

Modelo de Machine Learning para predecir Alteración del Orden en el Sistema Penitenciario Argentino

Grupo M – Integrantes:

- Lucia Cortes,
- María Fernanda Farias,
- Favio Ruggieri,
- Alejandro Gomez,
- Sergio Salanitri,
- Karina Calvo

Fecha: 01/10/2025

Diplomatura en Ciencia de Datos y Análisis Avanzado

Índice:

1. Resumen Ejecutivo
2. Definición del Problema y Relevancia
3. Datos y Metodología
 - 3.1 Fuente del dataset
 - 3.2 Características de los datos
 - 3.3 Limpieza, tratamiento de valores faltantes e ingeniería de variables
 - 3.4 Conclusiones de informe EDA
4. Modelado y Evaluación
 - 4.1 Algoritmos probados y criterio de selección
 - 4.2 Métricas obtenidas
 - 4.3 Comparativa de modelos y justificación de la elección final.
5. Impacto en el Negocio
6. Conclusiones y Recomendaciones
7. Anexos
8. Bibliografía consultada

1. Resumen ejecutivo

Resumen Ejecutivo (1 página): Describir el problema a resolver, la solución propuesta, el impacto esperado en el negocio y los resultados clave, incluyendo métricas principales.

Problema: Actualmente en la práctica se aplican métodos de evaluación de riesgo penitenciario, que incluyen: historial de conducta en prisión, perfil criminológico, factores personales (edad, nivel educativo, vínculos familiares), observación directa del personal penitenciario.

Se determinan los perfiles de riesgo mediante evaluación objetiva de la información judicial, legajo, antecedentes criminológicos, informes interdisciplinarios y sistemas de clasificación interna. Esto permite protocolizar controles reforzados (alojamiento, visitas, comunicaciones, traslados) y coordinar acciones del Servicio Penitenciario Federal (SPF) para neutralizar la capacidad de daño y proteger el orden y la seguridad del establecimiento.

Propuesta: se propone un modelo de ML que estime el riesgo de alteración del orden (leve y grave) de personas privadas de libertad en penales de Argentina, esta herramienta basada en inteligencia artificial podría servir para priorizar intervenciones, optimizar la utilización de recursos de seguridad y tratamiento preventivo.

Público:

- Interno: autoridades del Servicio Penitenciario, analistas de datos.
- Externo: Ministerio de Justicia, áreas de DD.HH.

KPIs de impacto esperado:

- Precision $\geq 50\%$;
- Recall $\geq 50-70\%$;
- F1 score $\geq 60\%$;
- AUPRC $\geq 3 \times$ prevalencia;
- AUROC $\geq 0.80-0.85$, menos importante para evaluar el modelo.

Impacto esperado: se espera reducir incidentes y optimizar recursos.

Resultados clave obtenidos del modelo: LightGBM alcanzó F1 macro de 0.64, sensibilidad de 70% y precisión del 50% para la clase 2 .

2. Definición del Problema y Relevancia: Presentar un contexto breve y la motivación del proyecto, definir un objetivo concreto y medible, y explicar la importancia estratégica para el negocio

2. Definición del problema y caso de negocio:

Mediante resolución del Ministerio de Seguridad Argentino de desarrollo el “*Sistema Integral de Gestión para Personas Privadas de la Libertad de Alto Riesgo*”. Este Sistema busca aislar, controlar y gestionar a internos de alta peligrosidad, evitando que continúen con su actividad criminal desde prisión, reduzcan la seguridad de otros presos o corrompan al personal.

La incorporación al Sistema lo decide el SPF, tras un proceso previo y puede darse en tres momentos:

- 1) Al ingreso al SPF, si ya se constata el perfil de alto riesgo.
- 2) Durante la detención, al detectarse nueva información fundada.
- 3) Tras un incidente grave o ante la falta de respuesta a otras medidas previas.

