

Article

Predicción de robos de vehículos en la ciudad de Bogotá por medio del lenguaje de Python

Johann Forero¹

¹ Universidad Libre de Colombia; johanns-foreros@unilibre.edu.co

Received: 20/03/2023; Accepted: 10/05/2023; Published: 30/06/2023

Abstract: Insecurity is still the daily life of Bogotá and car thefts seem to be increasing. In this way, it became evident in the last security report, in the last months of 2021, the cases of car theft increased by 340 new cases, in addition to the complaints of citizens who have suffered the theft of their cars while driving or, even, at the doors of their houses while parking. According to the figures provided by the Security Secretariat, in the last months of 2021, 3,331 cases of vehicle theft have been reported in the city, which represents an increase of 6.4% compared to the figures for the same period in 2020. The situation is even more dramatic when this year's figures are compared with those of a completely normal year, such as 2019, where the variation between the two is 9% more. Looking at the figures broken down by month, throughout this year, in June there were 76 more cases than on the same date in 2020, that is, the total number of stolen cars in the sixth month of the year was 344. In July, the difference with last year was 29 cases, for a total of 319 thefts. In August and September, the figures were also better than those of 2020, with year-on-year differences of 21 and 27 cases, respectively. October and November were stable months, since the number of vehicle thefts has been the same in both years, with 370 and 340 for all said periods with python, a pre-processing of a data based on previous years and in the localities with the highest percentage of car thefts is carried out, granted by the page Secretaría Distrital de Seguridad, Convivencia y Justicia.

Keywords: Robo, Automovil, data, pre-procesamiento, alistamiento.

1. Introducción

Este proyecto busca solucionar un problema en la sociedad que cada día es más preocupante que se trata del robo de vehículos con un sistema de predicción [1] para poder tener una idea de cuanta puede llegar a ser la posibilidad de que su vehículo llegue a ser hurtado [2], reconocer cuales son las localidades más inseguras en la ciudad, donde corre más peligro, modalidades de robo, horas de los robos, la solución propuesta aquí se enmarca dentro de seguridad social [3], los datos recolectados fueron por la Secretaría Distrital de Seguridad, Convivencia y Justicia.

LOCALIDADES Y UPZ	ENE-ENE 2020	ENE-ENE 2021	Variacion % ENE - ENE2021	ENE 2020	ENE 2021	Variacion % ENE2021
▶ ANTONIO NARIÑO	9	8	-11,11	9	8	-11,11
▶ BARRIOS UNIDOS	8	3	-62,50	8	3	-62,50
▶ BOSA	19	21	10,53	19	21	10,53
▶ CANDELARIA	5	1	-80,00	5	1	-80,00
▶ CHAPINERO	8	0	-100,00	8	0	-100,00
▶ CIUDAD BOLÍVAR	13	20	53,85	13	20	53,85
▶ ENGATIVÁ	50	25	-50,00	50	25	-50,00
▶ FONTIBÓN	16	7	-56,25	16	7	-56,25
▶ KENNEDY	58	50	-13,79	58	50	-13,79
▶ LOS MÁRTIRES	3	12	300,00	3	12	300,00
▶ PUENTE ARANDA	22	14	-36,36	22	14	-36,36
▶ RAFAEL URIBE URIBE	18	4	-77,78	18	4	-77,78
▶ SAN CRISTÓBAL	15	3	-80,00	15	3	-80,00
▶ SANTA FE	4	1	-75,00	4	1	-75,00
▶ SUBA	19	12	-36,84	19	12	-36,84
▶ TEUSAQUILLO	7	10	42,86	7	10	42,86
▶ TUNJUELITO	8	3	-62,50	8	3	-62,50
▶ USAQUÉN	9	4	-55,56	9	4	-55,56
▶ USME	6	4	-33,33	6	4	-33,33
TOTAL GENERAL	297	202	-31,99	297	202	-31,99

Tabla 1: Localidades con mayor robo de vehículos en Bogotá

En enero de 2022 registró 277 casos. Además, 43 % de aquellos hurtos de cometieron con arma de fuego. Por su lado, a lo largo de le primer mes del año se hurtaron 375 motocicletas en Bogotá, 61 más que en el mismo lapso del año anterior. Las ciudades donde se manifestó un más grande número de hurtos de motocicletas fueron Kennedy, Suba y Engativá.

El estudio del Dane reveló que, en Bogotá, la percepción de inseguridad pasó del 69 % en 2020, al 77,8 % en 2021, siendo las calles de la ciudad, el transporte público y los puentes peatonales, los lugares más temidos por los ciudadanos.

