|  |
| --- |
| **Diagnóstico Temprano de Cáncer de Pulmón**  **‘‘Lung GuardIA’’** |
| Capstone |
|  |
| **Profesor:** Antonio Gallardo |
| **Alumnos:** David P. & Yoselin C. |
|  |

**Índice**

[**I. Descripción del Proyecto APT 3**](#_heading=h.ophwd13y0oci)

[**II. Relación con las competencias del perfil de egreso 3**](#_heading=h.r9x6d5je0tc0)

[**III. Relación con intereses profesionales 4**](#_heading=h.h12k1dh70htr)

[**IV. Factibilidad del proyecto 5**](#_heading=h.rldvq7nd4h1s)

[**V. Objetivos 6**](#_heading=h.3n293lq45ei2)

[**VI. Propuesta metodológica 7**](#_heading=h.rrxfdzjmi35c)

[**VII. Abstract 8**](#_heading=h.nhd4tpomhjab)

# Descripción del Proyecto APT

Lung GuardIA tiene como objetivo desarrollar un sistema predictivo que permita el diagnóstico temprano de cáncer de pulmón mediante el uso de algoritmos de machine learning. Este sistema se diseñará específicamente para su implementación en hospitales, facilitando a los médicos oncólogos y neumólogos la toma de decisiones clínicas más rápidas y precisas.

La relevancia de este proyecto en el campo laboral de la informática es significativa, ya que combina dos áreas cruciales: la salud y la tecnología. Al aplicar inteligencia artificial en la medicina, se busca mejorar la precisión en los diagnósticos de cáncer de pulmón, una de las enfermedades más mortales en Chile y el mundo. Además, el proyecto apunta a optimizar los recursos hospitalarios, reduciendo el tiempo dedicado a diagnósticos manuales y permitiendo un enfoque preventivo que puede salvar vidas.

Este enfoque tecnológico es fundamental en un entorno donde la transformación digital en salud es una tendencia creciente, lo que hace que este proyecto tenga un impacto positivo en la industria y proporcione una solución a un problema crítico.

# Relación con las competencias del perfil de egreso

Lung GuardIA se relaciona directamente con diversas competencias del perfil de egreso de la carrera de Ingeniería en Informática, particularmente en áreas clave como la gestión de proyectos, pruebas de certificación y el desarrollo de soluciones de software:

Pruebas de certificación: El proyecto contempla la validación de modelos predictivos utilizando buenas prácticas de la industria, como la validación cruzada y la evaluación mediante métricas estándar como precisión, sensibilidad y especificidad. Esta competencia es fundamental para garantizar que el sistema predictivo funcione de manera confiable en entornos clínicos.

Gestión de proyectos informáticos: A lo largo del desarrollo del proyecto, se aplicarán principios de planificación y control, siguiendo una metodología ágil y el marco CRISP-DM. El secuenciador y el Canvas MVP guiarán el desarrollo del sistema, asegurando que las decisiones técnicas y organizativas estén alineadas con los objetivos del hospital.

Modelado de datos: La creación de modelos predictivos escalables es uno de los componentes esenciales del proyecto. Los datos médicos serán utilizados para construir un modelo que se adapte a las necesidades de los hospitales, permitiendo predicciones precisas en un entorno que puede evolucionar con el tiempo.

Desarrollo de software: El sistema será desarrollado utilizando técnicas de integración continua y despliegue en entornos de producción, asegurando que el sistema pueda ser fácilmente implantado en un hospital y utilizado por el personal médico. Esto incluye el uso de tecnologías como Flask, Python y herramientas de machine learning para asegurar la calidad y escalabilidad del software.

Estas competencias no solo son necesarias para el desarrollo de Lung GuardIA, sino que también preparan para enfrentar desafíos futuros en el ámbito profesional, donde la integración de tecnología en sectores como la salud es cada vez más crítica.

