

Predicción de Ventas de productos alimenticios

BOOTCAMP DATA SCIENCE (MARZO 2024) – CODING DOJO

A solid orange horizontal bar at the bottom of the slide.

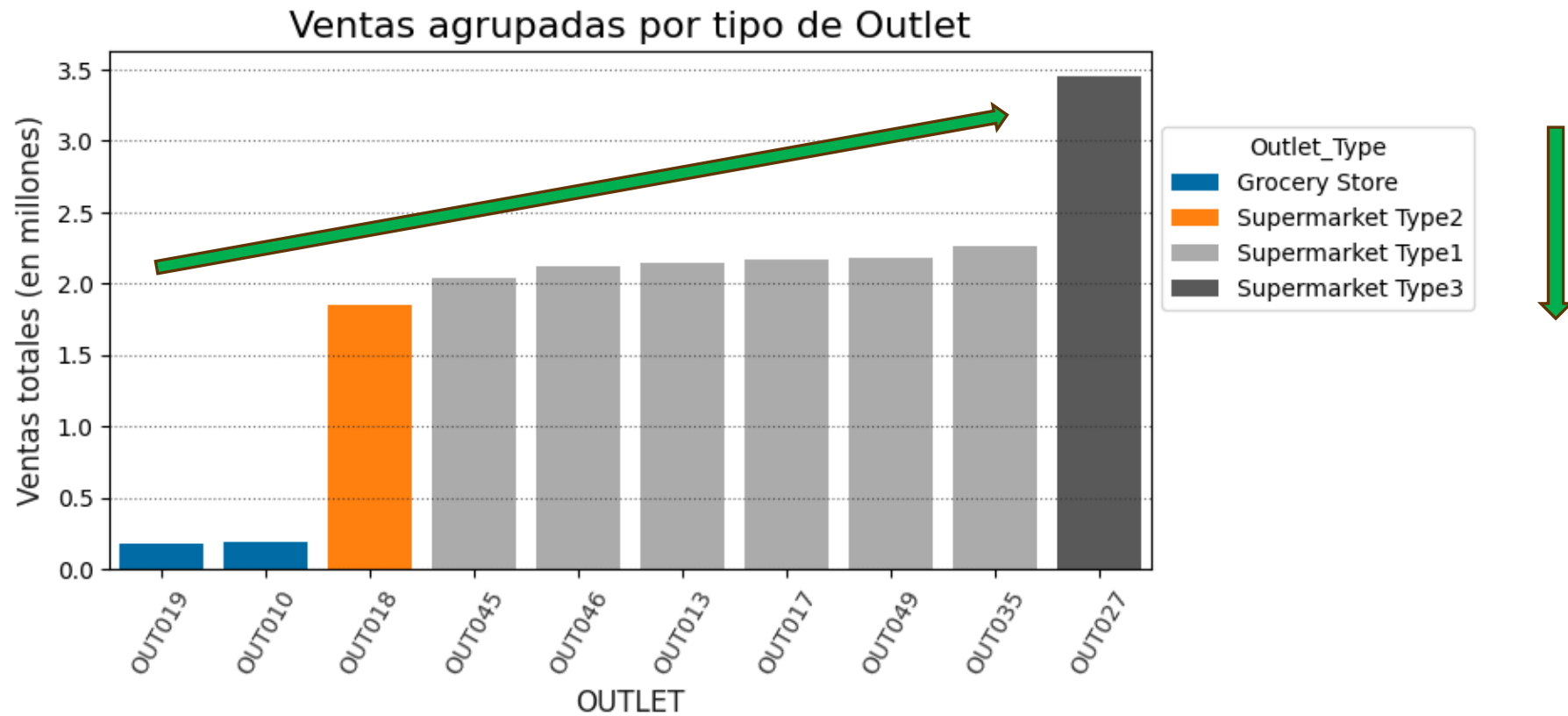
Objetivos

- Ayudar al cliente (comerciante) a predecir las ventas de productos alimenticios en distintas tiendas.
- Mejorar la comprensión de las propiedades de los productos
- Determinar factores que afectan las ventas en los distintos puntos de ventas