Esta manera delictiva, como las demás presentes en la ciudad de Bogotá, tiene un mapa de frecuencia dividido por horas, días y regiones donde es frecuente la comisión del delito. Ejemplificando, conforme con la información de las autoridades, en la noche, en medio de las 7p. m.

y la 1 a. m., es el horario en el cual más se muestran los casos de robo de coches. En lo corrido de enero a noviembre, en esta franja horaria se reportaron 1.372 robos de coches, con una más grande incidencia los martes y miércoles, que en lo cual va del 2021 han dejado un saldo de 239 y 236 casos, respectivamente, para dichos días.

La segunda franja horaria más peligrosa en temas de hurto de coches está en medio de las 12 de la noche y las 6 a. m. Allí, los días martes y viernes son los que más números rojos reportan, con registros de 142 y 133 casos para todos dichos días entre enero y noviembre de 2021. En el conteo anual, en esta franja el número de coches robados sube a 800.

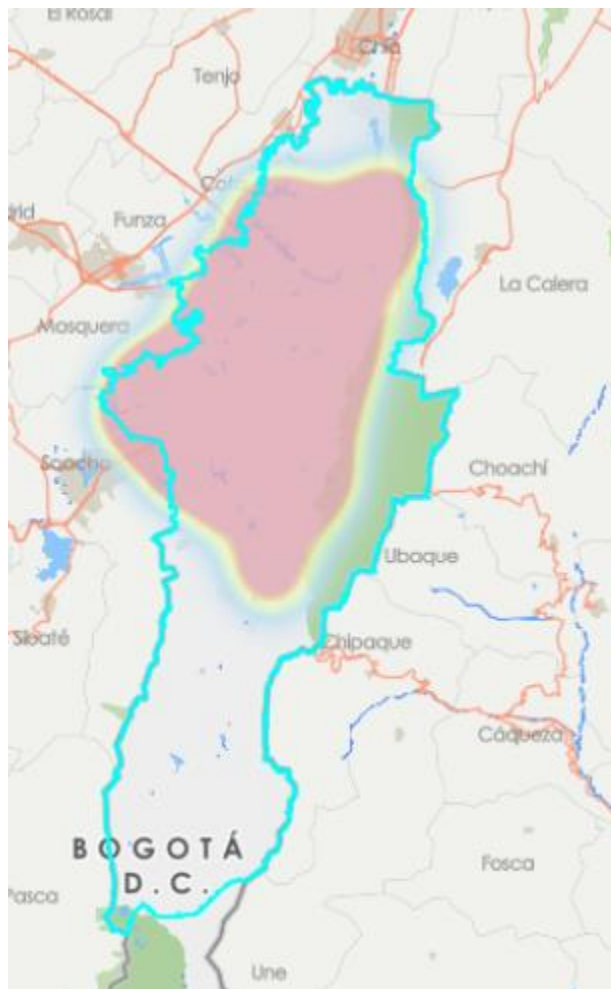


Ilustración 1: Mapa de calor de las zonas más afectadas por los robos

Según reporte de la Secretaría Distrital de Estabilidad, en el primer mes del año de 2022 el robo a vehículos incrementó un 40 % comparado con el mismo mes del año pasado, y las ciudades donde se muestra más grande número de hurtos a vehículos son Kennedy, Engativá y Suba.

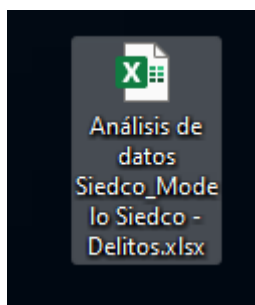
A lo largo de enero, según las cifras de la Secretaría de Estabilidad, Convivencia y Justicia de Bogotá, se reportaron 277 vehículos robados, ante los 189 casos que se registraron en el mismo mes de 2021.

Según la entidad, de los 277 hurtos registrados en las primeras cuatro semanas de este año, en 120 hubo intimidación y actos de maltrato.

La implementación de la nueva tecnología para poder evitar este tipo de robos es primordial y es por esto por lo que se realiza este Proyecto para poder ayudar a la Sociedad a cuidar sus objetos personales, con la ayuda de Python.

