

# **Fortalecimiento del SAT mediante aprendizaje de máquinas en la detección de violencia atípica**

**Economía y Derecho**

**Juan Diego Heredia Niño**

Facultad de Economía  
Universidad de los Andes

22 de octubre de 2025



# Objetivo de estudio

---

¿Pueden los **modelos de aprendizaje automático** mejorar la detección de **eventos atípicos de violencia** asociados al conflicto armado, y así fortalecer el **Sistema de Alertas Tempranas** de la Defensoría del Pueblo?



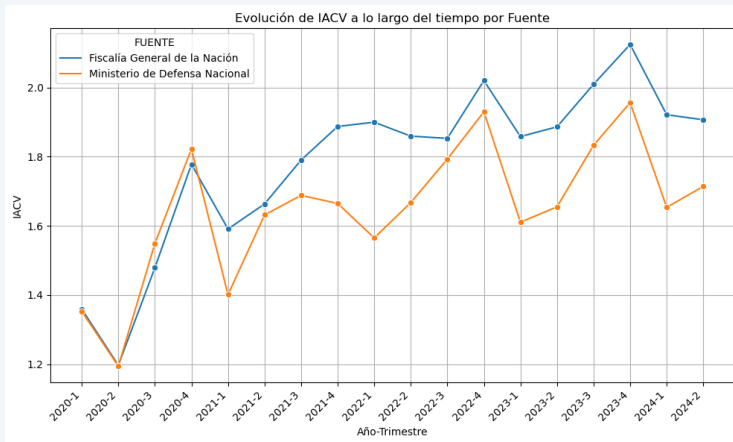
- El **Sistema de Alertas Tempranas (SAT)** es una herramienta institucional de monitoreo y advertencia que identifica situaciones de riesgo frente a violaciones de derechos humanos en el marco del conflicto armado en Colombia.
- Estas situaciones de riesgo se pueden entender como **eventos atípicos de violencia** o **picos de violencia** en determinados territorios y periodos de tiempo.
- La idea de esta tesis es **cuantificar y pronosticar estos eventos** a través de técnicas de aprendizaje automático, con el fin de **complementar y fortalecer el SAT**.



# Motivación

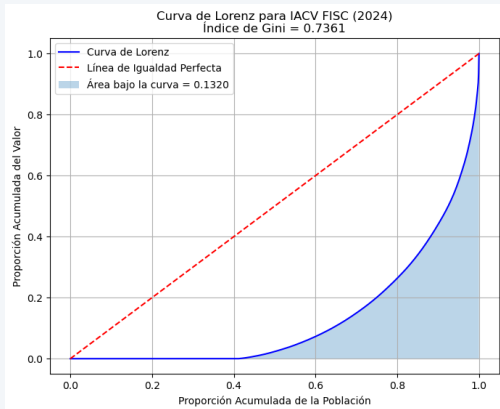
---

# Evolución de la Violencia en Colombia

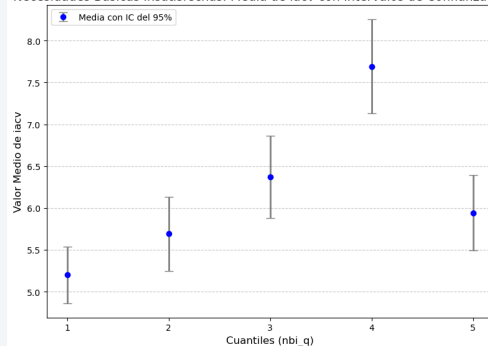


*Fuente: Datos MinDefensa. Elaboración propia*

# Vulnerabilidad municipal a la violencia



Necesidades Básicas Insatisfechas: Media de iacv con Intervalos de Confianza (95%)



*Fuente: Datos Panel CEDE y MinDefensa. Elaboración propia*



## ¿Por qué importa esta pregunta?

- La violencia atípica tiene **alto impacto social** y **poca capacidad de respuesta**.
- Las herramientas actuales (SAT) son cualitativas y pueden no detectar patrones emergentes.

## ¿Qué propone esta tesis?

- Usar **aprendizaje automático** para predecir eventos críticos.
- Complementar el SAT con una herramienta cuantitativa y replicable.





# Resultados preliminares

---

# ¿Ayuda el aprendizaje automático?



## Hallazgo principal (preliminar)

Los modelos de aprendizaje automático **mejoran la capacidad de detección de violencia atípica** frente a métodos tradicionales:

- **Mayor precisión global** (hasta 85.4 % con Random Forest).
- **Mejor equilibrio** entre sensibilidad y precisión ( $F1 \approx 0.75$ ).
- **AUC elevado** ( $\approx 0.89$ ), lo que permite **priorizar municipios según su riesgo estimado**, abriendo la puerta a un sistema de alerta más focalizado y preventivo.

