Comparison:

* CosineAnnealingLR: The training loss decreases gradually with each epoch, indicating that the learning rate is being reduced gradually. This suggests that the model is able to find better minima over time by reducing the learning rate.
* MultiStepLR: The training loss decreases more abruptly at the beginning and then gradually with each epoch. This suggests that reducing the learning rate at specific milestones (as set by the user) has helped the model converge faster and find better minima.
* StepLR: The training loss remains constant throughout all epochs. This suggests that reducing the learning rate after a fixed number of epochs is not enough to help the model converge.

Conclusion:

Overall, all three learning rate schedulers were able to reduce the training loss, but CosineAnnealingLR and MultiStepLR appear to be more effective than StepLR in this particular case.

CosineAnnealingLR showed a gradual decrease in loss over time, which suggests that it allowed the model to find better minima.

MultiStepLR also helped the model converge faster and find better minima, by reducing the learning rate at specific milestones.On the other hand, StepLR did not seem to have any effect on the model's performance in this case, as the training loss remained constant throughout all epochs.

However, it is important to note that the effectiveness of a learning rate scheduler can vary depending on the dataset and model architecture, and it is always recommended to try out different schedulers to find the one that works best for a particular task.

Top of Form

Bottom of Form