

PRUEBA TÉCNICA DE SELECCIÓN



**VP CUMPLIMIENTO, ADC ANALÍTICA Y
CIENCIA DE DATOS.**

Cargo: Analítico 4

INTRODUCCIÓN

En el presente documento se expone la prueba de selección para el cargo de analítico 4 en la VP de Cumplimiento. Se busca que la realización de la prueba se autónoma y se muestren los conocimientos del aspirante. Cualquier copia o uso indebido de información hará la pérdida automática de la misma.

Objetivo de la prueba.

El objetivo de esta prueba es evaluar la capacidad para abordar un problema de clasificación de texto con poca información, incluyendo el flujo completo de carga, procesamiento, transformación, entrenamiento, y despliegue de un modelo, siguiendo buenas prácticas de MLOps.

Descripción del problema.

Dentro de la organización existen controles que buscan mitigar los riesgos de cumplimiento para diferentes programas internos. Estos controles vienen de manera escrita y se clasifican en 5 categorías de riesgo que pertenecen al programa SAC. El problema es que esa identificación se hace de manera manual y por lo tanto este proceso puede generar errores, manualidades y muchos reprocesos. Por lo cual se busca generar un modelo que sea capaz de clasificar los controles en las diferentes categorías de riesgo, permitiendo una automatización del proceso para la reducción de fricciones y tener una mayor certeza en la identificación correcta de los riesgos para cada control.

En el archivo, "*insumo_prueba.xlsx*" se puede encontrar la base de datos, con la siguiente información:



control: INT, Identificador del control, descripción del control que se realiza.

Riesgo 1: Si el control mitiga este riesgo (SI/NO).

Riesgo 2: Si el control mitiga este riesgo (SI/NO).

Riesgo 3: Si el control mitiga este riesgo (SI/NO).

Riesgo 4: Si el control mitiga este riesgo (SI/NO).

Riesgo 5: Si el control mitiga este riesgo (SI/NO).

Debe diseñar y construir una solución partiendo de la información dada y proponer una solución analítica E2E cumpliendo prácticas de MLOps.

Instrucciones para la prueba:

Deberás crear un pipeline de Machine Learning completo con los siguientes pasos:

1. Carga de Datos (Data Ingestion)
2. Preprocesamiento y Transformación de Datos (Data Preprocessing & Transformation)
3. Entrenamiento del Modelo (Model Training)
4. Validación del Modelo (Model Validation)
5. Despliegue (Model Deployment)
6. Documentación y Buenas Prácticas (Documentation and Best Practices)

Entregables:

Documento en donde se explique el proceso seguido para resolver la prueba. Debe incluir el detalle de todas las etapas (técnicas, hallazgos,



supuestos) que se hayan tenido para solucionar esta prueba (máximo 10 Páginas).

Códigos documentados que respalden el ejercicio analítico (ideal un repositorio en GitHub).

Presentación de los resultados del modelo, como lo considere (.ppt .pdf .pbi) y Blueprint o arquitectura de la solución.

Los documentos y archivos generados deben ser enviados a los correos del cual se le envió esta prueba antes de la fecha indicada.