

Predicción de asesinatos de líderes sociales en Colombia

Nicolás García y Carolina Dueñas

Agosto 2, 2022

[Carpeta trabajo final GitHub](#)

Abstract

El presente artículo busca dar cuenta de la probabilidad de que un líder social que ha sido amenazado, sea asesinado en Colombia, a partir de datos novedosos construidos por la Fundación Ideas para la Paz. Los resultados muestran que el mejor modelo de predicción es un modelo logit lasso, el cual encuentra el λ óptimo (1.02) para maximizar la sensibilidad (34.6%). Finalmente se recomienda introducir variables que estén directamente relacionadas con la seguridad, en especial índices de presencia de fuerza pública en los municipios, sistemas de vigilancia, entre otros, y variables que estén relacionadas con el contexto social como, la posición ideológica del individuo, el grado de confianza con la comunidad y la participación ciudadana, con el objetivo de mejorar la predicción del modelo.

Introducción y justificación

Desde la firma del Acuerdo de Paz, se han registrado alrededor de 1,327 líderes sociales y firmantes de paz asesinados (Indepaz, 2022) por múltiples actores, como paramilitares, grupos armados y actores desconocidos.

Según Human Rights Watch (2021), las zonas donde tuvieron lugar los asesinatos de líderes sociales tienen algunos elementos en común: se trata de municipios predominantemente rurales, donde operan grupos armados, hay presencia de economías, tala y minería ilegal, con altas tasas de homicidios, altos niveles de pobreza, y que han sido históricamente golpeados por el conflicto. Además, han sido escenario de la desmovilización de actores armados como las FARC.

En este mismo sentido, Marín (2021) evidencia que los acuerdos comunitarios del Programa Nacional integral de Sustitución de Cultivos Ilícitos (PNIS) influyeron en el asesinato de líderes sociales, y que la probabilidad de un asesinato mensual aumentó en 1.5 puntos porcentuales. En términos porcentuales, este efecto refleja un aumento en 546% y 167% en la tasa y en la probabilidad de asesinato, respectivamente. Sin embargo, aún se desconoce si existen otros factores contextuales y demográficos (como el género, la edad, la etnia o la profesión), que aumentan las probabilidades de que un(a) líder social sea asesinado(a) después de recibir una amenaza.

Protection International (2021) señala que, tras una amenaza, se debe hacer una suposición sobre la disposición del agresor de actuar según sus intereses, así como la evaluación de la amenaza desde un enfoque de género e interseccional “para analizar su impacto diferencial, teniendo en cuenta condiciones interseccionales como el sexo, la identidad de género, la edad, la pobreza, etc” (p.12). Por tanto, la organización internacional considera que es de carácter urgente desarrollar una investigación aplicada para comprender la interrelación entre las amenazas, agresiones y asesinatos contra los defensores de los derechos humanos en diferentes escenarios.

Según la Unidad Nacional de Protección (UNP), el 49% de las personas protegidas son líderes sociales que corren algún tipo de peligro. Lograr identificar los municipios o características sociodemográficas que inciden en que un líder tenga una mayor probabilidad de ser asesinado permitiría brindar recomendaciones de política pública para direccionar esfuerzos hacia determinadas poblaciones en municipios particulares. Igualmente, le permitiría a la UNP priorizar a ciertas personas, o incluso crear un sistema de alertas tempranas más efectivo.

Datos

Los datos utilizados que dan cuenta de las amenazas y asesinatos contra líderes sociales fueron contruidos por la [Fundación Ideas para la Paz \(FIP\)](#) a partir de la revisión de prensa nacional, de donde se extrajo la información básica de cada líder social para los años de 2010 hasta el 2022. El *Anexo 1* describe las características demográficas de los líderes sociales que se incluyeron en el presente estudio, como el género y el tipo de actor (social, indígena, comunal, afrodescendiente, entre otros).

Por otro lado, se incluyeron variables de los municipios de los líderes, las cuales podrían tener relación con su situación de seguridad, como lo son la presencia de cultivos ilícitos, el número de hectáreas deforestadas, el número de víctimas de minas antipersona y munición sin explotar (común en zonas de conflicto en Colombia), y la presencia de programas estatales como los son el Programa Nacional de Sustitución de Cultivos (PNIS), y los Programas de Desarrollo con Enfoque Territorial (PDET). Cabe resaltar que estas variables fueron obtenidas de diferentes fuentes de acceso público, como el Departamento Nacional de Planeación y el Sistema Integrado de Monitoreo de Cultivos Ilícitos (SIMCI). Todas estas características podrían guardar relación con la probabilidad de asesinato, en tanto que, según la literatura revisada, son los líderes sociales que abogan por la sustitución de cultivos, las causas ambientales y luchan contra la corrupción quienes más se sitúan en el blanco de los actores armados.

Por último, incluimos variables relacionadas con el desempeño de los municipios e indicadores de seguridad, ya que la presencia estatal ha estado históricamente correlacionada con una situación de seguridad favorable. En este sentido, los líderes provenientes de los municipios más alejados y con menor capacidad institucional tendrían una mayor probabilidad de ser asesinados, por la misma falta de capacidad de las instituciones encargadas de mantenerlos a salvo.

