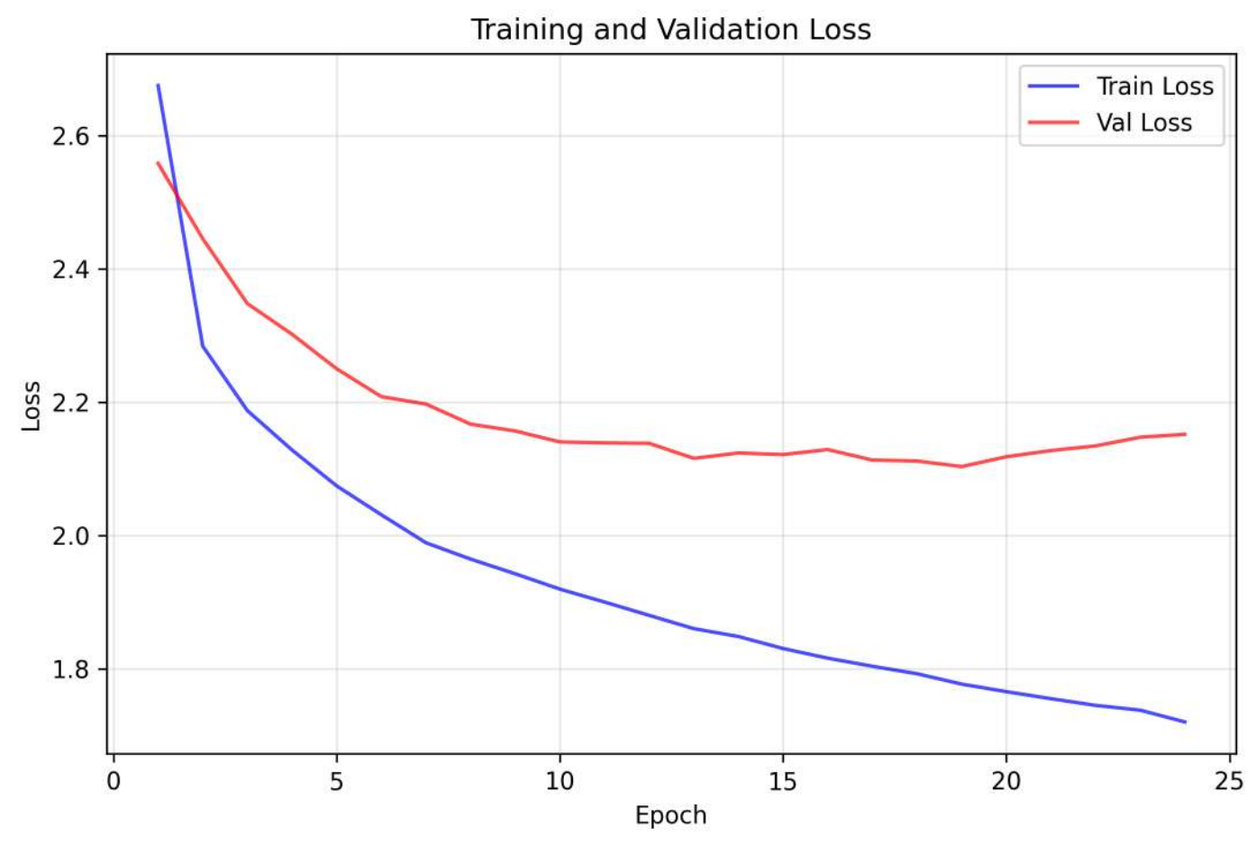
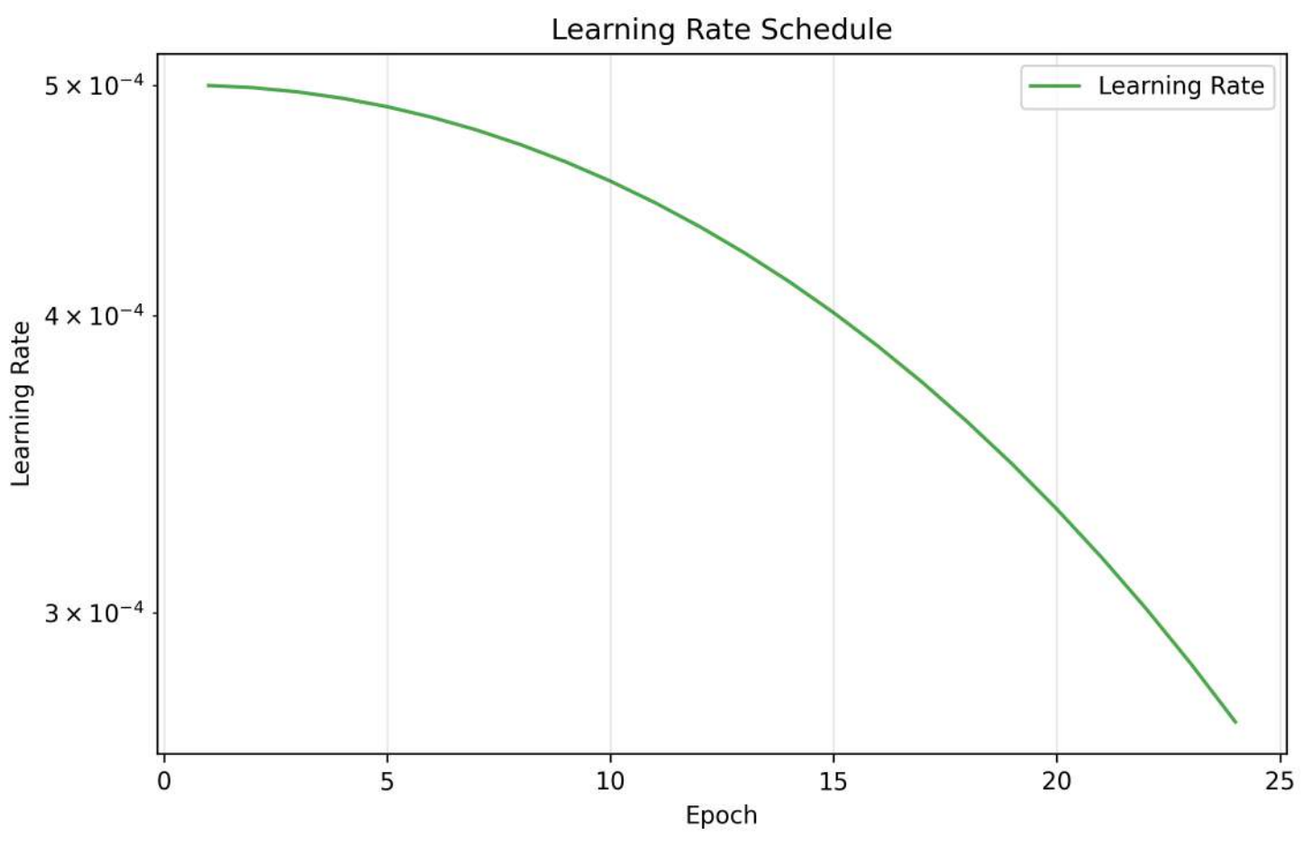

Training and Validation Loss
Train Loss
Val Loss
Loss
Epoch
2.6
2.4
2.2
2.0
1.8
0
5
10
15
20
25


Learning Rate Schedule
Learning Rate
Learning Rate
Epoch
5×10^{-4}
4×10^{-4}
3×10^{-4}
0
5
10
15
20
25