



Análisis para la toma de decisiones  
estratégicas.



# GeoInsights

# DEFINICIÓN DEL PROBLEMA



## OBJETIVO

**Descubrimiento de nuevos puntos de venta.**

Desarrollo de un modelo de predicción que a partir de inputs determine si una tienda Oxxo tiene alto potencial de éxito.

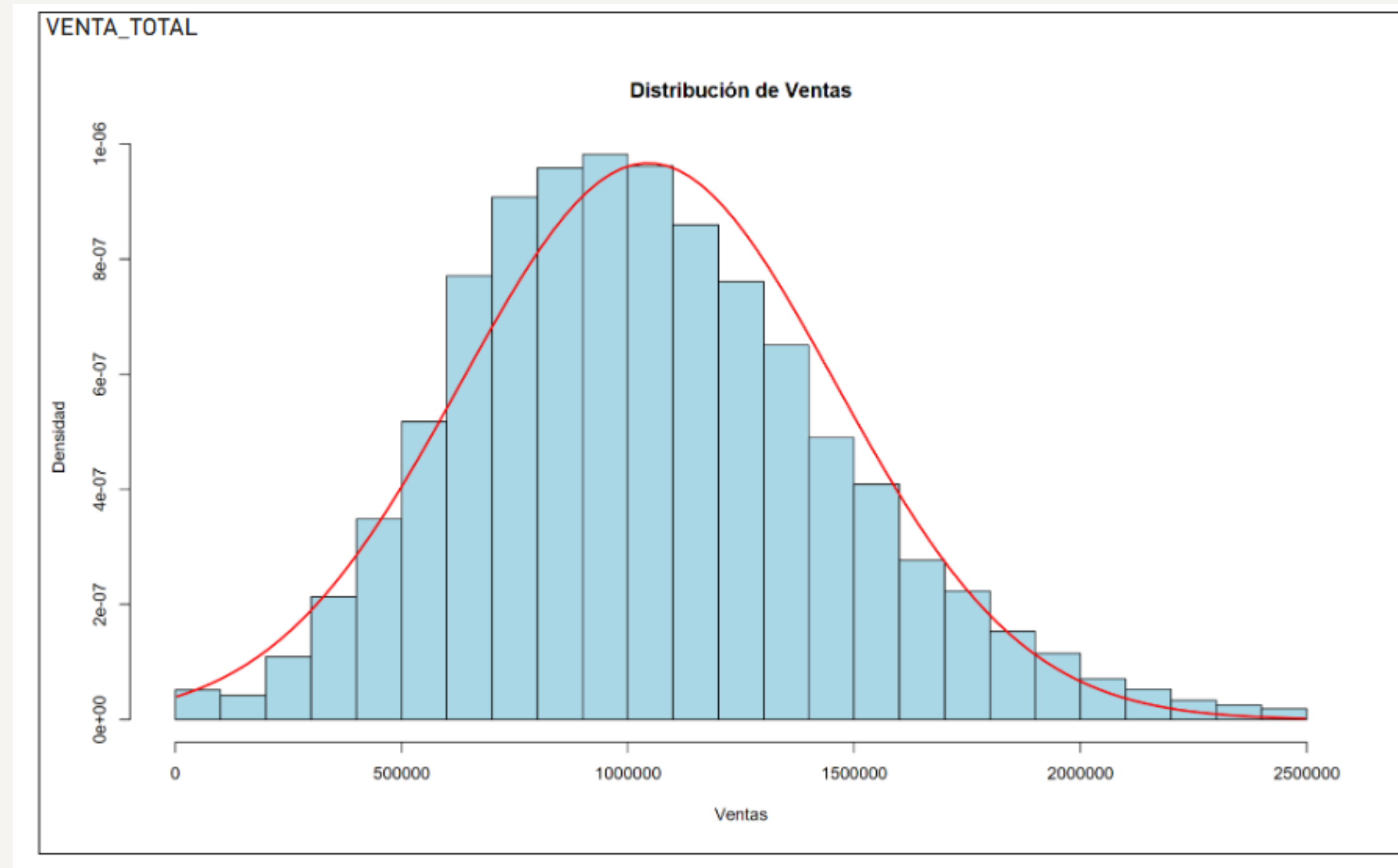
## INPUTS

- **Latitud:** Latitud de la tienda.
- **Longitud:** Longitud de la tienda.
- **MT2:** Medida del área del piso de ventas y cuarto frío de exhibición al cliente.
- **Plaza\_CVE:** Nombre de la plaza asignada a la tienda.
- **Entorno:** Entorno en el que se ubica la tienda



