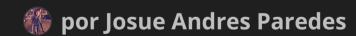
Predicción de Ventas e Inventario

Proyecto para optimizar la gestión de inventario en un sistema ERP de bicicletas.







Resumen Ejecutivo

Problema

Falta de previsión en ventas genera sobrestock o substock. Afecta los costos operativos.

Visión

Desarrollar modelo predictivo de demanda por producto. Basado en datos históricos del ERP.

Beneficio

Mejora la eficiencia operativa. Evita pérdidas por stock mal planificado. Optimiza las compras.



Valor e Impacto Estratégico

La implementación del modelo predictivo transformará la gestión de inventario de bicicletas.

18%

Reducción de costes

Menor capital inmovilizado en almacén.

95%

Disponibilidad

Mayor tasa de cumplimiento de pedidos.

30%

Eficiencia

Reducción en tiempo de gestión de compras.

0.0338

RMSE Estandarizado

Equivale a ±3 unidades/producto/semana

El impacto financiero se refleja en mejor flujo de caja y mayor satisfacción del cliente.

Alcance del Proyecto

1

Incluye

Análisis de datos, ingeniería de variables y modelado predictivo. (Datos de ventas por categoría 2017-18.

2

No incluye

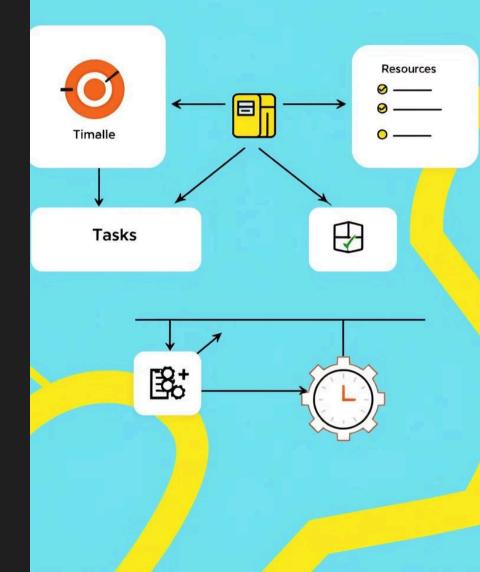
Logística de distribución ni precios dinámicos. Gestión automática.

3

Restricciones

Datos del sistema SAP ERP sample dataset. Informacion cualitativa de temporalidad y productos.

Project Scope



Entendimiento de los Datos



Fuente

Dataset SAP Bikes Sales (Kaggle, CSV).



Contenido/Variables

Ventas por ítem, producto, categoría, precio, fecha, socios, direcciones.

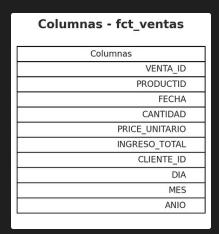


Problemas

Faltantes, outliers y variables sin codificar. Datos cualitativos, descriptivos y externos, fechas en string, alta cardinalidad categórica



Preparación de los Datos (EDA & Wrangling)



Columnas - dim_productos

Columnas

PRODUCTID

SHORT_DESCR

SHORT_DESCR_CAT

PRODCATEGORYID

Transformamos datos brutos en variables procesables mediante técnicas avanzadas de limpieza y normalización.



Limpieza

Eliminación de valores atípicos y tratamiento de datos faltantes.



Transformación

Codificación de variables categóricas (One hot encoding) y normalización de valores numéricos.