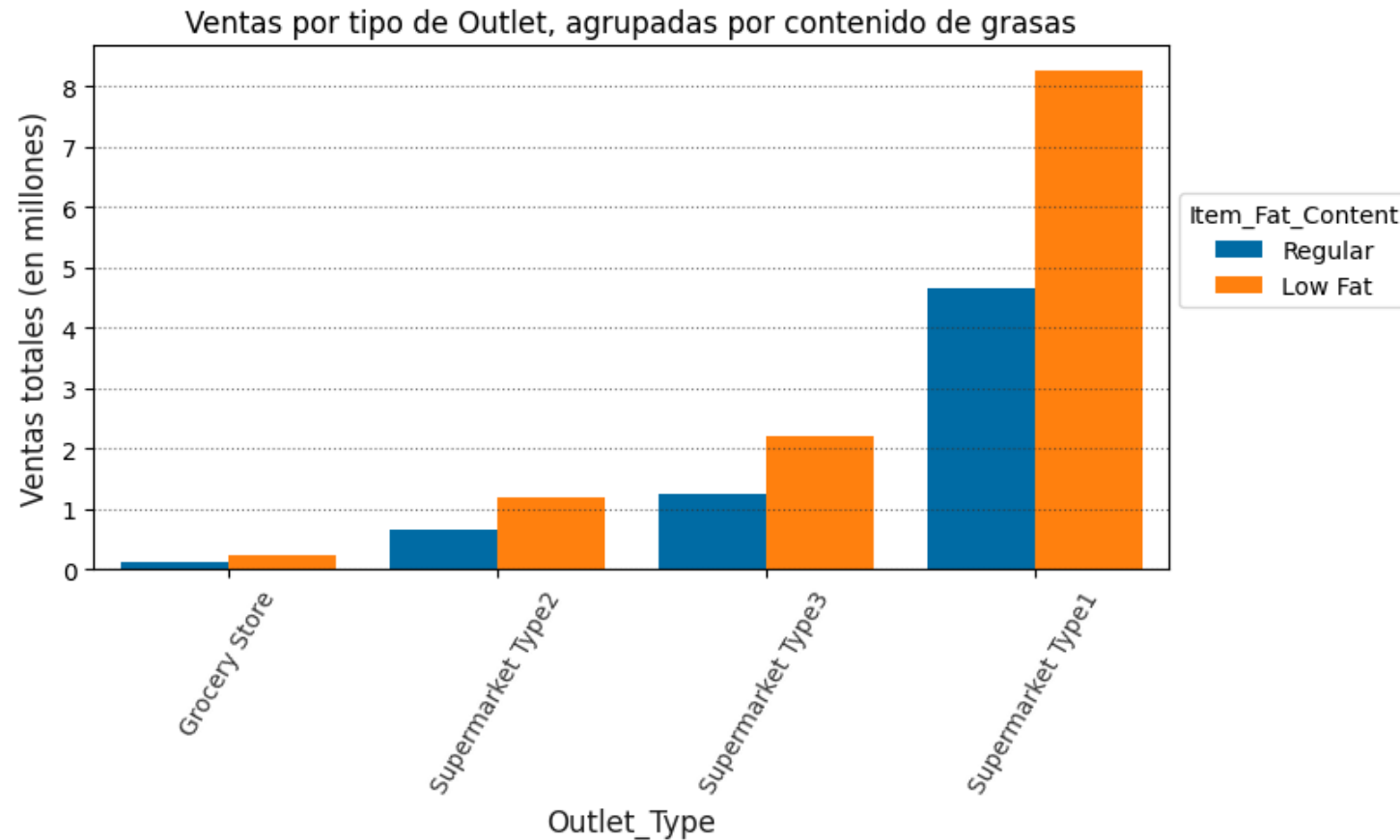
Datos

Nombre de la variable	Descripción
Item_Identifier	Identificador único del producto
Item_Weight	Peso del producto
Item_Fat_Content	Si el producto es bajo en grasa o regular
Item_Visibility	Porcentaje de área total de exposición de todos los productos para una tienda asignada al producto concreto
Item_Type	Categoría a la que el producto pertenece
Item_MRP	Precio máximo de venta al público (precio de catálogo) del producto
Outlet_Identifier	Identificador único de la tienda
Outlet_Establishment_Year	Año en que se estableció la tienda
Outlet_Size	El tamaño de la tienda en relación al área total
Outlet_Location_Type	El tipo de área donde se encuentra la tienda
Outlet_Type	Si el punto de venta es una tienda de comestibles o algún tipo de supermercado
Item_Outlet_Sales	Ventas del producto en una tienda particular. Esta la variable objetivo a predecir.

