

Monitoreo de machine learning



Monitoreo de ML

- ¿Cómo vamos a saber si se está comportando como debería?
- ¿Cómo vamos a saber si los datos que se usaron para entrenar son similares a los que se están usando para predecir?



“Los modelos se empiezan a degradar en cuanto se ponen en producción”



“Los modelos se empiezan a degradar desde que se recolectan los datos para entrenarlos”



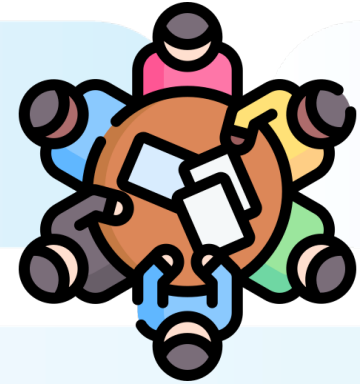
Fallos silenciosos

- Los fallos en el desarrollo de software pueden causar estruendo y lanzar excepciones, lo que provocará que las aplicaciones fallen de forma ruidosa.
- En el caso de ML ocurrir fallos silenciosos, donde los errores pasan desapercibidos pero afectan al hacer predicciones.
- Aunque el servicio siga funcionando y procesando peticiones, los errores están presentes y pueden afectar la precisión de las predicciones.



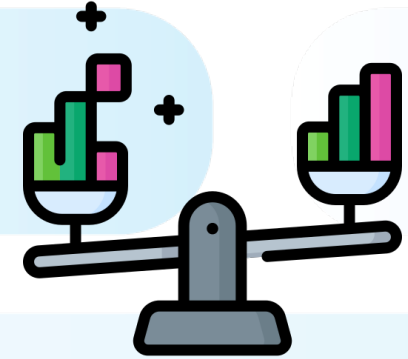
Interdisciplinaria

- DevOps se enfoca en la memoria consumida y el número de servidores que responden a las peticiones.
- Ingeniería de software se centra en la latencia y el número de predicciones por minuto.
- Ciencia de datos, el monitoreo se enfoca en las métricas del modelo y la calidad de los datos de entrada.
- El negocio, el desempeño del modelo se mide en la atracción de nuevos clientes, reducción de gastos de marketing y aumento de ventas.



Diferentes métricas

- Las métricas de negocio varían de empresa a empresa y de modelo.
- No existe un conjunto estándar de métricas de negocio debido a las diferencias en metas y objetivos comerciales.
- La diversidad dificulta el monitoreo. Es crucial seleccionar y definir unas que sean relevantes y alineadas con los objetivos de empresa.



Label lag - Ejemplo

- 50 predicciones de películas



Label lag - Ejemplo

- 50 predicciones de películas
 - 10 usuarios las ven inmediatamente
 - 15 al día siguiente
 - 5 la ven una semana después
 - 10 la ven eventualmente pero no desde su cuenta
 - 10 nunca la ven...



Label lag - Ejemplo

- 50 predicciones de películas: 10 usuarios las ven inmediatamente, 15 al día siguiente, 5 la ven una semana después, 10 la ven eventualmente pero no desde su cuenta y 10 nunca la ven...
- ¿Cómo calcularías la precisión de tu algoritmo en este caso?
- ¿Qué usuarios cuentas? los que vieron la película en los dos primeros días? ¿deberías incluir a todos los que la vieron una semana después?



Monitoreo de ML

- Monitorear modelos de machine learning en producción puede presentar varios problemas y desafíos.
- Es importante adoptar un enfoque interdisciplinario y estar consciente de las limitaciones de los modelos al evaluarlos.