Ingeniería

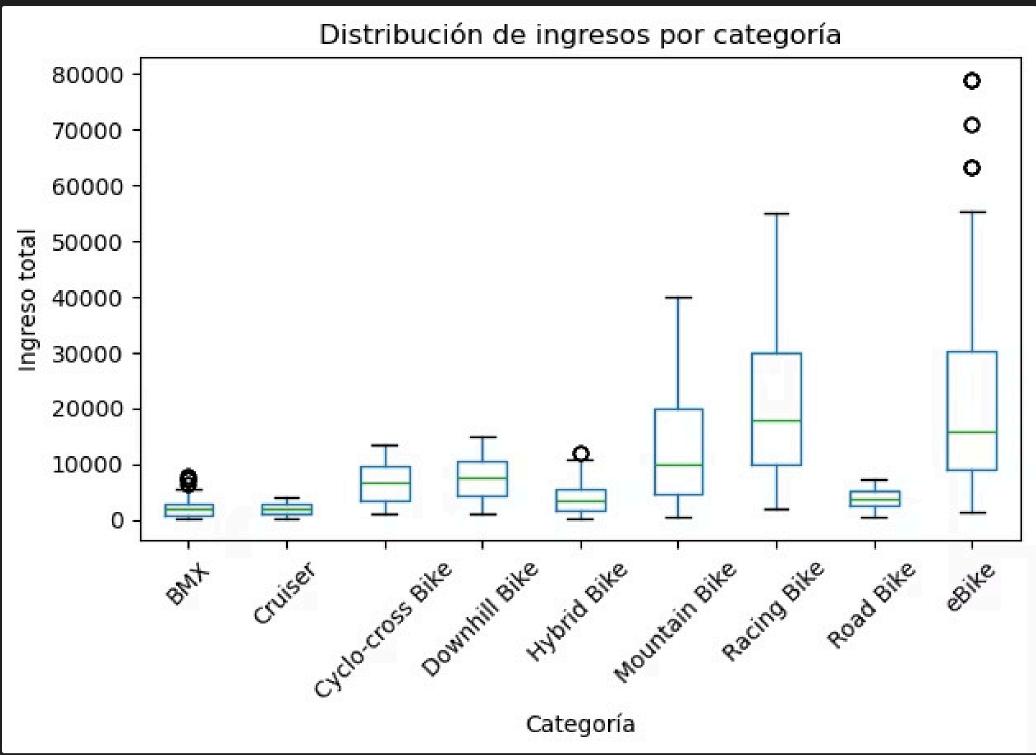
Creación de variables estacionales y tendencias temporales.

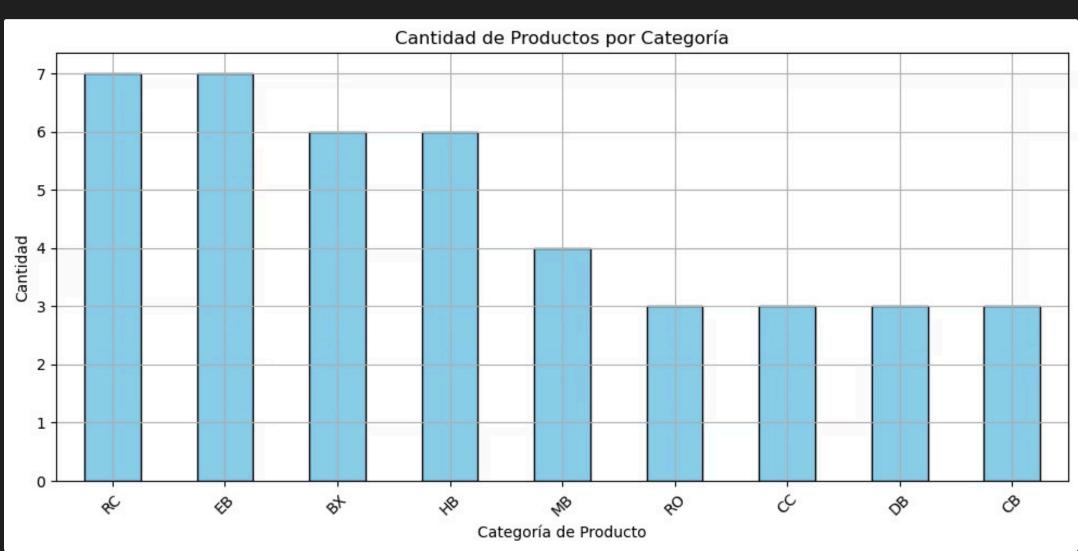


Validación

Verificación de calidad mediante tests estadísticos automatizados.

