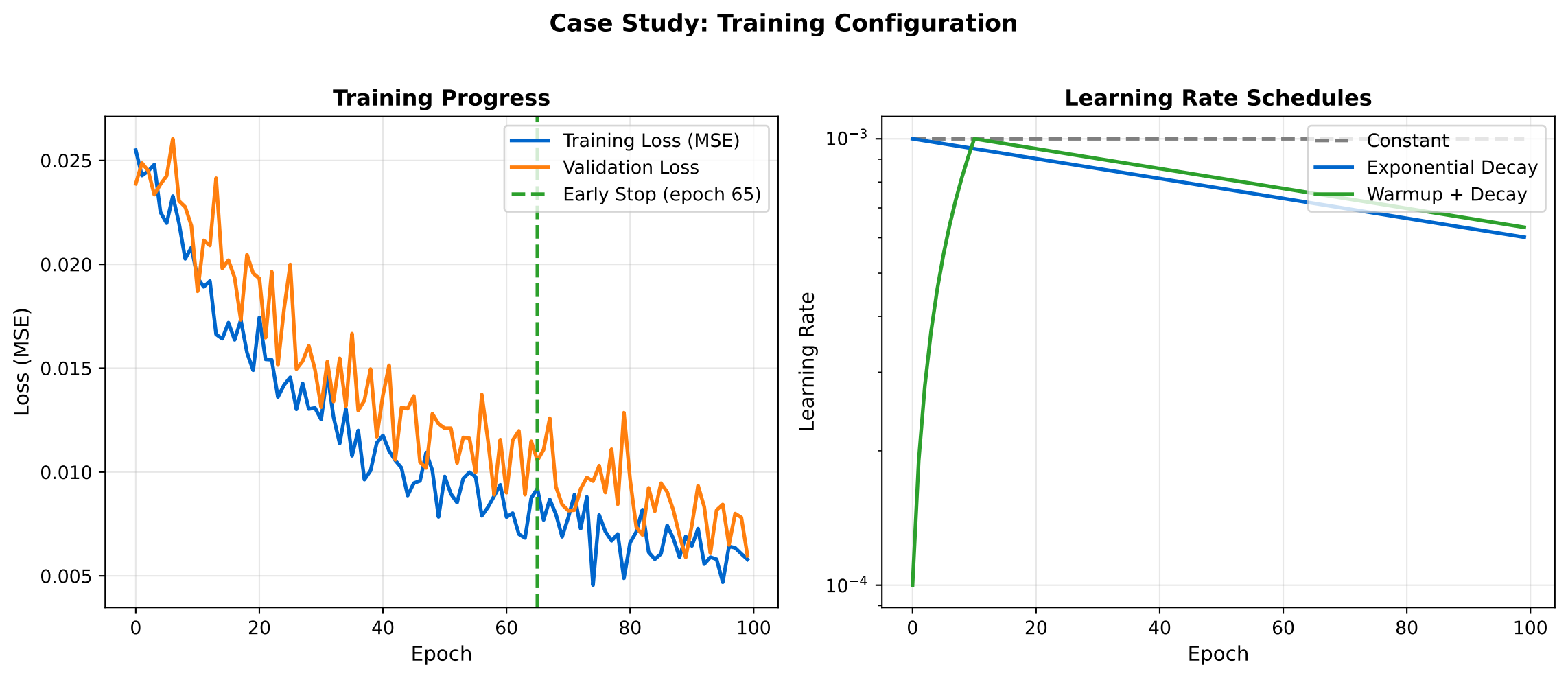
Case Study: Training Configuration
Training Progress
Training Loss (MSE)
Validation Loss
Early Stop (epoch 65)
Loss (MSE)
0.025
0.020
0.015
0.010
0.005
0
20
40
60
80
100
Epoch
Learning Rate Schedules
Constant
Exponential Decay
Warmup + Decay
Learning Rate
10^{-3}
10^{-4}
0
20
40
60
80
100
Epoch