

PyTorch/XLA Inference

Agenda

1. How to inference
2. What to check
3. Colab example

How to inference

1. Move model and inputs to xla device
2. Torch compile the model

What to check

1. Does your model emit a single graph