

# InfraccionesGDL

## Mejoramiento de la cultura vial.

Mentor: Abraham Alva

Oscar Flores | Joel Buenrostro | Alfredo Bello

Saturdays.Ai - Guadalajara - 4ta Edición - Noviembre 2022 - Equipo Morado

- 1. Descripción de la problemática**
- 2. Objetivo del proyecto**
- 3. Hipótesis**
- 4. Datos**
- 5. Descripción de la solución**
  - 5.1. Detalles sobre el modelo**
- 6. Resultados**
  - 6.1. Prototipo funcional**
- 7. Aprendizaje y siguientes pasos**
- 8. Recursos generados**

# Descripción de la problemática

# Descripción de la problemática

La **educación vial** de los conductores en la **Zona Metropolitana de Guadalajara** tiene tiempo deteriorándose y las futuras generaciones empezarán a seguir estos malos pasos, tales como estacionarse en lugares no permitidos como líneas amarillas, tomas para bomberos, lugares reservados para incapacitados, entre otros, causando problemas a peatones o ciudadanos que requieran de una rampa para poder subir las banquetas, o simplemente que deseen caminar por la calle. Mientras que, también, por otro lado, los automovilistas suelen omitir la tarifa de parquímetros en diferentes puntos de la ciudad, siendo acreedores a una infracción.

Los dos principales mercados dentro de la problemática:

- Los **conductores** en busca de un lugar para estacionarse que caen así en diferentes infracciones, por lo que el poder obtener dichos lugares donde normalmente han sido infraccionados otros conductores anteriormente, ayudará a los futuros conductores a evitar estas infracciones.
- **Secretaría de Movilidad** y los **oficiales viales de tránsito**, que podrá obtener un resultado de cuales son las zonas, fechas y horas con más infracciones y poder proponer soluciones a dicho problema vial, como más lugares de estacionamiento..

# Tipos de infracciones en la ZMG

El proceso de levantamiento de infracciones de movilidad del Municipio de Guadalajara se lleva a cabo mediante una aplicación móvil, donde se tiene como función levantar una infracción a los vehículos violando alguna de la normativa de movilidad del municipio, al igual la aplicación genera un ticket o infracción con los detalles del delito vial. Posteriormente, las infracciones son enviadas al sistema web..

- **Tipo 1:** Omitir la tarifa del parquímetro.
- **Tipo 2:** Invadir dos cajones o más de estacionamiento.
- **Tipo 3:** Obstruir una cochera.
- **Tipo 4:** Estacionarse en una intersección o en línea amarilla.
- **Tipo 5:** Estacionarse en lugares exclusivos, bomberos, policía, servicios médicos o personas con discapacidad.
- **Tipo 6:** Cordón o batería.
- **Tipo 7:** Vehículo inmovilizado.
- **Tipo 8:** Por insultar o agredir verbalmente a un oficial vial.
- **Tipo 9:** Por abandono de vehículo.
- **Tipo 10:** Exceder tiempo en los espacios de carga y descarga.
- **Tipo 11:** Agredir físicamente a un oficial vial.

# Objetivo del proyecto

# Objetivo del proyecto

Se busca crear una mejor educación vial de los conductores y ayudar a los transeúntes para que puedan transitar por las vías designadas, generando una mejor conciencia y cultura vial en el país, además de evaluar las tendencias en el cumplimiento vial y su vinculación con las estrategias en políticas públicas, todo esto, enfocado a buscar reducir el número de infracciones que cada año aumenta más y más.



# Hipótesis



# Hipótesis

Se buscará crear una mejor educación vial de los conductores y ayudar a los transeúntes para que puedan transitar por las vías designadas, generando una mejor conciencia y cultura vial en el país, además de evaluar las tendencias en el cumplimiento vial y su vinculación con las estrategias en políticas públicas. **La falta de espacios de estacionamiento y la falta de conocimiento de los parquímetros digitales causa el gran número de infracciones registradas año tras año en la Zona Metropolitana de Guadalajara.**

## **OBJETIVOS** **DE DESARROLLO SOSTENIBLE**



Ya que el desconocimiento o negligencia de los lugares correctos para estacionarse tiene un efecto negativo en los peatones y automovilistas, las leyes no tienen un efecto deseado y los automovilistas reinciden cometiendo infracciones degradando en ciertas zonas la movilidad peatonal.

