

PROYECTO DE APRENDIZAJE AUTOMÁTICO

Tema: Detección temprana de situaciones de trata de personas (foco en Tierra del Fuego)

Resumen

Se propone un modelo de clasificación binaria para estimar la probabilidad de que una intervención corresponda a trata (1) o no (0). La motivación es priorizar recursos y reducir falsos negativos, por lo que la métrica foco es el Recall, complementado con F1 macro, ROC-AUC y PR-AUC. Se realizará un ETL reproducible (limpieza, imputación, manejo de alta cardinalidad y variables temporales), partición 80/20 estratificada y validación cruzada ($k=5$). Como línea base se usarán Regresión Logística y Árbol de Decisión, y se explorará Random Forest y/o Gradient Boosting. Se documentarán sesgos/limitaciones y un análisis de transferencia para Tierra del Fuego como evaluación adicional de generalización.

Contexto y relevancia

El proyecto aborda la detección temprana de posibles situaciones de trata de personas a partir de registros administrativos de la Oficina de Rescate y Orientación a las Víctimas en Argentina (período 2020–2023) [@ datarawoficina-rescate-orientaciones-202001-202308.csv.csv](#). El foco aplicado es Tierra del Fuego (TDF), una provincia con baja densidad poblacional, estacionalidad turística y logística insular, donde las intervenciones son menos frecuentes y, por ende, más difícil de modelar únicamente con datos locales. La relevancia social del problema es alta: identificar tempranamente casos probables permite priorizar recursos y mejorar la derivación de alertas hacia equipos técnicos, siempre respetando el resguardo y la anonimización de datos sensibles.

Formulación del problema y tipo

El problema se formula como una tarea de clasificación supervisada binaria: predecir si una intervención registrada corresponde a un caso de trata (1) o no (0). La variable objetivo es un indicador binario (es_trata), mientras que las variables predictoras incluyen atributos categóricos (por ejemplo, provincia, vía de ingreso de la consulta, tema/subtema, tipo de consultante) y

numéricos (por ejemplo, edad aparente). Para la evaluación se priorizará el recall (sensibilidad) con el objetivo de reducir falsos negativos.

Objetivo general y objetivos específicos

Objetivo general

Desarrollar un modelo de clasificación capaz de anticipar, con alta sensibilidad, si una intervención puede corresponder a trata de personas, a fin de priorizar la atención y derivación en contextos con recursos limitados (caso TDF).

Objetivos específicos

1. Realizar un análisis exploratorio (EDA) para comprender distribución de variables, ausentes y potenciales sesgos.
2. Definir y aplicar una estrategia de preprocesamiento: imputación de faltantes, codificación one-hot de categóricas y, cuando corresponda, estandarización.
3. Entrenar y comparar modelos base (Regresión Logística, Árbol de Decisión) y modelos de conjunto (Random Forest, Gradient Boosting).
4. Optimizar hiperparámetros con validación cruzada estratificada priorizando la métrica de recall, junto con precisión, F1 y ROC-AUC.
5. Evaluar la capacidad de transferencia del mejor modelo nacional sobre el subconjunto de TDF para medir desempeño en un escenario periférico y de baja frecuencia.
6. Documentar consideraciones éticas y de privacidad de datos (anonimización, uso responsable y limitaciones).

Dataset (resumen)

Para entrenamiento/validación se usa el conjunto nacional. TDF se reserva como test externo de transferencia.

Fuente principal	Registros administrativos de la Oficina de Rescate (2020–2023)
Cantidad de instancias (aprox.)	≈ 7.800 a nivel nacional; en TDF: 27.
Variable objetivo	es_trata (1 = trata; 0 = no trata).
Principales variables	provincia; tema/subtema; vía de ingreso; origen/ámbito; tipo y género del consultante; es_anonima; consultante_edad_aparente.
Tipo de datos	Mayormente categóricos (codificación one-hot) y algunos numéricos.

Dataset para entrenar el modelo

Se detalla el conjunto de datos utilizado para el entrenamiento del modelo, la construcción de la etiqueta y las decisiones de preparación que garantizan su calidad y reproducibilidad.

Campo	Tipo	Uso/Notas
es_trata	Binaria (0/1)	Etiqueta objetivo: 1 si tema='trata'; 0 en caso contrario.
subtema	Categórica (alta cardinalidad)	Agrupar categorías raras (otros_≤N) o aplicar target encoding con CV.
vía_ingreso	Categórica	Canal de ingreso (Línea 145, derivación, etc.).
provincia	Categórica	Efectos territoriales; útil para análisis de transferencia.
origen	Categórica (media/alta card.)	Ámbito/fuente de la intervención.
consultante_tipo	Categórica	Rol del consultante.
consultante_genero	Categórica	Tratar faltantes; usar 'desconocido'.
consultante_edad_aparente	Numérica	Validar rango; imputación por mediana si falta.
es_anonima	Categórica/Binaria	Normalizar (0/1).
fecha_ingreso → mes/año/temporada	Derivadas	Estacionalidad y contexto temporal.

