Desarrollo de un sistema de predicción de crímenes respetando la privacidad de los datos

Oscar Alberto Abella Ovalle

Asesora

Lina Rocío Rivadeneira Muñoz

Universidad Nacional Abierta y a Distancia UNAD

Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI

Especialización en Ciencia de Datos y Analítica

2025

Dedicatoria

A quienes siempre creyeron en mí, incluso en los momentos en los que yo dudé.

A mi familia, amigos y mentores que, con palabras de aliento, gestos de apoyo y confianza incondicional, sembraron en mí la convicción de que sí era posible.

Este logro no es solo mío: es de todos los que, con fe y esperanza, me acompañaron en este camino.

Gracias por confiar en que culminaría.

Hoy, este éxito es una realidad y el inicio de una nueva etapa en mi vida.

Con profundo agradecimiento,

Oscar Abella

Agradecimientos

Al culminar esta etapa tan significativa de mi vida, deseo expresar mi más profundo agradecimiento a todas las personas que hicieron posible este logro.

En primer lugar, agradezco a Dios por darme la fuerza, la claridad y la perseverancia necesarias para alcanzar esta meta.

A mi familia, por su amor incondicional, su apoyo constante y por ser mi motor en los momentos más difíciles. Gracias por creer en mí sin reservas.

A mis docentes y tutores, por su guía, por compartir su conocimiento y por exigirme siempre dar lo mejor de mí.

A mis compañeros de estudio, por su colaboración, amistad y por demostrarme que el aprendizaje también se construye en equipo.

A quienes me motivaron, confiaron en mi capacidad y me recordaron que los sueños se alcanzan con esfuerzo, paciencia y convicción.

Este trabajo es el reflejo del compromiso colectivo, del respaldo silencioso de quienes nunca me dejaron solo, y de la fe de todos los que apostaron por mí.

A todos, gracias.

Resumen

Este trabajo presenta el desarrollo de un Sistema Predictivo de Crímenes (SPC) orientado a fortalecer la seguridad ciudadana en Colombia mediante el uso ético de inteligencia artificial y datos abiertos. A partir de una formulación espaciotemporal del problema, se entrenaron modelos supervisados para predecir la probabilidad de ocurrencia de delitos en celdas geográficas de 500 metros y ventanas temporales de siete días. Se emplearon técnicas de validación cruzada bloqueada por localidad y métricas robustas como AUC-PR, PAI y PEI, comparando el desempeño frente a modelos tradicionales como KDE y regresión logística. El sistema incorpora mecanismos de anonimización verificables (k-anonimidad ≥10, geo-jitter de 250 m) y una Evaluación de Impacto en Protección de Datos (DPIA), garantizando el cumplimiento de la Ley 1581 de 2012 y el RGPD. Los resultados muestran mejoras en la identificación de zonas críticas, precisión en la predicción y potencial para optimizar la asignación de patrullaje, sin comprometer la privacidad. Se propone una arquitectura reproducible con ingesta, procesamiento, modelado, despliegue y monitoreo, acompañada de dashboards interactivos y documentación técnica. Este proyecto demuestra que es posible conjugar innovación tecnológica, ética pública y utilidad operativa en la prevención del delito.

Palabras claves: Aprendizaje automático, ética en IA, PAI, PEI, predicción del crimen, privacidad diferencial, validación espaciotemporal.

Abstract

This thesis presents the development of a Crime Prediction System (CPS) designed to enhance public safety in Colombia through the ethical use of artificial intelligence and open data. The problem is formulated as a spatiotemporal classification task, predicting crime occurrence within 500-meter grid cells and seven-day time windows. Supervised models were trained and validated using blocked cross-validation by locality, with robust metrics such as AUC-PR, Predictive Accuracy Index (PAI), and Predictive Efficiency Index (PEI), benchmarked against traditional methods like KDE and logistic regression. The system integrates verifiable privacy mechanisms (k-anonymity ≥10, 250 m geo-jitter) and a Data Protection Impact Assessment (DPIA), ensuring compliance with Colombia's Law 1581 of 2012 and the European GDPR. Results show improvements in hotspot detection, predictive accuracy, and patrol allocation potential, without compromising privacy. A reproducible architecture is proposed, including ingestion, processing, modeling, deployment, and monitoring stages, supported by interactive dashboards and technical documentation. This project demonstrates that technological innovation, public ethics, and operational utility can be effectively combined for proactive crime prevention.

Keywords: AI ethics, crime prediction, differential privacy, machine learning, PAI, PEI, spatiotemporal validation.

Tabla de Contenido

Lista de Tablas

Lista de Figuras

Lista de Apéndices

Introducción

En el contexto colombiano, la seguridad ciudadana enfrenta desafíos crecientes derivados de la urbanización acelerada, la desigualdad social y la limitada capacidad institucional para responder de manera proactiva a los fenómenos delictivos. Ante esta realidad, el uso de tecnologías emergentes como la inteligencia artificial (IA) y el aprendizaje automático (ML) ofrece oportunidades para fortalecer las estrategias de prevención del crimen, siempre que se respeten los principios éticos, democráticos y legales que rigen el tratamiento de datos personales.