## Advertencia

Estos resultados son **preliminares** y están sujetos a revisión en etapas posteriores.



# Revisión de literatura

---



## ¿Qué dice la literatura?

- **Econometría tradicional:** relaciones causales entre violencia y pobreza, choques económicos, ausencia estatal.  
*Goldstone et al. (2010), Hegre et al. (2013), Dube et al. (2019), Levitt & Rubio (2000), Bourguignon et al. (2003)*
- **Aprendizaje automático:** mejora la precisión, detecta patrones no lineales, predice zonas violentas.  
*Muchlinski et al. (2016), Hegre et al. (2019), Bazzi et al. (2022)*
- **Colombia:** énfasis en posconflicto, choques económicos, patrones territoriales.  
*Meisel & Vega (2021), Rojas Guerrero & Grautoff (2019)*

# Limitaciones y contribución



## ¿Qué limitaciones enfrentan?

- Dificultad para predecir **eventos atípicos o escaladas inesperadas**.
- Falta de **integración institucional** y problemas de interpretabilidad.
- Calidad de datos limita la eficacia en amenazas o extorsión.

## ¿Qué aporta esta tesis?

- Se enfoca en detectar **violencia atípica**, no solo patrones promedio.
- Evalúa modelos por su **capacidad para anticipar escaladas de violencia**.
- Integra herramientas cuantitativas al **SAT de la Defensoría del Pueblo**.



# Definiciones

---

# Violencia agregada: el índice IACV



## ¿Qué es el IACV?

El **Índice Agregado de Casos Violentos (IACV)** resume la intensidad de violencia a nivel municipal por trimestre.

**Propósito:** Medida de seguridad pública agregada

| Delito     | Pena (años) | Peso (%) |
|------------|-------------|----------|
| Homicidio  | 19.0        | 17.04    |
| Extorsión  | 11.5        | 10.31    |
| Secuestro  | 16.0        | 14.35    |
| Terrorismo | 15.0        | 13.45    |
| Masacres   | 50.0        | 44.84    |

# Fórmula del IACV



## Definición

$$IACV_{t,m} = \frac{0,17 \cdot \text{Hom}_{t,m} + 0,10 \cdot \text{Ext}_{t,m} + 0,14 \cdot \text{Sec}_{t,m} + 0,13 \cdot \text{Terr}_{t,m} + 0,44 \cdot \text{Masc}_{t,m}}{\text{Población}_{t,m}}$$

## Ponderaciones

- Reflejan el impacto social relativo de cada tipo de hecho.



# ¿Qué otros tipos de violencia definimos?



## Índice de Amedrentamiento (IA)

- Amenazas
- Tentativas de asesinato y atentados
- Desplazamiento forzado
- Hostigamiento

**Propósito:** Medir clima de miedo

## Índice de Gobernanza Criminal (IGC)

- Confinamientos
- Retenes ilegales
- Paros armados
- Extorsión

**Propósito:** Medir control territorial

# ¿Qué es la violencia atípica?



Inspirado en Bazzi et al. (2022)

Se define como un **nivel de violencia que supera el umbral promedio más una desviación estándar**, con respecto a su historial reciente.

## Criterio formal

$$\text{Violencia Atípica}_t = \begin{cases} 1 & \text{si } IACV_t \geq I\bar{ACV}_{t-1:t-12} + \sigma_{t-1:t-12} \\ 0 & \text{en otro caso} \end{cases}$$

- Detecta picos inusuales a partir de **comportamientos históricos del municipio**.
- Sirve como variable objetivo para los modelos de predicción.



# Datos

---

# Fuentes de Datos



Los datos de violencia provienen de tres fuentes principales:

- **Fiscalía General de la Nación:** Registros mensuales de homicidio, extorsión, secuestro, terrorismo y masacres a nivel municipal, mensual (2014-2024).
- **Ministerio de Defensa Nacional:** Serie histórica de los mismos delitos con mayor cobertura temporal a nivel municipal, mensual(1997-2024)
- **Jurisdicción Especial para la Paz (JEP):** Datos sobre presencia de grupos armados y eventos violentos (2017-2024), incluyendo desplazamientos, hostigamientos y paros armados.