La *Figura 1* demuestra que, en general, los líderes sociales amenazados no son en su gran mayoría mujeres sino, por el contrario, hombres o personas LGTBI. En cuanto a las características municipales, los municipios que no pertenecen a programas estatales como el PDET y el PNIS son los que más asesinatos y amenazas registran. Sin embargo, factores como las hectáreas de coca, lla deforestación y la cantidad de víctimas de minas antipersona sí parecen ser factores determinantes en el asesinato de líderes. Por último, las variables que no muestran grandes diferencias son seguridad ciudadana y medición de desempeño municipal.

Figura 1.

Characteristic	0, N = 2,181 ¹	1, N = 1,196 ¹
genero		
0	1,728 (79%)	1,040 (87%)
1	453 (21%)	156 (13%)
PDET		
0	1,325 (61%)	548 (46%)
1	856 (39%)	648 (54%)
PNIS		
0	1,938 (89%)	902 (75%)
1	243 (11%)	294 (25%)
ha_coca	270 (1,330)	877 (3,073)
seguridad_convivencia	83 (10)	86 (8)
MDM_2020	59 (16)	54 (12)
map_muse	0.67 (3.15)	1.92 (6.67)
tc_loss_ha	454 (1,238)	833 (2,704)
¹ n (%); Mean (SD)		

Esta distribución de los asesinatos según las variables mencionadas puede brindar unos primeros indicios sobre las variables que podrían resultar más significativas en el modelo.

Modelo

Como se mencionó anteriormente el objetivo de la presente investigación es predecir la probabilidad de que un líder social que ha sido amenazado, sea asesinado en Colombia, la predicción se realizó mediante un análisis a nivel de personas en donde la variable dependiente es una *dummy* que toma el valor de 1 cuando el líder social es asesinado después de ser amenazado, y 0 cuando ha sido amenazado, pero no es asesinado.

Inicialmente para el presente análisis se unieron todas las variables descritas en el anexo 1 en una sola muestra con el fin de corregir las clases de las variables, imputar valores y en general hacer la limpieza de los datos. Seguido de esto se dividió la base de datos en las submuestras de entrenamiento, evaluación (para evaluar el threshold) y prueba.

Con la base de entrenamiento se analizaron nueve (9) modelos de clasificación, el primero de ellos es un modelo logit (logit_caret) que arroja la mayoría de métricas para los modelos de clasificación; este modelo arrojó una sensibilidad de 0.91, una especificidad de 0.18, una precisión de 0.65 y un ROC de 0.64.

El siguiente modelo fue nombrado “logit_lasso_sens”, el cual calcula el hiperparámetro (lambda) óptimo que maximiza la sensibilidad del modelo, en este caso el modelo arrojó un lambda de 1.02 con una sensibilidad máxima del 99%. El tercer modelo “logit_lasso_roc”, el cual calcula el hiperparámetro (lambda) óptimo que maximiza el ROC (área bajo la curva) del modelo, en este caso el modelo arrojó un lambda de 0.082 con un ROC máximo del 0.64. El siguiente modelo se realizó con el fin de equilibrar la muestra por medio de la función upsample, en este caso el número de homicidios se incrementó dando un hiperparámetro (lambda) óptimo de 0.109. El siguiente modelo se realizó utilizando la función downsample con el fin de equilibrar la muestra, en este caso el número de no pobre disminuyó dando un hiperparámetro (lambda) óptimo de 0.190. También se realizó una combinación de árboles predictores por medio de la metodología de random forest encontrando una sensibilidad de 0.88, una especificidad de 0.35, una precisión de 0.69 y un ROC de 0.76. Finalmente se realizaron dos modelos boosting, adaboost y xgboost.

Luego de evaluar todos los anteriores modelos en la muestra de prueba, el modelo “logit_lasso_sens” obtuvo una mayor sensibilidad y un mayor número de verdaderos positivos, por lo tanto, fue el modelo seleccionado para el presente análisis.

$$\text{homicidio} = \beta_1 + \beta_2 \text{genero} + \beta_3 \text{PDET} + \beta_4 \text{PNIS} + \beta_5 \text{ha_coca} + \beta_6 \text{seguridad_convivencia} \\ + \beta_7 \text{MDM_2020} + \beta_8 \text{map_muse} + \beta_9 \text{tc_loss_ha}$$

Es importante mencionar que este modelo es el que mejor predice entre todos los que se analizaron en la presente investigación, sin embargo, es posible que al explorar otras formas funcionales y/o combinaciones entre modelos se consiga una mejora en la precisión del modelo.

