

■ Agente IA - Seguro de Accidentes Personales

Autores: **Solange Areco y Milagros Lencina**

Trabajo Práctico - Challenge One

Fecha: Septiembre 2025

1. Introducción

Este proyecto implementa un agente de inteligencia artificial que evalúa la asegurabilidad de personas para seguros de accidentes personales. El sistema combina un modelo de Machine Learning con reglas de negocio determinísticas y permite calcular una prima estimada según tablas locales de Argentina.

2. Datasets utilizados

- ENNyS2 (Encuesta Nacional de Nutrición y Salud, Argentina 2018-2019) - datos.gob.ar - Licencia Pública.
- Kaggle Insurance Dataset - Kaggle.com - Licencia CC0 (uso libre).
- Dataset sintético generado con Pandas/Numpy para ampliar diversidad de casos.

3. Pipeline del modelo

1. Ingestión: carga robusta de CSV (UTF-8, latin-1, cp1252). 2. Limpieza y features: edad, peso, talla, IMC, tabaquismo, enfermedades. 3. Entrenamiento: RandomForestClassifier con balance de clases. 4. Evaluación: métricas de accuracy, ROC AUC y matriz de confusión. 5. Exportación: modelo guardado con joblib para reutilización en Streamlit.

4. Reglas de negocio aplicadas

- Edad ≥ 55 y fumador \rightarrow No asegurable.
- Informe PDF con 'NO APTO' \rightarrow No asegurable.
- Dos o más enfermedades graves (diabetes, hipertensión, renal, respiratoria) \rightarrow No asegurable.
- En caso de rechazo, se aplica penalización $\times 1.5$ en la prima final.

5. Métricas del modelo

El modelo RandomForest alcanza un accuracy ≈ 0.98 y un ROC AUC ≈ 0.999 en validación. La matriz de confusión y métricas detalladas están integradas en la interfaz de Streamlit.

6. Análisis ético y sesgos

El uso de variables como edad, sexo y enfermedades crónicas puede introducir sesgos en la decisión de asegurabilidad. Se recomienda monitorear fairness y aplicar métricas adicionales para evitar discriminación no justificada desde el punto de vista actuarial.

7. Conclusiones

El agente IA cumple con los objetivos planteados: predicción binaria de asegurabilidad, probabilidad de riesgo y cálculo de prima estimada en base a normativa argentina. Como pasos futuros, se sugiere ampliar las fuentes de datos, mejorar el análisis ético y desplegar el sistema en la nube para pruebas con usuarios reales.