

Prediciendo el crimen en Bogotá

Alvaro J. Riascos, Mateo Dulce, Juan Sebastián Moreno y Francisco Gómez.

Entre enero de 2012 y septiembre de 2015, todos los homicidios y un cuarto de los crímenes reportados en Bogotá ocurrieron en el 2% de los segmentos de vía de la ciudad. Sin embargo, estos sectores recibieron menos del 10% del tiempo de patrullaje policial y recursos públicos limitados (Blattman et.al 2017). Entender la formación y dinámica de estos puntos, denominados puntos calientes o *hotspots*, es indispensable para la gestión óptima de los recursos policiales y lograr un mayor impacto en la seguridad y bienestar ciudadano.

Esta nota de política resume los resultados preliminares del proyecto “Diseño y validación de modelos de analítica predictiva de fenómenos de seguridad y convivencia para la toma de decisiones en Bogotá” (véase Acerca del Estudio) y algunas de sus aplicaciones: (1) despliegue óptimo de la policía por cuadrante y turno, (2) ubicación y monitoreo óptimo de las cámaras de seguridad de la ciudad y (3) ubicación óptima de los Centros de Atención Inmediata CAI y/o estaciones de policía.

Para entender la dinámica del crimen los investigadores debemos superar varios retos: (1) los homicidios son distintos a otros crímenes (e.g., hurtos, riñas, etc.) y menos frecuentes, (2) los datos de crímenes son sesgados y sub reportados (e.g., robo de celulares), (3) donde hay más policías se descubren más crímenes, lo que puede hacer que los modelos utilizados para desplegar los recursos policiales refuercen el sesgo (i.e., sesgo de retroalimentación) en los crímenes registrados, y (4) las recomendaciones de los modelos pueden ser sesgadas en contra de algunas poblaciones (i.e., sesgo de discriminación).

Principales resultados

- El crimen es un fenómeno social que puede ser caracterizarlo a lo largo de ciertas dimensiones: espacio temporalmente, según la distribución de la infraestructura física de la ciudad, entorno ambiental y visual, etc. Por ejemplo, es posible predecir, con un día de anticipación, que en una área de menos del 5 % de la ciudad, ocurrirán aproximadamente el 42 % de los crímenes de ese día.

- Es posible generar predicciones razonables de ocurrencia de crímenes en la ciudad y de esta forma hacer una asignación más eficiente de los recursos policiales. Estos modelos son hoy el estándar en diferentes departamentos de policía en las principales ciudades del mundo. De acuerdo al ítem anterior, en un patrullaje eficiente, el 5 % del área de la ciudad debería de tener aproximadamente el 42 % de atención policial.

- Este tipo de modelos y/o algoritmos que permiten la toma de decisiones con implicaciones sobre alguna población, traen nuevos retos como son el potencial sesgo de discriminación. Reconocer estos problemas, que no son ajenos al comportamiento humano sin el uso de algoritmos, le da transparencia al proceso de toma de decisiones y arroja luces sobre las formas de corregirlo.

- Contar con un modelo de predicción de crimen razonable sugiere aplicaciones obvias de la máxima importancia para la seguridad de los ciudadanos. En particular, es posible priorizar el monitoreo de las cámaras de seguridad de la ciudad y determinar la ubicación óptima de CAIs, estaciones de policía, entre otros.

Acerca de los autores

Alvaro Riascos, Profesor Facultad de Economía de la Universidad de los Andes y Co director de Quantil.

Mateo Dulce, Estudiante de Doctorado en Aprendizaje de Maquinas y Políticas Publica en la universidad de Carnegie Mellon.

Juan Sebastián Moreno, Director Asociado del área de Minería de Datos en Quantil.

Francisco Gómez. Profesor Asociado en el Departamento de Matemáticas de la Universidad Nacional de Colombia, en Bogotá.

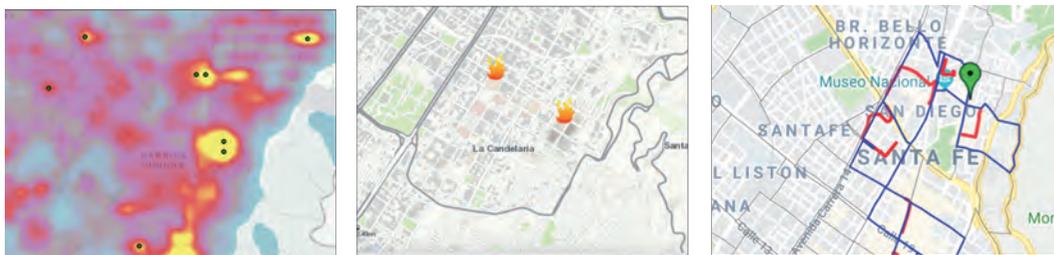
Datos y Modelos

Para abordar estos retos y proponer políticas óptimas, el estudio usa técnicas de inteligencia artificial y grandes volúmenes de información estructurada y no estructurada.¹