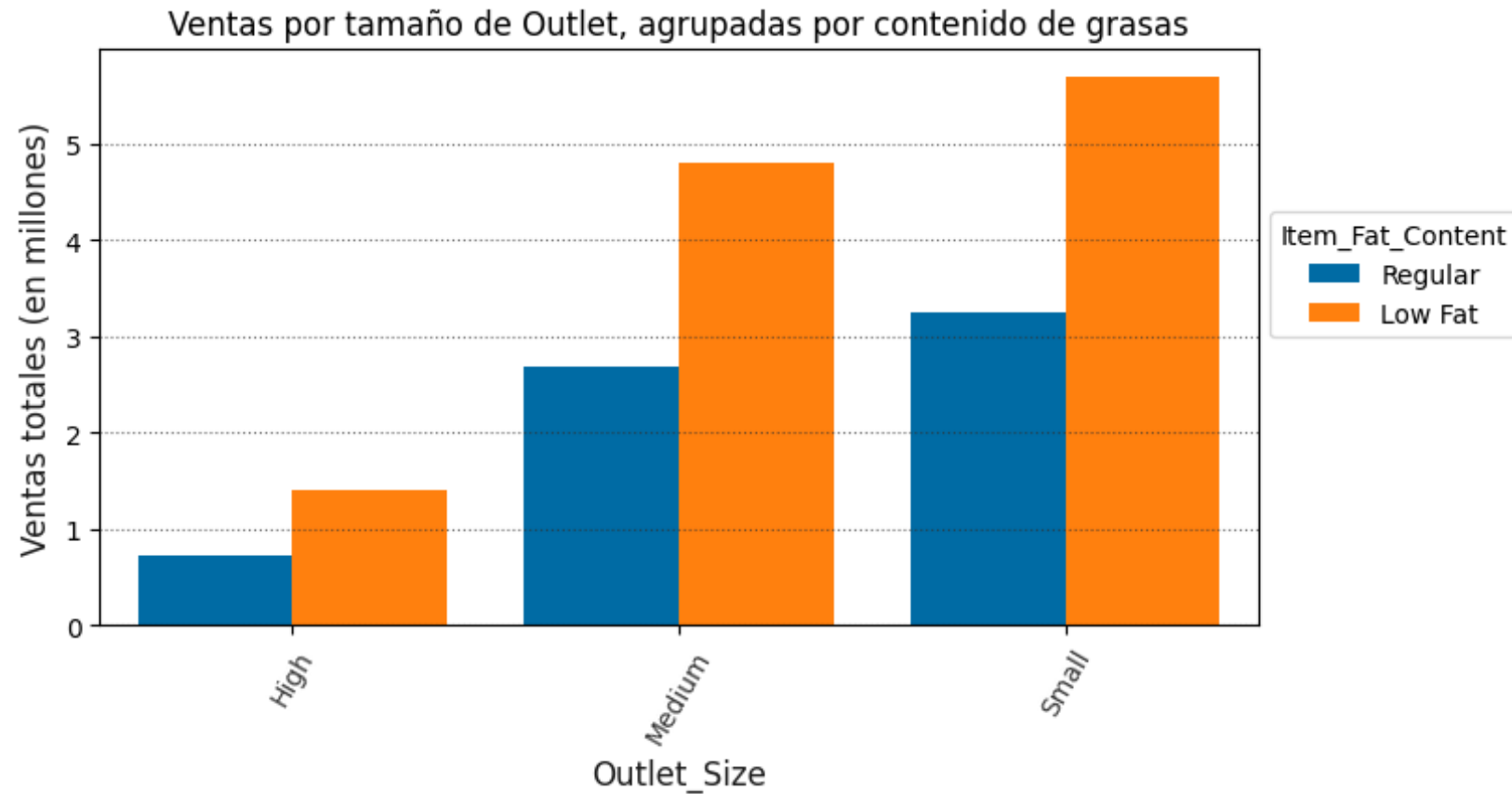
Análisis – Ventas por Outlet



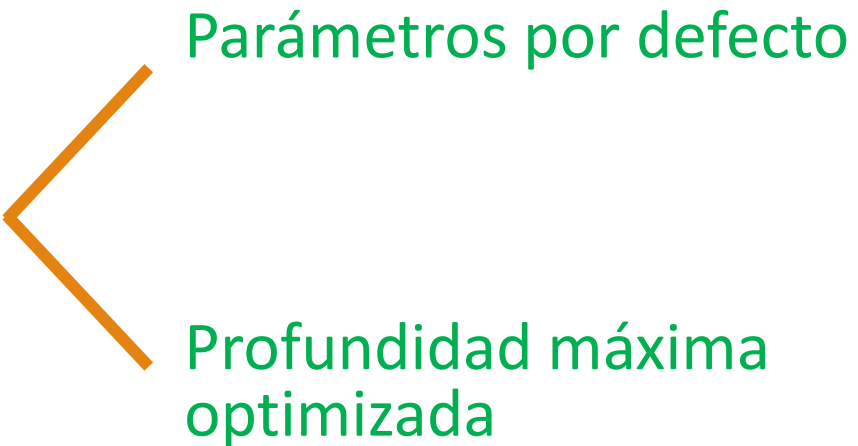
Análisis – Ventas por Tipo de Outlet



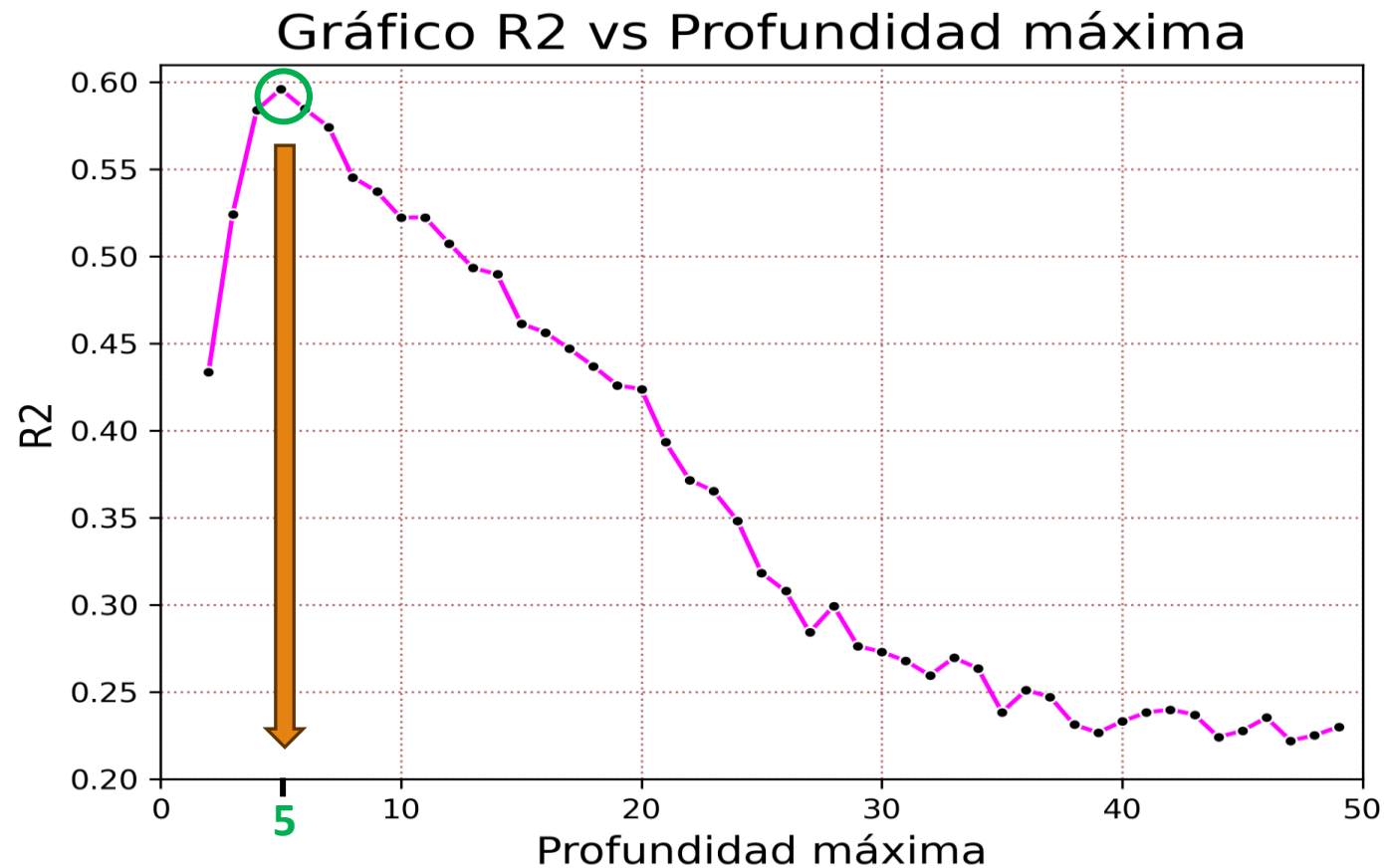
Análisis – Ventas por Tamaño de Outlet



Modelos probados

- Regresión Lineal
 - Árbol de Decisión
- 
- Parámetros por defecto
- Profundidad máxima optimizada
- The diagram consists of two orange lines forming a right-pointing chevron shape. The top line connects the 'Árbol de Decisión' text to the 'Parámetros por defecto' text. The bottom line connects the 'Árbol de Decisión' text to the 'Profundidad máxima optimizada' text.

Optimización de parámetros



Resultados

MODELO	R2		RECM	
	Entrenamiento	Prueba	Entrenamiento	Prueba
Regresión Lineal	0.672	0.383	985	1.304
Árbol de decisión (por defecto)	1.000	0.243	0.000	1.444
Árbol de decisión (Óptimo)	0.604	0.596	1.082	1.055

- R2: valor más alto es mejor
- RECM: valor más bajo es mejor

Recomendación de Modelo

Árbol de decisión
Optimizado
(Profundidad de 5 niveles)