# Relación con intereses profesionales

Lung GuardIA está estrechamente vinculado a los intereses profesionales en el campo de machine learning, inteligencia artificial y data science. La aplicación de tecnologías avanzadas para abordar problemas complejos en el sector salud, específicamente en la detección temprana de enfermedades como el cáncer de pulmón, permite explorar y desarrollar modelos predictivos, una área de gran relevancia profesional.

Este proyecto ofrece la oportunidad de adquirir y consolidar competencias en el desarrollo de soluciones tecnológicas innovadoras, donde la inteligencia artificial juega un papel crucial. Además, permite contribuir al ámbito médico, optimizando la precisión de los diagnósticos y reduciendo el tiempo necesario para realizar estos, generando un impacto positivo tanto en la eficiencia hospitalaria como en el bienestar de los pacientes.

El interés también radica en el potencial del proyecto para avanzar en la intersección entre la informática y la medicina, un sector en constante crecimiento que demanda profesionales con conocimientos en el desarrollo y gestión de soluciones basadas en datos.

# Factibilidad del proyecto

Lung GuardIA es factible de realizarse dentro del marco de la asignatura debido a varios factores que respaldan su desarrollo:

1. **Recursos disponibles:** Se cuenta con acceso a datos clínicos relevantes y a tecnologías de machine learning, como bibliotecas de Python (por ejemplo, Scikit-learn y TensorFlow), que son fundamentales para el desarrollo del sistema predictivo. Además, el entorno de trabajo incluye herramientas de desarrollo como Flask para la creación de aplicaciones web.
2. **Metodología clara**: Se aplicará la metodología CRISP-DM, que ofrece un marco estructurado para la gestión de proyectos de data science. Esta metodología facilitará la planificación y ejecución del proyecto, asegurando que se sigan las etapas necesarias desde la comprensión del negocio hasta el despliegue del sistema.
3. **Planificación temporal:** El proyecto se desarrollará en un plazo de tres meses, dividido en fases bien definidas: investigación y diseño en el primer mes, desarrollo y pruebas iniciales en el segundo mes, y finalmente la integración y pruebas piloto en el tercer mes. Esta planificación asegura que se dispondrá del tiempo necesario para abordar cada fase adecuadamente.
4. **Colaboración con instituciones:** El proyecto se llevará a cabo en colaboración con un hospital, lo que proporciona un entorno real para la implementación y validación del sistema. La participación de profesionales médicos en el proceso asegurará que el sistema cumpla con las necesidades prácticas y que se reciba retroalimentación valiosa.
5. **Identificación de riesgos y estrategias de mitigación:** Se han identificado posibles desafíos, como la calidad de los datos disponibles y la resistencia al cambio por parte del personal médico. Para abordar estos aspectos, se implementarán estrategias como la capacitación del personal en el uso del sistema y la realización de pruebas piloto para evaluar y ajustar el sistema antes de su implementación completa.

La combinación de estos factores demuestra que el proyecto Lung GuardIA es viable y está bien posicionado para ser desarrollado de manera efectiva dentro del marco académico establecido.

# Objetivos

Lung GuardIA tiene como finalidad desarrollar un sistema predictivo para el diagnóstico temprano de cáncer de pulmón. Para alcanzar esta meta, se establecen los siguientes objetivos claros y coherentes:

1. **Desarrollar un sistema predictivo:** Crear un modelo de machine learning que analice datos clínicos para predecir el riesgo de cáncer de pulmón en pacientes. Este sistema deberá ser capaz de identificar patrones en los datos que permitan detectar la enfermedad en etapas tempranas.
2. **Integrar el sistema con plataformas hospitalarias:** Asegurar que el sistema se integre de manera fluida con los sistemas existentes en el hospital, permitiendo a los médicos acceder fácilmente a las predicciones y recomendaciones generadas por el modelo.
3. **Mejorar la precisión del diagnóstico:** Establecer que el sistema alcance al menos un 80% de precisión en las predicciones, lo que contribuirá a la reducción de falsos positivos y negativos en los diagnósticos de cáncer de pulmón.
4. **Reducir el tiempo de diagnóstico:** Lograr que el tiempo requerido para diagnosticar cáncer de pulmón se reduzca en un 20% en comparación con los métodos tradicionales, permitiendo una atención más rápida y efectiva para los pacientes.
5. **Capacitar al personal médico:** Desarrollar un programa de capacitación para médicos y personal administrativo, asegurando que estén familiarizados con el uso del sistema predictivo y comprendan cómo interpretar los resultados de manera efectiva.

