COMP4901B Large Language Model

Lam Yeung Kong Sunny

Assignment 1

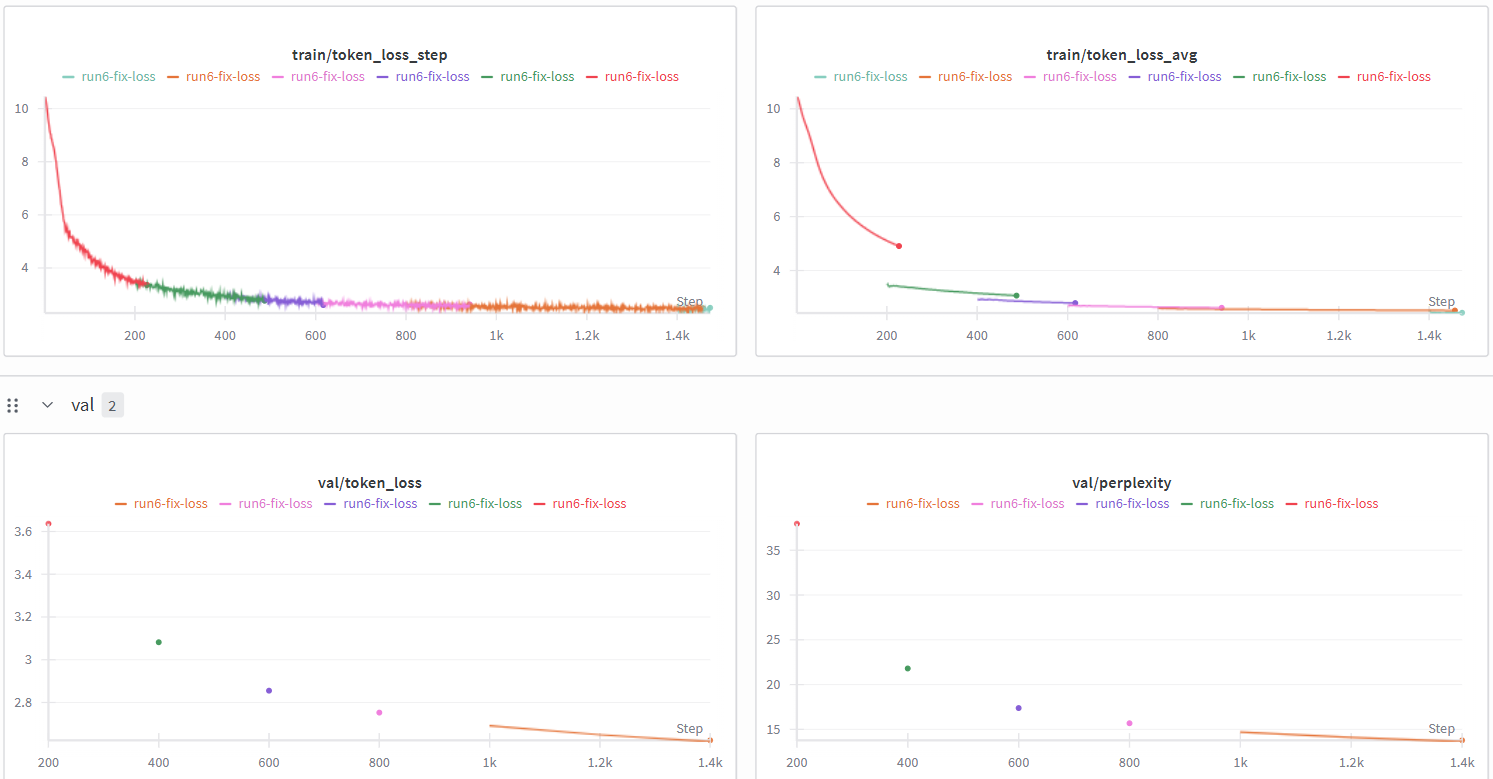
Task 1

A black text on a white background

AI-generated content may be incorrect.

Task 2

* Training setup
  + GPU
  + effective batch size : 512
  + Micro batch size: 32
  + sequence length: 256
  + LR: 0.001
  + Warmup ratio: 0.1
  + Steps: 1475 = 1 epoch



A screenshot of a computer

AI-generated content may be incorrect.