# Fuentes de Datos



Para analizar la relación entre violencia y factores estructurales, se integran:

- **Panel Municipal del CEDE** (2005–2023): Información demográfica, socioeconómica e institucional, como **pobreza, acceso a servicios y programas para víctimas**.
- **Cultivos ilícitos de coca** (1999–2023, Observatorio de Drogas de Colombia): Financiación de grupos armados y la **dinámica del conflicto**.
- **Luminosidad nocturna (VIIRS Nighttime Light)** (2012–2023): Indicador proxy de **actividad económica local**.



## Objetivo del modelo

Predecir la **probabilidad de que ocurra un evento de violencia atípica** en un municipio en un trimestre determinado.

## ¿Por qué clasificación y no regresión?

- El interés está en una variable binaria: violencia atípica (sí/no).
- La clasificación permite estimar **probabilidades asociadas a riesgo**.
- Los modelos pueden ser **ajustados (fine-tuned)** para priorizar sensibilidad o precisión.
- Fáciles de integrar a esquemas de priorización o alertas.

Estimar  $P(\text{Violencia Atípica}_t = 1 | \vec{X})$

$$\text{ViolenciaAtípica}_{t,m} = f(\text{IACV}_{t-1:12,m}, \text{IGC}_{t-1:4,m}, \text{IA}_{t-1:4,m}, X_{t,m}, )$$

- **Variables explicativas:**

- **Factores estructurales:**

- Desigualdad, presencia estatal, economías ilegales, empleo, presencia grupos armados, educación, ubicación, etc.
    - Si tiene o no cultivos de coca en el periodo pasado
    - Si tiene o no presencia de un grupo criminal en el periodo pasado

- **Rezagos temporales de violencia:**

- IACV, IGC e IA

## Descripción

- Modelo lineal penalizado que selecciona automáticamente las variables más relevantes.
- Supone que **cada variable influye de forma constante y predecible** sobre el riesgo de violencia atípica: si una variable aumenta, el riesgo cambia proporcionalmente.
- Favorece interpretabilidad y evita sobreajuste.

| % Acierto | % Sensibilidad | % Especificidad | AUC   | F1    |
|-----------|----------------|-----------------|-------|-------|
| 72        | 74.3           | 70.7            | 0.811 | 0.660 |



# Random Forest



## Descripción

- Conjunto de árboles de decisión entrenados sobre muestras aleatorias.
- Captura relaciones **no lineales** y **interacciones complejas** entre variables.
- No requiere supuestos paramétricos fuertes.

| % Acierto | % Sensibilidad | % Especificidad | AUC   | F1    |
|-----------|----------------|-----------------|-------|-------|
| 85.4      | 60.4           | 99.7            | 0.898 | 0.750 |

**¡Gracias!**

# Referencias



- Banco Interamericano de Desarrollo. (2017). *Los costos del crimen y la violencia*. Banco Interamericano de Desarrollo. <https://doi.org/10.18235/0006383>
- Bazzi, S., Blair, R. A., Dube, O., Gudgeon, M., & Peck, R. (2022). *The promise and pitfalls of conflict prediction: Evidence from Colombia and Indonesia*. *The Review of Economics and Statistics*, 104 (6), 1246–1262. [https://doi.org/10.1162/rest\\_a\\_01172](https://doi.org/10.1162/rest_a_01172)
- Bourguignon, F., Núñez, J., & Sánchez, F. (2003). *What part of the income distribution does inequality affect crime? The case of Colombia*. Documento CEDE, Universidad de los Andes.
- Dube, O., & Vargas, J. F. (2013). *Commodity price shocks and civil conflict: Evidence from Colombia*. *Review of Economic Studies*, 80(4), 1384–1421. <https://doi.org/10.1093/restud/rdt009>
- Feldmann, A., & Hinojosa, V. (2009). *Terrorism in Colombia: Logic and sources of a multidimensional and ubiquitous phenomenon*. *Terrorism and Political Violence*, 21 (1), 42–61. <https://doi.org/10.1080/09546550802544694>
- Levitt, S., & Rubio, M. (2000). *Understanding crime in Colombia and what can be done about it*. Documento de trabajo, Universidad de Chicago/Fedesarrollo.
- Londoño, J. L., & Guerrero, R. (1999). *Violencia en América Latina: Epidemiología y costos*. Banco Interamericano de Desarrollo.
- Mejía, D., & Restrepo, P. (2011). *The economics of the war on illegal drug production and trafficking*. Documento CEDE, Universidad de los Andes.
- Rubio, M. (2003). *El rapto de la pesca milagrosa: Breve historia del secuestro en Colombia*. Documentos CEDE No. 2003–36, Universidad de los Andes.
- Sánchez, F., & Chacón, M. (2005). *Conflicto, Estado y descentralización: Disputa armada por el control local, 1974–2002*. Documento CIDER/Universidad de los Andes – LSE.