

Optimización con inteligencia artificial del despacho de riego solar en cultivos de pistacho

Autor

Ing. Tomas Corteggiano

Director del trabajo

Dr. Carlos Larisson

Este plan de trabajo ha sido realizado en el marco de la asignatura Gestión de Proyectos entre marzo y abril de 2025.

Tabla de contenido

1. Breve resumen del trabajo realizado hasta la fecha	2
2. Avance en las tareas	4
3. Cumplimiento de los requerimientos	5
4. Gestión de riesgos	6

IMPORTANTE: No borrar las consignas en cada una de las cuatro secciones de este documento, de forma tal que el jurado tenga claro qué es lo solicitado en cada caso, así como el significado de los símbolos y colores utilizados.

Revisión	Cambios realizados	Fecha
1.0	Creación del documento	27/10/2025

1. Breve resumen del trabajo realizado hasta la fecha

Elabore un detalle del estado del proyecto a la fecha. **Utilicé como mínimo dos páginas completas y como máximo tres páginas.** Explique muy brevemente en qué consiste su Trabajo Final, **aunque esa información esté más detallada en el Plan de Trabajo al cual su Jurado también tiene acceso.** Incluya imágenes y tablas según considere apropiado. **Indique con claridad por qué estima que podrá completar todos los faltantes (o al menos la gran mayoría) antes del inicio del Taller de Trabajo Final.**

El objetivo principal de este trabajo es el diseño e implementación de la capa de gestión y optimización para un sistema de riego solar inteligente en cultivos de pistachos. El núcleo del proyecto es un software capaz de procesar telemetría, simular el estado del sistema mediante un gemelo digital y generar automáticamente planes de despacho de riego que maximicen la eficiencia energética e hídrica.

A la fecha (25 de octubre de 2025), el proyecto ha completado exitosamente su hito técnico más crítico y se encuentra en la fase final de implementación del módulo de optimización. Entre los 3 hitos más importante logrados hasta la fecha se tiene:

1. Desarrollo del gemelo digital híbrido
2. Implementación y deployment en servidor del *pipeline* de MLOps para adquisición, preprocesamiento, entrenamiento y deployment de modelo de IA estimador de residuales.
3. Implementación de la optimización global

Gemelo digital híbrido - Completado

El desafío central del proyecto es tomar decisiones óptimas (que válvulas abrir y cuando) basándose en mediciones limitadas (pocos puntos de medición real de presión y caudal en la salida de la bomba). Para lograr esto, se requiere un modelo predictivo de muy alta fidelidad (un “gemelo digital”) que describa cómo se comporta la red hidráulica.

Se determinó que un enfoque puramente físico (modelo de “caja blanca”, como EPANET) sufre sesgos debido a incertidumbres en los parámetros (ej. rugosidad de tuberías) y simplificaciones del modelo. Por otro lado, un modelo puro de IA (“caja negra”) no es robusto, requiere demasiados datos y puede generar predicciones físicamente inconsistentes.

La solución implementada fue un **modelo híbrido o “caja gris”**, que fusiona ambos enfoques:

1. **Sistema físico:** Se implementó un simulador hidráulico basado en primeros principios (WNTR/EPANET) que proporciona una línea de base robusta y físicamente consistente. Ese componente captura la dinámica principal del sistema de riego.
2. **Corrector de residuales:** Se desarrolló un modelo de aprendizaje profundo cuya única tarea es aprender y predecir el error (el “residual”) del modelo físico. De esta forma el residual simplemente es definido como la resta entre el caudal real medido y la salida del simulador para ese mismo estado de las variables de entrada del sistema (fase de entrenamiento usando datos históricos procesados por mi *pipeline*). El modelo entonces aprende a estimar el error de la simulación de base y corrige las sobre y subestimaciones del mismo.

El modelo final de alta fidelidad, es la suma de la predicción física más la corrección del modelo de IA.

Este hito representa un avance de suma importancia y representa la finalización del modelo directo, el cual es prerequisite técnico indispensable para la estimación del estado y alimentación de la optimización global.

