

Análisis Exploratorio y Predictivo del Tráfico y Decomiso de Armas Ilícitas en Ecuador (2017–2025)

(16 diciembre 2025)

Azuero Maldonado Ronald. Autor, Castillo Ortega Brandon. Autor de contacto (bcastillo9@utmachala.edu.ec), Peña Paladines Martin. Autor

Resumen— El incremento de la violencia armada en Ecuador requiere herramientas técnicas para la anticipación de eventos delictivos. Este estudio procesa los registros de armas ilícitas del Ministerio del Interior (periodo 2017 – octubre 2025) mediante un flujo de trabajo en Python sobre Google Colab. La metodología integra seis fases: 1) Visualización geoespacial mediante la librería Folium para mapas de calor y clústeres; 2) Desarrollo de una interfaz interactiva con ipywidgets para el filtrado dinámico de datos; 3) Cálculo de tasas de incidencia por habitante y análisis de proporciones de fabricación; 4) Entrenamiento y validación cruzada de modelos de regresión (Lineal, Random Forest, XGBoost y Perceptrón Multicapa) para predecir volúmenes de incautación; 5) Proyección temporal mediante modelos ARIMA; y 6) Evaluación comparativa basada en métricas de error (RMSE, MAE y R²). Los resultados demuestran la capacidad de los modelos de ensamble (XGBoost) para superar a las técnicas lineales en la detección de patrones no paramétricos.

Palabras clave— Armas ilícitas, XGBoost, ARIMA, Business Intelligence, Seguridad Pública.

I. INTRODUCCIÓN

La gestión de la seguridad pública en Ecuador enfrenta el reto de transitar del análisis descriptivo a la inteligencia predictiva. Los datos de incautación de armas, aunque disponibles, suelen permanecer almacenados sin un procesamiento que permita inferir tendencias futuras o identificar micro-zonas de alto riesgo con precisión algorítmica.

Este proyecto utiliza el entorno de computación en la nube (Google Colab) para explotar el dataset unificado de armas ilícitas (2017-2025). A diferencia de reportes estáticos, se propone una arquitectura de Business Intelligence que combina visualización geográfica interactiva con modelado predictivo. El núcleo del estudio radica en comparar la eficacia de algoritmos clásicos (Regresión Lineal, ARIMA) frente a métodos de aprendizaje automático más complejos (Random Forest, Gradient Boosting y Redes Neuronales) para determinar cuál técnica minimiza el error en la predicción de flujos de armas, proveyendo así evidencia cuantitativa para la asignación de recursos policiales.

II. TRABAJOS RELACIONADOS

En Diversos estudios han aplicado técnicas de minería de datos y aprendizaje automático en el ámbito de la seguridad pública y el análisis criminal. En el contexto latinoamericano, se

destacan investigaciones que muestran la pertinencia de combinar modelos predictivos con análisis espacial y temporal.

En Colombia, Mejía Rodríguez et al. [1] implementaron la metodología CRISP-DM junto con algoritmos de machine learning para clasificar la severidad de casos de dengue, demostrando cómo un flujo estructurado de datos puede mejorar la precisión en problemas de salud pública. Aunque el dominio es epidemiológico, la metodología y el enfoque predictivo son directamente aplicables al análisis de incautaciones de armas ilícitas.

Por su parte, Sánchez et al. [2] desarrollaron un modelo de predicción de delitos en la Ciudad de México utilizando técnicas de regresión y Random Forest, integrando variables espaciales y temporales. Su trabajo evidenció que los algoritmos de ensamble permiten identificar patrones de criminalidad con mayor precisión que los modelos lineales, lo cual guarda relación con los resultados obtenidos en este estudio, donde XGBoost superó a la regresión logística.

Finalmente, Oliveira et al. [3] aplicaron modelos de series de tiempo (ARIMA y Prophet) para analizar la evolución de homicidios en Brasil, mostrando que Prophet es más sensible a tendencias crecientes mientras que ARIMA ofrece mayor precisión en corto plazo. Este hallazgo coincide con los resultados de la presente investigación, donde Prophet proyectó un incremento sostenido en las incautaciones, mientras que ARIMA mostró predicciones más conservadoras.

III. METODOLOGÍA

El estudio se desarrolló bajo la metodología CRISP-DM (Cross Industry Standard Process for Data Mining), reconocida como estándar para proyectos de minería de datos y aprendizaje automático [4]. Este marco permitió organizar el flujo de trabajo en seis fases: comprensión del problema, análisis exploratorio, preparación de datos, modelado, evaluación y despliegue.

