High learning rate leading to increasing training and validation loss in 3rd epoch, which means model is moving away from the local minima in the opposite direction to the expected gradient. Need to fix it with lr\_scheduler or gradient clipping.

A screen shot of a graph

AI-generated content may be incorrect.

A screenshot of a graph

AI-generated content may be incorrect.