- Normalización de texto en categóricas (minúsculas, recorte de espacios, sin tildes).
- Construcción explícita del target: es_trata = (tema == 'trata').
- Tratamiento de faltantes: categóricas → 'desconocido'; numéricas → mediana.
- Edad aparente ≤0 o >100 → NA antes de imputar por mediana.
- Control de alta cardinalidad en subtema (otros_≤N o target encoding con CV).
- Ingeniería de variables temporales: mes, año y temporada.

Modelos candidatos y criterios

Se propone iniciar con una línea base interpretable (Regresión Logística) y un modelo no lineal de reglas (Árbol de Decisión). Luego, se considerarán ensambles como Random Forest y Gradient Boosting. Se utilizará validación cruzada estratificada y ajuste de hiperparámetros priorizando recall, sin descuidar precisión, F1 y ROC-AUC.

Evaluación y partición

- Realizar split 80/20 estratificado sobre el conjunto nacional, excluyendo TDF para preservar el test externo de transferencia. Usar random_state fijo.
- Sobre el 80% aplicar validación cruzada estratificada ($k=5$) para seleccionar modelo e hiperparámetros, priorizando Recall.
- Reportar también F1 (macro), Precisión, ROC-AUC y (si es posible) PR-AUC.
- Informar la matriz de confusión del 20% de test nacional y el umbral usado (0.5 por defecto), evaluando su ajuste para mejorar recall/F1.
- *El umbral operativo se seleccionará por curva PR para maximizar Recall con F1 macro aceptable.*
- Considerar class_weight='balanced' o ajuste de umbral para reducir falsos negativos.

Cuadro de experimentos (para completar)

#	Modelo	Hiperparámetros (clave)	Recall	F1 (macro)	ROC-AUC	PR-AUC	Umbral
1							
2							
3							

Consideraciones éticas y limitaciones

El dominio implica datos sensibles: se trabajará con información agregada y/o debidamente anonimizada, evitando cualquier intento de reidentificación. El modelo es un apoyo a la gestión, no reemplaza el criterio profesional. Se registrarán limitaciones (desbalance de clases, variabilidad territorial, posibles errores de registro) y se comunicará la incertidumbre.

Análisis de transferencia a Tierra del Fuego

Se evaluará el mejor modelo nacional sobre el subconjunto de TDF ($n \approx 27$) como test externo de generalización. Esta evaluación no reemplaza el test nacional 20%; lo complementa.

Reproducibilidad (entorno y versiones)

Se fija random_state=42 para particiones y validación. Entorno previsto: Python 3.x; pandas 2.x; scikit-learn 1.x.

Anexo – Diccionario de campos

Campo	Descripción (resumen)	Tipo / Observaciones
tema	Motivo principal de la intervención	Texto; fuente del target es_trata
es_trata	Etiqueta objetivo (1=trata; 0=no trata)	Binaria; construida desde tema
subtema	Detalle específico del motivo	Categórica; alta cardinalidad
vía_ingreso	Canal de ingreso/derivación	Categórica; señal predictiva
provincia	Provincia del caso	Categórica; útil para transferencia
origen	Ámbito/fuente de la intervención	Categórica; media/alta cardinalidad
consultante_tipo	Rol del consultante	Categórica
consultante_genero	Género del consultante	Categórica; tratar faltantes (usar “desconocido”)
consultante_edad_a_parente	Edad estimada del consultante	Numérica; validar rangos e imputar si falta
es_anonima	Si la consulta es anónima	Binaria; normalizar 0/1
fecha_ingreso	Fecha de la intervención	Fecha; deriva mes/año/temporada
mes	Mes de ingreso (derivado)	Entero 1–12
año	Año de ingreso (derivado)	Entero
temporada	Clasificación estacional	Texto; definida por reglas (alta/baja)
nro_registro_intern o	Identificador interno	Eliminar para evitar fuga/ID

Conclusión

El proyecto propone un clasificador binario para priorizar la detección temprana de posibles casos de trata en Argentina, entrenado con datos nacionales 2020–2023 y con Tierra del Fuego reservado como test externo de transferencia. El ETL estandariza tipos, trata faltantes y controla la alta cardinalidad de subtema, incorporando variables temporales. La evaluación seguirá un split 80/20 estratificado y CV k=5, con foco en Recall para minimizar falsos negativos, complementado por F1 (macro), ROC-AUC y PR-AUC y ajuste de umbral. Los primeros análisis muestran señal en vía_ingreso, subtema y provincia, lo que sustenta la viabilidad del enfoque. Se reconocen límites (desbalance de clases, tamaño reducido en TDF y posibles sesgos de registro) y se asumen criterios de anonimización y uso responsable. En la siguiente etapa se compararán modelos (Regresión Logística, Árbol, Random Forest), se fijará el umbral operativo y se reportarán resultados en test nacional y transferencia a TDF.

