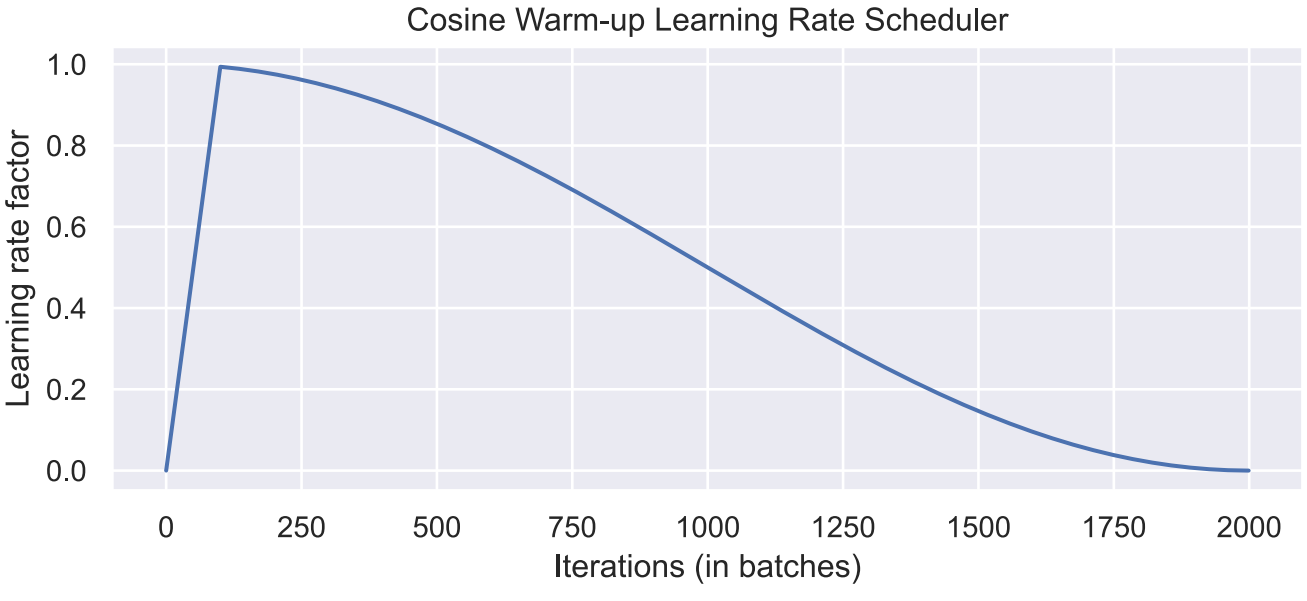
Cosine Warm-up Learning Rate Scheduler
Learning rate factor
1.0
0.8
0.6
0.4
0.2
0.0
0
250
500
750
1000
1250
1500
1750
2000
Iterations (in batches)