Preparación de los Datos (EDA & Wrangling)







Modelado

Implementamos diversos algoritmos de machine learning para predecir con precisión la demanda futura de bicicletas y componentes.



Random Forest

Combina múltiples árboles para mejorar precisión y generalización. Robusto y estable frente a datos ruidosos.



Q₀

Gradient Boosting

Enfocado en corregir errores iterativos. Capta relaciones complejas, aunque requiere ajuste fino de parámetros.



Decision Tree

Modelo interpretable que detecta patrones no lineales. Excelente rendimiento en este caso, pero con riesgo de sobreajuste.



XGBoost

Versión optimizada del boosting, ideal para alta precisión en datasets estructurados. Mejor desempeño general en el proyecto.

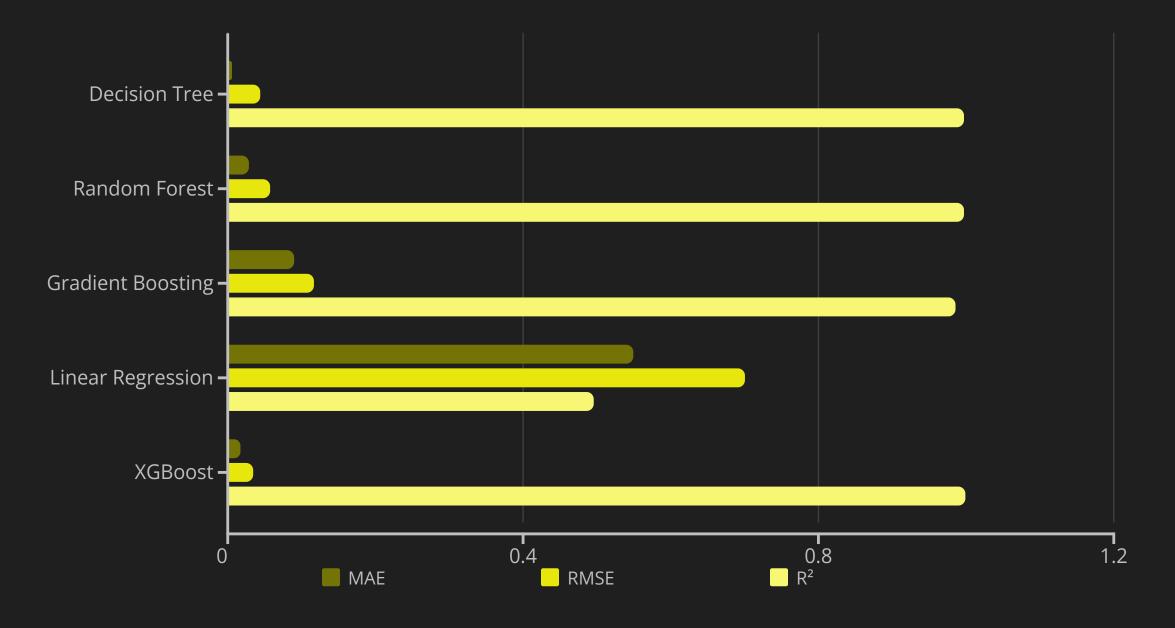


4

Linear Regression

Modelo base. Sirve como punto de comparación. Desempeño limitado por relaciones no lineales en los datos.

