PACS-AI Assist: Asistente predictivo con IA para diagnóstico en imágenes médicas

Capstone – Proyecto Integrador (APT)

Nombre de los integrantes: Diego Castañeda

Millaray Llanquin

Scarlett Socías

TABLA DE CONTENIDOS

[**Fase 1: Autoevaluación – Definición de Proyecto APT 3**](#_heading=h.a9o6yiqfdbkx)

[1. Abstract (English) 3](#_heading=h.4ue763ethsfx)

[Resumen (Español) 3](#_heading=h.5ag2u19uwzh8)

[2. Descripción del Proyecto APT 4](#_heading=h.eaqw4xerkr82)

[3. Relación con Competencias del Perfil de Egreso 5](#_heading=h.zdr7phhdas8g)

[4. Relación con Intereses Profesionales 6](#_heading=h.vav2hgp4aylt)

[5. Factibilidad 6](#_heading=h.v1nb24drlmz4)

[6. Objetivos (SMART) 7](#_heading=h.o9z3mt7nuixv)

[7. Metodología 9](#_heading=h.82x4eslb6vbh)

[8. Plan de Trabajo 10](#_heading=h.m4rthnnvvdio)

[9. Evidencias 11](#_heading=h.qtkog55celld)

[10. Conclusions (English) 12](#_heading=h.qllgfjh8i4ji)

[11. Reflection (English) 12](#_heading=h.veytcmjt1ymj)

# Fase 1: Autoevaluación – Definición de Proyecto APT

## 1. Abstract (English)

This project proposes PACS-AI Assist, an AI-powered assistant designed to integrate with hospital Picture Archiving and Communication Systems (PACS) using the DICOM standard. The system processes CT scan images, applies a CNN-based model to detect pulmonary nodules, and presents results through a user-friendly interface with highlighted regions and malignancy probability scores. The solution demonstrates relevance to the Informatics Engineering field by integrating software development, image processing, cloud deployment, and data governance, with direct applicability to health technology. Feasibility is ensured by starting with a simulated environment (Orthanc PACS, open DICOM datasets)weeks.

## Resumen (Español)

Este proyecto propone PACS-AI Assist, un asistente impulsado por IA diseñado para integrarse con sistemas hospitalarios PACS mediante el estándar DICOM. El sistema procesa imágenes de tomografía computarizada, aplica un modelo CNN para detectar nódulos pulmonares y presenta los resultados en una interfaz amigable, destacando las regiones de interés y mostrando probabilidades de malignidad. La solución es relevante para Ingeniería en Informática al integrar desarrollo de software, procesamiento de imágenes, despliegue en la nube y gobierno de datos, con una clara aplicación en el área de la salud. La factibilidad se asegura iniciando con un entorno simulado (Orthanc PACS, datasets DICOM abiertos).

## 2. Descripción del Proyecto APT

**Título:** PACS-AI Assist – Asistente de diagnóstico predictivo con Inteligencia Artificial.

**Problema:** Los radiólogos deben analizar grandes volúmenes de imágenes médicas en poco tiempo, lo que dificulta la detección temprana de nódulos pulmonares. Este escenario genera riesgo de diagnósticos tardíos y sobrecarga laboral, lo que hace necesario contar con herramientas tecnológicas que apoyen el proceso clínico e integren Inteligencia Artificial con los sistemas ya existentes en los hospitales.

**Solución propuesta:** Se desarrollará un sistema integrado a un PACS simulado (Orthanc) a través del estándar **DICOMweb**, capaz de procesar estudios de tomografía computarizada. El sistema incluirá un **servicio de inferencia basado en redes neuronales convolucionales (CNN)** para identificar posibles nódulos pulmonares y calcular probabilidades de malignidad. Los resultados se mostrarán en una **interfaz web con superposición gráfica (overlays)** que permita resaltar las regiones de interés. Además, se implementará una **base de datos en PostgreSQL** para almacenar los resultados, métricas y auditoría de los análisis.

**Impacto laboral:** Este proyecto fortalece competencias del área de salud digital, inteligencia artificial aplicada y desarrollo full stack, además de promover el uso de estándares de interoperabilidad como **DICOM**. También aporta valor en el campo profesional, al demostrar cómo la informática puede apoyar la toma de decisiones médicas, mejorar la eficiencia y generar soluciones innovadoras con impacto social.

