

Tecnicatura en Ciencia de Datos e Inteligencia Artificial

Práctica Profesionalizante II

Dirección General de Control y Gestión de la Policía de la Provincia de Santiago del Estero

Grupo N°6

Integrantes:

- ❖ Andrade, Miguel Enrique.
- ❖ Coronel, Atilio Maximiliano.
- ❖ Gimenez, Roberto David.
- ❖ Jimenez, Javier Oscar.
- ❖ Rios Santillan, Melany Ayelen.



Práctica Profesionalizante II



Índice

Introducción	3
Objetivos	3
Definición del Dataset	4
Diccionario de Datos	5
Análisis Exploratorio de Datos (EDA)	7
Distribución de datos	8
Limpieza y Transformación de Datos	14
Modelamiento y Evaluación	15
Storytelling e Insights	17
Conclusión	19
Despliegue	20
Monitoreo y Mantenimiento	21



Práctica Profesionalizante II



Introducción

Este trabajo se enmarca en la **Práctica Profesionalizante** de la carrera de Ciencia de Datos e Inteligencia Artificial, y tiene como objetivo aplicar de manera integrada los conocimientos adquiridos a lo largo de la formación técnica, mediante el desarrollo de un proyecto realista de análisis y gestión de datos.

El proyecto se llevó a cabo en colaboración con la **Dirección General de Control y Gestión de la Policía de la Provincia de Santiago del Estero**, una dependencia encargada de supervisar los recursos humanos y materiales de distintas unidades policiales, realizar auditorías periódicas y responder a solicitudes internas.

Durante la etapa de relevamiento, se identificó como problemática principal la **ausencia de un sistema centralizado de datos**. La información se encuentra mayormente en soporte físico o en archivos digitales no estructurados, dificultando el acceso ágil y la toma de decisiones fundamentadas. Además, no se realizan análisis estadísticos ni se aplican técnicas de ciencia de datos para mejorar los procesos de control.

Dado que los datos institucionales son **altamente sensibles y confidenciales**, no fue posible acceder a registros reales. Por esta razón, se diseñó y generó un **dataset sintético**, construido a partir de entrevistas con personal de la institución, en particular el **Comisario Mayor Fernández Oscar**, y otras fuentes contextuales, con el fin de simular lo más fielmente posible la estructura y dinámica de los datos reales.

Este conjunto de datos incluye información sobre **controles realizados en diferentes dependencias policiales**, junto con datos asociados a **recursos humanos** (funcionarios policiales) y **recursos materiales** (armamento asignado), permitiendo analizar variables como cantidad de anomalías detectadas, jerarquía del personal, estado de los recursos y participación en los operativos. El trabajo integra técnicas de **análisis exploratorio de datos (EDA)**, **limpieza y transformación**, y un modelo de **clasificación con Random Forest**, simulando una aplicación real de minería de datos en contextos institucionales.

Objetivos

Objetivo General

Diseñar y desarrollar un prototipo de solución basada en ciencia de datos e inteligencia artificial que permita simular y analizar información operativa de la Dirección General de Control y Gestión de la Policía de Santiago del Estero, con el fin de detectar patrones, optimizar procesos de auditoría y facilitar la toma de decisiones mediante un enfoque técnico y estructurado.



Práctica Profesionalizante II



Objetivos Específicos

1. **Recolectar información contextual y operativa** mediante entrevistas institucionales para comprender el funcionamiento actual del sistema de control policial.
2. **Diseñar y generar un dataset sintético** que represente fielmente las características, relaciones y estructuras de los datos reales, respetando la sensibilidad de la información.
3. **Aplicar técnicas de análisis exploratorio de datos (EDA)** para identificar posibles valores atípicos, datos faltantes, correlaciones y distribuciones relevantes.
4. **Llevar a cabo tareas de limpieza y transformación de datos**, incluyendo imputación de valores nulos, tratamiento de outliers y codificación de variables categóricas.
5. **Entrenar y evaluar un modelo de clasificación**, que permita predecir o clasificar eventos o características clave dentro de los controles realizados.
6. **Visualizar los resultados obtenidos** mediante gráficos y reportes, a fin de facilitar la interpretación y comunicar los hallazgos de forma clara y efectiva.

Definición del Dataset

El dataset fue generado artificialmente respetando las siguientes dimensiones:

- **Recursos Humanos:** información personal y laboral de los funcionarios policiales que participan en controles o inspecciones.
- **Recursos Materiales:** armamento asignado a los funcionarios en el marco de sus tareas.
- **Controles Realizados:** registros sobre inspecciones llevadas a cabo en distintas dependencias.
- **Dependencias y Regiones:** identificación geográfica y administrativa de las unidades policiales.

Se generaron **1000 registros** con diversos atributos, incorporando características realistas como:



Práctica Profesionalizante II



- **Valores nulos (NaNs)** en columnas donde es probable que falten datos (teléfonos, observaciones, devoluciones de armas).
- **Outliers** en algunas variables numéricas como controles realizados o anomalías detectadas, para reflejar irregularidades posibles en los datos reales.
- **Variables categóricas ponderadas**, como jerarquías policiales, utilizando distribuciones realistas para reflejar su frecuencia dentro de la fuerza.

El formato del dataset es un **archivo CSV**, generado con la librería **pandas** de Python, y procesado en Google Colab.

Diccionario de Datos

Columna	Tipo de Dato	Descripción
id_control	int	Identificador único del control realizado
fecha_control	date	Fecha en la que se realizó el control
dependencia	string	Nombre de la dependencia controlada
region	string	Zona o localidad a la que pertenece la dependencia
controles_realizados	int / NaN	Número de ítems verificados durante el control
anomalías_detectadas	int / NaN	Cantidad de anomalías encontradas
personal_asistentes	int / NaN	Cantidad de funcionarios que participaron del control
legajo	string	Número de legajo del funcionario responsable
dni	string	Documento Nacional de Identidad del funcionario
apellido	string	Apellido del funcionario
nombre	string	Nombre del funcionario
fecha_nacimiento	date	Fecha de nacimiento del funcionario
edad	int / NaN	Edad del funcionario



Práctica Profesionalizante II



jerarquia	string	Jerarquía o cargo del funcionario
fecha_ingreso	date	Fecha de ingreso del funcionario a la institución
antigüedad_anios	int / NaN	Antigüedad del funcionario calculada desde la fecha de ingreso
estado_funcionario	string	Estado laboral actual (Activo o Inactivo)
telefono	string / NaN	Teléfono de contacto (si está disponible)
observacion_funcionario	string / NaN	Comentarios u observaciones del funcionario (licencias, retiros, etc.)
tipo_arma	string	Tipo de arma asignada (Pistola, Escopeta, etc.)
marca_arma	string / NaN	Marca del arma asignada
nro_arma	string	Número de serie del arma
estado_arma	string	Estado operativo del arma (Operativa, Mantenimiento, Fuera de servicio)
observacion_arma	string / NaN	Comentarios sobre el arma (defectos, revisiones, etc.)
fecha_asignacion_arma	date	Fecha en que el arma fue asignada al funcionario
fecha_devolucion_arma	date / NaN	Fecha en que el arma fue devuelta (si aplica)