Se consideran de *alto riesgo* los internos que presentan uno o varios de estos factores:

- Riesgo de fuga alto o muy alto
 - Riesgo comunitario, es decir, ejercer violencia o causar daño en caso de fuga, dirigir o participar en delitos desde prisión, corromper funcionarios con fines criminales, entorpecer investigaciones judiciales.
- Pertenencia o liderazgo en organizaciones criminales nacionales o transnacionales, grupos organizados complejos o bandas violentas.

Este proyecto propone utilizar el Modelo de ML desarrollado como herramienta para seleccionar reclusos con perfiles de alto riesgo basado en la predicción de alta probabilidad de provocar alteración del orden en el penal.

El objetivo del modelo es un F1 score mayor al 60% con una sensibilidad superior al 70% para la categoría 2 y mayor al 50% para la categoría 3.

El impacto esperado es una reducción significativa en incidentes y optimización del uso de recursos humanos, con foco en la priorización de internos con mayor riesgo (alineado con la estrategia institucional del Sistema Integral de Gestión Penitenciaria).

● **Datos y Metodología:** Detallar la fuente y características de los datos, incluyendo tamaño, variables clave y calidad, incluir conclusiones de informe EDA; describir los procesos aplicados como limpieza, tratamiento de valores faltantes e ingeniería de variables; y explicar la metodología seguida, como CRISP-DM u otra similar.

3. Datos y Metodología

3.1 Fuente del Dataset

Base pública de Estadísticas de Política Criminal – Ministerio de Justicia de la Nación (<https://www.argentina.gob.ar/justicia/derechoshumanos/estadisticas>)

3.2 Características de los datos

El tamaño del dataset original era 1.436.508 registros y 88 características; tras limpieza de datos quedaron 971672 filas y 36 columnas.

Diccionario mínimo:

Nombre de las variables del dataset:

```
[
'censo_anio',
'provincia_sneep_id',
'establecimiento_id',
'edad',
'genero_id',
'nacionalidad_id',
'estado_civil_id',
'nivel_instruccion_id',
'ultima_situacion_laboral_id',
'capacitacion_laboral_al_ingresar_id',
'situacion_legal_id',
'delito1_id',
'delito2_id',
'delito3_id',
'delito4_id',
'delito5_id',
'horas_trabajo_remunerado_id',
'participacion_programa_laboral',
'participacion_programa_educativo_id',
'participacion_actividades_deportivas',
'recibio_atencion_medica_ult_anio',
'recibio_visitas_ultimo_anio',
'participo_alteracion_orden_ult_anio_id',
'tipo_infraccion_disciplinaria_id',
'calificacion_conducta_id',
'tentativa_fugas_evasiones_id',
'tentativa_suicidio',
'fue_lesionado_id',
'duracion_condena_anios',
'tipo_condena',
'es_reincidente_id',
'tiene_periodo_progresividad_id',
'tuvo_salidas_transitorias_id',
'incorporado_reg_semi_libertad_id',
'participa_programa_pre_libertad',
'participa_programa_prision_discontinua_id',
'participa_programa_semi_detencion_id',
'tuvo_reduccion_pena_id']
```

3.3 Limpieza, tratamiento de valores faltantes e ingeniería de variables

- Agrupación de categorías (instrucción, nacionalidad, estado civil).
- Limpieza de duplicados.
- Eliminación de columnas descriptivas redundantes y otras que no reviste interés al problema de negocio.
- Tratamiento de valores nulos, se encontraron valores nulos en 7 variables (participacion_programa_laboral, tentativa_suicidio, participacion_actividades_deportivas, recibio_atencion_medica_ult_anio, recibio_visitas_ultimo_anio, tipo_condena, participa_programa_pre_libertad) se reemplazaron por valor 0, ya que en el archivo original los campos vacíos son categorizados como 0. Luego se decidió eliminar las variables tipo_condena y participa_programa_pre_libertad debido a que más del 50% de los datos eran campos vacíos.
- **Outliers:** las únicas variables cuantitativas son edad y duración de la condena, por eso no se aplicó tratamiento de outliers

- Eliminación de filas con categoría 0 (correspondiente a campos vacíos) en la columna **participo_alteracion_orden_ult_anio_id**, que va a ser nuestro **target**, esto representa un 5,3% de los datos. Quedando 1301434 filas y 36 columnas.
- Codificación de la variable target en: categorías 1= No participó, 2=Leve (Con daños, o sin heridos ni rehenes), 3=Grave (Con heridos o muertos, o con rehenes).

