

# モデルトレーニング

- 実際のトレーニングを開始する前に、コード、環境、データが正しいことを確認してください。
- トレーニング環境の最大実行タイムアウトを確認してください。クラウドプラットフォームの残高を確認してください。トレーニングが中断されないようにしてください。
- 中断が発生した場合にタスクを再開できることを確認してください。中間結果が失われないようにしてください。
- コードに問題がある場合、それはトレーニング中に必ず発生します。
- データがクリーンでない場合、それは後の段階に確実に影響を与えます。
- 反復的なアプローチを使用してください。桁違いにスケールしてください。数百万のトークンでトレーニングし、次に数十億のトークンでトレーニングし、そして数兆のトークンでトレーニングしてください。
- トレーニングが終了したら、何が起こるでしょうか？トレーニングの最終的な処理は正しく行われていますか？
- Llama2 で起こったような他人の失敗から学んでください。
- 結果を修正するのではなく、問題の根本原因に対処してください。