# ESTUDIO DE LOS DATOS

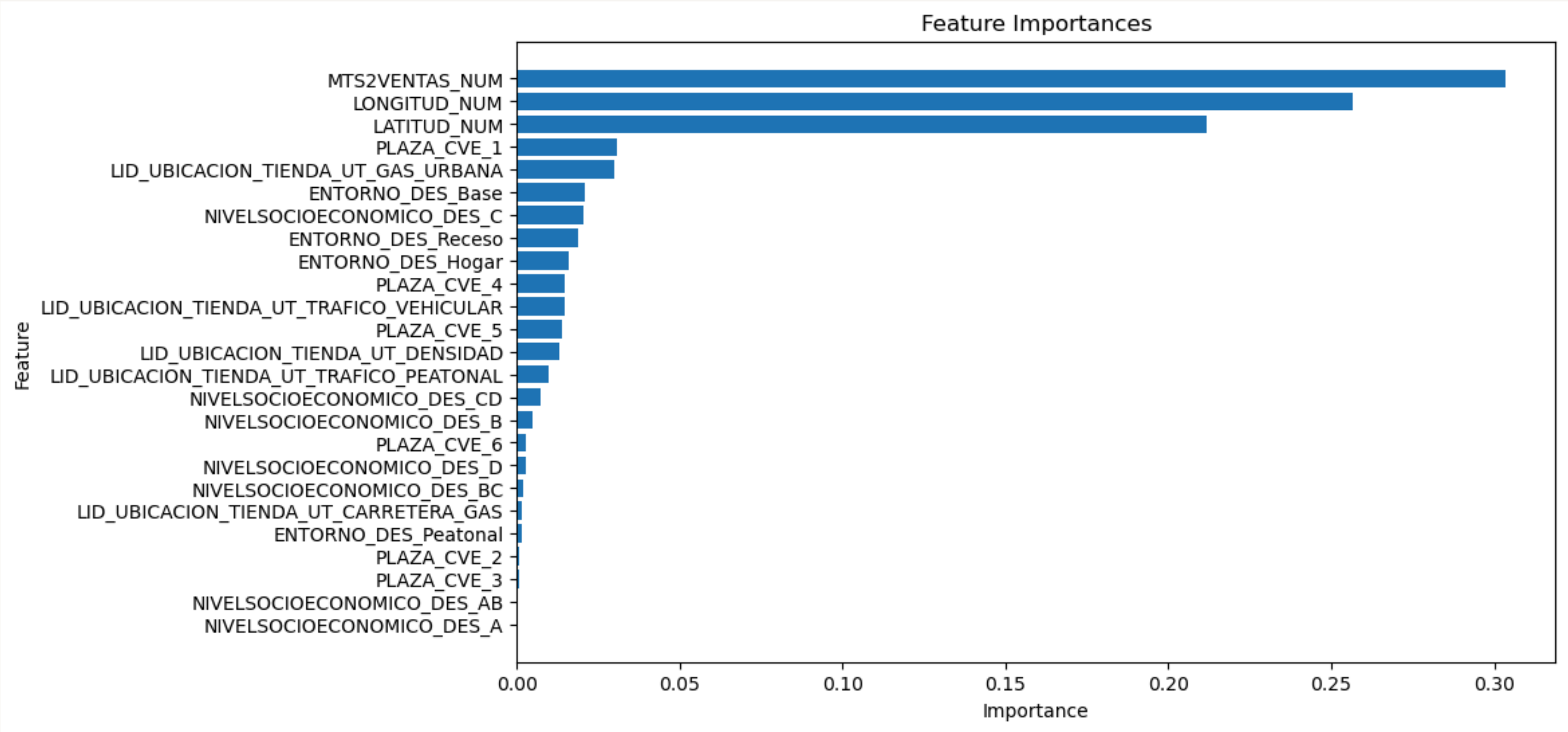
## MODELO GENERADO POR RANDOM FOREST



## MODELO DE DISTRIBUCIÓN NORMAL

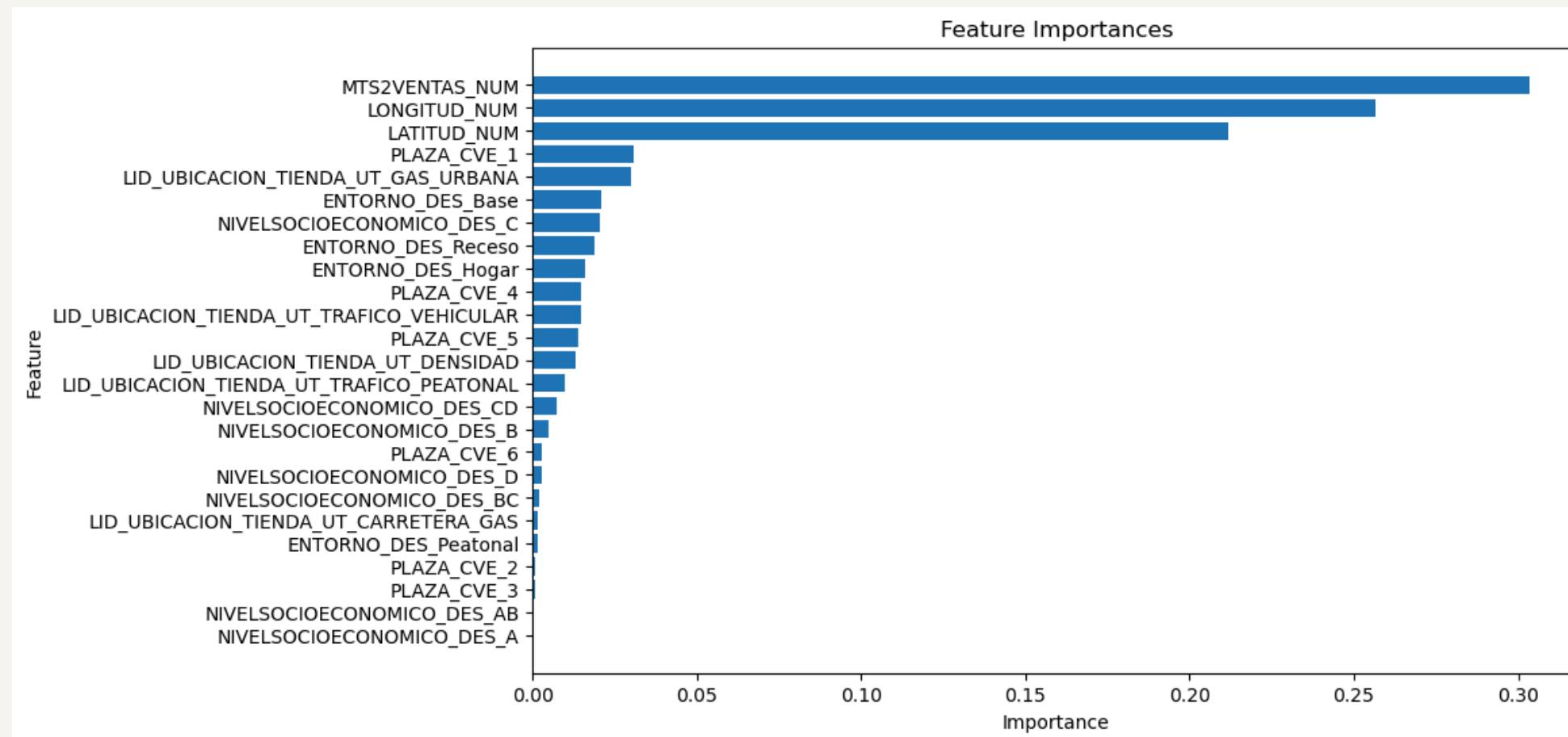
# ESTUDIO DE LOS DATOS

## MODELO GENERADO POR RANDOM FOREST



# ESTUDIO DE LOS DATOS

## MODELO GENERADO POR RANDOM FOREST



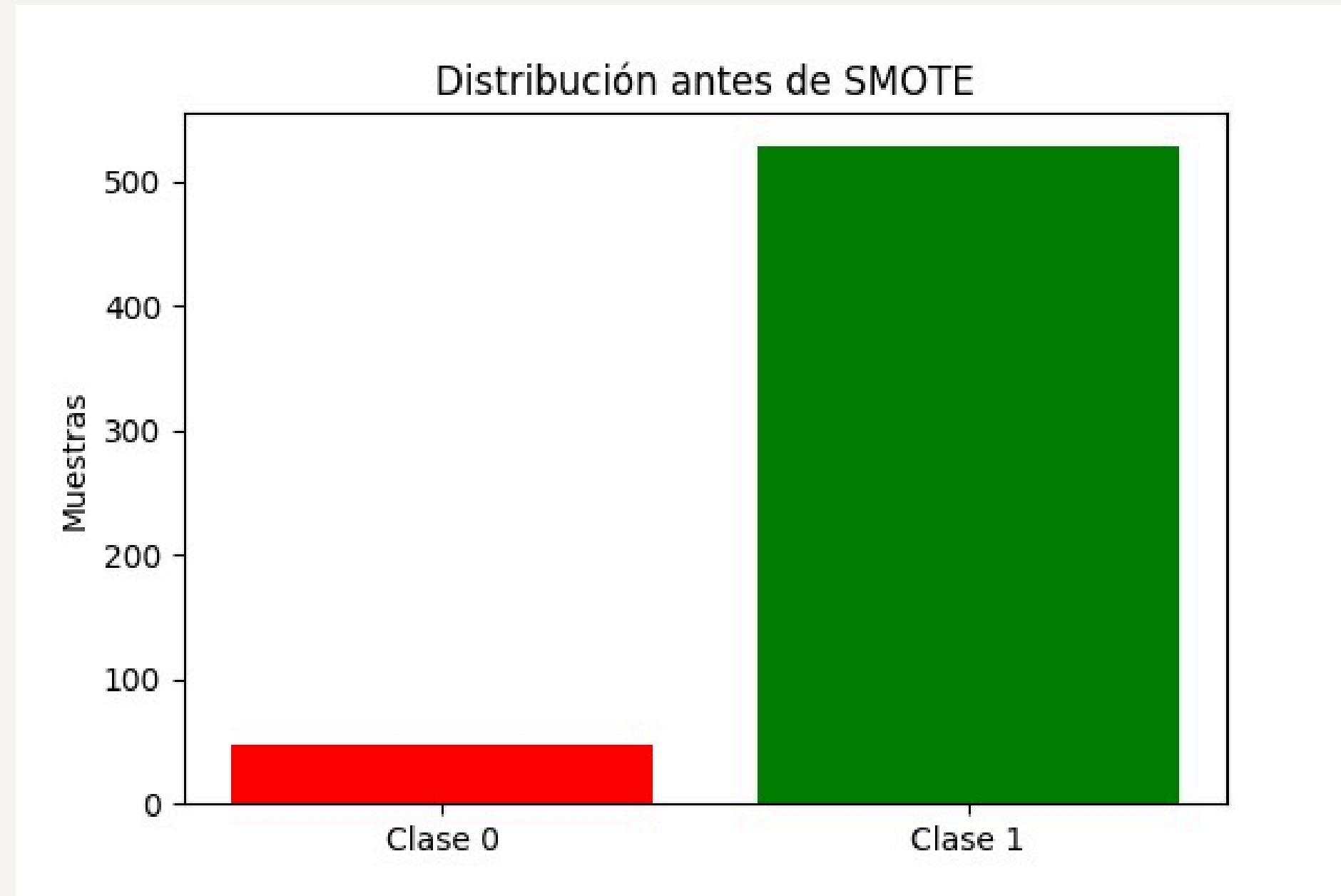
## Variables que se mantuvieron

- MT2VENTAS\_NUM
- LONGITUD\_NUM
- LATITUD\_NUM
- UBICACION\_TIENDA
- ENTORNO\_DES
- NIVELSOCIOECONOMICO\_DES
- PLAZA\_CVE



# ESTUDIO DE LOS DATOS

## MODELO GENERADO POR RANDOM FOREST



**$R^2$ : 0.88 | Recall: 0.05**

### Distribución de clases y sesgo del modelo

La mayoría de los datos (90%) corresponden a tiendas exitosas, lo que genera un desequilibrio en la distribución.

Al clasificar como clase 1 a los OXXO con más del 75% de probabilidad de éxito, se introduce un sesgo que limita la capacidad del modelo para predecir fracasos con precisión.