3.4 Conclusiones de informe EDA

Encontramos variables con alta correlación ($>0,85$), pero por ahora al no pensar utilizar un modelo de regresión logística y no tener problema de multidimensionalidad (base de datos grande), en un primer momento no se van a agrupar variables ni eliminarlas.

La prevalencia global de la categoría 2 fue de 13,7% y en la categoría 3 del 3,7%. La prevalencia de alteración del orden en cada año censado fue cambiando a lo largo de los años, se decidió trabajar con un subset más reciente de los datos (2010-2022) para trabajar con menos cantidad de datos y eliminar los primeros años en los cuales la categoría 2 del target supera el 20%.

La metodología seguida se enmarca en el estándar CRISP-DM, incluyendo fases de comprensión de negocio, comprensión de datos, preparación, modelado, evaluación y despliegue. Se identificaron 10 variables clave (edad, tiempo de condena, reincidencia, participación en programas, visitas recibidas, tipo de delito, historial disciplinario, pabellón de alojamiento, estudios cursados y situación procesal). Se reportaron limitaciones de calidad de datos como valores faltantes.

● **Modelado y Evaluación: Indicar los algoritmos probados y los criterios de selección, mostrar las métricas obtenidas (precisión, recall, F1-score, AUC, entre otras), y realizar una comparativa de modelos con justificación de la elección final.**

4. Modelado y Evaluación

4.1 Algoritmos probados y criterio de selección

Debido a que tenemos aproximadamente 1,3 millones de datos y 36 características, donde la mayoría son categóricas pensamos en un modelo de Random Forest, robusto, fácil de entrenar y maneja muchas características sin problemas. Para hacerlo más preciso utilizamos modelos más

modernos basados en gradient boosting: LightGBM (soporta variables categóricas de forma nativa, entrenamiento muy rápido) y CatBoost (diseñado para muchas variables categóricas, no necesita encoding manual, más lento que LightGBM pero a veces más estable en categorías con pocos registros).

Esquema de validación:

Para el entrenamiento se utilizó 2010-2018 (60% del dataset), para validación de 2019-2021 (30% del dataset) y para el testeo 2022 (10%)

Luego del procesamiento del dataset completo se divide por años de censo (temporal holdout), se entrena hasta 2018 y se valida desde 2019-2021 en forma Rolling window con múltiples validaciones consecutivas:

Fold 1: Train 2010–2018 → Val 2019

Fold 2: Train 2010–2019 → Val 2020

Fold 3: Train 2010–2020 → Val 2021

Ventajas del Holdout temporal: representa mejor la realidad de despliegue (siempre se predice hacia adelante), permite detectar estabilidad en el tiempo (muy importante en fenómenos sociales/judiciales que cambian con políticas, normativa o gestión penitenciaria).

Para la predicción se utilizan los datos de 2022, test, no utilizados ni para entrenamiento ni validación.

4.2 Métricas obtenidas

Modelo	Balanced Accuracy	F1 Macro	Precisión Clase 2	Recall Clase 2	F1 Clase 2	Precisión Clase 3	Recall Clase 3	F1 Clase 3
LightGBM	0.734	0.647	0.53	0.70	0.60	0.29	0.58	0.39
CatBoost	0.737	0.635	0.51	0.70	0.59	0.26	0.59	0.36

Ambas alcanzan una sensibilidad del 70% para categoría 2, lo que es crucial porque prioriza identificar internos con riesgo, con una precisión del 50%.

En categoría 3 (graves), los dos modelos logran sensibilidad mayor al 55%, aunque con baja precisión.

Ambos modelos tienen un área bajo la curva ROC mayor al 85% en todas las categorías.

En cuanto al AUC de las Curvas Precision-recall para LGBM y CatBoost fueron del 58-59% para categoría 2 y 39-42% para la categoría 3, respectivamente, siendo mayor a 4 veces la prevalencia de la categoría 2 y más de 10 veces la prevalencia para la categoría 3.