El crimen se concentra en el espacio: en zonas con pocos servicios públicos, de tolerancia, de entretenimiento nocturno, y áreas visualmente

deterioradas y abandonadas. También se concentra en el tiempo: hay días y horas más propensas al crimen y, para algunos crímenes como el hurto a viviendas, hay evidencia de patrones repetitivos (por ejemplo, porque los ladrones aprenden los riesgos y vulnerabilidades de los lugares donde realizan hurtos). Estos hechos estilizados los capturamos con modelos probabilísticos que identifican los puntos calientes de la ciudad, Figura 1 (véase Barreras et.al 2016 y Dulce et.al 2020).

Figura 1: Identificación de zonas calientes o hotspots de crimen

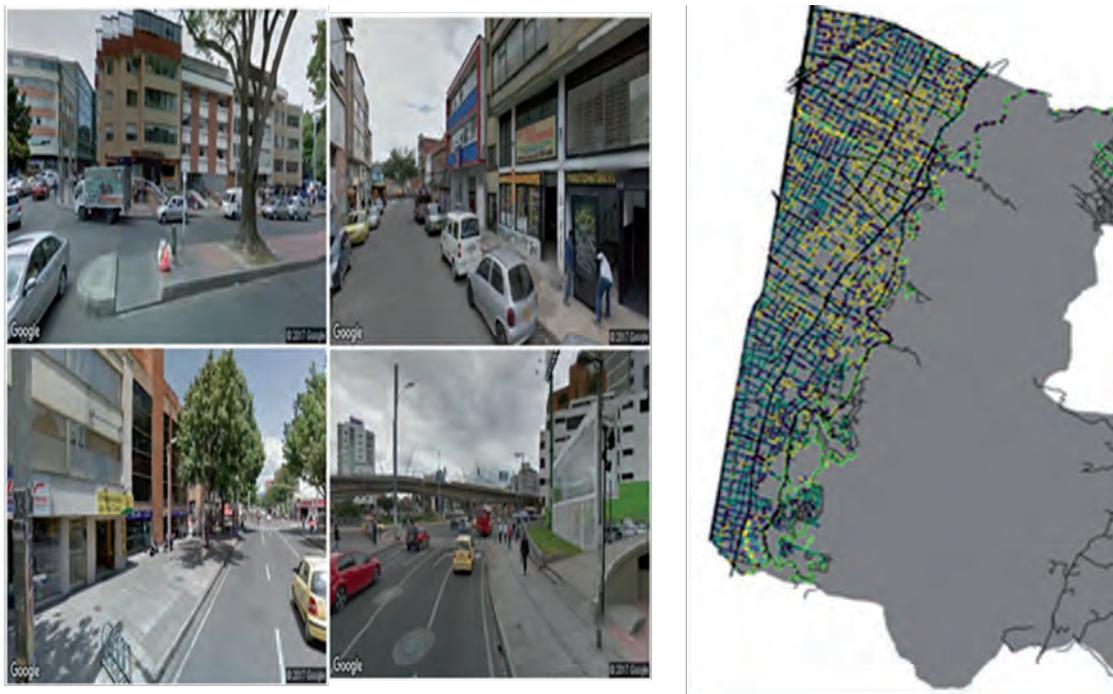


Panel izquierdo: intensidad del crimen en un día y turno policial. Panel del centro: identificación de los dos lugares más propensos al crimen. Panel derecho: identificación de los segmentos de vía más comprometidos.

La Figura 2 muestra imágenes típicas en nuestra base de datos (panel izquierdo) y la representación de la localidad de Chapinero (panel derecho). El panel derecho es el resultado de agrupar en un número reducido de

grupos las características más sobresalientes de las imágenes (vías principales, zonas verdes, luminosidad, etc.) Estas características finales entran en los modelos para predecir los puntos calientes.

Figura 2: Convirtiendo imágenes en información utilizable para predecir el crimen



Panel de la izquierda, imágenes de la ciudad descargadas usando Google Street View. Panel de la derecha representación de estas después de utilizar diferentes técnicas de inteligencia artificial (véase Dulce et.al. 2020). En particular, se agrupan las imágenes en un número reducido de grupos de acuerdo con las principales características de estas (vías principales, zonas verdes, luminosidad, etc.).