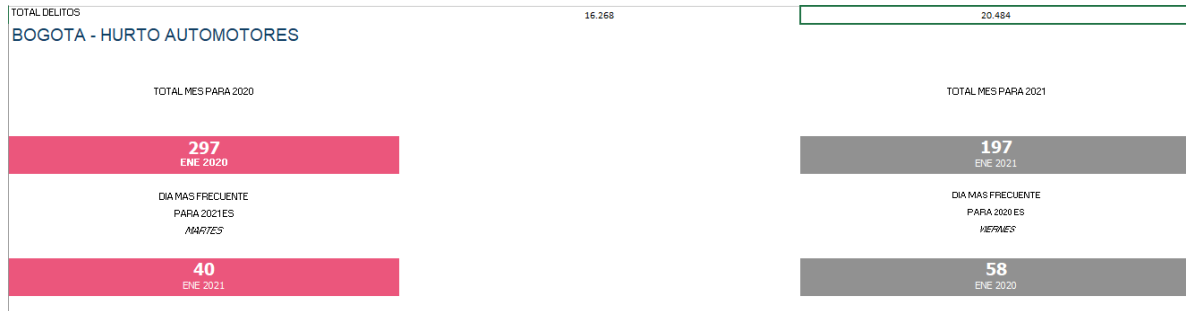
2. Materiales y Métodos

Los datos utilizados para crear el algoritmo de predicción se obtuvieron de la página web Secretaría Distrital de Seguridad, Convivencia y Justicia y se descargaron como un archivo de Excel (Figura 2)



Este es el archivo que se utilizó antes del preprocesamiento dentro del código

Año	Mes	Delito
2021	ENE	HURTO AUTOMOTORES
Ver Mapa		
Datos Tabulados		
Análisis General de Datos para BOGOTA		
Año 2021		
DELITOS	ENE-ENE 2019	ENE-ENE 2020
DELITOS SEXUALES	366	498
EXTORSION	116	67
HOMICIDIOS	83	79
HOMICIDIOS EN AT	26	29
HURTO A COMERCIO	1.449	1.403
HURTO A ENTIDADES FINANCIERAS	6	8
HURTO A PERSONAS	8.583	11.499
HURTO A RESIDENCIAS	850	905
HURTO AUTOMOTORES	286	297
HURTO DE BICICLETAS	593	821
HURTO DE CELULARES	4.529	5.575
HURTO MOTOCICLETAS	332	351
HURTO PIRATERIA TERRESTRE	5	8
LESIONES EN AT	279	332
LESIONES PERSONALES	1.703	1.524
COMPORTAMIENTO POR LOCALIDADES PARA HURTO AUTOMOTORES		
LOCALIDADES Y UPZ	ENE-ENE 2020	ENE-ENE 2021
ANTONIO NARIÑO	9	8
BARRIOS UNIDOS	8	3
BOSA	19	21
CANDELARIA	5	1
CHAPINERO	8	0
CIUDAD BOLÍVAR	13	20
ENGATIVÁ	50	25
FONTIBÓN	16	7
KENNEDY	58	50
LOS MÁRTIRES	3	8
FUENTE ARANDA	22	14
RAFAEL URIBE URIBE	18	4
SAN CRISTÓBAL	15	3
SANTA FE	4	1
SUBA	19	12
TEUSAQUILLO	7	10
TUNJUELITO	8	3
USAQUÉN	9	3



3. Resultados

Se crea el archivo con el nombre de preprocesamiento y se importan las librerías necesarias que en este caso son: pandas para poder leer los datos de la hoja de Excel y para seleccionar y filtrar de manera sencilla tablas de datos en función de posición, valor o etiquetas. fusionar y unir datos.

Numpy: da soporte para crear vectores y matrices grandes, junto con una gran colección de funciones matemáticas de alto nivel para operar con ellas.

matplotlib.pyplot: es una biblioteca para la generación de gráficos a partir de datos contenidos en listas o arrays en el lenguaje.

```

1 import pandas as pd
2 import numpy as np
3 import matplotlib.pyplot as plt
4 df = pd.read_excel("Análisis de datos Siedco_Modelo Siedco - Delitos.xlsx")
5 print(114,15)
6 df = df.fillna(0)
7 df.fillna(0, inplace=True)
8 nan_replaced = df.replace(np.nan,0)
9 print(nan_replaced)
10 print("El tamaño de la data original es : ", df.shape)
    
```

Dentro de la misma línea de código se crea la variable *df* que es la encargada de leer e imprimir los datos de la hoja de Excel, a esa variable se le ejecuta un método el cual es *fillna(0)* que este se hace cargo de los datos nulos que los imprima como un 0.

```

114 15
      Año      Mes \
0      2021     ENE
1      Ver Mapa      0
2      Datos Tabulados  0
3      Análisis General de Datos para BOGOTA  0
4      0      0
..      ...      ...
109     USME      6.0
110     TOTAL GENERAL  297.0
111     Fecha Corte Vigencia 02/03/2021
112     Mensaje Hurto Personas  0
113 Fuente: Cálculos propios con información de SI...  0
    
```

Se utiliza la función *df.tail(10)* para imprimir las últimas *n* filas del objeto según la posición, en este caso se hizo que se devolviera las ultimas 10 filas.

