

11/10/2026



Grado: 8

Materia: MLOps · Ingeniería de Datos

Antes de empezar

Reentrenar a mano cada semana no escala. Un reentrenamiento continuo, disparado por datos nuevos o por una alerta de deriva, debe validar el candidato y solo promocionarlo si supera al modelo en producción.

QUÉ APRENDERÁS

- Diseñar canalizaciones que reentrenen, validen y desplieguen modelos de forma automatizada y segura ante datos nuevos o deriva.

Automatización de reentrenamiento

Renovar el modelo sin esperar

El reentrenamiento automático cierra el círculo: cuando el mundo cambia, el modelo se renueva solo, tras un control de calidad automático que decide si el recién llegado merece el trono.

VIDEO

MLOps: qué es y cómo funciona en la práctica

Cómo automatizar el ciclo de vida del modelo, del reentrenamiento al despliegue controlado.



Escanea para verlo en Lumbre

<https://www.youtube.com/watch?v=BlBcNFNGQk0>

Por qué automatizar

- Los datos evolucionan y el modelo se desfasa.
- Reentrenar a mano es lento y propenso a olvidos.
- La automatización mantiene el sistema fresco.

Detonadores del reentrenamiento

Se dispara por calendario, por volumen de datos nuevos acumulados, o por una alerta de deriva. Cada estrategia se ajusta al ritmo de cambio del dominio.

La canalización de reentrenamiento

Un pipeline end-to-end: ingiere datos actualizados, los prepara, entrena, evalúa y empaqueta. Automatizar cada eslabón elimina intervención manual.

Validación obligatoria

- El candidato se mide contra el modelo vigente.
- Solo se promueve si supera el umbral de calidad.
- Nada entra en producción sin aprobación automática.

Puerta de calidad (gate)

Una barrera automática detiene el despliegue si métricas o pruebas de robustez fallan. Es el control que conviene la automatización de peligrosa en fiable.

Reentrenar con datos frescos

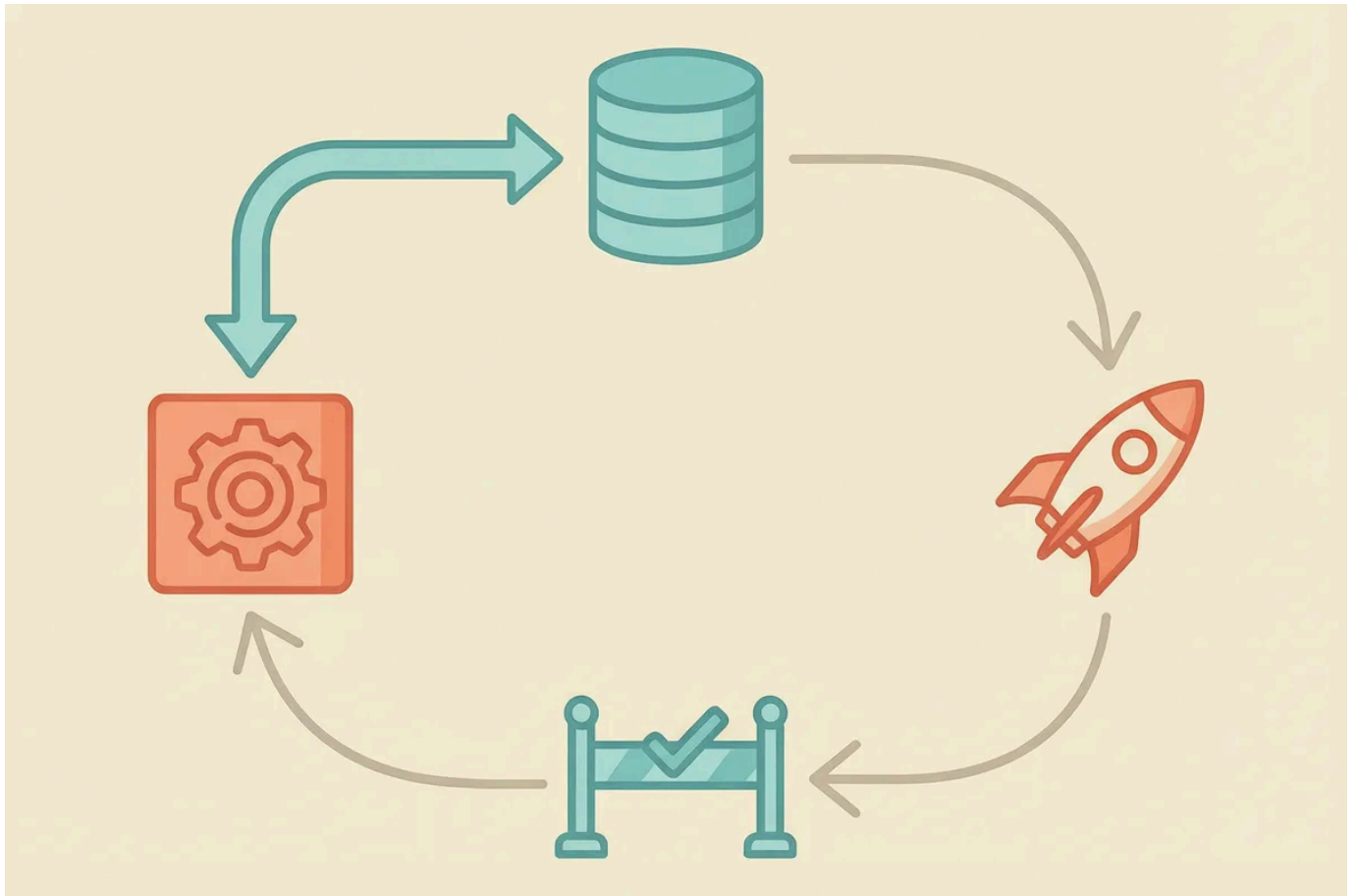
Usar ventanas recientes mantiene al día el conocimiento, pero hay que vigilar la deriva de etiquetas y el desequilibrio que los datos nuevos puedan introducir.