Plataforma MLOps - Completada

Un modelo predictivo tiene valor limitado si no está integrado en un ciclo de vida automatizado que le permita aprender y ejecutarse periódicamente. Por ello, se implementó una plataforma completa de **Machine Learning Operations (MLOps)** utilizando Apache Airflow y MLflow.

Esta plataforma automatiza todo el *pipeline* del Gemelo Digital Híbrido (con modelo de corrección):

1. **Ingesta y Limpieza de Datos:** Consolida y limpia los datos de telemetría del campo.
2. **Entrenamiento y Versionado:** Entrena periódicamente el corrector de IA con los datos más recientes. Cada modelo, sus métricas (RMSE, R^2) y artefactos (gráficos de residuales) son versionados en MLflow.
3. **Validación y Despliegue:** El nuevo modelo es validado rigurosamente (ej. debe mejorar el RMSE en >5% vs. el modelo físico). Si supera la validación, se despliega automáticamente para ser consumido por la API del sistema.

La finalización de estos dos hitos ha mitigado uno de los mayores riesgos técnicos del proyecto y completa la "Optimización de Entrenamiento del Modelo de IA".

Optimización global - En proceso de finalización

Con el Gemelo Digital Híbrido actuando como una "función de aptitud" (fitness function) precisa, el proyecto ha transitado a la fase de "Optimización Global". Esta fase se divide en dos problemas distintos:

- A. Estimación del estado completo (incluyendo no solo el caudal total medido sino la distribución interna corregida a partir de la inferencia del modelo de IA):
 - a. Se ha implementado una solución funcional. Este módulo utiliza el modelo híbrido como oráculo y emplea un optimizador (`scipy.optimize.minimize_scalar`) para encontrar un **factor de corrección global óptimo**. Este factor ajusta los parámetros del modelo físico (los coeficientes de los emisores) hasta que el caudal simulado coincide perfectamente con el caudal objetivo predicho por el modelo híbrido. Esto permite al sistema calibrar el estado real de la red en tiempo real.
- B. Optimización del Despacho (El Problema de Programación) - En Fase de Implementación:

- a. El trabajo de análisis y diseño de este módulo está 100% completo.
 1. **Clasificación del Problema:** Se ha identificado formalmente como un **Problema de Programación Mixta-Entera No Lineal (MINLP)**. Esto se debe a que las variables de decisión son binarias (válvula abierta/cerrada) y la función de evaluación es altamente no lineal.
 2. **Selección de Método:** Los problemas MINLP son NP-hard. Los solvers tradicionales no pueden garantizar una solución dentro del requerimiento funcional de RNF 3.1 (< 1 minuto). Por lo tanto, se ha optado estratégicamente por una **metaheurística**.
 3. **Método Elegido:** Se ha seleccionado un **Algoritmo Genético (GA)** como la implementación principal, dada su robustez para problemas de programación complejos.

Estimación de finalización

El proyecto se encuentra en una posición inmejorable. Los hitos más riesgosos (el modelo predictivo y la plataforma MLOps) están 100% completados y validados. El trabajo restante (WBS 3.4 y 3.5) consiste en implementar el optimizador GA, cuyo diseño teórico y plan de implementación están finalizados. Las 90 horas asignadas a la "Redacción de la memoria final" son más que suficientes para documentar estos logros sustanciales y el diseño detallado del optimizador. Se estima que todos los requerimientos funcionales y entregables estarán completados antes del inicio del Taller de Trabajo Final

2. Avance en las tareas

a) Indicar a continuación para cada una de las tareas su estado de situación según su criterio, utilizando verde si considera que es satisfactorio, amarillo si considera que es insatisfactorio por sobrecostos y/o demoras, y rojo si lo considera muy insatisfactorio por sobrecostos y/o demoras.

Si a la fecha de completar este informe no está previsto que la tarea haya comenzado entonces deje la celda correspondiente en blanco, sin pintarla con ningún color.