	Año	Mes	Delito	Localidad	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9	Unnamed: 10
104	SANTA FE	4.0	1.0	-75.0	4.0	1.0	-75.0	0	0	0	0
105	SUBA	19.0	12.0	-36.842105	19.0	12.0	-36.842105	0	0	0	0
106	TEUSAQUILLO	7.0	10.0	42.857143	7.0	10.0	42.857143	0	0	0	0
107	TUNJUELITO	8.0	3.0	-62.5	8.0	3.0	-62.5	0	0	0	0
108	USAQUÉN	9.0	3.0	-66.666667	9.0	3.0	-66.666667	0	0	0	0
109	USME	6.0	4.0	-33.333333	6.0	4.0	-33.333333	0	0	0	0
110	TOTAL GENERAL	297.0	197.0	-33.670034	297.0	197.0	-33.670034	0	0	0	0
111	Fecha Corte Vigencia	02/03/2021	0	0	0	0	0	0	0	0	0
112	Mensaje Hurto Personas	0	0	0	0	0	0	0	0	0	0
113	Fuente: Cálculos propios con información de SI...	0	0	0	0	0	0	0	0	0	0

También se utilizó la función *df.shape* el cual no imprime la dimensionalidad del DataFrame

```
1 df.shape
(114, 15)
```

Se utilizó la función *df.describe()* que nos genera las estadísticas del DataFrame

	Año	Mes	Delito	Localidad	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9	Unnamed: 10	Unnamed: 11
count	114	114	114	114	114	114	114	114	114	114	114	114
unique	78	44	46	47	50	36	59	27	24	24	24	19
top	0	0	0	0	0	0	0	0	0	0	0	0
freq	24	48	44	53	44	71	44	79	79	79	80	95

Por otra parte se utilizó la función *df.head()* que hace lo contrario a la función *df.tail*, esta función imprime las primeras *n* filas, pero en este caso se imprimieron las primeras 25 filas.

```

1 import pandas as pd
2 df = pd.read_excel("Análisis de datos Siedco_Modelo Siedco - Delitos.xlsx")
3 df.head(25)
    
```

	Año	Mes	Delito	Localidad	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9
0	2021	ENE	HURTO AUTOMOTORES	(Todos los Valores de Columna)	NaN	NaN	NaN	NaN	NaN	NaN
1	Ver Mapa	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	Datos Tabulados	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	Análisis General de Datos para BOGOTA	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5	Año	2021	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6	DELITOS	ENE-ENE 2019	ENE-ENE 2020	ENE-ENE 2021	Dif ENE-ENE 2021 vs ENE-ENE 2020	% Var ENE-ENE 2021 y ENE-ENE 2020	Dif ENE-ENE 2021 vs ENE-ENE 2019	% Var ENE-ENE 2021 y ENE-ENE 2019	ENE 2019	ENE 2020
7	DELITOS SEXUALES	366.0	498.0	338.0	-160.0	-0.321285	-28.0	-0.076503	366.0	498.0

Se utiliza la función `df.drop` (["aquí se colocó el nombre de las columnas que se necesitaban eliminar para no imprimir esos datos nulos o datos vacíos"])

```

1 dataset = df.drop(["Unnamed: 4", "Unnamed: 5", "Unnamed: 6", "Unnamed: 7", "Unnamed: 8", "Unnamed: 9", "Unnamed: 10", "Unnamed: 11", "Unnamed: 12", "Unn
2 dataset
    
```

	Año	Mes	Delito	Localidad
0	2021	ENE	HURTO AUTOMOTORES	(Todos los Valores de Columna)
1	Ver Mapa	0	0	0
2	Datos Tabulados	0	0	0
3	Análisis General de Datos para BOGOTA	0	0	0
4	0	0	0	0
...
109	USME	6.0	4.0	-33.333333
110	TOTAL GENERAL	297.0	197.0	-33.670034
111	Fecha Corte Vigencia	02/03/2021	0	0
112	Mensaje Hurto Personas	0	0	0
113	Fuente: Cálculos propios con información de SI...	0	0	0