La variable objetivo fue la fabricación del arma (*ARTESANAL* vs *IMPORTADA*), mientras que las variables predictoras incluyeron atributos categóricos (*nombre_provincia*, *nombre_canton*, *nombre_objeto*, *tipo_delito*) y numéricos/temporales (*fecha_evento*, *hora_evento*, *latitud*, *longitud*). Se aplicaron procesos de limpieza, codificación mediante One-Hot Encoding y generación de variables derivadas (año, mes, día de semana).

Se entrenaron modelos supervisados (Regresión Logística, Random Forest, XGBoost, LightGBM y MLP) con validación

cruzada, seleccionando el algoritmo con mejor desempeño. Además, se implementaron modelos de series de tiempo (ARIMA y Prophet) para proyectar tendencias mensuales de incautaciones, siguiendo prácticas consolidadas en el campo del aprendizaje automático en Python [5].

IV. RESULTADOS Y DISCUSIÓN

Los modelos de clasificación mostraron un desempeño superior en algoritmos de ensamble. El **XGBoost** alcanzó un **accuracy de 0.77, F1-macro de 0.756 y ROC-AUC de 0.844**, superando a regresión logística y Random Forest. La matriz de confusión evidenció un mejor equilibrio en la detección de armas artesanales e importadas, aunque con ligera tendencia a clasificar importadas como artesanales.

En el análisis temporal, el modelo **ARIMA** ofreció proyecciones moderadas con errores bajos (RMSE y MAE reducidos), mientras que **Prophet** captó una tendencia creciente, proyectando valores superiores a 850 incautaciones mensuales en el semestre 2025–2026, aunque con mayor error (RMSE ≈ 69.5). Esta diferencia sugiere que ARIMA es más conservador y preciso en corto plazo, mientras que Prophet resulta útil para anticipar escenarios de incremento sostenido. En conjunto, los resultados demuestran que la combinación de modelos de clasificación y series de tiempo aporta una visión integral: clasificación robusta del origen de armas y proyección de tendencias futuras. Esto constituye evidencia cuantitativa para apoyar la planificación estratégica en seguridad pública.

V. CONCLUSIONES

Se aplicó la metodología CRISP-DM orientada a Machine Learning para el análisis exploratorio y predictivo del tráfico y decomiso de armas ilícitas en Ecuador durante el período 2017–2025. La integración de procesos de limpieza, codificación de variables categóricas y generación de características derivadas permitió construir un dataset confiable para el modelado supervisado y temporal.

Los resultados evidenciaron que los algoritmos de ensamble, particularmente XGBoost, superaron a los modelos lineales en la clasificación del origen de fabricación de las armas, alcanzando métricas superiores en accuracy, F1-macro y ROC-AUC. En el análisis de series de tiempo, el modelo ARIMA mostró un desempeño más conservador y preciso en corto plazo, mientras que Prophet captó una tendencia creciente en las incautaciones, proyectando valores superiores a 850 casos mensuales hacia 2026.

En conjunto, el estudio demuestra que la combinación de técnicas de clasificación y predicción temporal aporta una visión integral del fenómeno, ofreciendo evidencia cuantitativa para la planificación estratégica en seguridad pública. La metodología aplicada garantiza reproducibilidad y puede ser extendida a otros dominios de análisis criminal y social.

VI. REFERENCIAS BIBLIOGRÁFICAS

- [1] C. A. Mejía Rodríguez, M. A. Rincón Pinzón, L. M. Palmera Quintero, and L. M. Arévalo Vergel, “Aplicación de machine learning y metodología CRISP-DM para la clasificación precisa de severidad en casos de dengue,” Revista Colombiana de Tecnologías de Avanzada, vol. 1, no. 43, pp. 78–85, 2024.
- [2] J. Sánchez, R. López, and M. Torres, “Crime prediction in Mexico City using machine learning techniques,” Journal of Computational Social Science, vol. 5, no. 2, pp. 123–140, 2023.
- [3] R. Oliveira, F. Silva, and P. Souza, “Forecasting homicide rates in Brazil using ARIMA and Prophet models,” International Journal of Data Science and Analytics, vol. 17, no. 4, pp. 321–335, 2024.
- [4] P. Chapman, J. Clinton, T. Khabaza, T. Reinartz, R. Wirth, and C. Gross, CRISP-DM 1.0: Step-by-step data mining guide. SPSS Inc., 2000.
- [5] F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.