    G2 --> H

    G3 --> H

    H --> I1[Evaluación agente RL<br>Recompensa acumulada, ROI, IC, FCA]

    H --> I2[Evaluación modelos auxiliares<br>MAE, MSE en crecimiento y consumo]

    H --> I3[Evaluación cualitativa<br>Encuestas y entrevistas a expertos, métricas UX]

### 7.1.3. Cronograma

---

config:

  theme: neo

---

gantt

    title Cronograma del Plan de Tesis (Jun 2025 – Dic 2026)

    dateFormat  YYYY-MM-DD

    axisFormat  %b %Y

    section Fase 1: Diseño

    Recolección y análisis preliminar             :done, a1, 2025-06-01, 30d

    section Fase 2: Entrenamiento

    Entrenamiento del agente                      :done, a2, 2025-06-01, 30d

    Ajustes iterativos del modelo                 :done, a3, 2025-07-01, 15d

    section Fase 3: Validación

    Validación en simulaciones                    :active, a4, 2025-07-15, 30d

    Evaluación con expertos                       :a5, 2025-08-15, 30d

    section Fase 4: Resultados

    Redacción de resultados                       :a6, 2025-09-15, 30d

    section Fase 5: Finalización

    Elaboración y presentacion del primer borrador de la tesis           :a7, 2025-10-15, 60d

## 7.2. Matriz de consistencia

| **Objetivos** | **Hipótesis** | **Variables** | **Técnicas y métodos** | **Instrumentos** |
| --- | --- | --- | --- | --- |
| **Objetivo General** Desarrollar un modelo de RL que optimice dietas en acuicultura según condiciones ambientales y productivas. | El modelo de RL mejora la eficiencia alimentaria y el ROI respecto a métodos tradicionales. | Variable dependiente:  • ROI  • FCA  Variables independientes:  • Temperatura, oxígeno, salinidad, pH  • Peso del camarón, consumo, tipo de alimento | • Reinforcement Learning (RL) (DDPG, Q-learning, PPO)  • Integración con modelos predictivos auxiliares  • Simulación y validación comparativa | • Dataset histórico 2020–2025 (Vitapro)  • Simulador acuícola  • Ficha de validación de expertos |
| **OE1. Analizar datos históricos y actuales para construir modelos predictivos auxiliares.** | El uso de modelos predictivos mejora la precisión del entorno simulado. | • Crecimiento proyectado• Consumo estimado  • Parámetros ambientales diarios | • Análisis exploratorio de datos  • Aprendizaje supervisado (regresión, redes neuronales, XGBoost) | • Registros históricos de CEAs  • Modelos de predicción entrenados offline |
| **OE2. Desarrollar un modelo de RL (Q-learning, DDPG o PPO) y seleccionar el mejor.** | DDPG o PPO presentan mejor desempeño frente a Q-learning en entornos acuícolas variables. | • Tipo de algoritmo RL  • Tasa de convergencia  • Calidad de la política generada | • Entrenamiento iterativo de agentes RL  • Métricas de recompensa y estabilidad | • Scripts en Python (OpenAI Gym, TensorFlow/PyTorch)  • Simulador de cultivo |
| **OE3. Validar el sistema con expertos y métricas productivas.** | La validación por expertos confirma la utilidad práctica del sistema propuesto. | • Evaluación cualitativa  • Satisfacción del experto  • Nivel de adopción | • Entrevistas estructuradas  • Evaluación de prototipo en entorno simulado | • Cuestionario de validación  • Registro de decisiones del modelo |