En subcelda inferior izquierda colocar:

- ** si los recursos u horas utilizadas fueron o están siendo muy inferior a lo planificado.
- * si los recursos u horas utilizadas fueron o están siendo inferior a lo planificado.
- \$ si los recursos u horas utilizadas fueron o están siendo de acuerdo a lo planificado.
- \$\$ si los recursos u horas utilizadas fueron o están siendo superior a lo planificado.
- \$\$\$ si los recursos u horas utilizadas fueron o están siendo muy superior a lo planificado.

En subcelda inferior derecha colocar:

- -- si la tarea se ejecutó o se está ejecutando mucho más rápido de lo previsto
- - si la tarea se ejecutó o se está ejecutando más rápido de lo previsto
- = si la tarea se ejecutó o se está ejecutando en el tiempo previsto.
- + si la tarea se ejecutó o se está ejecutando con demoras.
- ++ si la tarea se ejecutó o se está ejecutando con demoras muy significativas.

IMPORTANTE: Indicar con borde grueso las tareas que forman parte del camino crítico

1.1 Planificación detallada inicial (10 h)	1.1 Reuniones periódicas (30 h)	1.3. Administración del backlog (10 h)	
*	=	\$	+
2.1 Arquitectura y stack (25 h)	2.2. Diseño del modelo BD	2.3. Investigación de APIs (15 h)	2.4. Diseño de la API (25 h)
\$	=	\$	+
3.1. Configuración entorno (20 h)	3.2. Módulo de ingesta (40 h)	3.3. Integración API pronóstico	3.4. Motor de decisión (60 h)
*	+	\$	=
3.5. Algoritmos optimización (70 h)	3.6. Desarrollo de la API (35 h)	3.7. Lógica certificación (10 h)	
\$\$\$	=	*	=

4.1. Pruebas unitarias (20 h)	Pruebas funcionales (30 h)	4.3. UAT con Ing. Palomas (10 h)	4.4. Despliegue (10 h)
5.1. Guía de uso API (20 h)	5.2. Documentación técnica (20 h)	5.3. Informe de validación (10 h)	5.4. Cierre formal (5 h)
6.1. Redacción de la memoria (90 h)			

3. Cumplimiento de los requerimientos

a) Indicar a continuación para cada uno de los requerimientos el estado de situación según su criterio, utilizando verde si considera que ya se ha cumplido, amarillo si considera que aún no se ha cumplido pero se podrá cumplir, y rojo si considera que aún no se ha cumplido y tiene dudas si se podrá cumplir.

Si considera que es necesario modificar los requerimientos respecto a los indicados en la planificación inicial entonces incluya acá los requerimientos actualizados, **marcando en negrita** aquellos que son nuevos o se han modificado.

RF1.1: el sistema deberá generar un plan de riego base para las próximas 24 horas y, posteriormente, reevaluar y ajustar dicho plan en intervalos periódicos en función de las condiciones en tiempo real (potencia solar y nivel freático del pozo).
RF1.2: el sistema deberá enviar automáticamente los comandos de accionamiento a las válvulas de acuerdo al plan de riego activo y sus ajustes dinámicos.
RF1.3: al final de cada ciclo, el sistema deberá calcular el déficit de riego (diferencia entre lo planificado y lo ejecutado). Este déficit deberá ser considerado como una demanda hídrica adicional en la planificación del día siguiente.
RF1.4: el sistema deberá registrar en una base de datos persistente todas las acciones de riego, mediciones de sensores, ajustes del plan y pronósticos generados.

RF1.5.1: el modelo de optimización deberá aceptar como entradas: irradiación solar, la telemetría en tiempo real del campo (incluyendo el nivel freático del pozo) y los parámetros del cultivo
RF1.5.2: (modificado) el núcleo del módulo deberá resolver un problema de optimización cuyo objetivo principal sea maximizar la eficiencia operativa y la sostenibilidad, permitiendo al usuario ponderar estratégicamente entre (1) la uniformidad del despacho, (2) la minimización del tiempo de bombeo y (3) la minimización del estrés sobre el acuífero , sujeto a la restricción de cumplir con el perfil de riego objetivo.
RF1.5.3: la salida del módulo deberá ser un cronograma de despacho de riego dinámico, detallando para cada válvula los intervalos de tiempo de apertura y cierre para el próximo horizonte de control.
RF1.5.4: el plan generado deberá respetar en todo momento las restricciones operativas, incluyendo la potencia solar instantánea disponible, los horarios de no riego y el nivel freático mínimo de seguridad del pozo.
RNF 3.1: cada ciclo de optimización periódica deberá completarse en menos de 1 minuto.
RD4.1: se deberá entregar la documentación técnica de la arquitectura del software.
RD4.2: se deberá entregar la memoria final del proyecto, siguiendo el formato establecido por la universidad.

