

Informe de Análisis de Ventas de Vehículos en Argentina

Contexto Empresarial y Problema

Este análisis se realiza sobre un dataset de ventas de vehículos en Argentina, con el objetivo de identificar patrones que permitan predecir si un vehículo vendido es nuevo o usado.

Esto podría mejorar la toma de decisiones en concesionarias para orientar campañas de marketing y planificación de stock.

Reducción de Dimensionalidad

Se utilizó el método de selección de características SelectKBest con la prueba Chi-cuadrado (χ^2), que evalúa la dependencia estadística entre cada variable independiente y la variable objetivo (Vehiculo_Nuevo). Previamente se discretizaron las variables numéricas usando KBinsDiscretizer y se aplicó codificación one-hot a las variables categóricas.

Modelado Predictivo

Se utilizaron dos algoritmos de clasificación para predecir si un vehículo es nuevo:

- Regresión Logística
- Random Forest Classifier

Ambos son apropiados para problemas de clasificación binaria. La regresión logística es interpretativa, mientras que Random Forest es más robusto ante no linealidades y valores atípicos.

Evaluación de Modelos

Se aplicó validación cruzada (5-fold), obteniendo una precisión promedio de:

- Regresión Logística: 99.55%
- Random Forest: 99.95%

Ambos modelos alcanzaron una exactitud casi perfecta, con métricas de precisión, recall y F1-score superiores al 99%.

Esto indica que el conjunto de datos tiene características altamente predictivas, en especial el kilometraje y el año del modelo del vehículo.

Informe de Análisis de Ventas de Vehículos en Argentina

Conclusiones y Aplicación Práctica

- Las variables más relevantes fueron: Kilometraje y Modelo del Vehículo.
- Random Forest mostró mayor estabilidad (menor desviación estándar).
- La variable objetivo (Vehiculo_Nuevo) está altamente correlacionada con el kilometraje (cercano a 0).
- El modelo puede ayudar a:
 - Identificar si un vehículo será vendido como nuevo o usado.
 - Optimizar campañas comerciales según características del vehículo.
 - Planificar inventario y logística.

Milanesio Ignacio José