

# Bootcamp de MLOps

## Proyecto Final

### Rúbrica de retroalimentación:

- Definición del problema y objetivos (20%):
  - Claridad en la descripción del problema a resolver.
  - Establecimiento de objetivos claros y alcanzables.
  - Documentación detallada de la infraestructura necesaria para llevar a cabo el entrenamiento, despliegue y monitoreo del modelo o aplicación de ML.
  - Creación de diagramas que representen el flujo de datos en la infraestructura propuesta.
  - Justificación de la elección de despliegue basada en el objetivo final.
- Gestión de Experimentos y Modelos (15%):
  - Implementación de un sistema eficaz para el tracking de experimentos.
  - Registro adecuado de parámetros, métricas y artefactos de los experimentos.
  - Versionado efectivo de modelos.
  - Implementación de un proceso claro para la selección y promoción de modelos.
  - Documentación clara de los experimentos y sus resultados.
- Orquestación y Pipelines de ML (20%):
  - Diseño e implementación de pipelines de ML reproducibles y escalables.
  - Automatización efectiva de los flujos de trabajo de preprocesamiento, entrenamiento y evaluación.
  - Manejo adecuado de dependencias entre tareas en los pipelines.
  - Implementación de mecanismos para la paralelización y distribución de tareas.
  - Integración efectiva de los pipelines con sistemas de gestión de recursos.
- Integración continua incluyendo pruebas (15%):
  - Implementación de un pipeline de CI/CD robusto para proyectos de ML.
  - Uso de IaC para el despliegue de la infraestructura de entrenamiento o inferencia.
  - Ejecución de pruebas automatizadas.

- Automatización efectiva del proceso de entrenamiento, evaluación y despliegue de modelos.
- Despliegue del modelo (15%)
  - Despliegue del modelo en el modo más pertinente (batch, online, streaming)
  - Ejemplo de consulta de predicciones
  - Documentación clara de la API del modelo
- Monitoreo de Modelos (15%):
  - Implementación de un sistema robusto de logging para modelos en producción.
  - Configuración de alertas para detectar degradación del rendimiento del modelo.
  - Implementación de mecanismos para la detección de data drift y concept drift.
  - Diseño e implementación de dashboards para visualizar métricas clave del modelo.

### **Extra:**

- Uso de infraestructura real a través de laC
- Identificación de áreas de mejora y debilidades en la implementación del proyecto
- Versionamiento efectivo de datos.
- Establecimiento de procesos para la actualización y reentrenamiento de modelos basados en el monitoreo.

Los porcentajes son solo representativos de la relativa importancia de cada sección; la retroalimentación es solo para tu beneficio y mejora profesional.

Recuerda que a final de cuentas lo que importa es lo que sabes y puedes hacer.