Resultados

El modelo seleccionado (logit_lasso_sens) calcula el hiperparámetro (lambda) óptimo que maximiza la sensibilidad, en este caso el modelo arrojó un lambda de 1.02. De igual manera, al evaluar el modelo en la muestra de prueba y al verificar la matriz de confusión se pudo observar que este modelo predice 230 valores verdaderos positivos; a continuación, se evidencia la matriz de confusión para cada uno de los modelos analizados en el presente estudio.

logit			lasso			lasso_Sens		
homicidio	NO	SI	homicidio	NO	SI	homicidio	NO	SI
NO	33	403	NO	25	411	NO	1	435
SI	47	191	SI	37	201	SI	8	230

lasso_thresh			lasso_upsample			lasso_downsample		
homicidio	NO	SI	homicidio	NO	SI	homicidio	NO	SI
NO	200	236	NO	165	271	NO	166	270
SI	155	83	SI	134	107	SI	131	107

forest			adaboost			xgboost		
homicidio	NO	SI	homicidio	NO	SI	homicidio	NO	SI
NO	40	396	NO	120	316	NO	14	422
SI	114	124	SI	120	118	SI	25	213

En la muestra de prueba, este modelo obtuvo el valor de la sensibilidad más alto, el cual fue de 34.6%; este cálculo se realizó por medio de los valores de la matriz de confusión para cada uno de los modelos analizados, como se puede observar en la siguiente tabla (Anexo al presente documento se encontrará un archivo en formato .csv en donde se pueden observar las predicciones los modelos mencionados).

Modelo	Sensibilidad
logit	32,15%
lasso	32,84%
lasso_Sens	34,58%
lasso_thresh	26,01%
lasso_upsample	27,73%
lasso_downsample	28,38%
forest	23,84%
adaboost	27,18%
xgboost	33,54%

Conclusiones y recomendaciones

El modelo *Logit lasso sens* obtuvo un desempeño predictivo favorable; este podría ser un punto de partida para dar recomendaciones de política pública a los organismos del Estado encargados de la seguridad de los líderes sociales, implementando métodos de *Machine Learning* para otorgarle prelación a ciertos individuos que se encuentran amenazados. Esto podría no solamente salvar muchas vidas, sino optimizar de la mejor manera los recursos del Estado, los cuales son escasos.

Sin embargo, con el objetivo de mejorar la predicción del modelo, sería ideal introducir variables que estén directamente relacionadas con la seguridad, en especial índices de presencia de fuerza pública en los municipios, sistemas de vigilancia, entre otros, de igual manera es recomendable hacer uso de variables que estén relacionadas con el contexto social como, la posición ideológica del individuo, el grado de confianza con la comunidad, la participación ciudadana, etc.

Finalmente, se recomienda explorar otras formas funcionales y/o combinaciones entre modelos con el fin de mejorar la precisión del modelo; como por ejemplo evaluar los modelos “adaboost” y “xgboost” con la submuestra upsample o downsample.

Bibliografía

HRW (2021). Left Undefended: Killings of Rights Defenders in Colombia's Remote Communities. Recuperado de:

Indepaz (2022). Boletín Informe Estadístico de Homicidios de Líderes Sociales y Defensores de DDHH. Recuperado de: <https://indepaz.org.co/category/informes/>

Martín, L. (2021). El asesinato de líderes sociales: efecto no intencionado de la sustitución de cultivos de uso ilícito. Bogotá: CESED. Recuperado de: <https://cesed.uniandes.edu.co/el-asesinato-de-lideres-sociales-efecto-no-intencionado-de-la-sustitucion-de-cultivos-de-uso-ilicito/>

Protection International (2021). Understanding Death Threats Against Human Rights Defender: Reflection Paper. Recuperado de: https://www.protectioninternational.org/sites/default/files/understanding_death_threats_against_human_rights_defenders_aa_-_en.pdf

ANEXOS

Anexo 1

Variable	Definición	Fuente de datos
Líderes amenazados	Número de líderes sociales que han recibido una amenaza	Base de prensa, Fundación Ideas para la Paz
Líderes asesinados	Número de líderes sociales que han sido asesinados	Base de prensa, Fundación Ideas para la Paz
Hectáreas reportadas de cultivos ilícitos	Número de hectáreas de coca a nivel municipal registradas por UNODC mediante reconocimiento aéreo.	SIMICI
Hectáreas deforestadas	Número de hectáreas de pérdida de bosque a nivel municipal	Global Forest Watch, Maryland
Medición de desempeño municipal	Índice que mide y compara el desempeño municipal entendido como la gestión de las Entidades Territoriales y la consecución de resultados de desarrollo.	Departamento Nacional de Planeación (DNP)
Medición de seguridad y convivencia municipal	Índice que mide y compara la seguridad y convivencia municipal.	Departamento Nacional de Planeación (DNP)
Municipio PNIS	Variable <i>dummy</i> que indica si el municipio pertenece al Programa Nacional de Sustitución de Cultivos Ilícitos (PNIS)	United Nations Office on Drugs and Crime (UNODC)
Municipio PDET	Variable <i>dummy</i> que indica si el municipio pertenece al Programa de Desarrollo con enfoque territorial (PDET)	Agencia de Renovación del Territorio (ART)