Aprendizaje incremental frente a desde cero

- Reentrenar desde cero: más estable, más caro.
- Actualizar incrementalmente: barato, pero riesgo de olvidar lo viejo.
- Elegir según frecuencia y coste.

Despliegue gradual

El modelo validado entra por canario o azul-verde, sirviendo poco tráfico primero. Si degrada, se revierte automáticamente.



El reentrenamiento continuo cierra el círculo: ingiere datos, entrena, valida en la gate y despliega solo si el candidato supera al vigente.

Registro y trazabilidad

Cada corrida queda registrada con sus datos, parámetros y resultados. Si algo falla aguas abajo, se sabe exactamente qué versión se desplegó y por qué.

Rollback automático

- Detectar empeoramiento tras desplegar.
- Volver a la versión previa sin esperar a humanos.
- La reversión es tan importante como el avance.

Riesgos de la automatización ciega

Automatizar sin puertas de calidad propaga errores a escala: un dato contaminado se convierte en un mal modelo en producción en horas. La gate no es opcional.

Coste y frecuencia

Reentrenar mucho gasta cómputo; poco deja el modelo oxidado. La frecuencia óptima equilibra frescura y coste según cuánto cambia el mundo.

Errores frecuentes

- Automatizar el reentrenamiento sin una gate que compare al candidato con el modelo vigente.
- Disparar el reentrenamiento con tanta frecuencia que se gasta cómputo sin ganancia real.
- No tener rollback automático y dejar un modelo degradado sirviendo producción.

Ejemplo resuelto: un pipeline que se renueva solo

Un modelo de demanda se reentrena cada lunes con los datos de la semana. El pipeline ingiere la ventana fresca, entrena un candidato y lo mide contra el vigente; solo si mejora el error medio lo despliega por canario, y si las métricas caen, revierte a la versión anterior sin intervención humana.

1. detectar el detonador: nuevos datos o una alerta de deriva
2. ingerir la ventana de datos actualizada y prepararla
3. entrenar el modelo candidato con los parámetros registrados
4. validarlo contra el vigente en la puerta de calidad
5. desplegar por canario con rollback automático si degrada

¿Para qué sirve en la realidad?

Servicios de previsión, precios dinámicos o recomendación que operan a escala no pueden reentrenar a mano: un pipeline con puertas de calidad mantiene el modelo fresco sin exponer la producción a candidatos peores.

1. La «puerta de calidad» en un pipeline de reentrenamiento automático existe para:

- ☐ acelerar el entrenamiento
- ☐ impedir que un modelo peor que el vigente pase a producción
- ☐ reemplazar el versionado
- ☐ eliminar la monitorización

2. Un reentrenamiento completamente automático sin puertas de validación es siempre seguro y confiable.

- ☐ Verdadero
- ☐ Falso

ACTIVIDAD INTERACTIVA DISPONIBLE EN LUMBRE

3. Reentrenamiento**4. Ordena el ciclo de reentrenamiento automático:****Elementos:**

- reunir los datos actualizados
- entrenar el modelo candidato
- compararlo con el vigente y aprobarlo
- liberarlo de forma gradual

Ingerir	. . .
Entrenar	. . .
Validar	. . .
Desplegar	. . .

Entrega continua https://en.wikipedia.org/wiki/Continuous_delivery Práctica de liberar software de forma automatizada y fiable.

5. Une cada elemento del reentrenamiento automático con su función:

5.	detonador	-----
5.	pipeline	-----
5.	gate	-----
5.	rollback	-----

Opciones: decide cuándo disparar el reentrenamiento · canalización de extremo a extremo que ejecuta el ciclo · barrera que frena a un candidato peor · volver a la versión previa si empeora

Reentrenamiento

Reentrenamiento automático

Disparo

calendario

volumen de datos

alerta de deriva

Pipeline

ingerir

entrenar

evaluar

empaquetar

Control

gate de calidad

despliegue canario

rollback

Riesgos

automatización ciega

deriva de etiquetas

coste y frecuencia

El mapa mental interactivo es interactivo en la versión online.



Despliegue continuo https://en.wikipedia.org/wiki/Continuous_deployment Liberación automatizada de cada cambio validado a producción.

6. Tu pipeline reentrena cada noche pero de vez en cuando despliega un modelo peor. ¿Qué le falta: un mejor detonador, una gate o un rollback? Razona.

Referencia curricular

Grado:	8
Materia:	MLOps · Ingeniería de Datos

Versión interactiva completa (con actividades, videos y autoevaluación):



<https://univ.lumbre.lat/object/cdia-mlops-004-automatizacion-de-reentrenamiento>

Material educativo generado desde Lumbre · lumbre.lat