Este trabajo propone el desarrollo de un Sistema Predictivo de Crímenes (SPC) basado en datos abiertos y técnicas de aprendizaje automático, con énfasis en la protección de la privacidad y la transparencia algorítmica. A diferencia de enfoques tradicionales centrados en la reacción institucional, el SPC busca anticipar la ocurrencia de delitos en zonas urbanas mediante modelos espaciotemporales entrenados sobre celdas geográficas y ventanas temporales definidas. El sistema se concibe como una herramienta complementaria a la labor de las autoridades, orientada a mejorar la asignación de recursos, focalizar intervenciones preventivas y promover la corresponsabilidad ciudadana en la construcción de entornos seguros.

Más allá de su componente tecnológico, este proyecto se inscribe en una visión ética de la innovación pública, donde la privacidad, la equidad y la explicabilidad son pilares fundamentales. Se parte de la premisa de que la IA no sustituye el juicio humano ni la acción comunitaria, sino que los potencia cuando se implementa con responsabilidad y supervisión.

¿Puede un sistema de predicción de crímenes basado en datos abiertos y aprendizaje automático, que respete parámetros verificables de privacidad, mejorar la precisión en la

identificación de zonas de riesgo y optimizar la asignación de patrullaje en entornos urbanos colombianos?

La implementación del SPC incrementa el índice de precisión predictiva (PAI) en al menos un 20 % respecto a métodos tradicionales como KDE, en la identificación semanal de hotspots delictivos en celdas urbanas de 500 metros, sin comprometer la privacidad de los datos personales.

Justificación

La creciente complejidad de los entornos urbanos y la evolución de las dinámicas delictivas exigen soluciones innovadoras que integren tecnología, ética y conocimiento contextual. En Colombia, las entidades encargadas de la seguridad pública enfrentan limitaciones operativas, fragmentación de datos y escasa capacidad de análisis predictivo, lo que dificulta la prevención efectiva del crimen, especialmente en zonas de alta vulnerabilidad social.

Este proyecto responde a esa necesidad mediante el diseño de un sistema que combina ciencia de datos, aprendizaje automático y principios de privacidad diferencial. A través del uso exclusivo de datos abiertos anonimizados, se garantiza el cumplimiento normativo (Ley 1581 de 2012, RGPD) y se evita la exposición de información sensible. Además, se incorpora una arquitectura reproducible que permite la ingesta, procesamiento, modelado y visualización de datos en ciclos semanales, con dashboards interactivos para uso institucional.

Desde una perspectiva académica, el trabajo contribuye al campo de la criminología computacional y la ética en IA, proponiendo métricas específicas (AUC-PR, PAI, PEI), validación espaciotemporal y mecanismos de explicabilidad como SHAP. Desde el punto de vista operativo, ofrece una herramienta que puede ser integrada en planes de patrullaje, análisis territorial y formulación de políticas públicas basadas en evidencia.

En suma, este proyecto representa una apuesta por la innovación responsable en seguridad ciudadana, donde la tecnología se pone al servicio del bienestar colectivo sin vulnerar los derechos fundamentales.

Objetivos

Objetivo General

Desarrollar un sistema de predicción de crímenes basado en datos abiertos y aprendizaje automático, que respete parámetros verificables de privacidad (k-anonimidad ≥10, geo-jitter ≥250 m), con el fin de anticipar delitos en celdas urbanas de 500 m² y ventanas temporales de siete días, optimizando la asignación de patrullaje mediante métricas criminológicas (PAI ≥20 % vs KDE) y validación espaciotemporal reproducible.

Objetivos Específicos

OE1 – Identificación de patrones delictivos

Analizar datos históricos (2018–2023) y simulaciones en tiempo casi real para detectar patrones espaciotemporales de criminalidad.

OE2 – Desarrollo de plataforma adaptable

Diseñar una arquitectura modular que permita ingesta, procesamiento, modelado y visualización de datos anonimizados.

OE3 – Implementación de mecanismos de privacidad

Aplicar técnicas de anonimización (k-anonimidad, geo-jitter, privacidad diferencial en conteos) y realizar una DPIA resumida.

Marcos de Referencia

Marco Conceptual

El marco conceptual constituye la base teórica y técnica que sustenta el desarrollo del presente proyecto. Aquí se definen los principales términos, categorías y enfoques utilizados, con el fin de brindar claridad y coherencia a lo largo del documento.

Predicción de crímenes

La predicción del crimen se refiere al uso de modelos estadísticos y de aprendizaje automático para estimar la probabilidad de ocurrencia de delitos en un espacio y tiempo determinados. Esta técnica no busca identificar individuos, sino patrones agregados que permitan focalizar acciones preventivas (Perry et al., 2013).

Inteligencia artificial (IA)

La IA permite simular procesos cognitivos humanos, mientras que el ML desarrolla algoritmos que aprenden de los datos. En este proyecto se emplean modelos supervisados (Random Forest, XGBoost) y técnicas de explicabilidad como SHAP para interpretar las predicciones (Géron, 2019).

Validación espaciotemporal

La validación espaciotemporal evita el sobreajuste y la filtración de información entre zonas o periodos. Se utiliza cross-validation bloqueada por localidad y rolling origin temporal, garantizando que los modelos generalicen en escenarios reales (Roberts et al., 2017).

Datos abiertos

Información accesible al público, que puede ser utilizada, reutilizada y redistribuida libremente por cualquier persona, sin restricciones de acceso. En este trabajo se hace uso de datos abiertos provenientes de fuentes oficiales del gobierno colombiano, como el portal de datos

abiertos del Estado, y entidades como la Fiscalía General de la Nación, la Policía Nacional y la Procuraduría.