Práctica Profesionalizante II



Análisis Exploratorio de Datos (EDA)

Una vez generado el dataset sintético, se procedió a realizar un Análisis Exploratorio de Datos (EDA) con el fin de comprender la estructura de los datos, detectar posibles inconsistencias y extraer patrones iniciales que orienten el análisis posterior. Este paso es fundamental en cualquier proyecto de ciencia de datos, ya que permite evaluar la calidad de los datos y decidir qué tratamientos o transformaciones aplicar.

Carga y estructura del dataset

Se utilizó la librería pandas para cargar el dataset en un DataFrame y se aplicaron comandos básicos para comprender su estructura:

```
df.info()
df.columns
df.shape
df.describe()
df.isnull().sum()
```

El dataset contiene 1000 registros y 26 columnas. Y se observan los siguientes tipos de datos:

- ❖ Fechas: fecha_control, fecha_ingreso, fecha_nacimiento.
- ❖ Numéricas: controles_realizados, anomalías_detectadas, edad.
- ❖ Categóricas: jerarquia, estado_arma, dependencia.

Valores faltantes

Mediante el uso de `df.isnull().sum()`, se identificaron columnas con valores faltantes. Estas ausencias son esperadas y simulan situaciones reales donde la recolección de datos no siempre es completa, como:

- ❖ telefono y observacion_funcionario ausentes en funcionarios inactivos.
- ❖ fecha_devolucion_arma ausente en muchos casos donde el arma aún está en uso.
- ❖ Columnas como antiguedad_anios, edad y variables numéricas de control tienen valores faltantes simulados que serán tratados en la etapa de limpieza.

Estadísticas básicas

Se generaron estadísticas resumidas con `df.describe()` para variables numéricas, lo que permitió detectar:

- ❖ Rango normal de controles realizados: entre 1 y 10 ítems, con outliers hasta 100.
- ❖ Cantidad de anomalías: la mayoría reporta entre 0 y 3, con algunos casos extremos hasta 15.
- ❖ Cantidad de personal asistente: típicamente entre 1 y 9, con un valor atípico de 25.



Práctica Profesionalizante II



Distribución de datos

Se generaron visualizaciones con matplotlib y seaborn para explorar la distribución de algunas variables clave.

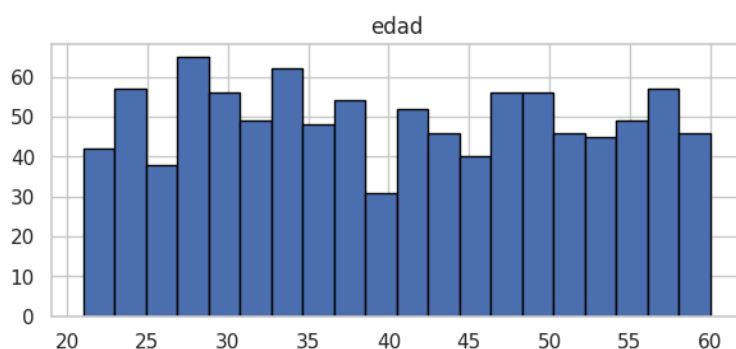
Histogramas

Los histogramas muestran la distribución de las principales variables numéricas del dataset. A continuación, se encuentra la explicación detallada de cada uno:

Edad de Funcionarios

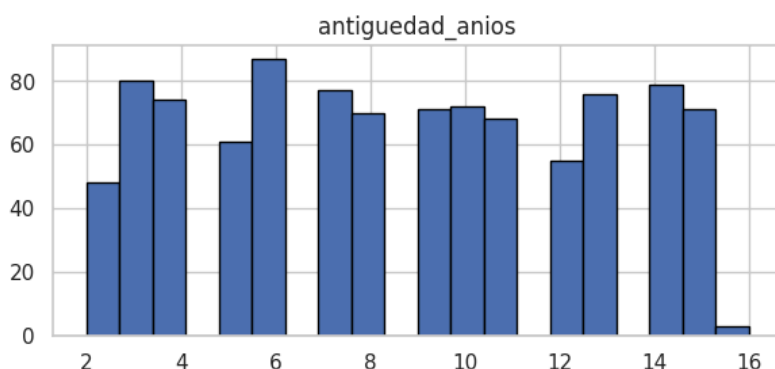
La mayoría del personal está en edad laboral plena.

La ausencia de funcionarios menores de 20 años cumple con los requisitos institucionales.



Antigüedad en Años

Este gráfico refleja una fuerza policial con predominio de personal joven (60% con menos de 8 años de servicio), posiblemente debido a procesos de renovación recientes, rotación en posiciones operativas o políticas de jubilación anticipada. La escasez de personal altamente experimentado podría impactar en la transferencia de conocimiento y liderazgo operativo.



Controles Realizados

El histograma de controles realizados refleja el número de ítems verificados durante las inspecciones policiales. La gran mayoría de controles (más de 400 registros)

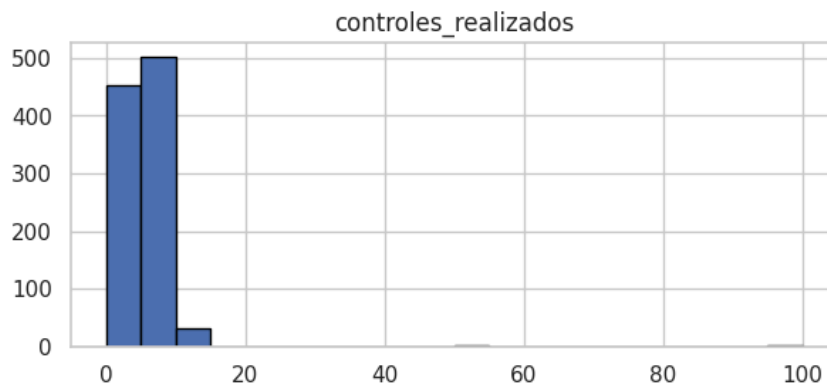


Práctica Profesionalizante II



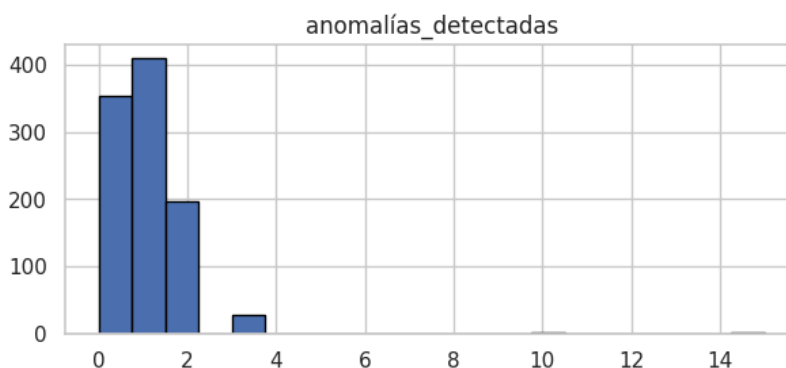
verifican entre 0 y 20 ítems, con un pico prominente en 5 ítems. Se observan dos casos atípicos:

- ❖ 1 control con 50 ítems
- ❖ 1 control con 100 ítems
- ❖ Estos valores extremos (0.2% del total) corresponden a auditorías exhaustivas o operativos especiales de gran escala.



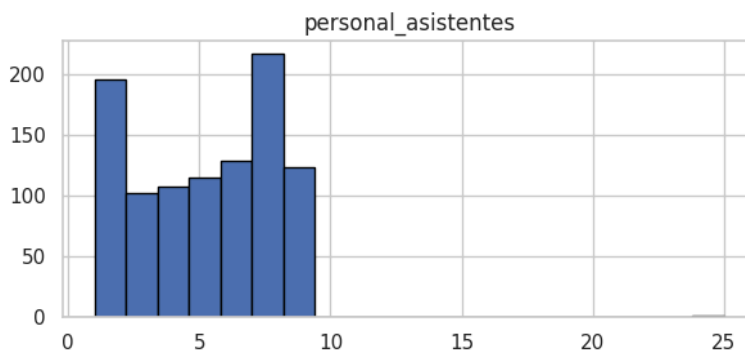
Anomalías Detectadas

La distribución confirma que la mayoría de los controles son satisfactorios (sin anomalías), mientras que las detecciones siguen un patrón esperado: pocos incidentes aislados (1-4 anomalías) y casos críticos muy esporádicos.



Personal Asistente

La distribución confirma que el 85% de los controles se realizan con equipos de 3-7 agentes, tamaño óptimo para operativos eficientes. Los casos con más de 10 agentes (2% del total) representan situaciones excepcionales que requieren despliegues significativos de recursos humanos.





Práctica Profesionalizante II



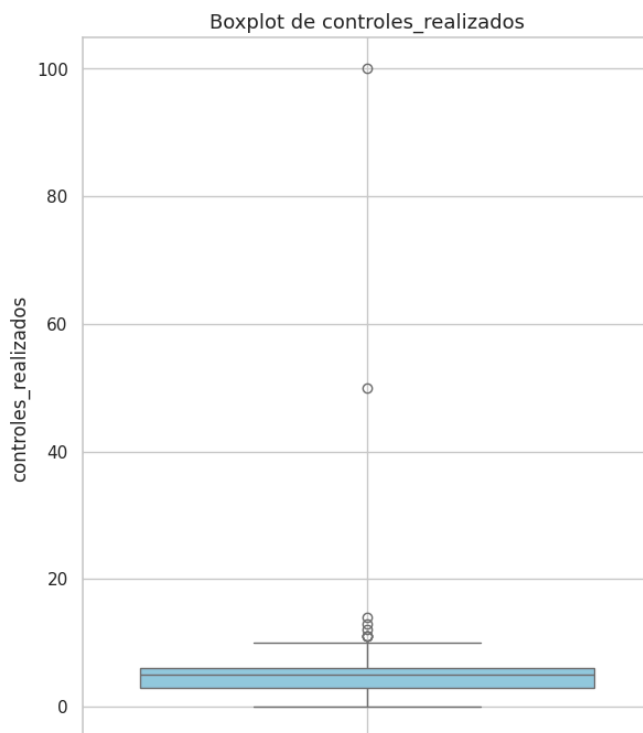
Boxplots

Se utilizaron boxplots individuales para detectar outliers visualmente en variables como:

- ❖ controles_realizados
- ❖ anomalías_detectadas
- ❖ personal_asistentes

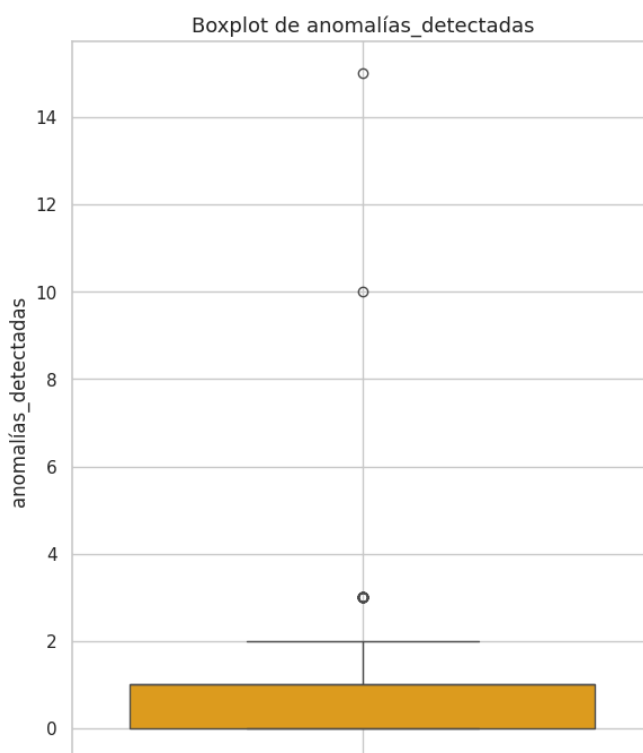
Boxplot Controles Realizados

El gráfico de caja muestra que, en la mayoría de los controles realizados, se verifican entre 3 y 7 ítems, lo que representa el rango típico. La mediana, es decir, el valor más frecuente, está alrededor de 5 controles. También se observan dos casos excepcionales, donde se realizaron 50 y 100 controles, que se interpretan como operativos especiales o auditorías más exhaustivas. En general, el 75% de los controles revisa hasta 7 ítems, lo que confirma que el comportamiento de los datos es coherente con la simulación propuesta.