# Datos

# Datos

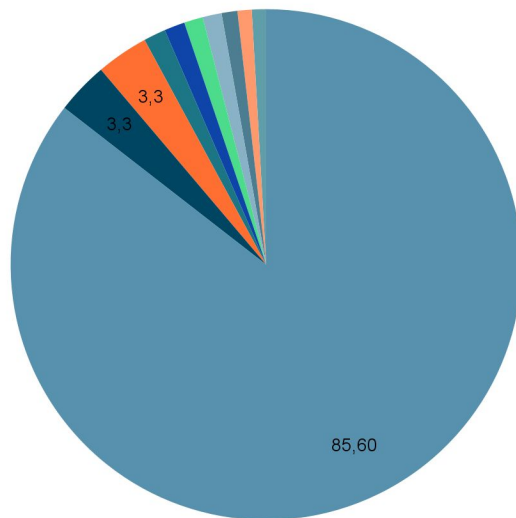
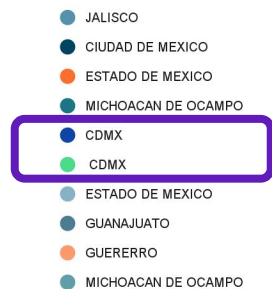
- **ESTADO:** Estado de la placa del automóvil infraccionado | Tipo String.
- **MARCA:** Marca del automóvil infraccionado | Tipo String.
- **CALLE:** Calle principal donde se realizó la infracción | Tipo String.
- **CRUCE:** Cruce más cercano de la calle principal donde se realizó la infracción | Tipo String.
- **FECHA:** Fecha de la infracción | Tipo Datetime.
- **HORA:** Hora de la infracción | Tipo Datetime.
- **LATITUD:** Latitud geográfica de la infracción | Tipo int.
- **LONGITUD:** Longitud geográfica de la infracción | Tipo int.
- **INFRACCIÓN:** Tipo de infracción cometida por el vehículo | Tipo String.

◆	ESTADO ◆	MARCA ◆	CALLE ◆	CRUCE ◆	FECHA ◆	HORA ◆	INFRACCION ◆	LATITUD ◆	LONGITUD ◆
17	JALISCO	MAZDA CX-3 5 PUERTAS (IMPORTADO)	duque de rivás	miguel blanco de Tejada	16/08/21	12:34:00	Clave 1 Omitir tarifa	20.67	-103.38
18	CIUDAD DE MEXICO	HONDA	Calle Miguel Lerdo de Tejada	duque de Rivas	16/08/21	12:38:00	Clave 1 Omitir tarifa	20.67	-103.38
19	JALISCO	MITSUBISHI OUTLANDER 4 PUERTAS IMP	Francisco de Quevedo	Miguel Lerdo de Tejada	16/08/21	12:42:00	Clave 1 Omitir tarifa	20.67	-103.38

# Limpieza de los datos

La gran desventaja de los datos obtenidos, es que, al ser campos abiertos para el registro de las infracciones, existen demasiados valores únicos en las principales columnas como lo son **ESTADO**, **MARCA** o **INFRACCIÓN**.

TOP 10 ESTADOS DE PLACAS MÁS INFRACCIONADAS



Random Fact:  
El 97% de las infracciones son de tipo 1, 4 o 5.

InfraccionesGDL

# Transformación de los datos

- La columna **ESTADO** y **MARCA** fueron transformadas a valores categóricos, asignándole un valor numérico a cada valor único de dichas columnas.
- En la columna **INFRACCIÓN** fueron removidos todos los caracteres y se conservaron los números, quedando así en tipo int con valores 1, 4 y 5.
- Se creó la columna **HORA\_NUM** para conservar únicamente la hora de la infracción (ejemplo, si la infracción fue 11:53:00, el valor conservado fue 11).
- Se creó la columna **MES** para guardar el mes de la infracción.

```
{'aguascalientes': 0, 'baja california norte': 1, 'baja california sur': 2,
'campeche': 3, 'chiapas': 4, 'chihuahua': 5, 'ciudad de mexico': 6,
'coahuila': 7, 'colima': 8, 'durango': 9, 'estado de mexico': 10,
'extranjero': 11, 'guanaxuato': 12, 'guerrero': 13, 'hidalgo': 14,
'jalisco': 15, 'michoacan': 16, 'morelos': 17, 'nayarit': 18,
'nuevo leon': 19, 'oaxaca': 20, 'puebla': 21,
'queretaro': 22, 'quintana roo': 23, 'san luis potosi': 24,
'sinaloa': 25, 'sonora': 26, 'tabasco': 27, 'tamaulipas': 28,
'tlaxcala': 29, 'veracruz': 30, 'yucatan': 31, 'zacatecas': 32}
```

