

Prueba Técnica para Científico de Datos

Objetivo:

Desarrollar un modelo de deserción estudiantil utilizando la base de datos proporcionada, aplicando técnicas de machine learning supervisado. El candidato deberá explorar los datos, realizar análisis preliminar y construir un modelo predictivo que permita identificar la probabilidad de deserción estudiantil. Los resultados deberán ser presentados en un dashboard que facilite la toma de decisiones.

Instrucciones

1. Exploración y análisis de datos:

- Realiza un análisis exploratorio de los datos (EDA) para entender las características principales del dataset proporcionado.
- Identifica posibles valores nulos o inconsistencias que deban ser tratadas.
- Haz una descripción general de las variables disponibles en el conjunto de datos y cómo pueden estar relacionadas con la deserción estudiantil.

2. Construcción del modelo de deserción:

- Selecciona un conjunto de algoritmos de machine learning adecuados para resolver el problema de deserción estudiantil (ej. Regresión Logística, Random Forest, SVM, etc.).
- Implementa al menos tres algoritmos y realiza un análisis comparativo de su desempeño.
- Justifica la elección del modelo final basado en métricas de desempeño.

3. Presentación de resultados:

- Crea un dashboard que permita visualizar los resultados del modelo y sea tal que sirva como insumo para toma de decisiones a la universidad. El dashboard debe incluir:
 - Distribución de las predicciones (probabilidades de deserción por estudiante).
 - Gráficos que faciliten la interpretación de los resultados por parte de usuarios no técnicos.
- Otros insights que ayuden a tomar decisiones

4. Entrega:

- El código debe estar documentado y escrito de manera clara y organizada.
- Se debe incluir un informe en formato PDF o Word que describa el proceso seguido, las decisiones tomadas, y las recomendaciones basadas en los resultados del modelo.
- El dashboard debe ser entregado en forma de una herramienta de visualización interactiva

Tiempo estimado para la prueba:

48 horas

Instrucciones de Entrega

- Sube tu código y documentación a un repositorio público de GitHub, proporciona el enlace y compártelo al correo yennijey@gmail.com . Además, tu dashboard debe estar accesible para revisión
- Envía un correo con el enlace del repositorio y cualquier instrucción adicional que consideres relevante.
- Si tienen alguna duda por favor enviarla a yennijey@gmail.com
- Siéntase libre de registrar los errores que cometa en el proceso de implementación; entendemos y aceptamos la incertidumbre y vemos oportunidades cuando los errores son comprendidos y superados. Piense en la deuda técnica y hazla explícita en la documentación.