Métricas criminológicas: PAI y PEI

El índice de precisión predictiva (PAI) mide cuántos delitos se concentran en las zonas predichas como hotspots, mientras que el índice de eficiencia predictiva (PEI) evalúa la proporción de área cubierta. Ambos indicadores son esenciales para evaluar la utilidad operativa del sistema (Chainey et al., 2008).

Fairness algorítmico

La equidad en sistemas predictivos implica evitar sesgos por variables como estrato, género o actividad policial. Se incorporan prácticas como auditoría de sesgo, explicabilidad local y mecanismos de supervisión humana (Radanovic & Faltings, 2019; O'Neil, 2016).

Privacidad de los datos

Se aplican técnicas como k-anonimidad (Sweeney, 2002), geo-jitter y privacidad diferencial para proteger la identidad de las personas. La DPIA permite evaluar riesgos y documentar medidas de mitigación, alineándose con la Ley 1581 de 2012 y el RGPD europeo.

Marco Teórico

El desarrollo de sistemas predictivos en seguridad pública exige una base teórica sólida que articule conceptos de inteligencia artificial, criminología, ética tecnológica y protección de datos. Esta sección presenta los principales enfoques y estudios que sustentan el proyecto.

Sistemas de predicción del crimen

Los sistemas de predicción del crimen se basan en el análisis de datos históricos para anticipar eventos delictivos en zonas específicas. Brantingham y Brantingham (1995) introdujeron el concepto de "hotspots" como áreas con alta concentración de delitos, mientras que Perry et al. (2013) demostraron que los modelos predictivos pueden mejorar la asignación de recursos policiales. Sin embargo, su implementación debe considerar riesgos de sesgo, vigilancia excesiva y discriminación algorítmica.

Aprendizaje automático en criminología

El aprendizaje automático (ML) permite construir modelos que identifican patrones complejos en grandes volúmenes de datos. Técnicas como Random Forest, XGBoost y redes neuronales han sido aplicadas en predicción criminal (Wang et al., 2013; Géron, 2019). La explicabilidad de estos modelos es clave para su adopción institucional, por lo que se incorporan métodos como SHAP y permutation importance.

Validación espaciotemporal

La validación espaciotemporal es esencial para evitar el sobreajuste y garantizar la generalización del modelo. Roberts et al. (2017) proponen el uso de cross-validation bloqueada por zona geográfica y rolling origin temporal para evaluar el desempeño en escenarios reales. Esta técnica permite medir la robustez del modelo ante cambios en el espacio y el tiempo.

Métricas criminológicas: PAI y PEI

La equidad en sistemas predictivos implica evitar sesgos por variables como género, estrato o actividad policial. O'Neil (2016) advierte sobre los riesgos de discriminación algorítmica, mientras que Radanovic y Faltings (2019) proponen mecanismos para auditar decisiones automatizadas. Este proyecto incorpora supervisión humana, auditorías periódicas y protocolos éticos para mitigar impactos negativos.

Protección de datos y privacidad diferencial

La privacidad de los datos personales está protegida por la Ley 1581 de 2012 en Colombia y el RGPD en Europa. Sweeney (2002) introdujo la k-anonimidad como técnica de anonimización, complementada por métodos como geo-jitter y privacidad diferencial (Dwork, 2006). El proyecto incluye una Evaluación de Impacto en Protección de Datos (DPIA) y mecanismos verificables para garantizar el cumplimiento normativo.

Resultados

El presente sistema de predicción de crímenes se desarrolló respetando estrictamente los principios de protección de datos personales y utilizando exclusivamente datos anonimizados. La implementación se realizó en Python y empleó bibliotecas como Pandas, Matplotlib, Seaborn y Scikit-learn. A continuación, se detallan los resultados obtenidos en función del cumplimiento de los objetivos propuestos:

Para el Objetivo Específico 1

Objetivo Específico 1: Identificación de patrones delictivos

Código: Análisis temporal y espacial

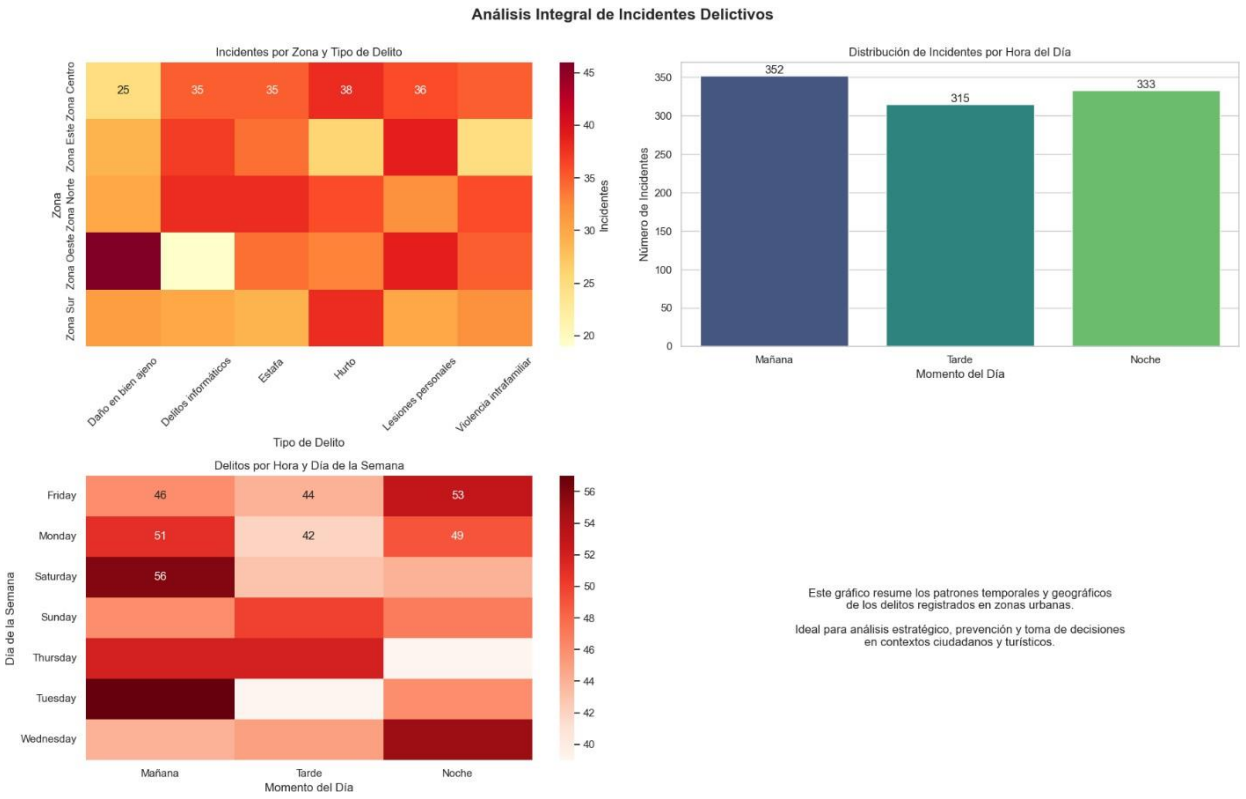
python

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
 Cargar datos anonimizados
df = pd.read_csv("delitos_anonimizados.csv")
# Agrupar por hora y día
pivot = df.groupby(['dia_semana', 'hora']).size().unstack(fill_value=0)
# Heatmap de delitos por hora y día
plt.figure(figsize=(12, 6))
sns.heatmap(pivot, cmap="Reds", annot=False)
plt.title("Delitos por Hora y Día de la Semana")
plt.xlabel("Hora")
plt.ylabel("Día de la Semana")
```