Precision@K (clases minoritarias):

Clase	LightGBM	CatBoost
2 (leve)	0.98 @1000	0.97 @1000
3 (grave)	0.60 @1000	0.60 @1000

Ambos logran alta precisión en los *top K* casos, lo que es útil si el objetivo es priorizar vigilancia en un subconjunto de internos.

4.3 Comparativa de modelos y justificación de la elección final.

LightGBM:

- Ligeramente mayor F1 macro (0.647), así como en las categorías 2 y 3
- Menor tiempo de entrenamiento.
- Mejor desempeño en clase 3 en precision@K.

CatBoost:

- Balanced accuracy ligeramente superior (0.737 vs 0.734).
- Más estable en recall de clase 2.
- Pero menor F1 macro y más costoso en cómputo.

Como elección final preferimos LightGBM, porque:

- Tiene mejor F1 macro (métrica clave dado el desbalance).
- Logra mejor precisión en clase 3 (críticamente importante para internos de alto riesgo).
- Ofrece entrenamiento más eficiente y escalable.

CatBoost sigue siendo útil como modelo alternativo de validación o en un esquema de ensemble.

● **Impacto en el Negocio:** Exponer los beneficios esperados, incluyendo ahorros, mejoras o automatización; identificar riesgos y limitaciones; y analizar la escalabilidad y los pasos futuros.

5. Impacto en el Negocio

Con la aplicación de este modelo de ML se pretende:

- Anticipar situaciones de gravedad (motines, intentos de fuga).
- Evaluar la incorporación de reclusos al *Sistema Integral de Gestión para Personas Privadas de la Libertad de Alto Riesgo*.
- Mejorar la planificación de recursos humanos (guardias, requisas, controles).
- Prevenir alteraciones del orden que pongan en riesgo la vida de internos y personal penitenciario.
- Ayuda a decidir qué internos están en condiciones de avanzar en el régimen progresivo de la pena (cerrado → semiabierto → abierto → libertad).
- Ubicar a cada preso en pabellones adecuados (de baja, media o alta conflictividad).
- Prevención y tratamiento. Al identificar internos con alto riesgo de conflictividad, se pueden aplicar programas específicos: tratamiento psicológico, mediación de conflictos, actividades educativas o laborales para canalizar conductas.
- Evitar abusos de disciplina indiscriminada, porque las sanciones se aplicarían en base a un perfil de riesgo objetivo y evaluado, no arbitrariamente.

El sistema combina seguridad física, procedimental y dinámica. Sus principales medidas son:

Área	Medidas principales
Alojamiento	Sectores especiales en cárceles de máxima seguridad o celdas individuales. Estrategia de concentración, dispersión parcial y separación. Evitar penales en zonas de influencia de sus grupos criminales.
Actividades	Sin contacto con población común. Visitas, trabajo, educación, religión y recreación dentro del sector. Rutinas diarias reguladas
Contacto con el exterior	Visitas limitadas, familiares directos, bajo controles electrónicos. Visitas a abogados en locutorio reservado. Comunicaciones telefónicas restringidas y Control de correspondencia

Traslados y salidas	Con autorización judicial y casos indispensables, rutas confidenciales. Custodia de Grupos Especiales del SPF y fuerzas externas. Uso de videoconferencia para audiencias. Salidas médicas con validación interdisciplinaria. Traslados de urgencia bajo protocolo de alto riesgo.
Seguridad física y electrónica	Cámaras inteligentes, registro de movimientos y recuentos. Controles biométricos de ingreso/egreso. Escáner corporal, detectores de metales, cercos energizados. Registros filmados de movimientos y recuentos. Modelo de círculos concéntricos de seguridad (múltiples barreras).
Seguridad procedimental	Registros de ingresos, egresos, recuentos y traslados. Control estricto de desplazamientos intramuros. Registros frecuentes a internos, celdas, pertenencias y personal. Planes de contingencia y simulacros para incidentes.
Seguridad dinámica	Relación personal-interno siempre profesional. Solo personal capacitado en alto riesgo interactúa con los internos. Reuniones periódicas de coordinación y reportes sistemáticos.
Gestión del Personal	Capacitación en prevención de manipulación y corrupción. Dedicación exclusiva al sistema. Identificación codificada (no se revelan datos personales). Rotación periódica para prevenir corrupción y estrés. Remuneración diferencial.

