Informe Proyecto MLOps: Clasificación de Calidad del Vino Blanco con Explicaciones IA

Objetivos

- Desarrollar un modelo de Machine Learning para clasificar la calidad del vino blanco usando Random Forest.
- Implementar un pipeline reproducible y versionado con MLflow para el entrenamiento y despliegue.
- Desplegar una aplicación web interactiva con Gradio que permita predicción individual y por batch (CSV).
- Incorporar explicaciones automáticas generadas mediante Google Gemini (Gen AI) para interpretar las predicciones.
- Documentar el ciclo de vida completo y verificar la trazabilidad de modelos.

Metodología

- Análisis exploratorio del conjunto de datos Wine Quality (vino blanco), con preprocesamiento y balanceo de clases.
- Entrenamiento de un modelo Random Forest, optimizando hiperparámetros y aplicando balance de clases.
- Uso de MLflow Projects para registrar experimentos, parámetros, métricas y guardar modelos en el Model Registry.
- Desarrollo de app Gradio con opciones para ingresar caracteristicas químicas manualmente o cargar archivos CSV para predicción masiva.
- Integración con API de Google Gemini para generar explicaciones en lenguaje natural basadas en las características y resultados.
- Registro de todas las explicaciones y resultados en MLflow como artifacts para trazabilidad.

Resultados

- El modelo alcanzó un desempeño adecuado en clasificación binaria, con métricas balanceadas de precisión y recall.
- El sistema de explicación automática provee interpretaciones detalladas y contextualizadas sobre la calidad del vino, ayudando a usuarios no expertos.
- El pipeline completo es reproducible y escalable, permitiendo nuevos experimentos y despliegues rápidos con MLflow.
- La app facilita la interacción y la toma de decisiones basada en datos y explicaciones generativas.

Reflexiones Éticas sobre Gen AI

- El uso de modelos generativos puede introducir sesgos o proporcionar explicaciones con incertidumbre; las explicaciones no deben ser tomadas como verdades absolutas.
- Es necesario acompañar siempre las predicciones con advertencias sobre la naturaleza probabilística del modelo y la IA generativa.
- Garantizar la transparencia en los procesos y permitir a los usuarios comprender limitaciones y aspectos éticos es fundamental.
- Se recomienda un monitoreo continuo y validación externa para asegurar que las explicaciones sean coherentes y no induzcan a error.
- La privacidad de datos, seguridad de claves API y protección del usuario final deben estar garantizadas durante todo el ciclo.