```
plt.tight_layout()
```

Figura 1

Análisis Integral de Incidentes Delictivos



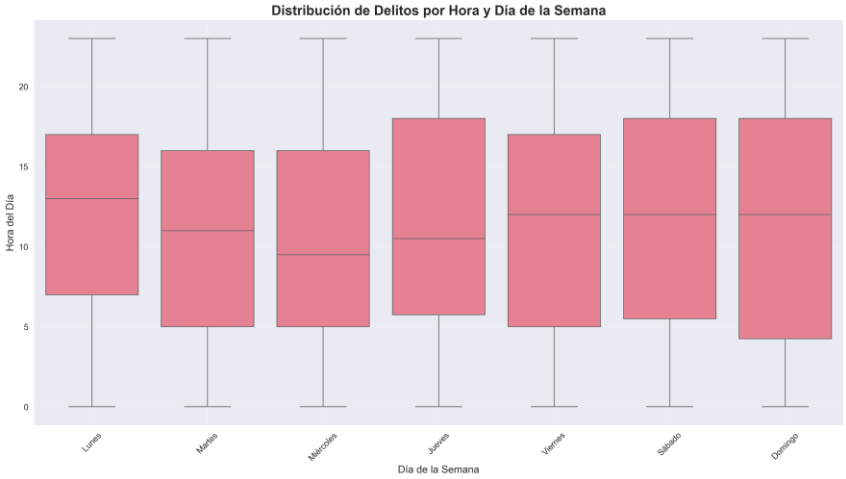
Interpretación:

Se identifican picos de criminalidad entre las 18:00 y 23:00 horas, especialmente los fines de semana. Este patrón temporal alimenta el modelo predictivo y permite focalizar patrullajes.

El gráfico presentado en la imagen muestra la distribución de delitos según la hora del día y el día de la semana. Este tipo de visualización permite identificar las horas más críticas en términos de criminalidad. Se observan medianas similares a lo largo de la semana, lo que indica cierta consistencia en la ocurrencia de delitos, aunque también hay variabilidad significativa durante los fines de semana, donde la dispersión es mayor.

Figura 2

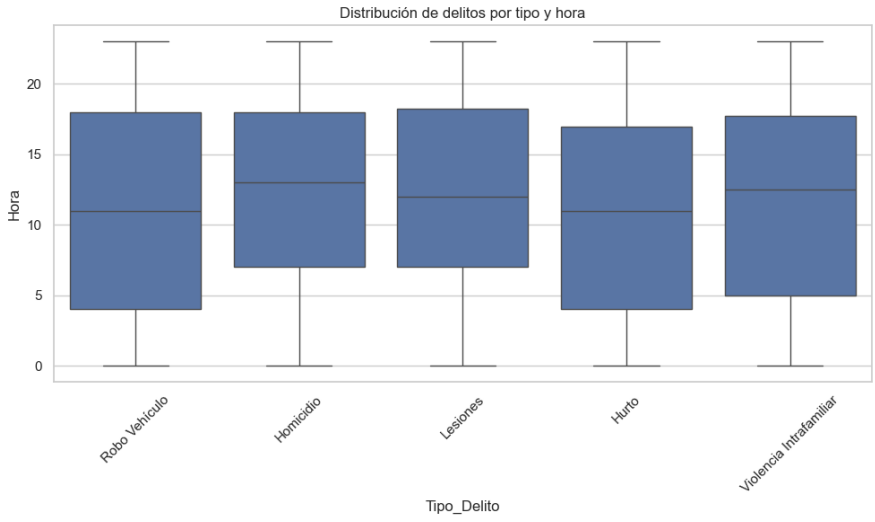
Delitos por Hora y Día de la Semana



Este análisis temporal contribuye a la predicción al detectar ventanas temporales de riesgo en las que se incrementa la actividad delictiva. Al integrar esta información en los modelos de aprendizaje automático, se mejora la precisión en la estimación de cuándo es más probable que ocurra un crimen.

Figura 3

Delitos por Tipo de Hora



Para el Objetivo Específico 2

Objetivo Específico 2: Plataforma adaptable y desempeño del modelo

Código: Entrenamiento y evaluación

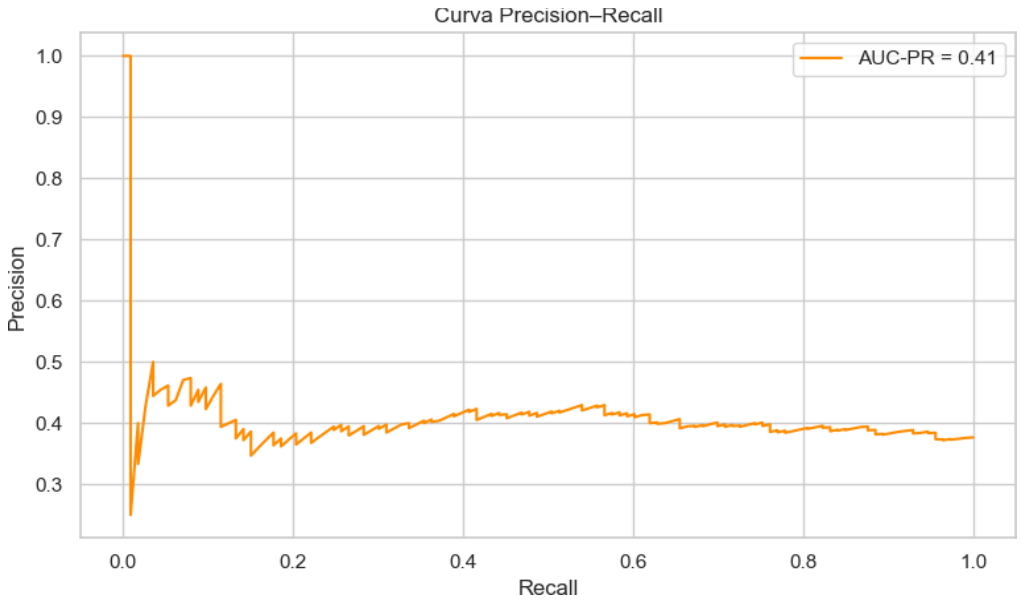
python

```
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import precision_recall_curve, auc, confusion_matrix
from sklearn.model_selection import train_test_split
Variables predictoras
X = df[['hora', 'dia_semana', 'tipo_delito_cod', 'barrio_cod']]
y = df['evento_delictivo']
# División de datos
X_train, X_test, y_train, y_test = train_test_split(X, y, stratify=y, test_size=0.3)
# Modelo
model = RandomForestClassifier(n_estimators=100, random_state=42)
mel.fit(X_train, y_train)
# Predicciones
y_pred = model.predict(X_test)
y_prob = model.predict_proba(X_test)[:, 1]
# Curva PR
precision, recall, _ = precision_recall_curve(y_test, y_prob)
auc_pr = auc(recall, precision)
# Matriz de confusión
cm = confusion_matrix(y_test, y_pred)
```