# Datos limpios y transformados

- **ESTADO:** Estado de la placa del automóvil infraccionado | Tipo String.
- **MARCA:** Marca del automóvil infraccionado | Tipo String.
- **FECHA:** Fecha de la infracción | Tipo Datetime.
- **AÑO:** Año de la infracción. | *Tipo int.*
- **MES:** Longitud geográfica de la infracción | *Tipo String.*
- **HORA:** Hora de la infracción | Tipo Datetime.
- **HORA\_NUM:** Hora de la infracción | *Tipo int.*
- **LATITUD:** Latitud geográfica de la infracción | Tipo int.
- **LONGITUD:** Longitud geográfica de la infracción | Tipo int.
- **ESTADO\_cat:** Valor numérico único por ESTADO. | *Tipo int.*
- **MARCA\_cat:** Valor numérico único por MARCA. | *Tipo int.*
- **INFRACCIÓN:** Tipo de infracción cometida por el vehículo | *Tipo int: Valores 1, 4 o 5*

ESTADO	MARCA	FECHA	ANO	MES	HORA	HORA_NUM	LATITUD	LONGITUD	ESTADO_cat	MARCA_cat	INFRACCION
2	jalisco	volkswagen	2021-08-16	2021	August	11:26:00	11	20.67	-103.38	15	65
3	jalisco	freightliner	2021-08-16	2021	August	11:27:00	11	20.67	-103.38	15	17
4	jalisco	volkswagen	2021-08-16	2021	August	11:38:00	11	20.67	-103.38	15	65

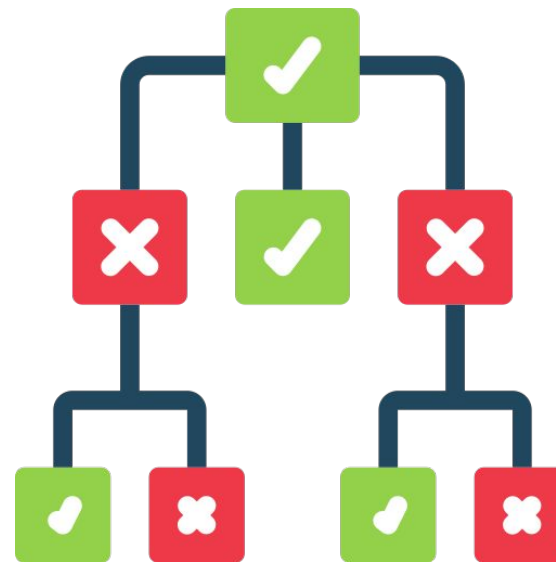
# Descripción de la solución

# Descripción de la solución

Crear un **modelo de predicción**, basado en históricos de 2021, para determinar el tipo de infracción que un automovilista puede recurrir. Esto utilizando principalmente los datos geográficos de la base de datos.

Los modelos probados fueron los siguientes:

- Árbol de Decisión
- Bosque Aleatorio
- Gradient Boosting



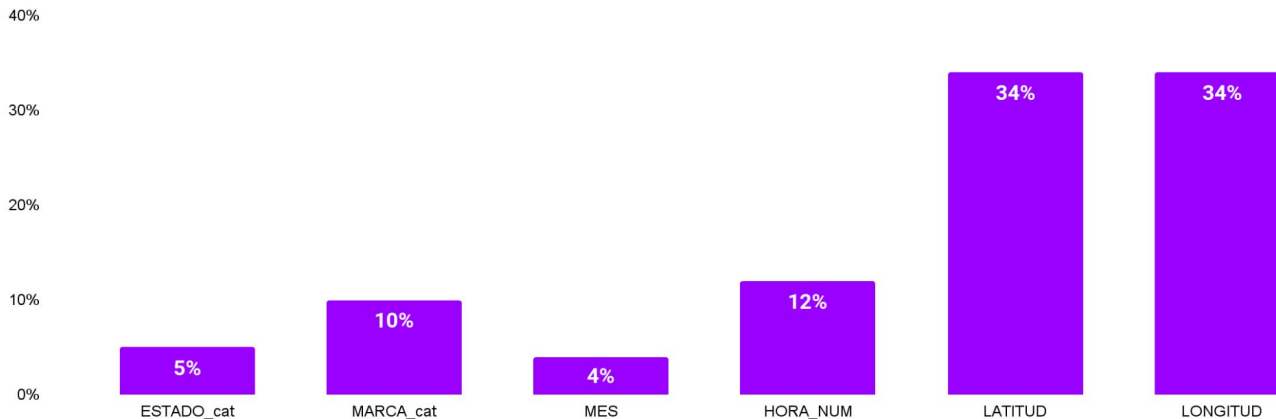


# Detalles del modelo

# Detalles del modelo

- Tipo de modelo: **Gradient Boosting**
- Features utilizadas: ESTADO\_cat, MARCA\_cat, MES, HORA\_NUM, LATITUD, LONGITUD
- Target: INFRACCIÓN

Model's Feature Importance



# Métricas de evaluación del modelo

		Predicted Values		
		Tipo 1	Tipo 4	Tipo 5
Real Values	Tipo 1	14,962	224	187
	Tipo 4	1,043	8,234	1,299
	Tipo 5	1,143	1,610	5,605

## Accuracy

$$\frac{14,962 + 8,234 + 5,605}{34,307} = \frac{28,801}{34,307} = 84.0\%$$

## Precision

$$\frac{14,962}{14,962 + 1,043 + 1,143} = \frac{14,962}{17,148} = 87.3\%$$

$$\frac{8,234}{224 + 8,234 + 1,610} = \frac{8,234}{10,068} = 81.8\%$$

$$\frac{5,605}{187 + 1,299 + 5,605} = \frac{5,605}{7,091} = 79.0\%$$

$$\frac{87.3\% + 81.8\% + 79.0\%}{3} = 82.7\%$$

## Recall

$$\frac{14,962}{14,962 + 224 + 187} = \frac{14,962}{5,373} = 97.3\%$$

$$\frac{8,234}{1,043 + 8,234 + 1,299} = \frac{8,234}{10,068} = 82.7\%$$

$$\frac{5,605}{1,143 + 1,610 + 5,605} = \frac{5,605}{8,358} = 67.1\%$$

$$\frac{97.3\% + 82.7\% + 67.1\%}{3} = 80.7\%$$

# Resultados

# Resultados: Backtesting 2022

**2022 Confusion matrix: XGBClassifier**

Los modelos fueron entrenados solamente con información de 2021, y las infracciones de 2022 fueron separadas para realizar una prueba de tipo backtesting, prediciendo la información que ya se tiene y que no sea solamente lo que se separó para la prueba del modelo.

	Tipo 1	Tipo 4	Tipo 5
Tipo 5	0.13	0.24	0.62
Tipo 4	0.1	0.77	0.13
Tipo 1	0.92	0.03	0.05

Predicted value

# Prototipo funcional

# Prototipo funcional

Como despliegue de prototipo funcional: un Dashboard para que los dos mercados principales, automovilistas y agentes viales, puedan consultar la información histórica desde 2021 hasta agosto de 2022 de las infracciones más comunes.

## InfraccionesGDL SaturdaysAI - 4ta Edición - Guadalajara Equipo Morado

La educación vial de los conductores en la ZMG suele ser mala al estacionarse en lugares no permitidos como líneas amarillas, tomas para bomberos, lugares reservados para discapacitados entre otros, causando problemas a peatones y gente que requiere de una rampa para poder usar las banquetas. Y por otro lado, las infracciones tienen un impacto económico, ya que los automovilistas suelen omitir la tarifa de parqueaderos en diferentes puntos de la ciudad.

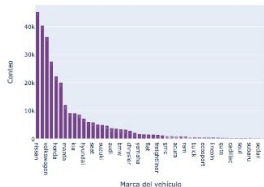
Tipo de Infracción	Descripción
Tipo 1	Omitir tarifa de parqueadero
Tipo 4	Estacionarse en intersección o línea amarilla
Tipo 5	Estacionarse en lugares exclusivos, bomberos, policía, servicios médicos o personas con discapacidad

Tipo de Infracción:

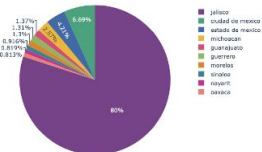
Tipo 1, 4 & 5

### Todas las infracciones

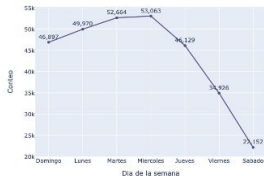
Top 50 marcas de vehículos infraccionados (2021 - 2022)



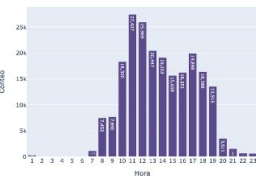
Top 10 Estados con más infracciones (2021 - 2022)



Infracciones por día (2021 - 2022)



Infracciones por hora del día (2021 - 2022)



### Mapa de Infracciones de la ZMG (2021 - 2022)



La ZMG recibe alrededor de 1,000 infracciones al día, de las cuales solo alrededor del 10% son pagadas en cada año. Este modelo está enfocado en dos principales temas: Uno para el ciudadano conductor, donde a partir del Estado de su placa, Modelo de su camion y hora donde se estacionará, y la zona y dirección del lugar, se podrá hacer una violación si entra en perfil estadístico de infracciones en la ciudad y empujará de conductores que el modelo aprenda. Por otro lado, beneficiará a la Fiscalía Pública, ya que a partir de este análisis se podrá hacer una proyección más precisa, basada en estadísticas, sobre las infracciones a futuro. A partir de estos pronósticos, se pueden crear o proponer políticas públicas o programas sociales que incentiven una mejor cultura vial en la ciudad. A partir del análisis de estos datos podemos determinar cuáles serían las zonas más eficientes donde agregar cajones de estacionamiento público o estacionamientos públicos. #LACiudadVial



InfraccionesGDL

<http://infraccionesgdl.herokuapp.com/>

# Aprendizajes y siguientes pasos

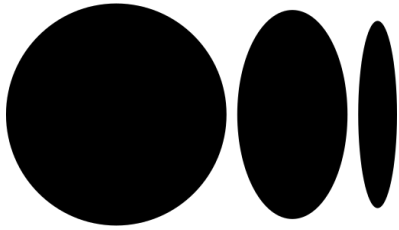


# Aprendizajes y siguientes pasos

- El número de infracciones registradas es **alto**. Además, es **poco o nulo** el conocimiento del parquímetro digital.
- **Explorar en campo** el conocimiento de los conductores acerca de las infracciones.
- Dar **seguimiento** al número de infracciones cometidas en los últimos meses.
- Empezar a buscar soluciones por parte de las secretarías y dependencias gubernamentales correspondientes, ya que, aparte, **solamente el 10%** de todas las infracciones son cubiertas, seguir infraccionando no está logrando mucho.
- Incluir en el **dashboard** un **predictor** del modelo, donde se puedan incluir datos de entrada, y obtener la infracción más probable que se podría obtener y así, buscar evitarla.

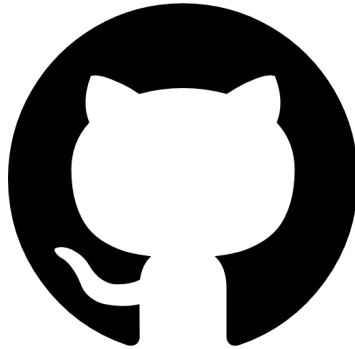
# Recursos generados

# Recursos Generados



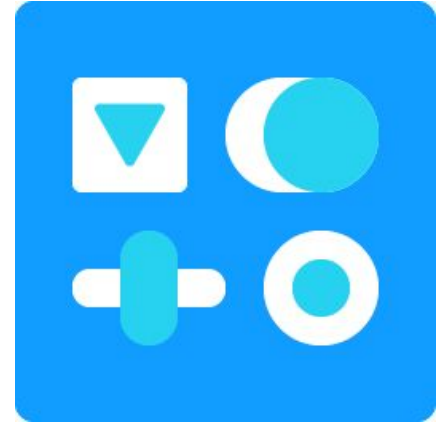
Artículo de medium

[LINK](#)



Repositorio de GitHub

[LINK](#)



Dashboard

[LINK](#)

InfraccionesGDL

# **InfraccionesGDL**

## **Mejoramiento de la cultura vial.**

Mentor: Abraham Alva

Oscar Flores | Joel Buenrostro | Alfredo Bello

Saturdays.Ai - Guadalajara - 4ta Edición - Noviembre 2022 - Equipo Morado