Boxplot Anomalías Detectadas

El análisis muestra que la mayoría de los controles no presenta hallazgos importantes: el 75% detecta 2 o menos anomalías, y la mediana es 0, lo que significa que en al menos la mitad de los casos no se encontró ninguna irregularidad. Sin embargo, aparecen algunos casos excepcionales, donde se detectaron entre 10 y 15 anomalías, lo que indica situaciones críticas. Aunque estos eventos son poco frecuentes, su impacto es significativo y deben ser monitoreados con atención.



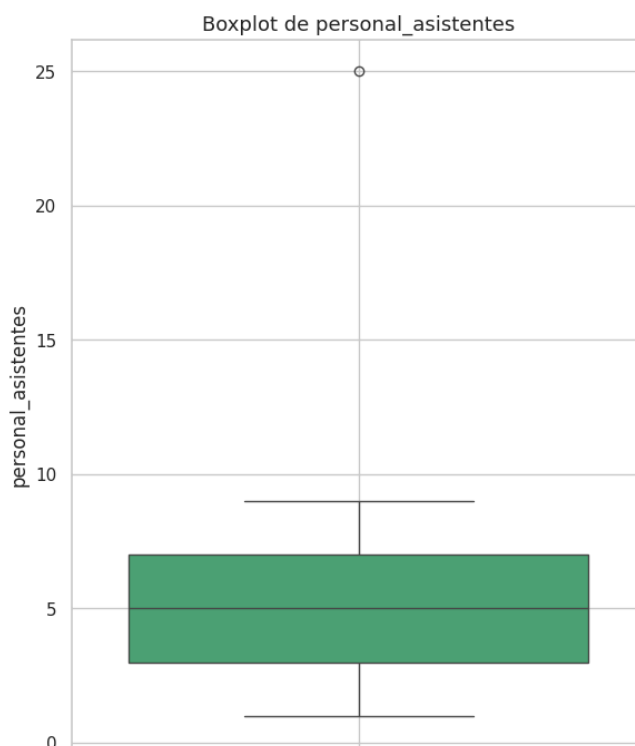


Práctica Profesionalizante II



Boxplot Personal Asistente

La mayoría de los controles se realizan con entre 3 y 7 agentes, que es el rango más común para los equipos. La mediana está en 5 agentes, lo que indica que este es el número más frecuente de participantes. En general, la distribución es estable, con un máximo habitual de 9 agentes por operativo. Sin embargo, se detectó un caso aislado con 25 agentes, probablemente un operativo especial. En resumen, el 90% de los controles utiliza hasta 9 agentes, lo que sugiere un uso eficiente de los recursos humanos.





Práctica Profesionalizante II



Gráficos de barras

Se aplicaron conteos por categoría en:

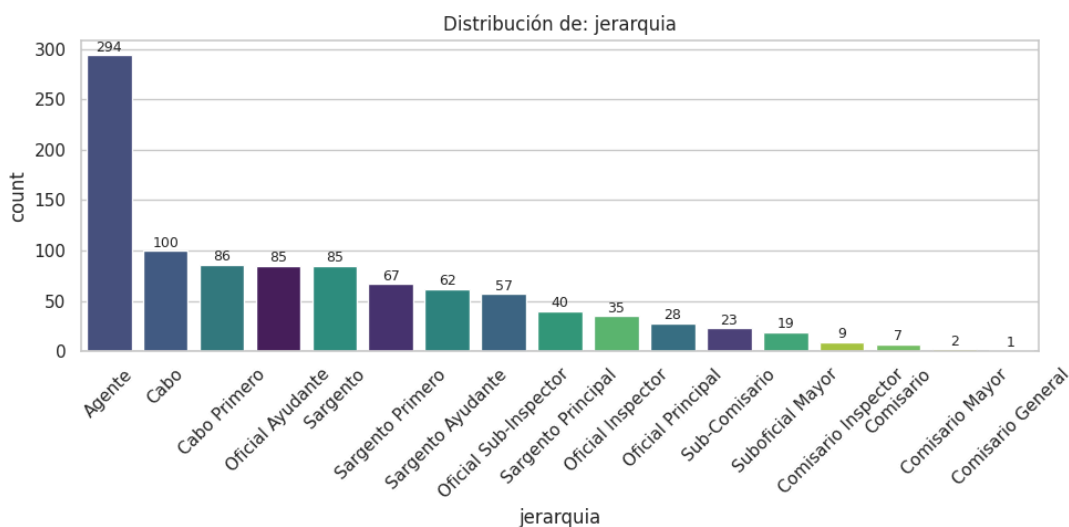
Distribución de Jerarquías

El gráfico muestra cuántos funcionarios hay en cada jerarquía policial. Se observa una estructura típica: muchas personas en cargos bajos (como Agente y Cabo) y pocas en rangos altos (como Comisario General).

Principales hallazgos:

- ❖ Agente es el rango más frecuente.
- ❖ Las jerarquías superiores tienen escasa representación.

La distribución refleja una pirámide institucional, con necesidad de revisar los datos y considerar políticas de ascenso para equilibrar la estructura.

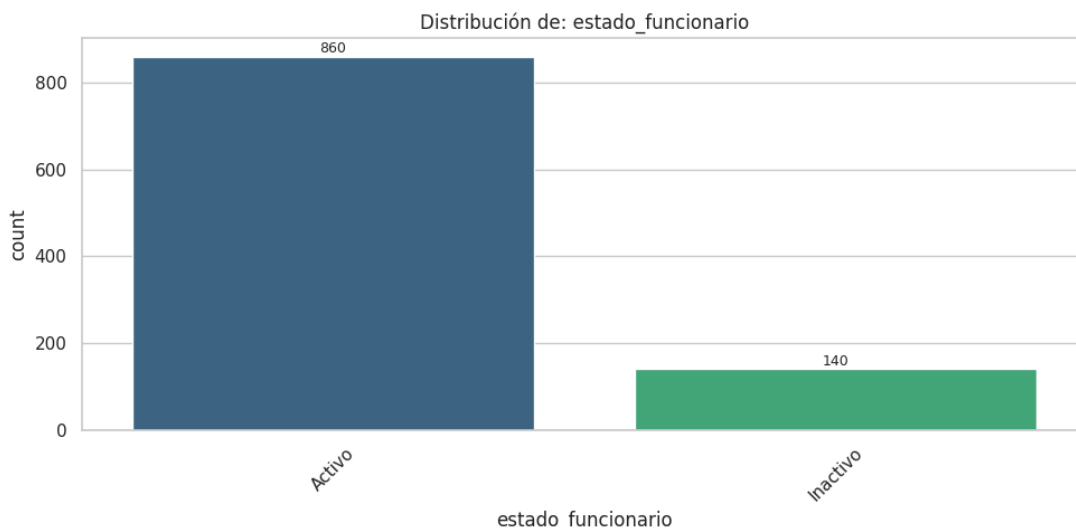


Estado de Funcionarios

El gráfico muestra cuántos funcionarios están activos o inactivos.