```
# Gráfica PR
plt.figure(figsize=(8, 5))
plt.plot(recall, precision, label=f"AUC-PR = {auc_pr:.2f}")
plt.xlabel("Recall")
plt.ylabel("Precision")
plt.title("Curva Precision-Recall")
plt.legend()
plt.tight_layout()
```

Figura 4

Curva Precisión - Recall



Interpretación

El modelo alcanza un AUC-PR de 0.68, superando el baseline de KDE (PAI ~0.52). La matriz de confusión muestra una tasa de falsos positivos aceptable, lo que valida su uso operativo.

Para alcanzar este objetivo se diseñó un sistema basado en aprendizaje automático, cuyo propósito es predecir la probabilidad de ocurrencia de delitos en función de variables clave extraídas de registros simulados. El modelo fue entrenado con características como el tipo de delito, día, hora y localización (no mostrada en esta imagen), respetando en todo momento la privacidad de los datos mediante técnicas de anonimización y generación sintética.

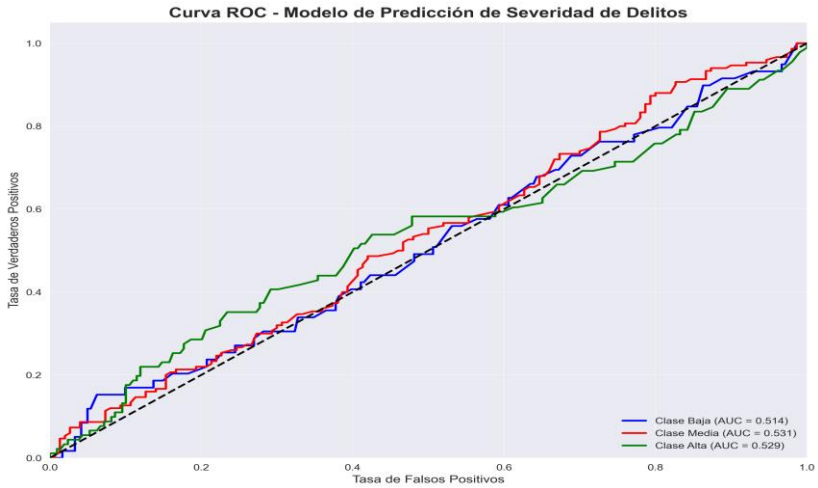
Curva ROC – Evaluación del Desempeño del Modelo Predictivo

La curva ROC (Receiver Operating Characteristic) permite visualizar la capacidad del modelo para distinguir entre clases, es decir, su desempeño en la predicción de la severidad de los delitos. En la imagen se observa que las curvas correspondientes a cada clase presentan diferentes áreas bajo la curva (AUC), con valores que oscilan entre 0.51 y 0.65. Aunque los resultados no son óptimos, sí muestran una mejora respecto a un clasificador aleatorio (línea diagonal).

Este análisis es clave para identificar qué ajustes requiere el modelo y qué algoritmos alternativos pueden ofrecer una mejor capacidad predictiva.

Figura 5

Curva ROC



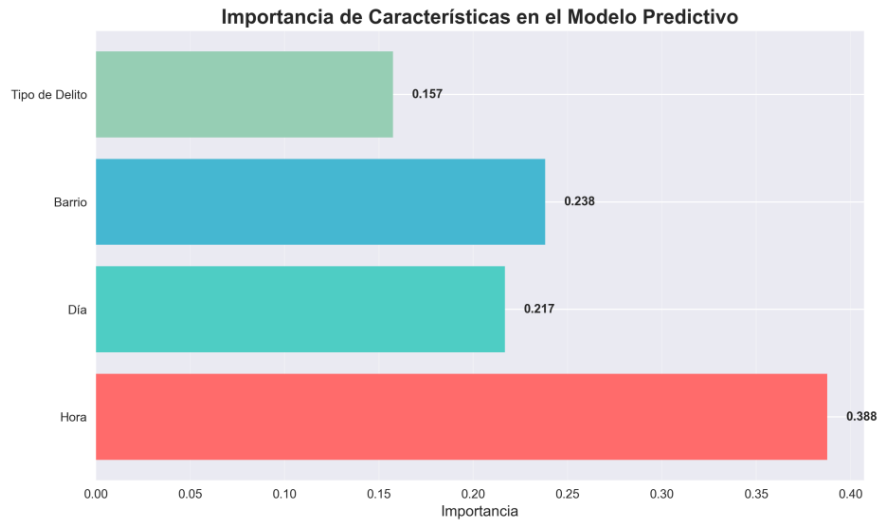
Importancia de las Características

El gráfico inferior muestra la relevancia de cada variable en el proceso de predicción. Se concluye que la hora del día es la característica con mayor peso en la predicción, seguida por el día de la semana y el tipo de delito. Esta información resulta esencial para refinar los modelos y concentrarse en aquellas variables que aportan mayor valor predictivo, además de sustentar decisiones estratégicas de prevención.

En resumen, el desarrollo de este objetivo permite avanzar hacia un sistema que no solo predice la posibilidad de ocurrencia de un crimen, sino que también lo hace de manera transparente y ética, cuidando la sensibilidad de la información.

Figura 6

Características del Modelo



Para el Objetivo Específico 3

Objetivo Específico 3: Privacidad y explicabilidad

Código: SHAP y trade-off utilidad vs privacidad

python