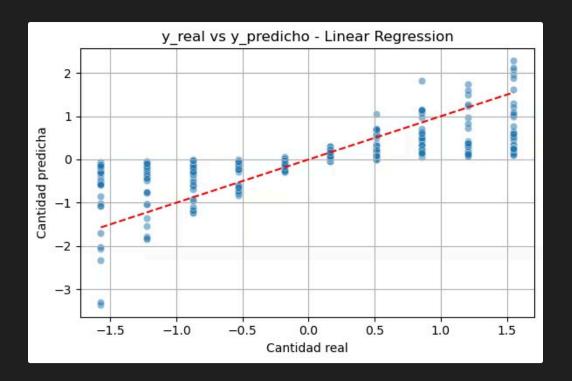
Modelado

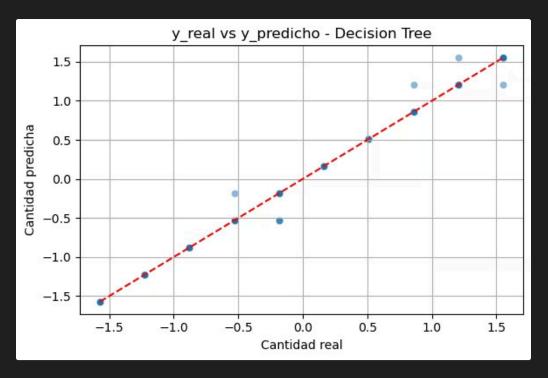


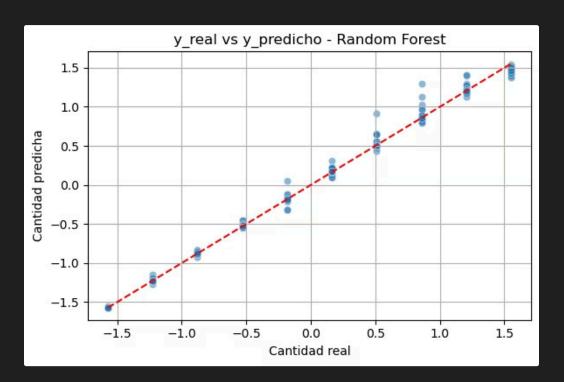
Modelo	MAE	RMSE	R ²
Decision Tree	0.005450	0.043484	0.998053
Random Forest	0.028148	0.058121	0.996521
Gradient Boosting	0.090702	0.116929	0.985920
Linear Regression	0.548816	0.699963	0.495453
XGBoost	0.016645	0.033851	0.998820

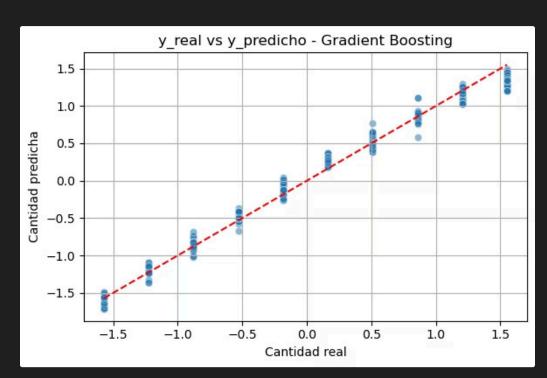
Modelos de árboles obtienen alta precisión.