Beneficios esperados: detección temprana de internos con riesgo de alteración del orden, reducción de incidentes, optimización de recursos humanos y automatización de procesos de clasificación.

Riesgos y limitaciones: posibilidad de falsos positivos, dependencia de la calidad de datos y costos de implementación tecnológica.

Escalabilidad: el modelo puede extenderse a sistemas penitenciarios provinciales y vincularse con sistemas judiciales para mejorar la gestión integral.

● **Conclusiones y Recomendaciones: Resumir los hallazgos principales y proponer los siguientes pasos sugeridos.**

6. Conclusiones y recomendaciones

La implementación de este modelo dentro del penal permitirá clasificar mejor a los reclusos y adoptar las medidas necesarias para anticipar motines, optimizar recursos y aplicar medidas preventivas personalizadas.

Resumen cuantitativo de resultados: el modelo LightGBM logró recall del 70% para clase 2 y 58% para clase 3 en la predicción.

Recomendación práctica: desplegar un prototipo piloto en un penal federal y medir impacto real.

Recomendación metodológica: evaluar un ensemble LightGBM + CatBoost para robustez.

● Anexos opcionales: Incluir gráficos adicionales, ejemplos de código, tablas extensas o diagramas que complementen la información del informe.

7. Anexos

Evaluaciones finales y métricas para los modelos LGBM y CatBoost, respectivamente:

```
--- Evaluación final en Holdout 2022 ---  
Balanced Accuracy: 0.7344990420821053  
F1 Macro: 0.6468742336071637
```

```
Classification Report:  
      precision    recall  f1-score   support  
  
     1       0.98       0.93       0.95       91652  
     2       0.53       0.70       0.60        7701  
     3       0.29       0.50       0.39        2357  
  
 accuracy          0.90       0.90       0.90      101710  
 macro avg          0.60       0.73       0.65      101710  
weighted avg          0.93       0.90       0.91      101710
```

