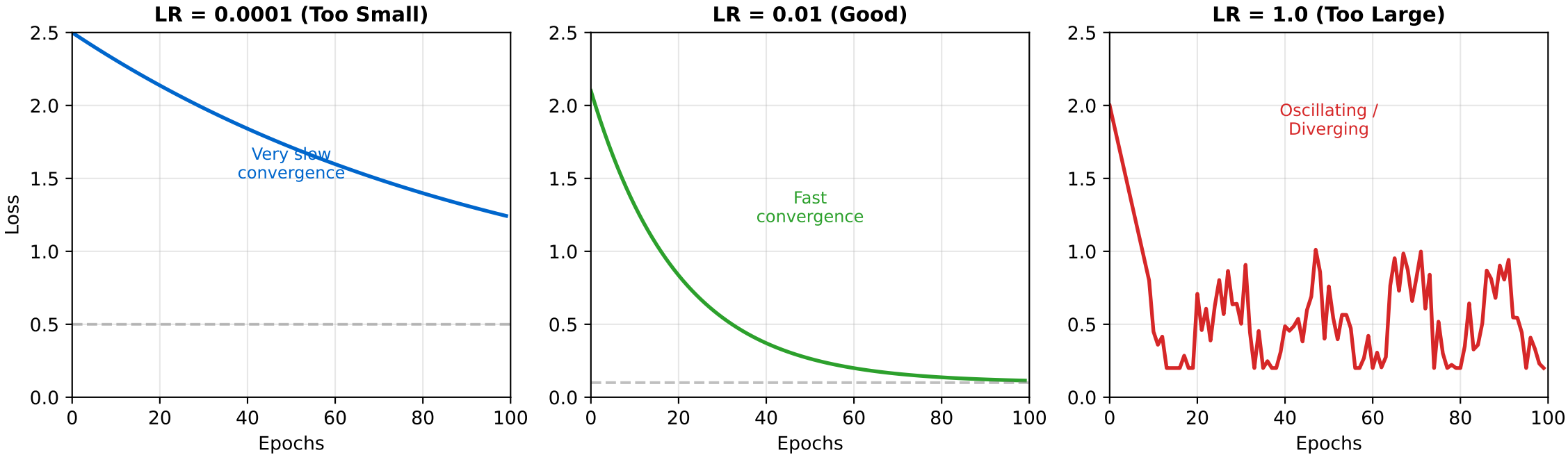
Effect of Learning Rate on Training
LR = 0.0001 (Too Small)
LR = 0.01 (Good)
LR = 1.0 (Too Large)
Loss
Epochs
Very slow convergence
Fast convergence
Oscillating / Diverging
0.0
0.5
1.0
1.5
2.0
2.5
0
20
40
60
80
100