



DESAFÍO INGELEAN

ANÁLISIS PREDICTIVO EMPRESARIAL PARA LA TOMA DE DECISIONES ESTRATÉGICAS

RESUMEN

INGE LEAN S.A.S. es una empresa de ingeniería con sede en Pereira (Risaralda), fundada en 2013, que brinda soluciones a medida en software, hardware, automatización industrial, inteligencia artificial y mantenimiento. Su misión es impulsar la eficiencia y competitividad de procesos industriales y comerciales en el Eje Cafetero

Actualmente, INGE LEAN enfrenta un reto en su canal de atención al cliente: los tiempos de respuesta son largos, las interacciones carecen de personalización y el equipo humano necesita atender el volumen de solicitudes relacionadas con sus soluciones tecnológicas. Como consecuencia, se pueden perder oportunidades comerciales y degradan la experiencia del usuario.

OBJETIVO DEL RETO

Diseñar una solución básica de análisis predictivo empresarial que utilice datos históricos simples (ventas, inventario o clientes) para generar predicciones iniciales sobre una variable clave del negocio, acompañadas de visualizaciones comprensibles y recomendaciones accionables.

RETO NIVEL BÁSICO – CHATBOT PARA ATENCIÓN AL CLIENTE

CONTEXTO DEL PROBLEMA

En un entorno empresarial en constante cambio, la toma de decisiones estratégicas no puede depender de la intuición. Las organizaciones requieren herramientas que les permitan anticipar comportamientos del mercado, prever riesgos financieros y responder de manera proactiva a las necesidades del cliente. Sin embargo, muchas empresas aún carecen de soluciones efectivas que aprovechen el poder de los datos para anticipar el futuro.

El análisis predictivo empresarial, apoyado en técnicas de inteligencia artificial, modelado estadístico y visualización de datos, brinda la posibilidad de transformar datos históricos en conocimiento útil para la acción. Este reto está diseñado para introducir a los campistas en el desarrollo de modelos predictivos aplicados a escenarios empresariales reales, permitiéndoles extraer valor a partir de la información disponible y generar soluciones de impacto desde el primer nivel.

NIVEL BÁSICO – RETO PARA CAMPISTAS

El objetivo de este nivel es construir un modelo predictivo sencillo utilizando datos históricos de una empresa, complementado con visualizaciones claras que ayuden a entender mejor los patrones y tendencias identificadas.

REQUISITOS TÉCNICOS

- Recolectar y limpiar un conjunto de datos proporcionado (por ejemplo: ventas mensuales, comportamiento de clientes o inventarios).



- Crear visualizaciones básicas que ayuden a interpretar los datos utilizando herramientas como Matplotlib, Seaborn o Power BI.
- Implementar un modelo predictivo básico (como regresión lineal, regresión logística o árboles de decisión) para anticipar una métrica empresarial relevante (ej. ingresos mensuales, rotación de clientes).
- Presentar un breve informe donde se expliquen las predicciones generadas, los hallazgos clave y recomendaciones simples para la empresa.

RESULTADO ESPERADO

Se espera una solución inicial que combine análisis de datos y visualización, junto a un modelo predictivo funcional capaz de generar estimaciones realistas sobre una métrica empresarial. Esta solución debe ser comprensible para una audiencia sin conocimientos técnicos profundos.

RECOMENDACIONES

- Asegúrense de entender bien qué representa cada variable en el conjunto de datos.
- Utilicen títulos, colores y etiquetas claras en las gráficas para facilitar la interpretación.
- Evalúen la precisión de sus modelos usando métricas simples como el error absoluto medio (MAE) o el coeficiente de determinación (R^2).
- Mantengan la presentación visual y el informe final organizados, claros y orientados a la toma de decisiones.

NIVEL EXPLORADOR – VISUALIZACIÓN DE DATOS

OBJETIVO

- Implementar un modelo predictivo inicial sobre un conjunto de datos proporcionado o construido por el equipo, y representar los resultados en forma visual y comprensible para usuarios no técnicos.