Estos objetivos no solo son coherentes con la disciplina de la ingeniería informática, sino que también abordan una situación crítica en el sector salud, contribuyendo a mejorar la calidad de atención médica en el diagnóstico de cáncer de pulmón.

# Propuesta metodológica

La metodología del proyecto Lung GuardIA se basa en el marco CRISP-DM (Cross-Industry Standard Process for Data Mining), que es un enfoque estructurado para el desarrollo de proyectos de minería de datos y machine learning. Esta metodología se divide en varias etapas que garantizan un desarrollo sistemático y eficiente del sistema predictivo:

1. **Comprensión del negocio:**

* Identificar las necesidades y requerimientos del hospital en relación con el diagnóstico de cáncer de pulmón.
* Reunir información sobre los flujos de trabajo existentes y cómo el nuevo sistema puede integrarse para mejorar la eficiencia.

1. **Comprensión de los datos:**

* Recolectar datos clínicos de pacientes, incluyendo historial médico, factores de riesgo, y resultados de exámenes previos.
* Evaluar la calidad de los datos disponibles y realizar una exploración inicial para identificar patrones y anomalías.

1. **Preparación de los datos:**

* Limpiar y transformar los datos recopilados, eliminando entradas duplicadas o erróneas y normalizando las variables relevantes para el modelo.
* Dividir los datos en conjuntos de entrenamiento y prueba, asegurando que el modelo se entrene y evalúe de manera efectiva.

1. **Modelado:**

* Seleccionar y aplicar algoritmos de machine learning adecuados, como regresión logística, árboles de decisión o redes neuronales, para desarrollar el modelo predictivo.
* Ajustar los parámetros del modelo y realizar validaciones cruzadas para optimizar su rendimiento.

1. **Evaluación**:

* Validar el modelo utilizando el conjunto de datos de prueba, analizando métricas como precisión, sensibilidad, y especificidad para evaluar su efectividad en la predicción del riesgo de cáncer de pulmón.
* Revisar los resultados en colaboración con médicos del hospital para asegurar que el modelo sea clínicamente relevante y aplicable.

1. **Despliegue:**

* Implementar el sistema en el entorno hospitalario, integrándolo con los sistemas existentes y asegurando que los médicos tengan acceso a las predicciones y recomendaciones del modelo.
* Establecer un programa de monitoreo continuo para evaluar el rendimiento del sistema en el tiempo y realizar ajustes según sea necesario.

# Abstract

El proyecto **Lung GuardIA** tiene como objetivo desarrollar un sistema predictivo para el diagnóstico temprano de cáncer de pulmón utilizando algoritmos de machine learning. Dado que el cáncer de pulmón es una de las principales causas de mortalidad en Chile y el mundo, este sistema se diseñará específicamente para su implementación en hospitales, facilitando a los médicos oncólogos y neumólogos la toma de decisiones clínicas más rápidas y precisas.

A través de la aplicación de inteligencia artificial, se busca mejorar la precisión en los diagnósticos y optimizar los recursos hospitalarios, reduciendo el tiempo dedicado a diagnósticos manuales y promoviendo un enfoque preventivo que puede salvar vidas.

El desarrollo del sistema se basará en el marco metodológico CRISP-DM, asegurando una planificación estructurada que abarque desde la comprensión de las necesidades del hospital hasta la integración del sistema en el entorno clínico.

Los resultados esperados incluyen una mejora en la precisión diagnóstica y una reducción significativa en el tiempo de diagnóstico, así como la capacitación del personal médico en el uso del sistema. Este proyecto representa una contribución significativa a la intersección entre la salud y la tecnología, abordando un problema crítico en el sector salud.