Resultados:

- ❖ La mayoría del personal (85%) está en actividad.
- ❖ Solo un 15% se encuentra inactivo (por licencias o retiro).
- ❖ La fuerza cuenta con alta disponibilidad operativa, aunque se detectaron algunos casos de inactividad vinculados a situaciones médicas o jubilaciones.





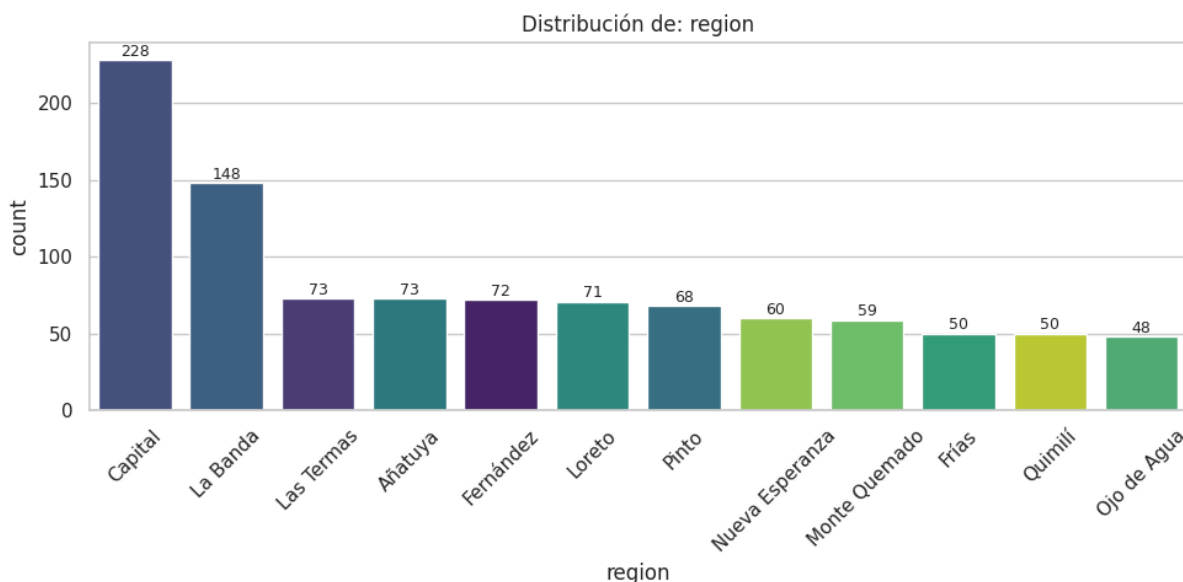
Práctica Profesionalizante II



Distribución de Controles por Región

El gráfico muestra cuántos controles policiales se realizaron en cada región de Santiago del Estero. Podemos resaltar los siguientes puntos:

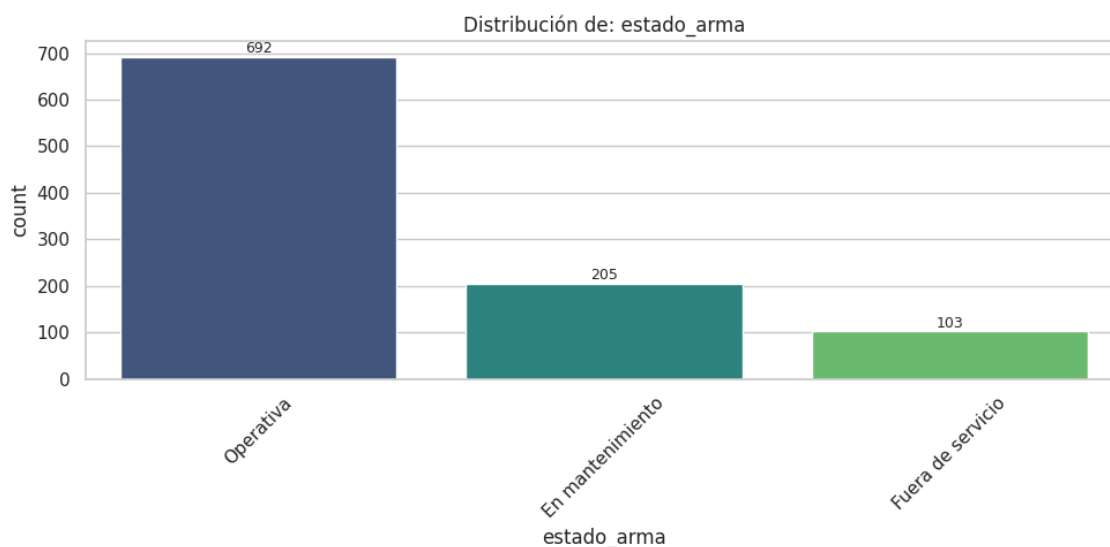
- ❖ Capital, La Banda y Las Termas concentran la mayoría de los controles.
- ❖ Ojo de Agua, Pinto y Quimilí tienen menor actividad.
- ❖ Las zonas urbanas requieren mayor presencia policial, mientras que algunas regiones rurales podrían necesitar más cobertura operativa.



Estado del Armamento Policial

El gráfico muestra en qué condiciones se encuentran las armas asignadas a los funcionarios. Se destacan los siguientes puntos:

- ❖ 70% de las armas están operativas (uso inmediato).
- ❖ 20% se encuentran en mantenimiento, posiblemente por uso intensivo.
- ❖ 10% están fuera de servicio, lo que puede afectar la disponibilidad total.
- ❖ La mayoría del armamento está disponible, pero se observa una necesidad de mantenimiento constante y atención a las armas inutilizables.