REQUISITOS TÉCNICOS

- Recolectar o usar un dataset predefinido (ventas, ingresos, inventario, etc.).
- Realizar limpieza básica y análisis exploratorio de datos (EDA).
- Implementar un modelo predictivo simple (como regresión lineal, árbol de decisión u otro algoritmo de regresión supervisada).
- Utilizar herramientas como Matplotlib, Seaborn, Plotly o Power BI para representar:
 - El comportamiento histórico de la variable.
 - Las predicciones generadas por el modelo.
- Elaborar un breve reporte técnico con las predicciones y al menos 2 recomendaciones estratégicas con base en los resultados.

CRITERIOS DE ÉXITO

- Calidad del modelo
El modelo entrega predicciones coherentes con los datos históricos
- Claridad de las visualizaciones
Gráficos comprensibles, con etiquetas y análisis básico



- Limpieza y estructura del dataset
Datos bien preparados y documentados
- Coherencia en las recomendaciones
Las conclusiones tienen sentido según los resultados
- Presentación del flujo de trabajo
Código organizado, notebook o documento claro

ENTREGABLES GENERALES

Dataset limpio y documentado

- Formato: CSV, Excel, o base estructurada desde API o simulación.
- Acompañado de descripción de variables y justificación de elección.
- Puede estar incluido en el notebook o en el repositorio.

Notebook o script de análisis

Debe incluir:

- Proceso de carga y limpieza de datos.
- Análisis exploratorio (EDA).
- Implementación del modelo predictivo.
- Métricas básicas de rendimiento.
- Visualizaciones con etiquetas y análisis.
- Herramientas sugeridas: Jupyter Notebook, Google Colab, VSCode, etc.

Gráficos y visualizaciones

- Mínimo 2 visualizaciones:
- Una del comportamiento histórico de la variable objetivo.
- Una que muestre las predicciones vs. los valores reales.
- Herramientas posibles: Matplotlib, Seaborn, Plotly, Power BI, etc.

Reporte breve (puede ser en markdown o PDF)

- Contenido:
- Explicación del enfoque.
- Resultados del modelo.
- Al menos dos recomendaciones estratégicas basadas en los resultados.
- O bien, esta información puede estar en el mismo notebook, como celdas explicativas.

Repositorio en GitHub o ZIP organizado

- Con el código, dataset, notebook y visualizaciones.
- Incluir README.md corto con instrucciones y descripción.

Pitch final (5 minutos)

- En vivo durante la etapa final del hackatón.
- Puede ser acompañado de slides o demo funcional.

**1. NOMBRE DEL EQUIPO**

Nombre creativo y representativo

2. NOMBRE DEL PROYECTO

DataVision

3. DESCRIPCIÓN GENERAL DEL PROYECTO (MÁX. 100 PALABRAS)

¿Qué problema empresarial aborda su solución?

¿Qué variable están intentando predecir?

¿Qué tipo de empresa o sector imaginaron?

Nuestro proyecto pretende optimizar el proceso de toma de decisiones estratégicas en una compañía del sector comercial a través del análisis predictivo. Empleamos información histórica de ventas, clientes y costos para anticipar los ingresos venideros, un factor crucial para valorar el desempeño y diseñar estrategias. La solución comprende visualizaciones nítidas que presentan tendencias, comparaciones y proyecciones, lo que simplifica la comprensión para usuarios no técnicos. Visualizamos una compañía del sector retail con presencia en diversas ciudades, que persigue incrementar su rentabilidad y potenciar su habilidad para prever variaciones en la demanda.

4. DATASET UTILIZADO

¿De dónde se obtuvo? (proporcionado / simulación / fuente externa)

¿Qué variables contiene?

¿Cuántos registros y columnas tiene?

¿Qué tipo de limpieza o preprocessamiento aplicaron?

El dataset fue proporcionado por la organización como base para el análisis. Adicionalmente, generamos 6,000 filas de datos simulados con el fin de ampliar el volumen y facilitar el entrenamiento del modelo predictivo, manteniendo coherencia con los valores originales. El dataset final contiene las siguientes columnas: Fecha, Ciudad, Categoría_Producto, Productos_Vendidos, Ingresos, Gastos, Clientes_Atendidos, Margen (%) (Margen de Ganancia) y Ticket_Promedio. En total, el dataset final cuenta con 7,500 registros y 9 columnas.