Justificación de Modificaciones (RF1.5.2): El requerimiento original se ha mejorado significativamente. La investigación reveló que una simple "minimización de agua y energía" es trivial en un sistema 100% solar (el costo de energía es cero). El nuevo requerimiento (implementado a través de la "Función de Costo Híbrida y Unificada") es estratégicamente superior, ya que permite a la usuaria final (Ing. Palomas) balancear objetivos agronómicos (uniformidad) , de eficiencia (minimizar tiempo) y de sostenibilidad (proteger el acuífero) , alineándose mucho mejor con las necesidades del negocio.

4. Gestión de riesgos

a) Indicar a continuación para cada uno de los riesgos el estado de situación según su criterio, utilizando verde si considera que el riesgo ya no se manifestará o es muy improbable que se manifieste, amarillo si considera que es posible que es improbable que el riesgo se manifieste o si se manifiesta estima que será fácilmente controlado, y rojo si considera que es muy probable que el riesgo se manifieste y que no pueda ser controlado fácilmente.

Si considera que es necesario modificar los riesgos respecto a los presentados en la planificación inicial entonces incluya acá los riesgos actualizados, **marcando en negrita** aquellos que son nuevos o se han modificado, e indicando para ellos los valores de S, O y RPN, junto con su respectiva justificación.

Riesgo 1: Falta de disponibilidad de datos históricos de calidad o en la cantidad necesaria para entrenar los modelos de IA.

Riesgo 2: Retraso significativo en el desarrollo de los algoritmos de optimización.

Riesgo 3: Problemas de compatibilidad o conectividad con el hardware existente (sensores, bombas, PLC).

Riesgo 4: Rechazo del plan de riego por parte del usuario final (Ing. Eugenia Palomas).

Justificación de Estados:

- **Riesgo 1 (Verde):** Este riesgo está completamente mitigado y resuelto. La finalización y validación de los Hitos 1 y 2 (MLOps) demuestra que los datos históricos fueron suficientes y de la calidad necesaria para entrenar exitosamente el corrector de residuales y construir el Gemelo Digital Híbrido.
- **Riesgo 2 (Amarillo):** Este riesgo se ha controlado. El análisis confirmó que el problema era computacionalmente complejo (MINLP). El plan de mitigación original (dividir en hitos, buscar guía) se ejecutó, llevando a la decisión estratégica de usar metaheurísticas (GA) en lugar de solvers tradicionales que no cumplirían el requisito de tiempo (< 1 min). El riesgo de "retraso" está controlado porque ya se seleccionó una vía de implementación viable (el GA) y su diseño está completo.
- **Riesgo 3 (Verde):** Este riesgo está resuelto. La plataforma MLOps está ingiriendo datos de telemetría en tiempo real. El modelo utiliza activamente las lecturas de los sensores (presión,

nivel freático) como entradas. Esto confirma que la conectividad y compatibilidad con el hardware de campo están solucionadas.

- **Riesgo 4 (Amarillo):** Este riesgo está controlado y es improbable que se manifieste. La decisión de implementar una "Función de Costo Híbrida y Unificada" es una mitigación directa. En lugar de forzar un plan rígido, el sistema permitirá a la Ing. Palomas (usuaria final) ajustar las ponderaciones de la optimización (ej. priorizar el acuífero vs. eficiencia). Esto le da control agronómico directo sobre la estrategia de riego, haciendo que el rechazo sea muy improbable.