

Desafío Técnico para Científico de Datos

Contexto:

Como parte del equipo de ciencia de datos de una gran cadena de tiendas. Tu objetivo es desarrollar un modelo predictivo que permita pronosticar las ventas diarias a nivel de producto (SKU) en cada una de las tiendas de la cadena. Este modelo es crucial para optimizar la planificación de inventarios, la gestión de promociones y la toma de decisiones estratégicas en la cadena de suministro.

Descripción del problema:

La cadena de tiendas enfrenta desafíos en la gestión eficiente de inventarios debido a la variabilidad en las ventas de productos, que se ve influenciada por múltiples factores como promociones, estacionalidad y eventos especiales. Un modelo preciso de predicción de ventas ayudará a reducir costos de almacenamiento, minimizar pérdidas por productos no vendidos y mejorar la disponibilidad de productos en las tiendas.

El desafío es desarrollar un modelo que pronostique las ventas diarias a nivel de SKU para cada tienda, considerando los factores mencionados.

Tareas:

1. Exploración y limpieza de datos:

- Realiza un análisis exploratorio del dataset proporcionado.
- Limpia y preprocesa los datos, manejando valores nulos, duplicados, y cualquier inconsistencia en los datos.
- Crea nuevas variables derivadas (features) que puedan ser útiles para la predicción.

2. Análisis de la estacionalidad y tendencias:

- Analiza la estacionalidad y las tendencias en las ventas para diferentes productos y tiendas.
- Identifica patrones recurrentes que podrían influir en las ventas, como los días de la semana, feriados, eventos especiales, etc.

3. Modelado Predictivo:

- Desarrolla un modelo de predicción de ventas a nivel de SKU.
- Experimenta con diferentes técnicas de modelado y selecciona el modelo que mejor se ajuste a los datos.

4. Interpretación de Resultados y Recomendaciones:

- Interpreta los resultados del modelo y explica cómo los diferentes factores (promociones, estacionalidad, etc.) influyen en las predicciones.
- Ofrece recomendaciones sobre cómo la empresa podría utilizar los pronósticos para mejorar la gestión de inventarios y las estrategias de promoción.

Entregables:

- Un informe detallado que incluya el análisis exploratorio de datos, la metodología de modelado, resultados del modelo, y recomendaciones.
- Un notebook de Jupyter con el código utilizado para la limpieza de datos, modelado predictivo, y evaluación del modelo.
- Gráficos y visualizaciones que respalden tus hallazgos y expliquen las predicciones.
- Presentación de 15 minutos:
 - Preparar una presentación en PowerPoint o Google Slides que muestre los entregable y explicando cómo los resultados pueden impactar al negocio.

Recursos proporcionados:

- Dataset con ventas históricas diarias a nivel de SKU.
- Información sobre promociones, feriados nacionales, y eventos especiales.