

## 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.





## HACKATÓN TALENTO TECH – PLANTILLA OFICIAL DEL RETO

### 1. NOMBRE DEL EQUIPO

Nombre creativo y representativo

Titanes

### 2. NOMBRE DEL PROYECTO

Breve, original, relacionado con análisis y predicción

DataTitanes

### 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?

R/. La mejora de la productividad en una fabrica teniendo en cuenta ciertas recomendaciones.

R/. Vimos la importancia de la variable unidades\_defectuosas.

R/. Una empresa de producción que puede ser de diferentes áreas: metalúrgica, alimentos, automotriz, podría ser de diferentes industrias que manejen una fábrica.

### 4. DATASET UTILIZADO

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

¿Qué variables contiene?

¿Cuántos registros y columnas tiene?

¿Qué tipo de limpieza o preprocesamiento aplicaron?

R/. El dataset fue proporcionado por el cliente.

R/. Contiene variables numéricas y variables categóricas.

R/. Se revisaron los datos faltantes, donde eran pocos y numéricos se ingresó el promedio de los datos, y en los categóricos se creó un nuevo dato para los faltantes sin afecta el dataset.



## 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?

R/. Regresión lineal.

R/. unidades\_defectuosas.

R/. No fue el esperado, dado que nos arrojó un coeficiente de determinación que tiende a 0, lo cual indica que las variables no presentan una relación lineal entre ellas.

R/. Que existiera una correlación entre las variables, la cual no existe.

## 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.)

R/. Gráfico de dispersión, de linealidad, diagrama de barras.

R/. Relación entre variables no-lineal; producción por operario, índice de defectos por máquina.

R/. No son predictivos.

R/. Matplotlib, Seaborn, Pandas, Numpy.

## 7. RECOMENDACIONES ESTRATÉGICAS

Redacten al menos 2 recomendaciones empresariales derivadas del análisis.

Deben estar justificadas con base en las predicciones y gráficas.

1. La eficiencia, no depende tanto de las personas o turnos, sino del control de las condiciones de máquina como: Temperatura, vibración, consumo y ciclo. Si la empresa quiere mejorar eficiencia, debe priorizar mantenimiento preventivo y monitoreo en tiempo real de estas variables.

2. Se evidencia una diferencia en producción de acuerdo a cada uno de los operarios, independiente de la jornada o la máquina. Se recomienda prestar importancia a los operarios con menor producción y prevenir algún accidente, deserción o mayor disminución de la productividad, se recomienda mayor capacitación y acompañamiento a estos operarios.

## 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?

R/. Nuestro enfoque se centra en un análisis profundo de los datos, para evaluar la productividad de la empresa, por medio del modelamiento matemático.

R/. Implementando más herramientas de análisis y una base de datos más completa de la producción de la empresa para un mejor desempeño en el análisis.

## 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)