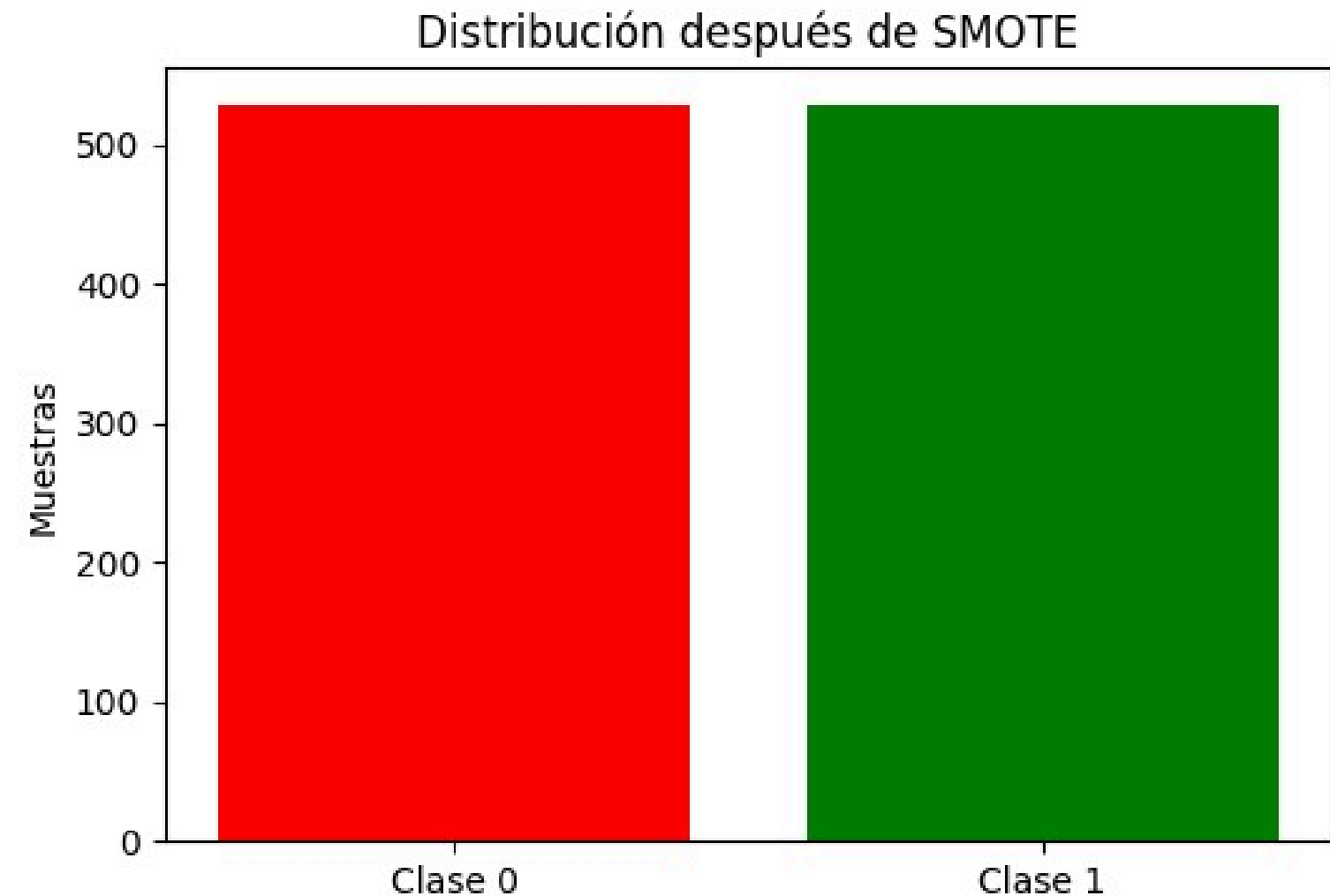


# ESTUDIO DE LOS DATOS

## Corrección del sesgo con SMOTE

Para evitar un modelo sesgado, aplicamos SMOTE, una técnica que genera datos sintéticos de tiendas sin éxito. Esto equilibró la distribución de clases y mejoró la capacidad del modelo para predecir, con mayor precisión, si un nuevo OXXO tendrá éxito o no.