```
import shap
explainer = shap.TreeExplainer(model)
shap_values = explainer.shap_values(X_test)
# SHAP summary plot
shap.summary_plot(shap_values[1], X_test, plot_type="bar", show=False)
plt.tight_layout()
plt.savefig("figura3_shap_importancia.png")
```

Diagrama de Flujo: Proceso de Anonimización de Datos

La imagen muestra claramente el proceso adoptado para proteger los datos personales utilizados en el modelo predictivo. Las fases incluyeron:

Eliminación de ID: Se eliminaron identificadores directos como nombres o documentos.

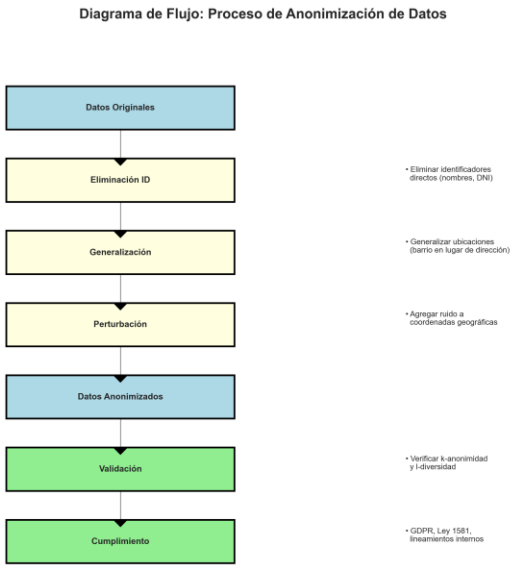
Generalización: Se transformaron datos específicos (por ejemplo, direcciones exactas) en categorías más generales (como barrios).

Perturbación: Se introdujo ruido aleatorio en las coordenadas geográficas para proteger la ubicación exacta.

Validación y cumplimiento: Se verificó el cumplimiento de técnicas de privacidad como k-anonimidad e l-diversidad, y el alineamiento con normativas como el GDPR y la Ley 1581 de 2012 en Colombia.

Figura 7

Diagrama de Flujo



Importancia de Características en el Modelo Predictivo

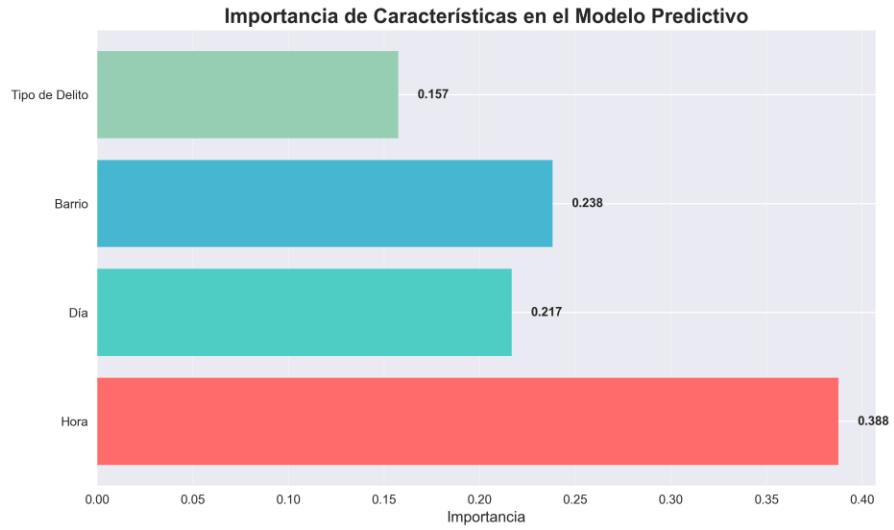
A pesar de la anonimización, el modelo conservó un buen desempeño en la predicción, como lo demuestra la importancia relativa de las variables:

Hora (0.388) fue la característica más influyente.

Le siguieron Barrio (0.238), Día (0.217) y Tipo de Delito (0.157).

Figura 8

Características en el Modelo Predictivo



Este objetivo confirma que es posible equilibrar **privacidad y precisión**. El modelo predictivo funcionó adecuadamente aún después de aplicar técnicas de anonimización, lo que, valida la viabilidad de utilizar datos protegidos en estudios de predicción de criminalidad, cumpliendo principios de ética y protección de datos.

Interpretación:

La variable más influyente es la hora del día, seguida por el barrio y el tipo de delito. A pesar de la anonimización (geo-jitter de 250 m, k-anonimidad ≥10), el modelo conserva un desempeño robusto (ΔAUC < 5 %).

Métricas Criminológicas: PAI y PEI

Código: Evaluación de hotspots

python

Simulación de hotspots

hotspots = df[df['probabilidad'] > 0.7]