## 3. Relación con Competencias del Perfil de Egreso

El proyecto **PACS-AI Assist** se vincula directamente con las competencias del perfil de egreso de Ingeniería en Informática, ya que integra conocimientos técnicos, de gestión y de aseguramiento de calidad en un contexto real.

* **Desarrollo de software e integración modular (backend y frontend):** se implementan microservicios en Python (FastAPI) para la inferencia de IA y una interfaz web en React que permite la visualización de los resultados con superposición de regiones detectadas.
* **Modelado e implementación de datos (PostgreSQL para resultados):** se diseña e implementa una base de datos relacional para almacenar diagnósticos, métricas y logs, asegurando trazabilidad y escalabilidad.
* **Gestión de proyectos TI (planificación incremental por releases):** se utilizan metodologías ágiles (Scrum ligero), dividiendo el proyecto en iteraciones que permiten entregar valor progresivamente y gestionar riesgos de forma oportuna.
* **Aseguramiento de calidad y pruebas:** se aplican pruebas unitarias e integrales para garantizar la confiabilidad del modelo CNN y la correcta comunicación entre los servicios.
* **Infraestructura y despliegue en entornos cloud simulados:** se configura un entorno de ejecución basado en contenedores y se simula la integración con un PACS mediante Orthanc, lo que permite validar la factibilidad del sistema en un entorno cercano a la realidad hospitalaria.

En conjunto, estas competencias reflejan cómo el proyecto consolida las capacidades profesionales del perfil de egreso, al aplicar conocimientos técnicos y de gestión en el desarrollo de una solución informática innovadora y de alto impacto en el sector salud.

## 4. Relación con Intereses Profesionales

El proyecto **PACS-AI Assist** se alinea directamente con mis principales intereses profesionales en el área de **inteligencia artificial aplicada**, el **desarrollo de soluciones full stack en la nube** y la integración de sistemas con estándares internacionales.

Este proyecto no solo me permite profundizar en técnicas de **machine learning y procesamiento de imágenes médicas**, sino que también refuerza mi capacidad para **diseñar arquitecturas de software escalables e interoperables**.

Además, fortalece mi perfil en el ámbito de la **salud digital (health-tech)**, un sector que actualmente experimenta un crecimiento sostenido y demanda profesionales capaces de unir el conocimiento técnico con la comprensión de los procesos clínicos. En este sentido, trabajar en un proyecto que integra IA con el estándar **DICOM** me acerca a las competencias necesarias para participar en equipos interdisciplinarios de salud e informática.

Finalmente, este proyecto representa un paso clave hacia mi proyección laboral como **desarrolladora especializada en soluciones tecnológicas de alto impacto social**, donde la combinación de **IA, cloud computing y análisis de datos** sea central para aportar valor en contextos reales.

## 5. Factibilidad

El proyecto **PACS-AI Assist** es factible de realizar en el marco de la asignatura, ya que se apoya en recursos accesibles y metodologías que permiten acotar su alcance. El uso de **Orthanc PACS** como entorno simulado y de **datasets DICOM públicos** posibilita trabajar con datos reales sin requerir acceso a sistemas hospitalarios productivos, lo que reduce barreras éticas y técnicas.

El alcance del proyecto se encuentra limitado a la construcción de un **MVP (Producto Mínimo Viable)** en un plazo de **10 a 12 semanas**, lo cual se ajusta al tiempo disponible en la asignatura. Este MVP contempla las etapas esenciales: integración con PACS, desarrollo del modelo de IA y presentación de resultados en la interfaz.

En cuanto a los **riesgos identificados**, se consideran principalmente el **alto volumen de datos** y el **rendimiento del modelo CNN**. Para mitigarlos, se planifica el uso de datasets reducidos en las primeras etapas, pruebas incrementales de desempeño y una arquitectura modular que permite escalar sin comprometer la funcionalidad del prototipo.

De esta manera, el proyecto equilibra **relevancia técnica, impacto social y viabilidad académica**, asegurando que pueda desarrollarse con éxito dentro del tiempo y recursos asignados.