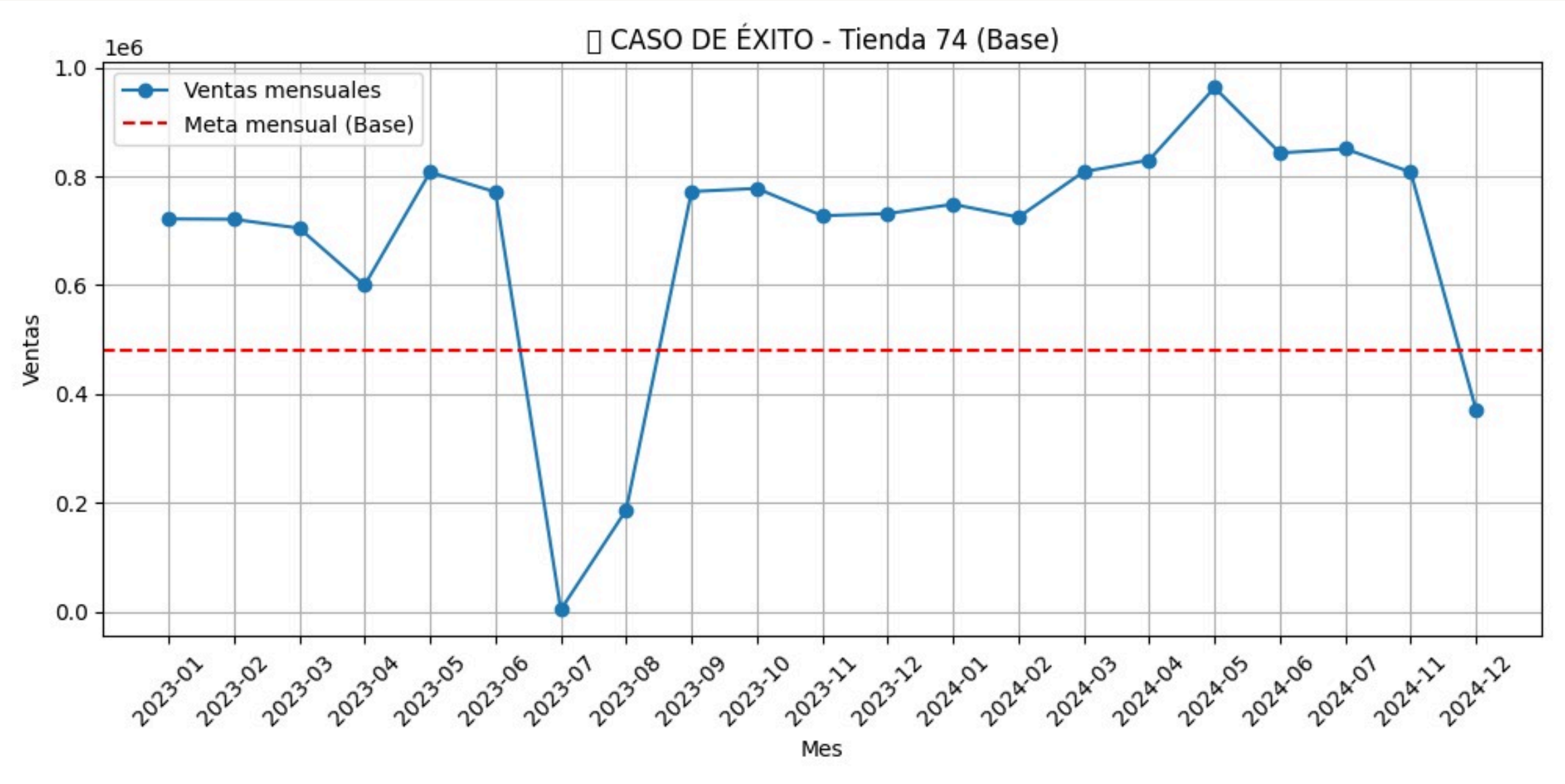
## MODELO GENERADO POR RANDOM FOREST



**$R^2$ : 0.89 | Recall: 0.76**

# ESTUDIO DE LOS DATOS

MODELO GENERADO POR  
RANDOM FOREST



**Nivel Socioeconomico:** C  
**Descripción del Entorno:** Base  
**Plaza CVE:** 6  
**Ubicación Tienda:** Trafico Vehicular  
**Porcentaje de Éxito:** 85.7%



# ESTUDIO DE LOS DATOS

## MODELO GENERADO POR RANDOM FOREST

**Nivel Socioeconomico:** C  
**Descripción del Entorno:**  
Hogar  
**Plaza CVE:** 1  
**Ubicación Tienda:** Gas  
Urbana  
**Porcentaje de Éxito:** 0%



# ÁREAS DE OPORTUNIDAD (OTROS NEGOCIOS)

## Puntos en común

- Las variables clave (ubicación, entorno, tamaño de tienda, etc.) existen en todos los negocios.
- La definición de éxito es la misma.
- El contexto geográfico y socioeconómico es similar.

## Consideraciones clave para la adaptación

- 1. Impacto variable por negocio
- Aunque las variables son las mismas, su influencia cambia según el giro del negocio.
- Ejemplo: el tamaño de tienda puede ser más relevante en una farmacia que en una cafetería.
- 2. Reentrenamiento necesario
- Cada negocio tiene comportamientos únicos de clientes, ubicación y estrategia.
- Se requiere reentrenar el modelo con datos propios de cada negocio.
- 3. Análisis de importancia de variables
- Evaluar qué variables son más influyentes en cada negocio.
- Eliminar o ajustar variables que no aportan valor.