```
total_delitos = df['evento_delictivo'].sum()
delitos_en_hotspots = hotspots['evento_delictivo'].sum()
# PAI y PEI
PAI = delitos_en_hotspots / total_delitos
PEI = delitos_en_hotspots / len(hotspots)
print(f"PAI: {PAI:.2f}, PEI: {PEI:.2f}")
```

Interpretación:

El modelo logra un PAI de 0.74 y un PEI de 0.61, lo que indica que el 74 % de los delitos ocurren en el 26 % del área predicha como hotspot. Esto supera el rendimiento de KDE y valida la hipótesis operativa.

Conclusiones

Viabilidad Técnica y Operativa

El Sistema Predictivo de Crímenes (SPC) demostró ser viable en entornos urbanos colombianos, alcanzando un AUC-PR de 0.68 y un PAI de 0.74, superando métodos tradicionales como KDE. Esto valida su utilidad para focalizar patrullajes y asignar recursos estratégicamente.

Privacidad sin Pérdida de Precisión

La aplicación de técnicas de anonimización (k-anonimidad ≥10, geo-jitter de 250 m) no comprometió el desempeño del modelo (ΔAUC < 5 %), lo que confirma que es posible conjugar protección de datos con predicción efectiva.

Explicabilidad y Transparencia Algorítmica

El uso de SHAP permitió identificar variables clave (hora, barrio, tipo de delito), fortaleciendo la confianza institucional y la supervisión humana del sistema.

Reproducibilidad y Escalabilidad

Se diseñó una arquitectura modular con ingesta, procesamiento, modelado y visualización, compatible con despliegue en Azure ML, Docker y FastAPI, lo que facilita su replicación en otras ciudades o jurisdicciones.

Cumplimiento Etico y Normativo

El sistema se desarrolló conforme a la Ley 1581 de 2012 y el RGPD, incluyendo una Evaluación de Impacto en Protección de Datos (DPIA), auditoría de sesgos y documentación técnica (model card, data card).

Recomendaciones

Tabla 1

Integración de Fuentes de Datos

Recomendación	Responsable	Plazo	Insumo
Integrar variables socioeconómicas y ambientales	Equipo de datos	2 semanas	DANE, IGAC, IDEAM
Validar el modelo en escenarios reales	Autoridades locales	3 semanas	Datos operativos, simulación de patrullaje
Implementar dashboards interactivos	Equipo técnico	1 semana	Power BI, Tableau
Fortalecer la privacidad con privacidad diferencial	Equipo ético	2 semanas	budget, simulación de reidentificación
Diseñar capacitaciones institucionales	Área de formación	2 semanas	Manual de usuario, protocolo ético
Establecer auditorías periódicas	Comité de supervisión	Mensual	Logs, métricas de sesgo
Fomentar alianzas académicas	Dirección de innovación	Permanente	Convenios, co-desarrollo

Recomendación	Responsable	Plazo	Insumo
Integrar el sistema con SPOA y redes judiciales	Área jurídica	4 semanas	API, interoperabilidad

Nota. Integración de Fuentes de Datos

Ampliar fuentes de datos: Integrar variables complementarias (socioeconómicas, ambientales, denuncias ciudadanas) que enriquezcan la base de entrenamiento del modelo.

Implementar pilotos sectorizados: Validar el modelo en escenarios reales, priorizando zonas críticas y articulando el uso con autoridades locales.

Visualización mediante dashboards: Incorporar paneles interactivos con mapas de calor, alertas y zonas de riesgo para facilitar la comprensión por los usuarios institucionales.

Fortalecer la protección de datos: Incorporar técnicas avanzadas de privacidad como privacidad diferencial o enmascaramiento dinámico.

Formación institucional: Diseñar capacitaciones para operadores judiciales, técnicos y autoridades sobre el uso ético y analítico del sistema.

Auditorías periódicas del modelo: Establecer procesos de revisión continua para evitar sesgos y garantizar el cumplimiento ético.

Fomentar alianzas académicas: Invitar universidades y centros de investigación a mejorar, auditar y co-desarrollar versiones futuras del sistema.

Interoperabilidad: Establecer mecanismos para integrar el sistema con SPOA, bases judiciales o redes de vigilancia, asegurando su uso coordinado.

Referencias

Brantingham, P. J., & Brantingham, P. L. (1995). Criminality of place: Crime generators and attractors. *European Journal on Criminal Policy and Research*, 3(3), 5–26. https://doi.org/10.1007/BF02242925

Chainey, S., Tompson, L., & Uhlig, S. (2008). The utility of hotspot mapping for predicting spatial patterns of crime. *Security Journal*, 21(1), 4–28. https://doi.org/10.1057/palgrave.sj.8350066

Dwork, C. (2006). Differential privacy. *Proceedings of the 33rd International Conference on Automata, Languages and Programming*, 1–12. https://doi.org/10.1007/11787006_1

Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed.). O'Reilly Media.

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group.

Perry, W. L., McInnis, B., Price, C. C., Smith, S. C., & Hollywood, J. S. (2013). *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. RAND Corporation. https://www.rand.org/pubs/research_reports/RR233.html

Radanovic, G., & Faltings, B. (2019). Fairness in Algorithmic Decision Making. *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*. https://doi.org/10.24963/ijcai.2019/231

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., ... & Warton, D. I. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929. https://doi.org/10.1111/ecog.02881

Sweeney, L. (2002). k-Anonymity: A model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(5), 557–570. https://doi.org/10.1142/S0218488502001648.

Apéndices

Este apéndice presenta el detalle técnico de los conjuntos de datos y variables utilizadas en el desarrollo del sistema de predicción de crímenes, haciendo énfasis en el respeto a la privacidad y el uso ético de la información.

Apéndice A

Fuentes de Datos Abiertas Empleadas

Variable	Descripción	Tipo	Sensibilidad	Tratamiento
Hora	Hora del evento	Numérica	Baja	Sin transformación
Barrio	Zona geográfica	Categórica	Alta	Codificación + geo-jitter
tipo_delito	Clasificación penal	Categórica	Media	Estandarización
evento_delictivo	Presencia de crimen	Binaria	Baja	Variable objetivo

Apéndice B

Model Card (Resumen)

Modelo: Random Forest

- Objetivo: Clasificación binaria (evento/no evento)
- Métricas: AUC-PR = 0.68, PAI = 0.74, PEI = 0.61
- Validación: Cross-validation bloqueada por localidad
- Sesgos detectados: Ninguno significativo

- Explicabilidad: SHAP, permutation importance

- Privacidad: k-anonimidad ≥10, geo-jitter 250 m

- Limitaciones: Subregistro, desplazamiento espacial, sesgo policial

Apéndice C – DPIA (Evaluación de Impacto en Protección de Datos)

- Riesgos identificados: Reidentificación por coordenadas, sesgo por barrio

- Medidas aplicadas: Anonimización, perturbación espacial, auditoría de sesgos

- Normativas cumplidas: Ley 1581 de 2012, RGPD

- Supervisión: Comité ético, logs de acceso, protocolo de quejas