## 6. Objetivos (SMART)

1. **Integrar Orthanc con un servicio de inferencia IA** en un plazo máximo de **4 semanas**, asegurando la correcta comunicación vía DICOMweb y validando el envío y recepción de imágenes médicas.  
   * Específico: conexión entre PACS simulado e IA.
   * Medible: pruebas de envío/recepción de estudios exitosas.
   * Alcanzable: uso de Orthanc y librerías disponibles.
   * Relevante: primer paso para habilitar la funcionalidad del sistema.
   * Tiempo: 4 semanas.
2. **Entrenar un modelo CNN** capaz de alcanzar un desempeño mínimo de **≥80% AUC** en detección de nódulos pulmonares, dentro de las **6 semanas** iniciales del proyecto.  
   * Específico: entrenamiento de modelo CNN con datasets públicos.
   * Medible: métricas de validación (AUC, sensibilidad, especificidad).
   * Alcanzable: acotado a datasets reducidos para MVP.
   * Relevante: núcleo del sistema de apoyo clínico.
   * Tiempo: 6 semanas.
3. **Implementar una interfaz web** que muestre overlays con las regiones sospechosas y la probabilidad de malignidad, logrando un prototipo navegable en **8 semanas**.  
   * Específico: desarrollo de frontend conectado al backend IA.
   * Medible: visualización correcta de imágenes con resaltados.
   * Alcanzable: uso de frameworks como React.
   * Relevante: facilita la interacción con el usuario final (radiólogo).
   * Tiempo: 8 semanas.
4. **Desplegar API, frontend y base de datos** en un entorno cloud simulado a la **semana 10**, garantizando su disponibilidad en un prototipo funcional.  
   * Específico: despliegue completo de la solución.
   * Medible: endpoints activos y frontend accesible.
   * Alcanzable: uso de Docker y servicios gratuitos o simulados.
   * Relevante: valida factibilidad en un entorno cercano a la práctica real.
   * Tiempo: 10 semanas.
5. **Entregar un informe técnico final** en la **semana 11**, documentando la arquitectura, resultados del modelo, métricas de desempeño, pruebas realizadas y consideraciones éticas.  
   * Específico: informe estructurado en formato académico.
   * Medible: incluye secciones exigidas por la asignatura.
   * Alcanzable: consolidación de entregables previos.
   * Relevante: evidencia formal del aprendizaje y resultados.
   * Tiempo: 11 semanas.

## 7. Metodología

Para el desarrollo del proyecto **PACS-AI Assist** se combinarán metodologías de **gestión ágil** y de **procesamiento de datos**, con el fin de asegurar un avance ordenado y medible en cada etapa.

* **CRISP-DM:** se utilizará como marco metodológico para el ciclo de datos, abarcando las fases de entendimiento del negocio, comprensión de los datos, preparación, modelado, evaluación y despliegue.
* **Scrum ligero:** se aplicará para la gestión ágil del proyecto, dividiendo el trabajo en sprints de 2 semanas con entregables incrementales (procesamiento de imágenes, integración IA, visualización y despliegue).

**Buenas prácticas de desarrollo:**

* Uso de control de versiones en **GitHub** con ramas, commits y revisiones de código.
* Organización del trabajo mediante tableros en **JIRA o Kanban**, facilitando el seguimiento de tareas.
* Integración y entrega continua (**CI/CD básico**) para mantener la calidad del código y detectar errores tempranamente.
* Documentación clara de **APIs** y componentes para garantizar la trazabilidad y la reproducibilidad del sistema.

**Evaluación de modelos:**

* Se evaluará el desempeño de la CNN con métricas estandarizadas como **matriz de confusión, AUC, sensibilidad y especificidad**, permitiendo medir tanto la precisión global como la capacidad del modelo para detectar casos positivos.

**Consideraciones éticas y de seguridad:**

* Se aplicará **anonimización de los datos médicos** utilizados en las pruebas.
* Se incluirán **disclaimers éticos**, aclarando que el sistema es un apoyo al radiólogo y no reemplaza la decisión clínica.

## 8. Plan de Trabajo

El proyecto **PACS-AI Assist** se desarrollará en un período de 12 semanas, dividiendo las actividades en fases concretas que permiten avanzar de manera incremental y controlada.

