

Entraînement du Modèle

- Assurez-vous que le code, l'environnement et les données sont corrects avant de commencer l'entraînement réel.
- Vérifiez le délai d'exécution maximum de l'environnement d'entraînement. Vérifiez l'équilibre de la plateforme cloud. Assurez-vous que rien n'interrompra l'entraînement.
- Assurez-vous que si une interruption se produit, nous pouvons reprendre la tâche. Les résultats intermédiaires ne seront pas perdus.
- S'il y a un problème avec le code, il se manifestera certainement pendant l'entraînement.
- Si les données ne sont pas propres, cela aura certainement un impact sur les phases ultérieures.
- Adoptez une approche itérative. Passez à des échelles de grandeur supérieures. Entraînez-vous sur des millions de tokens, puis des milliards de tokens, et enfin des billions de tokens.
- Si l'entraînement est terminé, que se passera-t-il ? La gestion finale de l'entraînement est-elle correctement effectuée ?
- Apprenez des erreurs des autres, comme celles commises avec Llama2.
- Au lieu de corriger le résultat, traitez la cause racine du problème.