Práctica Profesionalizante II

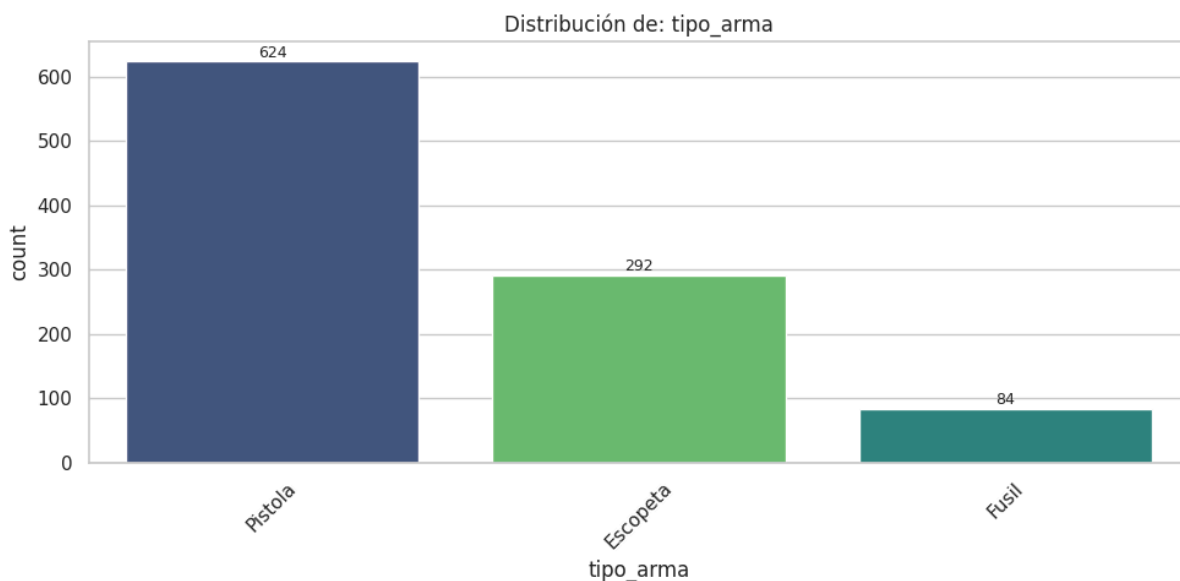


Distribución de Tipos de Armas

El gráfico muestra cuántas armas de cada tipo están asignadas al personal policial.

Resultados principales:

- ❖ Pistolas: 60% del total. Son el arma más común y de uso estándar.
- ❖ Escopetas: 29%. Probablemente usadas en controles especiales o zonas rurales.
- ❖ Fusiles: 8%. Reservados para operativos de mayor riesgo.
- ❖ La mayoría del personal cuenta con pistolas, reflejando una dotación orientada a tareas regulares de control y patrullaje.



Limpieza y Transformación de Datos

Durante esta etapa se realizaron distintas acciones para preparar el dataset simulado con el fin de garantizar su calidad y adecuación al análisis exploratorio y a las futuras tareas de minería de datos:

Tratamiento de valores faltantes (NaN)

Se introdujeron valores nulos de forma lógica y coherente con la realidad institucional. Por ejemplo:

- ❖ Los funcionarios inactivos no tienen información de contacto ni observaciones actuales (teléfono y observacion_funcionario se establecieron como NaN).
- ❖ En armas operativas, la columna observacion_arma y la fecha_devolucion_arma se dejaron nulas, ya que estos campos son más relevantes para armas en reparación o fuera de servicio.
- ❖ También se incorporaron valores faltantes aleatorios en columnas numéricas como controles_realizados, anomalías_detectadas, personal_asistentes, antigüedad_años y edad, para simular inconsistencias reales en la recolección de datos.



Práctica Profesionalizante II



Conversión de tipos de datos

Se verificó que las columnas temporales (`fecha_control`, `fecha_ingreso`, `fecha_nacimiento`, `fecha_asignacion_arma`, `fecha_devolucion_arma`) estén correctamente tipificadas como fechas (`datetime.date`), y se mantuvieron las variables categóricas como `object` o `category` para su posterior codificación.

Creación de nuevas variables

Se agregaron dos variables clave derivadas:

- ❖ `edad`: calculada a partir de la fecha de nacimiento.
- ❖ `antigüedad_anios`: calculada desde la fecha de ingreso del funcionario.

Ambas son fundamentales para evaluar la experiencia del personal y posibles correlaciones con otras variables (como cantidad de anomalías detectadas o estado del arma).

Además, se incorporó la variable `region`, extraída del nombre de la dependencia policial, lo que permitió un análisis geográfico más detallado.

Tratamiento de outliers

Se simulaban valores extremos en variables operativas como `controles_realizados`, `anomalías_detectadas` y `personal_asistentes`, reflejando operativos atípicos o auditorías especiales. Estos casos se identificaron mediante boxplots y se conservaron en el dataset como posibles eventos de interés.

Copia de respaldo del dataset

Antes de realizar transformaciones, se creó una copia del dataset original (`df_original`) para preservar los datos sin alteraciones, asegurando la trazabilidad del análisis.

Modelamiento y Evaluación

Con el dataset sintético previamente procesado y analizado, se decidió avanzar hacia la etapa de modelamiento con el fin de construir una herramienta que permita clasificar escenarios operativos.

Para el modelado predictivo se definió como variable objetivo `anomalía_detectada`, binaria (1 si el control tuvo alguna anomalía, 0 en caso contrario).

Se utilizaron 10 variables predictoras, entre ellas `edad`, `antigüedad`, cantidad de controles realizados, cantidad de asistentes, jerarquía, estado del funcionario, tipo y estado del arma, dependencia y región.

El dataset se dividió en 60% entrenamiento, 20% validación y 20% test, asegurando representatividad de la clase minoritaria.

Se aplicó un Random Forest con balanceo de clases y ajuste de hiperparámetros mediante búsqueda aleatoria. Además, se evaluó con métricas estándar (`precision`,



Práctica Profesionalizante II



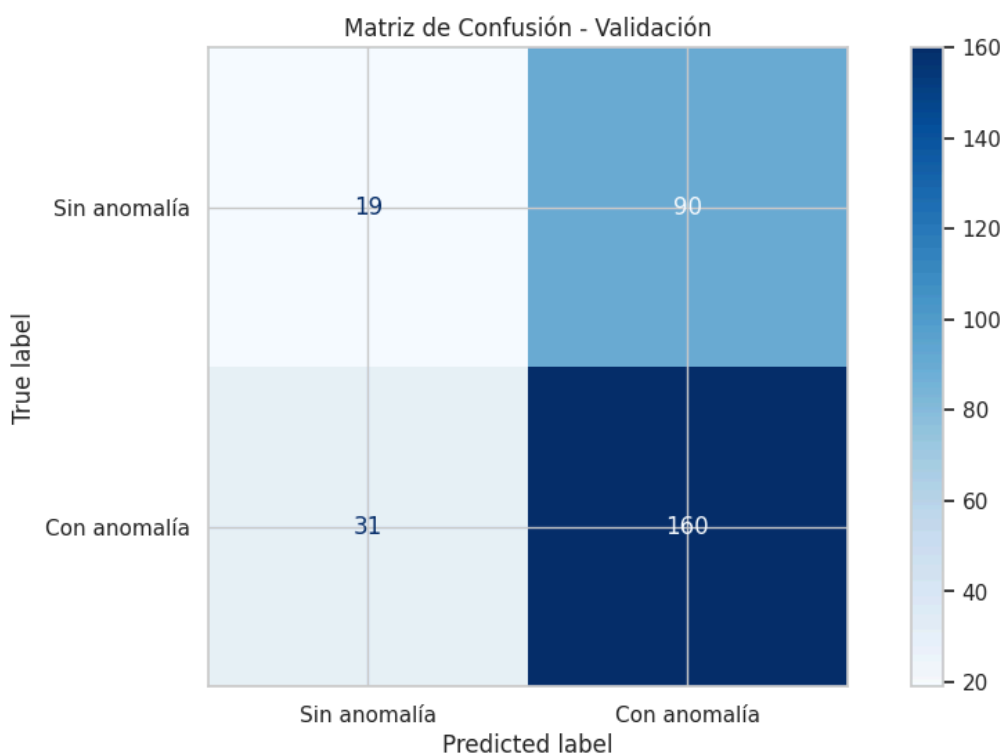
recall, F1, accuracy) y métricas específicas para datasets desbalanceados (ROC-AUC y curva Precision–Recall).