5. MODELO PREDICTIVO IMPLEMENTADO

¿Qué algoritmo utilizaron? (Regresión lineal, Árbol de decisión, etc.)

¿Qué variable predijeron?

¿Cómo fue el rendimiento del modelo? (Errores, precisión, etc.)

¿Qué supuestos hicieron?

Se implementó un modelo de regresión lineal para predecir la variable objetivo “Ingresos”, utilizando como variables independientes: Productos_Vendidos, Gastos y Clientes_Atendidos. El modelo fue entrenado con un conjunto de datos dividido en 80% para entrenamiento y 20% para prueba. El modelo logró un error cuadrático medio (MSE) de 967445873161.7156, lo que indica el promedio del error al estimar los ingresos futuros. Entre los supuestos realizados, se consideró la existencia de una relación lineal entre los ingresos y las variables independientes seleccionadas. Además, se generó una visualización comparativa entre los valores reales y los valores predichos de los ingresos (primeras 20 muestras) para evaluar visualmente la precisión del modelo.

6. VISUALIZACIONES

¿Qué tipo de gráficos generaron?

¿Qué muestran las gráficas?

¿Cómo se comparan las predicciones con los datos reales?

¿Qué herramienta usaron? (Matplotlib, Seaborn, Power BI, etc.)

Se generaron dos gráficos principales con la librería Matplotlib: un gráfico de barras comparativo que muestra los ingresos reales frente a los ingresos predichos para las primeras 20 muestras, permitiendo evaluar visualmente el desempeño del modelo de regresión lineal, y un gráfico de líneas con doble eje Y que representa la tendencia anual de los ingresos y gastos promedio, facilitando la identificación de patrones o variaciones a lo largo del tiempo. Estas visualizaciones permiten analizar la precisión del modelo y comprender la dinámica histórica de ingresos y gastos del negocio.

7. RECOMENDACIONES ESTRÁTÉGICAS

Redacten al menos 2 recomendaciones empresariales derivadas del análisis.

Deben estar justificadas con base en las predicciones y gráficas.

A partir del análisis predictivo y las visualizaciones obtenidas, recomendamos enfocar los esfuerzos comerciales en las ciudades y categorías con mayores ingresos promedio, ya que la tendencia anual muestra períodos específicos donde los ingresos superan significativamente los gastos, lo que podría potenciarse mediante campañas dirigidas. Además, con base en las predicciones del modelo, se sugiere optimizar el control de gastos en los meses o categorías con márgenes negativos, revisando costos operativos y estrategias de precios para mejorar la rentabilidad. Estas acciones están sustentadas en la comparación entre ingresos reales y predichos, así como en las tendencias históricas de la empresa.

8. VALOR DIFERENCIAL DE LA SOLUCIÓN (MÁX. 80 PALABRAS)

¿Qué hace único su enfoque?

¿Cómo podría escalarse o mejorarse esta solución en una siguiente etapa?

Nuestra solución combina predicciones claras y visualizaciones intuitivas, enfocándose en métricas estratégicas como ingresos, margen de ganancia y ticket promedio. Lo que la hace única es la simplicidad de su modelo predictivo, acompañado de gráficos comprensibles para usuarios no técnicos, facilitando decisiones rápidas. En una siguiente etapa, podría escalarse con modelos más avanzados (como Random Forest o redes neuronales), incluir datos en tiempo real y automatizar reportes interactivos con herramientas como Power BI o dashboards web.

9. ENLACES DE ENTREGA

- Repositorio de código (GitHub, GitLab, etc.):
- Notebook / Dashboard interactivo (si aplica):
- Video demo (opcional, máx. 3 minutos):

10. PITCH FINAL

- Estructura sugerida para la presentación:
- Problema y variable a predecir (30s)
- Análisis de datos y modelo implementado (1 min)
- Visualizaciones clave y resultados (1 min)
- Recomendaciones y posible impacto (30s)
- Cierre creativo o llamado a la acción (30s)

11. ENTREGABLES ESPERADOS

- Dataset limpio y documentado
- Notebook o script de análisis
- Mínimo 2 visualizaciones
- Reporte breve con 2 recomendaciones estratégicas
- Repositorio en GitHub o ZIP organizado
- Pitch final en vivo (opcional: video corto)