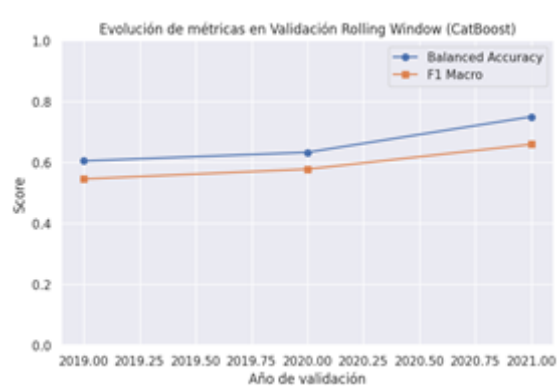
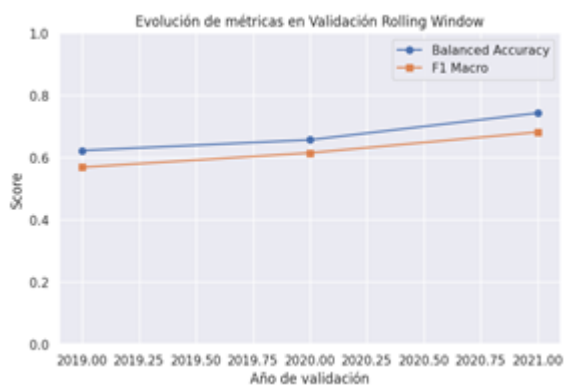
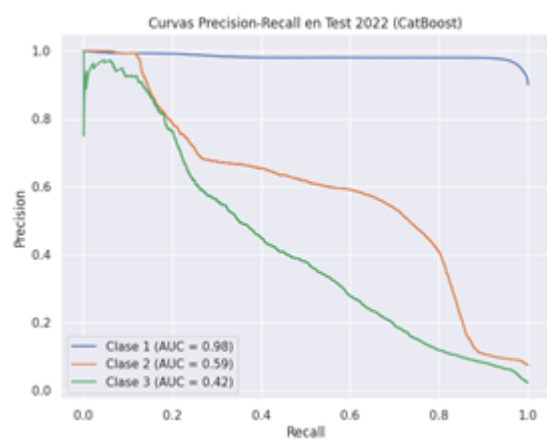
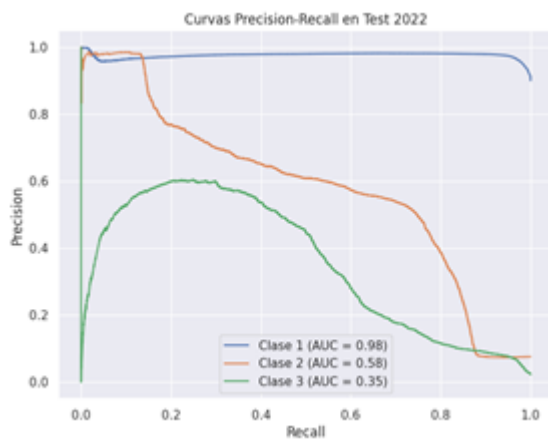
```
Confusion Matrix:  
[[84970  4310  2372]  
 [ 1383  5373   945]  
 [   496   497 1364]]
```

```
--- Evaluación final en Holdout 2022 ---  
Balanced Accuracy: 0.737112847304303  
F1 Macro: 0.6346448728980835
```

```
Classification Report:  
      precision    recall  f1-score   support  
  
     1       0.98       0.92       0.95       91652  
     2       0.51       0.70       0.59        7701  
     3       0.26       0.59       0.36        2357  
  
 accuracy          0.89       0.89       0.89      101710  
 macro avg          0.58       0.74       0.63      101710  
weighted avg          0.93       0.89       0.91      101710
```

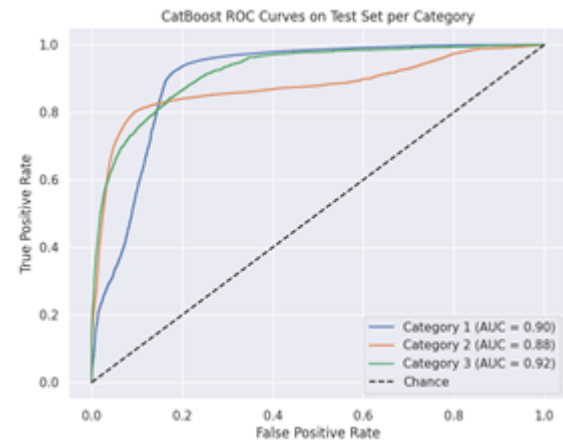
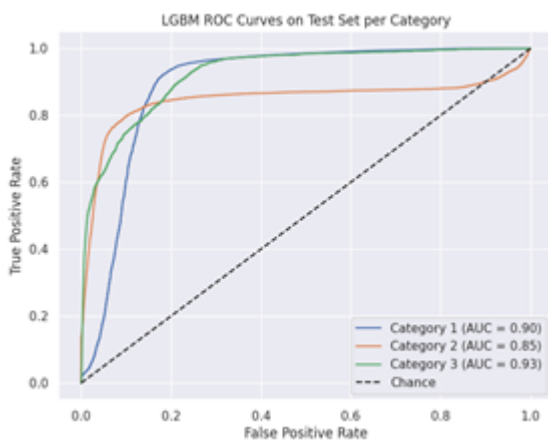
```
Confusion Matrix:  
[[84035  4655  2962]  
 [ 1319  5427   955]  
 [   486   481 1390]]
```

Curvas Precision-Recall para los modelos LGBM y CatBoost, respectivamente:

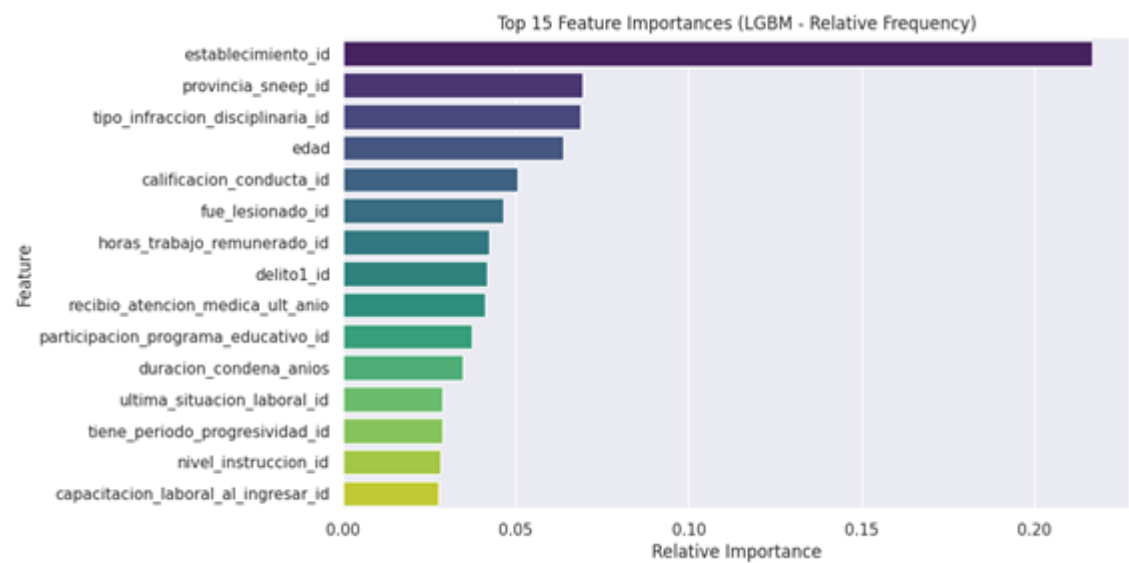


Estas gráficas muestran que los modelos mantienen su desempeño estable en el tiempo y no se degradan al acercarse a 2022.

Curvas ROC por categorías para los modelos LGBM y CatBoost, respectivamente:



Listado de las 15 características más importantes para el modelo LGBM:



• Precision@K Comparison for Categories 2 and 3:

	LGBM (category 2)	CatBoost (category 2)	LGBM (category 3)	CatBoost (category 3)
Precision@10	0.900	1.000	0.000	0.900
Precision@50	0.960	1.000	0.060	0.960
Precision@100	0.980	1.000	0.190	0.970
Precision@500	0.980	0.998	0.552	0.818
Precision@1000	0.981	0.974	0.600	0.604

8. Bibliografía consultada

- Ministerio de Seguridad, Resolución 35/2024, RESOL-2024-35-APN-MSG – Sistema Integral de Gestión para Personas Privadas de la Libertad de Alto Riesgo. [Texto completo | Argentina.gob.ar](https://www.argentina.gob.ar/seguridad)
- MINISTERIO DE SEGURIDAD Resolución 153/2025, RESOL-2025-153-APN-MSG- Medidas de control en la realización de visitas a internos en el “Sistema Integral de Gestión para Personas Privadas de la Libertad de Alto Riesgo en el Servicio Penitenciario Federal” [20250828](#)

- Ejecución de la pena privativa de la libertad -DECRETO 1.136/97- Reglamento de Comunicaciones y visitas de los Internos [Reglamento de Comunicaciones y Visitas_0.pdf](#)
- Ejecución de la pena privativa de la libertad Decreto 396/99- Reglamento Progresividad del Régimen Penitenciario y Programa de Prelibertad [EJECUCION DE LA PENA PRIVATIVA DE LA LIBERTAD](#)
- POLITICA PENITENCIARIA ARGENTINA -Decreto N° 18/97 - Reglamento de Disciplina para los Internos [Texto completo | Argentina.gob.ar](#)
- Procuración Penitenciaria de la Nación -La Recomendación es un instrumento utilizada ante problemas graves de vulneración de derechos fundamentales de las personas privadas de libertad. [Recomendaciones](#)