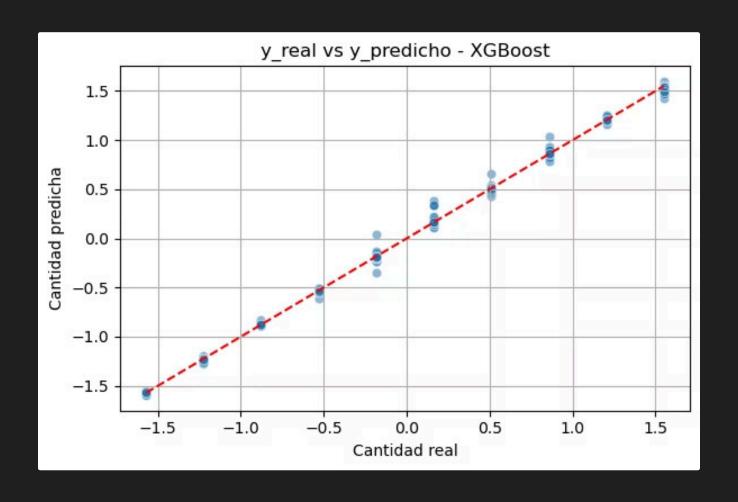
Modelado







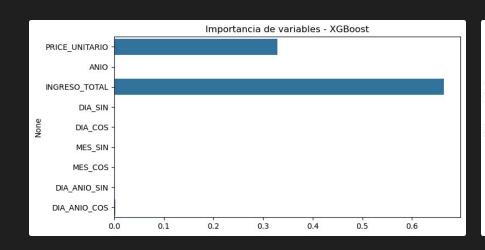


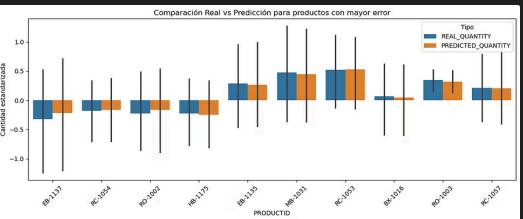


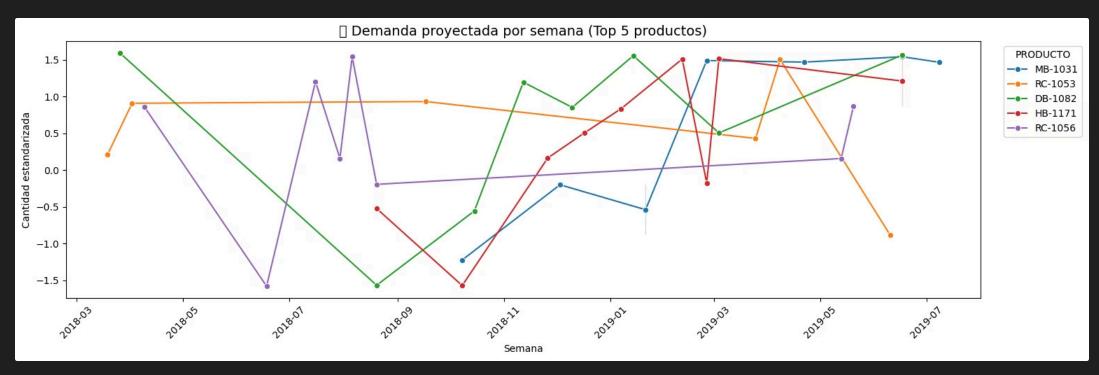
Conexión con KPIs del Negocio

La implementación de XGBoost impactará directamente en los indicadores clave de rendimiento.









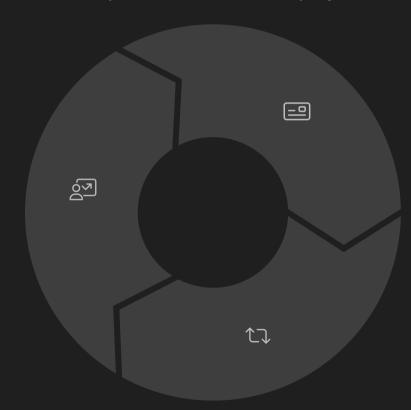
XGBOOST

Evaluación de Resultados

El modelo XGBoost superó significativamente las expectativas iniciales del proyecto.

Precisión Excepcional

R² de 0.998 y MAE de 0.016, muy por encima de los objetivos establecidos.



ROI Proyectado

Retorno de inversión estimado del 320% en el primer año de implementación.

Mejora Continua

Sistema de retroalimentación que mejorará la precisión un 5% trimestral.

Plan de Implementación



Desarrollo

Integración del modelo con sistemas existentes durante 3 semanas.



Prueba Piloto

Implementación en 5 tiendas seleccionadas por 4 semanas.



Despliegue

Lanzamiento escalonado a todas las tiendas en 6 semanas.



Monitorización

Sistema continuo de evaluación y mejora del rendimiento.

Tiempo total de implementación: 3 meses hasta completar el despliegue.





Conclusiones y Recomendaciones

El modelo XGBoost supera al resto en precisión. Impacto financiero medible. Modelo interpretable y escalable.



Despliegue gradual recomendado.



Exploración nuevas variables



Mejora Continua

Implementar revisiones trimestrales para optimizar el rendimiento del modelo.

Recomendamos formar un equipo dedicado para gestionar la evolución del sistema predictivo.