



DETECCIÓN DE FRAUDE DE PROVEEDORES MÉDICOS

Introducción:

- Entre el 5% y el 10% de las reclamaciones de seguros son fraudulentas.
- Muchas de estas no provienen de clientes, sino de los proveedores médicos.
- Detectar fraudes es complejo, costoso y requiere evidencias claras para iniciar investigaciones.

Objetivo: Desarrollar un modelo que priorice la investigación de proveedores sospechosos, brindando un puntaje de alerta y decisiones explicables para facilitar la acción.

Preparación de Datos:

Se analizaron datos de beneficiarios, hospitalizaciones, consultas externas y reembolsos.

Se realizaron transformaciones clave como:

- Conversión de fechas para calcular duraciones.
- Limpieza e imputación de datos faltantes.
- Balanceo de clases para mejorar la representación de casos fraudulentos.

Reglas Manuales para Identificar Fraude:

- Frecuencia de Tratamientos: Detecta sobretratamientos por proveedor.
- Reembolsos Totales por Paciente: Señala reclamaciones infladas.
- Distribución de Condiciones Crónicas: Evalúa tratamientos costosos asociados a ciertas condiciones.

- Priorizar Recursos: Proporciona un enfoque dirigido para investigar casos.
- Decisiones Explicables: Ayuda a justificar investigaciones ante reguladores.
- Mejor Eficiencia: Integra conocimiento experto con inteligencia artificial para resultados más confiables.

Modelos Utilizados:

- Random Forest: Modelo base por su robustez y simplicidad.
- XGBoost: Para mayor precisión y mejor manejo de datos desbalanceados.

Resultados Iniciales:

- Random Forest: 84.49% de precisión.
- XGBoost: 87.31% de precisión.

Mejoras mediante Reentrenamiento:

- Las reglas manuales se integraron como nuevas características.
- Los modelos fueron reentrenados, logrando:

- Random Forest: 92.79% de precisión.
- XGBoost: 94.27% de precisión.

