

Entrenamiento del Modelo

- Asegúrate de que el código, el entorno y los datos sean correctos antes de comenzar el entrenamiento real.
- Verifica el tiempo máximo de ejecución permitido en el entorno de entrenamiento. Revisa el balance de la plataforma en la nube. Asegúrate de que nada interrumpirá el entrenamiento.
- Asegúrate de que, si hay una interrupción, podamos reanudar la tarea. Los resultados intermedios no se perderán.
- Si hay un problema con el código, seguramente ocurrirá durante el entrenamiento.
- Si los datos no están limpios, seguramente afectarán las fases posteriores.
- Utiliza un enfoque iterativo. Escala por órdenes de magnitud. Entrena con millones de tokens, luego con miles de millones de tokens, y finalmente con billones de tokens.
- Si el entrenamiento termina, ¿qué sucederá? ¿Se ha manejado correctamente la finalización del entrenamiento?
- Aprende de los errores de otros, como los cometidos con Llama2.
- En lugar de corregir el resultado, aborda la causa raíz del problema.