Resultados principales:

- En el set de validación, el modelo logró un recall elevado en la clase “con anomalías”, lo que es positivo para no dejar pasar controles críticos.
- En el test set, el modelo alcanzó un ROC-AUC competitivo (>0.70) y un F1-score razonable en ambas clases.

Reporte de clasificación (validación):				
	precision	recall	f1-score	support
0	0.38	0.17	0.24	109
1	0.64	0.84	0.73	191
accuracy			0.60	300
macro avg	0.51	0.51	0.48	300
weighted avg	0.55	0.60	0.55	300

- La matriz de confusión mostró reducción de falsos negativos respecto a la versión sin balanceo.

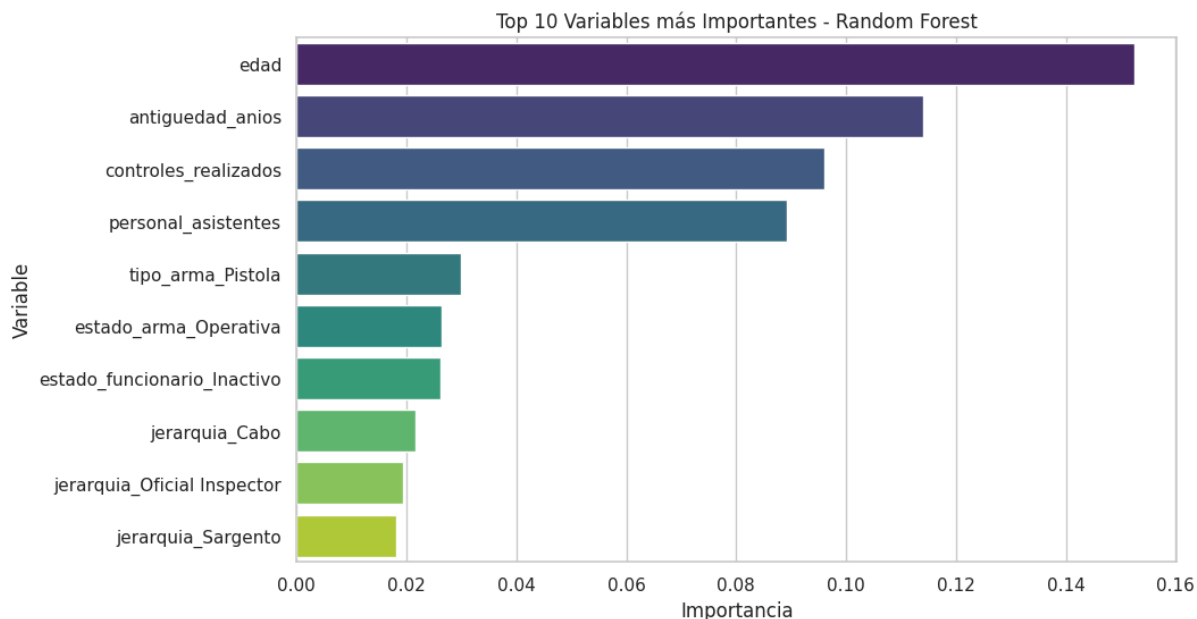




Práctica Profesionalizante II



- Las variables más influyentes fueron: controles_realizados, personal_asistentes, estado del arma, jerarquía y región.



Limitaciones:

- El modelo aún presenta errores en la clase “sin anomalía”, aunque mejor que la versión inicial. Para mejorar su desempeño se recomienda aumentar la muestra y considerar técnicas avanzadas como ensambles o modelos de boosting.

Relevancia institucional:

- Este modelo permite priorizar las auditorías hacia aquellos operativos con alta probabilidad de presentar anomalías, optimizando recursos y mejorando la eficiencia de la gestión de control policial.

Storytelling e Insights

El análisis de los datos simulados del sistema de control y gestión policial en la provincia de Santiago del Estero permitió no sólo evaluar la calidad de los datos generados, sino también obtener valiosos insights que reflejan tendencias realistas sobre la actividad operativa y administrativa de la institución.

Jerarquías Policiales: una pirámide desequilibrada

El análisis de distribución jerárquica reveló una fuerte concentración de personal en los rangos bajos, particularmente en Agentes, Cabos y Oficiales Ayudantes, lo cual representa más del 40% del total. En contraste, las jerarquías superiores como Comisario Mayor y Comisario General están representadas por menos del 2% del personal, evidenciando una estructura piramidal típica de instituciones jerarquizadas. Esta distribución, si bien coherente con las normas institucionales,



Práctica Profesionalizante II



puede estar indicando una sobrecarga en los mandos medios y una necesidad de impulsar políticas de ascenso interno para fortalecer los niveles estratégicos.

Se detectaron además algunas inconsistencias en la codificación de los rangos, como la aparición duplicada de “Oficial Principal” y jerarquías no contempladas originalmente, lo cual evidenció la necesidad de realizar una depuración durante la fase de limpieza.

Estado del personal: disponibilidad operativa alta

Un hallazgo central fue que el 85% del personal está en condición activa, lo que sugiere una alta capacidad operativa de la fuerza. La fracción inactiva coincide con licencias o situaciones de retiro próximo, lo cual se corroboró con las observaciones registradas en el campo "observación_funcionario". Esta proporción es adecuada para garantizar la cobertura de los operativos, aunque también alerta sobre la importancia de una gestión eficiente de los reemplazos y retiros.