¹ Las bases de datos empleadas son: el Sistema de Información Estadística, Delincuencial, Contravencional y Operativo (SIEDCO); el Sistema de Información Red de Desaparecidos y Cadáveres (SIRDEC) del Instituto Nacional de Medicina Legal y Ciencias Forenses; infraestructura física y equipamientos en Bogotá de Catastro; información climática de la ciudad del IDEAM; e imágenes que capturan el entorno visual obtenidas de Google Street View (Descargadas usando el API de Google Street View V3.0.)

Los problemas de retroalimentación y subreportes mencionados antes pueden producir modelos con predicciones que discriminan contra algunas poblaciones. En nuestro trabajo hemos evaluado hasta qué punto los modelos de predicción tratan diferente a los estratos bajos en comparación con los estratos altos. Si bien es cierto que los modelos pueden tener consecuencias discriminatorias, es posible utilizar técnicas estadísticas que corrijan estos sesgos (Urcuqui., et.al 2020).

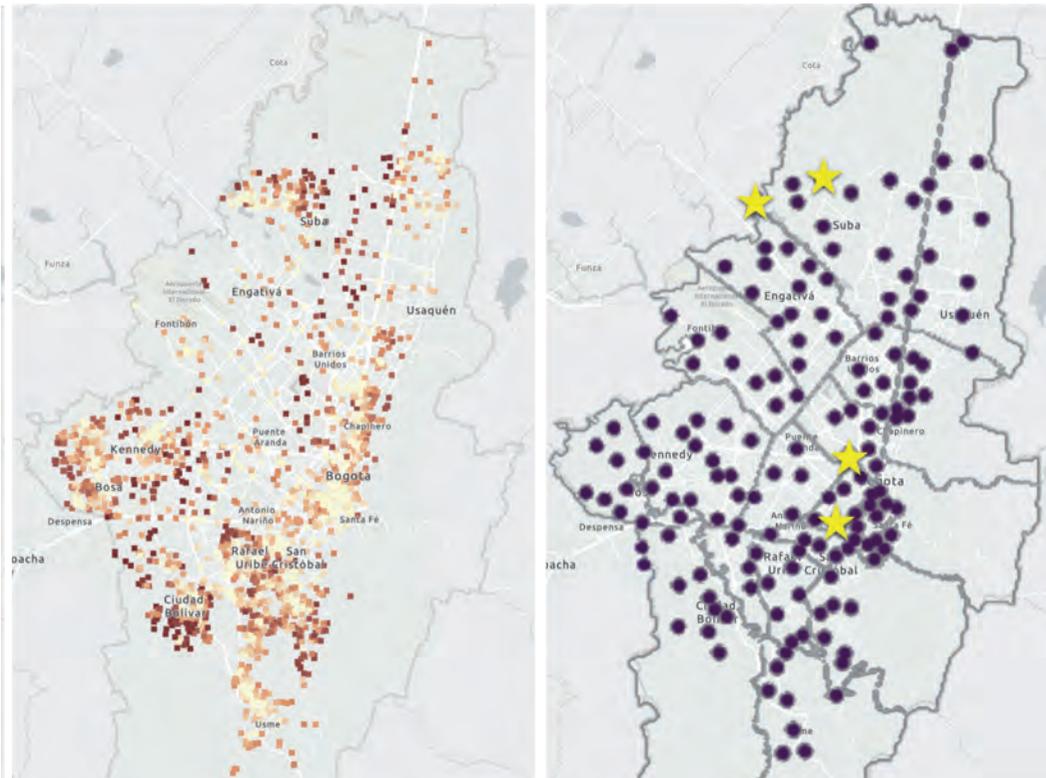
Otro reto es la baja frecuencia de crímenes como los homicidios. Para abordar este problema el estudio utiliza información adicional de crímenes relacionados (como por ejemplo, las riñas) para insertar los homicidios en un patrón más general (Moreno et.al 2020a y 2020b). De esta forma modelamos el patrón general (riñas) y de ahí deducimos características sobre los homicidios.

Aplicaciones adicionales

Algunas aplicaciones relevantes de los modelos descritos anteriormente son el monitoreo de las cámaras de seguridad de la ciudad y el análisis

de tiempos de respuesta de los CAIs o estaciones de policía a los eventos de crimen. En el primer caso, en Bogotá existen más de 3.500 cámaras de seguridad lo que hace su monitoreo detallado desde El Centro de Comando, Control, Comunicaciones y Cómputo (C4), sea muy difícil. Por eso es valioso usar un mecanismo de priorización para su monitoreo por día y periodo de tiempo. Esto puede hacerse usando el modelo de predicción de crimen. La Figura 3 muestra cómo se distribuyen algunas de las cámaras de Bogotá (i.e., cada punto representa una cámara y entre más rojo esté el punto, este representa una cámara que está en unas zonas más riesgosa). El panel derecho de la Figura 3 muestra cómo se distribuyen los CAIs de la ciudad (i.e. puntos morados) de tal forma que cubran la mayor área con el mismo tiempo de desplazamiento, de 3 a 4 minutos en bicicleta o moto. El problema es optimizar la ubicación de los CAIs ponderando por el riesgo de ocurrencia de los crímenes. El panel derecho muestra como sería la ubicación óptima de cuatro nuevos CAIs cuando se incorpora el modelo de predicción de crimen al análisis de tiempos de respuesta (i.e., estrellas amarillas en la figura).