**Cronograma de actividades:**

* **Semanas 1–2:** Configuración del entorno simulado con **Orthanc PACS** y revisión de **datasets DICOM públicos**.
* **Semanas 3–4:** Preprocesamiento de imágenes médicas y almacenamiento estructurado en **PostgreSQL**, asegurando trazabilidad.
* **Semanas 5–6:** Entrenamiento de un modelo **CNN** y comparación con un modelo baseline, evaluando métricas de desempeño (AUC, sensibilidad, especificidad).
* **Semanas 7–8:** Desarrollo del **MVP de la API** y del **frontend web**, incluyendo overlays para mostrar las regiones de interés.
* **Semana 9:** Ejecución de pruebas de calidad (QA), validación de resultados y ajustes en los modelos y la interfaz.
* **Semanas 10–11:** Despliegue del sistema completo (API, frontend y base de datos) en un entorno cloud simulado, acompañado de una demo funcional.
* **Semana 12:** Ajustes finales en base a retroalimentación docente y preparación del informe técnico.

**Recursos necesarios:**

* Datasets abiertos (LUNA16, NIH Chest CT).
* Software: Python, TensorFlow/Keras, FastAPI, React, PostgreSQL, Docker.
* Herramientas de gestión: GitHub, JIRA/Kanban.

**Facilitadores:**

* Acceso a datasets públicos.
* Experiencia previa en IA y desarrollo web.
* Trabajo incremental con entregables parciales.

**Obstaculizadores:**

* Complejidad en el preprocesamiento de imágenes médicas.
* Limitaciones de tiempo para el entrenamiento de modelos complejos.
* Posible sobrecarga de recursos en despliegue cloud gratuito.

Este plan asegura un **avance progresivo y verificable**, equilibrando los objetivos técnicos con la disponibilidad de recursos y el tiempo establecido en la asignatura.

## 9. Evidencias

Las siguientes evidencias permitirán demostrar el logro de las actividades y objetivos del proyecto **PACS-AI Assist**:

* **Repositorio Git con commits e issues:** evidenciará el proceso de desarrollo, las ramas utilizadas, la trazabilidad de cambios y la gestión de tareas.
* **Pipelines de preprocesamiento y dataset documentado:** mostrarán la preparación de las imágenes médicas, asegurando reproducibilidad y claridad en el manejo de los datos.
* **Modelo CNN con métricas reproducibles:** permitirá validar el rendimiento del sistema mediante indicadores como AUC, sensibilidad y especificidad, demostrando su aporte técnico.
* **API FastAPI documentada:** garantizará la interoperabilidad del sistema, mostrando endpoints funcionales y documentación clara con OpenAPI/Swagger.
* **Interfaz web con overlays:** evidenciará la integración entre IA y la visualización clínica, mostrando regiones de interés resaltadas y probabilidades de malignidad.
* **Informe técnico y presentación:** consolidarán los resultados obtenidos, la arquitectura implementada, los riesgos abordados y las proyecciones del proyecto.

Estas evidencias, en conjunto, aseguran que el proyecto cumpla con los **criterios de calidad, transparencia y aplicabilidad profesional**, además de permitir una retroalimentación efectiva en cada fase.

## 10. Conclusions (English)

The **PACS-AI Assist** project confirms the feasibility of integrating Artificial Intelligence into medical imaging workflows using DICOM standards. By applying a CNN model, the system can provide reliable support to radiologists in the early detection of pulmonary nodules. However, transparency, rigorous validation, and ethical disclaimers remain essential to ensure trust and responsible use.

Furthermore, the modular architecture enables continuous improvement, scalability, and adaptation to real-world hospital PACS environments. This demonstrates the project’s potential not only as an academic exercise but also as a practical foundation for future innovations in health technology.

## 11. Reflection (English)

Working on **PACS-AI Assist** significantly strengthened my ability to integrate AI models with interoperable healthcare systems. I gained a deeper understanding of the importance of medical data standards such as **DICOM**, and I became more aware of the complexity of clinical validation processes.

This experience also enhanced my skills in backend and frontend integration, AI deployment, and agile project management. At the same time, it revealed areas for further growth, such as advanced **MLOps practices**, clinical data governance, and effective collaboration with healthcare professionals.

Overall, the project allowed me to connect my technical knowledge with real-world applications, preparing me for future professional challenges in the intersection of **informatics, AI, and healthcare**.