114 rows x 4 columns

En esta línea de código se imprimió la información del dataset

```
[ ] 1 dataset.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 114 entries, 0 to 113
Data columns (total 4 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Año         114 non-null   object
1   Mes         114 non-null   object
2   Delito      114 non-null   object
3   Localidad   114 non-null   object
dtypes: object(4)
memory usage: 3.7+ KB
```

En esta parte del código se utilizó la función `dataset.drop(["Elimine filas o columnas especificando los nombres de las etiquetas y los ejes correspondientes, o especificando directamente los nombres de índice o columna"])` y en el código se escribió mes para eliminar esa columna

```
1 X = dataset.drop(["Mes"], axis=1)
2 Y = dataset["Delito"]

1 X
```

	Año	Delito	Localidad
0	2021	HURTO AUTOMOTORES	(Todos los Valores de Columna)
1	Ver Mapa	0	0
2	Datos Tabulados	0	0
3	Análisis General de Datos para BOGOTA	0	0
4	0	0	0
...
109	USME	4.0	-33.333333
110	TOTAL GENERAL	197.0	-33.670034
111	Fecha Corte Vigencia	0	0
112	Mensaje Hurto Personas	0	0
113	Fuente: Cálculos propios con información de SI...	0	0

114 rows x 3 columns

En la variable Y se guardó el delito y se imprimió

```
1 Y
```

```
0   HURTO AUTOMOTORES
1   0
2   0
3   0
4   0
...
109  4.0
110  197.0
111  0
112  0
113  0
Name: Delito, Length: 114, dtype: object
```

En esta parte use `dataset.describe` para que imprimiera una breve descripción de la data


```
1 dataset.describe()
```

	Año	Mes	Delito	Localidad
count	114	114	114	114
unique	78	44	46	47
top	0	0	0	0
freq	24	48	44	53

En las últimas partes del código se imprime la precisión del modelo

```
1 from sklearn.metrics import precision_score
2 precision = precision_score(y_test, y_pred)
3 print('Precisión del modelo:')
4 print(precision)
```

Precisión del modelo:
0.9868421052631579

Y también se imprime la exactitud

```
1 from sklearn.metrics import accuracy_score
2 exactitud = accuracy_score(y_test, y_pred)
3 print('Exactitud del modelo:')
4 print(exactitud)
```

Exactitud del modelo:
0.9824561403508771

Y al final del código, se imprime la sensibilidad

```
1 from sklearn.metrics import recall_score
2 sensibilidad = recall_score(y_test, y_pred)
3 print('Sensibilidad del modelo:')
4 print(sensibilidad)
```

Sensibilidad del modelo:
0.9868421052631579

4. Discusión

En los resultados se encontró que es muy probable que es casi toda Bogotá se roben algún vehículo en el código no se especifica el tipo de vehículo, ni marca ni modelo, esto solo demuestra la inseguridad de la ciudad, que, a pesar del trabajo implementado de la policía, cada día se registran nuevos robos de vehículos.

El hurto de vehículos en la capital colombiana sigue siendo uno de los preferidos por ladrones, según las cifras proporcionadas por la Secretaría de Seguridad, Convivencia y Justicia de Bogotá, las recomendaciones que nos brindan para que no sigan robando vehículos son:

- No alejarse del transporte con el motor encendido, ya sea por unos minutos.
- Evitar parquear en regiones solitarias, sin buena iluminación.
- Instala un sistema de alarma o dispositivo antirrobo.
- Utiliza una cámara de estabilidad en tu transporte.
- No dejes objetos de costo a la vista una vez que parquees tu coche.
- Sitúa una cámara de estabilidad que tenga vista hacia el área donde se estaciona regularmente el transporte.
- No compres autopartes de dudosa procedencia, debido a que esto incentiva la criminalidad y te convierte en un cómplice de este delito.

References

- [1] E. Sosa Erdozain, «Predicción del grado de superhidrofobicidad en superficies microtexturizadas, mediante el uso de inteligencia artificial en Python.,» Escuela Técnica Superior de Ingenieros Industriales UPM, madrid, 2021.
- [2] M. V. B. QUINTERO, «ANÁLISIS DE LA COMERCIALIZACIÓN DE VEHÍCULOS DE SALVAMENTO POR PARTE DE LAS ASEGURADORAS EN LA CIUDAD DE BOGOTÁ,» UNIVERSIDAD DE CIENCIAS APLICADAS Y AMBIENTALES U.D.C.A., Bogotá, 2017.
- [3] S. C. R. González, « “La ciudad más insegura del mundo”: homicidio y crimen en Bogotá, 1988-1994,» Bogotá, Revista Criminalidad, 2017, pp. 49-64.



© 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).