Figura 3: Cámaras de seguridad y ubicación óptima de CAIs



Panel de la izquierda, imágenes de la ciudad descargadas usando Google Street View. Panel de la derecha representación de estas después de utilizar diferentes técnicas de inteligencia artificial (véase Dulce et.al. 2020). En particular, se agrupan las imágenes en un número reducido de grupos de acuerdo con las principales características de estas (vías principales, zonas verdes, luminosidad, etc.).

Referencias

Barreras, Francisco and Díaz, Carlos and Riascos, Álvaro y Ribero, Mónica. 2016. A comparison of different crime prediction models for Bogotá. Documentos CEDE No 34. Facultad de Economía, UNIANDES.

Dulce, M., Ramirez, S., A. Riascos. 2018. Efficient allocation of law enforcement resources using predictive police patrolling. <http://arxiv.org/abs/1811.12880>. NIPS 2018 workshop proceedings. Machine Learning for the Developing World: Achieving Sustainable Impact.

Moreno, J., Dulce, M., Riascos, A., Castaño, J., y P. Rodríguez. 2020a. A Manifold Learning Data Enrichment Methodology for Homicide Prediction. Por salir en Proceedings 7th International Conference on Behavioural and Social Computing.

Urcuqui, Ch., Moreno, J., Montenegro, C, Riascos, A., Dulce, M., 2020. Accuracy and Fairness in a Conditional Generative Adversarial Model of Crime Prediction. Por salir en Proceedings 7th International Conference on Behavioural and Social Computing.

Blattman, Ch., Green, D., Ortega, D. y Tobon S. 2017. Pushing crime around the corner? Estimating experimental impacts of large-scale security interventions. National Bureau of Economic Research.

Dulce, M., Rodriguez, P., Moreno, J., Riascos, A y J. Camargo. 2020. Street Level Images to Improve Accuracy and Address Biases in a Self-Exciting Point Process Model of Crime. Trabajo en proceso.

Moreno, J., Quintero, S. Sanchez C., Riascos, A., y L. Nonato. 2020b. Homicide prediction using sequential features from Graph Signal Processing. Trabajo en proceso.

Acerca del estudio

Esta Nota de Política reporta los avances del proyecto: Diseño y validación de modelos de analítica predictiva de fenómenos de seguridad y convivencia para la toma de decisiones en Bogotá, financiado por Colciencias con recursos del Sistema General de Regalías, BPIN 2016000100036. Los cooperantes del proyecto son la Secretaria de Seguridad, Convivencia y Justicia del Distrito de Bogotá, La Universidad Nacional de Colombia y la firma Quantil S.A.S. Las opiniones y afirmaciones en este documento no representan ni comprometen las personas o instituciones mencionadas.

Recomendaciones de política

- Recolectar todo tipo de datos descriptivos de fenómenos sociales en la ciudad y apalancarse en técnicas modernas de inteligencia artificial puede generar muchos conocimientos sobre estos fenómenos aun sobre problemas tan complejos como la seguridad ciudadana.
- En el caso particular del fenómeno de predicción del crimen, de este proyecto muestran que es posible contar con herramientas de predicción que mitiguen problemas de discriminación y a su vez sean útiles para mejorar la eficiencia de los recursos policiales.
- Este tipo de análisis tiene muchas externalidades positivas: aplicaciones diversas como las mencionadas en el documento (priorización de cámaras y ubicación optima de CAIs), pero además disciplina las discusiones de los temas sociales más complejos y promueve una cultura de datos y su análisis.

Comité editorial

Leopoldo Fergusson, director del CEDE
Hernando Zuleta, profesor de la Facultad de Economía
Darío Maldonado, director de Investigaciones, Escuela de Gobierno
María Cecilia Dedios, profesora de la Escuela de Gobierno

En esta edición

Leopoldo Fergusson, editor.
David Bautista, diagramación.



Escuela de Gobierno
Alberto Lleras Camargo

Facultad de Economía

CEDE
Center for Economic and Development Research