



## **Tecnicatura Superior en Ciencia de Datos e Inteligencia Artificial**

**Materia: Aprendizaje Automático**

**Examen parcial Octubre 2025**

**Entrega 1: Descripción y Formulación del Objetivo**

**Proyecto: Auditoría de Prestaciones Extrahospitalarias con  
Aprendizaje Automático**

**Estudiante: Nancy Julieta Cassano**



## 1. Descripción del proyecto y contexto

El presente proyecto se desarrolla en el marco del Ministerio de Salud de la Provincia de Tierra del Fuego, Antártida e Islas del Atlántico Sur, específicamente en la Dirección de Prestaciones Médicas y Programas Sociales, área responsable de autorizar o rechazar solicitudes de cobertura para prestaciones extrahospitalarias, entre otras, entendidas como aquellas prácticas médicas que no pueden resolverse dentro del sistema público provincial pero sí se encuentran disponibles en el ámbito privado local.

Cada solicitud es evaluada por el equipo de auditoría administrativa, social y médica, considerando factores como la cobertura social del paciente, la situación socioeconómica, la disponibilidad de la práctica en el sistema público y la completitud de la documentación presentada.

El proceso actual se realiza de forma manual, dependiendo casi por completo del criterio individual de cada auditor y de la disponibilidad de tiempo, lo que puede generar demoras, variabilidad en los criterios de evaluación y mayor carga administrativa.

El objetivo del proyecto es incorporar herramientas de aprendizaje automático supervisado que permitan predecir si una solicitud de prestación extrahospitalaria corresponde ser autorizada o no, para asistir a los equipos de auditoría en la toma de decisiones y contribuir a la eficiencia, estandarización y transparencia del proceso.

## 2. Formulación del problema

El problema se plantea como una tarea de clasificación supervisada binaria, en la cual el modelo debe predecir la variable objetivo **autorización** con dos posibles valores:

**1 (Autorizar):** cuando la solicitud cumple los criterios establecidos por la normativa vigente.

**0 (No autorizar):** cuando la solicitud presenta faltantes, incumple condiciones o no corresponde a una prestación cubierta.

El modelo será entrenado sobre un dataset sintético de 500 registros, generado a partir de las reglas de negocio reales de la Dirección de Prestaciones Médicas.

El conjunto de variables incluirá, entre otras:

- Cobertura social (sí/no)
- Completitud de la documentación
- Lugar de residencia
- Disponibilidad de la práctica en el sistema público

- Urgencia médica
- Edad del paciente
- Nivel socioeconómico
- Diagnóstico codificado (CIE-10)
- Resultado final (autorizado / no autorizado)

### **3. Objetivo general**

Desarrollar un modelo de aprendizaje automático supervisado que prediga la decisión de autorización o no autorización de una prestación médica extrahospitalaria, en base a variables administrativas, médicas y socioeconómicas, para asistir el proceso de auditoría de la Dirección de Prestaciones Médicas y Programas Sociales del Ministerio de Salud de Tierra del Fuego.

### **4. Objetivos específicos**

- Analizar y depurar el dataset para garantizar la calidad de los datos y su adecuación al problema de clasificación.
- Explorar las relaciones entre variables y su influencia sobre la decisión de autorización.
- Entrenar y comparar dos modelos supervisados: Regresión Logística y Árbol de Decisión.
- Evaluar los modelos utilizando métricas como exactitud, precisión, recall, F1-score y matriz de confusión.
- Seleccionar el modelo con mejor equilibrio entre interpretabilidad y desempeño, considerando su aplicabilidad institucional.
- Generar visualizaciones y reportes que faciliten la comprensión de los resultados por parte del equipo auditor.

### **5. Tipo de problema**

El proyecto corresponde a un problema de clasificación binaria supervisada, ya que la variable objetivo es categórica (autorizar / no autorizar) y el modelo se entrena con ejemplos previamente etiquetados.

El aprendizaje se basa en inferir patrones y relaciones entre las variables de entrada y la decisión final de autorización.

## 6. Modelos propuestos

Se seleccionan dos modelos de aprendizaje automático vistos en la materia:

- **Regresión Logística:** modelo estadístico interpretable, útil para estimar la probabilidad de autorización y analizar el peso de cada variable en la decisión.
- **Árbol de Decisión:** modelo no paramétrico que permite representar gráficamente las reglas de decisión, lo que facilita su comprensión y aplicación por parte del equipo auditor.

Ambos modelos serán entrenados y evaluados sobre el mismo conjunto de datos, con el fin de comparar su desempeño y justificar la elección del modelo más adecuado para futuras implementaciones institucionales.

## 7. Relevancia del estudio

La aplicación de técnicas de aprendizaje automático en la auditoría de prestaciones médicas puede aportar beneficios significativos:

- Estandarización de los criterios de evaluación, reduciendo la variabilidad entre auditores.
- Priorización automatizada de casos con mayor probabilidad de autorización, optimizando los tiempos de respuesta.
- Transparencia y trazabilidad de las decisiones públicas en el ámbito de la salud.
- Aprovechamiento del conocimiento institucional, transformando datos históricos en herramientas de apoyo a la gestión.

El proyecto representa un caso concreto de integración entre Ciencia de Datos e Inteligencia Artificial aplicada a la salud pública, alineado con las recomendaciones de la Agencia de Acceso a la Información Pública (AAIP) para una inteligencia artificial responsable, explicable y éticamente fundamentada.