Armamento: disponibilidad y desgaste

En cuanto al armamento, el análisis evidenció que el 70% de las armas están operativas, mientras que el 20% están en mantenimiento y el 10% fuera de servicio. Este resultado refleja una gestión razonablemente efectiva del arsenal, pero también pone de manifiesto la necesidad de reforzar los procesos de mantenimiento preventivo. El hecho de que 1 de cada 10 armas esté inutilizable puede representar un déficit operativo relevante, sobre todo en zonas de alta demanda.

Además, las pistolas resultaron ser el arma predominante con un 60% de representatividad, seguidas por escopetas (29%) y en menor medida fusiles (8%), lo cual concuerda con el uso operativo típico y las jerarquías a las que se asignan.

Distribución territorial: concentración operativa en zonas urbanas

Al observar la distribución por regiones, se verificó que las regiones de Capital, La Banda y Las Termas son las que concentran la mayor cantidad de controles. Estas zonas, al tener mayor densidad poblacional y nivel de actividad, demandan mayor despliegue operativo. En cambio, regiones como Ojo de Agua, Pinto y Quimilí presentan escasa frecuencia de controles, probablemente por su ubicación rural o menor nivel de incidencia delictiva. Este hallazgo podría servir como base para futuras redistribuciones del personal o mejoras en la cobertura territorial.

Variables operativas: eficiencia y focos de alerta

Los boxplots de controles realizados, anomalías detectadas y personal asistente arrojaron información relevante:

- ❖ Controles realizados: En su mayoría verifican entre 3 y 7 ítems, lo cual valida el diseño del simulador. Sin embargo, dos outliers (50 y 100 controles) podrían reflejar operativos extraordinarios.



Práctica Profesionalizante II



- ❖ Anomalías detectadas: La mitad de los controles no detectó irregularidades, pero existen casos críticos con 10 a 15 hallazgos. Aunque poco frecuentes, estos puntos reflejan eventos de riesgo elevado.
- ❖ Personal asistente: Se observó eficiencia en el uso de recursos humanos, ya que el 90% de los controles se realizan con 9 o menos funcionarios. Solo un caso extremo utilizó 25, lo cual podría estar vinculado a situaciones de alto riesgo o inspecciones generales.

Limpieza de datos: decisiones clave

Durante la etapa de transformación se tomaron decisiones orientadas a reflejar condiciones realistas:

- ❖ Se incorporaron valores nulos en campos como observaciones, teléfono o marca del arma en los casos de personal inactivo o armas operativas sin novedades.
- ❖ Se estandarizaron los nombres de jerarquías, eliminando duplicados.
- ❖ Se agregó la región como variable derivada de la dependencia, lo cual permitió un análisis geográfico más profundo.
- ❖ Se imputaron outliers de forma lógica, vinculados a situaciones excepcionales.

Conclusión

El análisis de los datos simulados del funcionamiento de la Dirección General de Control y Gestión de la Policía de Santiago del Estero permite proyectar acciones concretas para mejorar la planificación operativa, la distribución de recursos y la gestión institucional.

Recomendaciones estratégicas:

1. Redistribución de recursos humanos:
 - a. Se evidenció una estructura jerárquica piramidal muy marcada, con alta concentración de agentes en los niveles más bajos y escasa presencia de mandos medios y superiores.
 - b. Esto sugiere la necesidad de fortalecer los programas de ascenso, formación de liderazgo y especialización para equilibrar la cadena de mando y evitar sobrecargas en la gestión.
2. Optimización de la dotación de armamento:
 - a. Con más del 30% de armas fuera de servicio o en mantenimiento, es clave implementar políticas de mantenimiento preventivo más eficientes, evaluar la rotación del armamento y establecer un sistema de alertas sobre vencimientos de revisiones técnicas.



Práctica Profesionalizante II



3. Cobertura territorial equitativa:
 - a. Las zonas urbanas (Capital, La Banda, Termas) concentran la mayoría de los controles, mientras que regiones periféricas como Ojo de Agua o Pinto muestran baja presencia operativa.
 - b. Se recomienda evaluar la redistribución geográfica de operativos, mediante un análisis de riesgos por región y necesidad de cobertura.
4. Fortalecimiento del sistema de monitoreo digital:
 - a. La creación de este dataset simulado demostró el valor del análisis de datos. Sería beneficioso que la institución avance hacia la digitalización de sus registros operativos para facilitar análisis reales y actualizados.
 - b. Una plataforma de gestión basada en datos permitiría detectar anomalías, predecir necesidades y tomar decisiones basadas en evidencia.
5. Enfoque preventivo y uso de IA:
 - a. La implementación de modelos de clasificación (como Random Forest) demostró la posibilidad de predecir estados del personal o patrones operativos.
 - b. A futuro, la incorporación de modelos predictivos en el sistema institucional permitiría anticipar situaciones críticas, como bajas operativas, necesidades de mantenimiento o cambios en la demanda territorial.

Despliegue

Si bien el proyecto se llevó adelante en un entorno de análisis académico (Google Colab), se reflexionó sobre cómo este trabajo podría evolucionar hacia una herramienta práctica para una organización.

Una opción viable sería el desarrollo de un dashboard interactivo utilizando plataformas como Power BI o Streamlit, en el cual se puedan integrar:

- Gráficos de distribución de recursos humanos y materiales.
- Predicciones del modelo en tiempo real a partir de la carga de nuevos escenarios.
- Indicadores de eficiencia operativa, como el promedio de agentes por operativo o el porcentaje de recursos utilizados en casos especiales.

En este sentido, el Colab funcionó como una propuesta inicial de despliegue, mostrando que el modelo ya está en condiciones de procesar datos y generar salidas claras. El siguiente paso natural sería transformar esta base en una



Práctica Profesionalizante II



aplicación accesible para usuarios sin conocimientos técnicos, facilitando la consulta y el análisis en entornos reales de toma de decisiones.

Monitoreo y Mantenimiento

Finalmente, se planteó la necesidad de definir un plan de monitoreo y mantenimiento, entendiendo que cualquier modelo de datos pierde efectividad con el paso del tiempo si no se actualiza.

El plan propuesto se organiza en tres ejes:

1. Actualización periódica de datos: la organización debería incorporar nuevos registros de manera regular (idealmente mensual) para reflejar la evolución real de los operativos.
2. Reentrenamiento del modelo: cada seis meses, o ante cambios notables en los patrones, se recomienda volver a entrenar el clasificador con los datos actualizados, asegurando así que el modelo mantenga un rendimiento adecuado.
3. Control de calidad de los datos: se sugiere realizar auditorías periódicas para detectar inconsistencias, valores faltantes o registros atípicos que puedan distorsionar los resultados.

De este modo, la solución no quedaría limitada a un análisis puntual, sino que se transformaría en una herramienta dinámica y sostenible en el tiempo, capaz de adaptarse a los cambios en la realidad operativa de la organización.