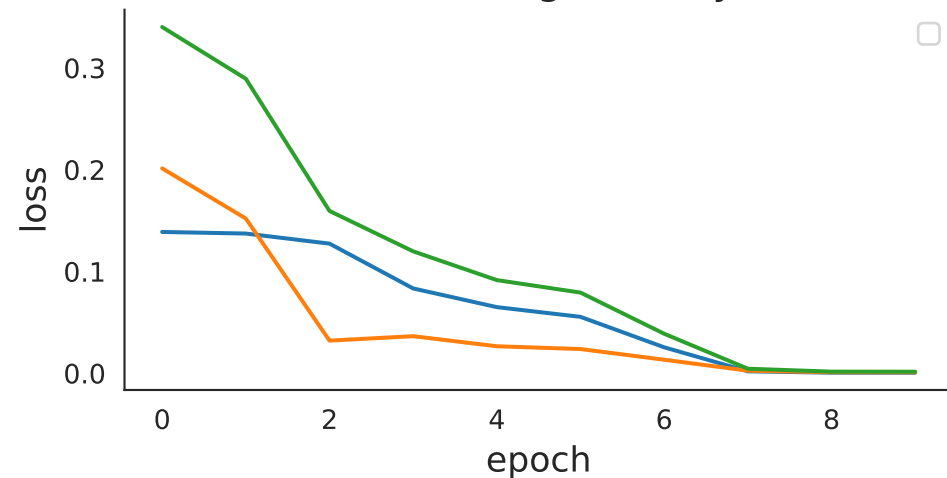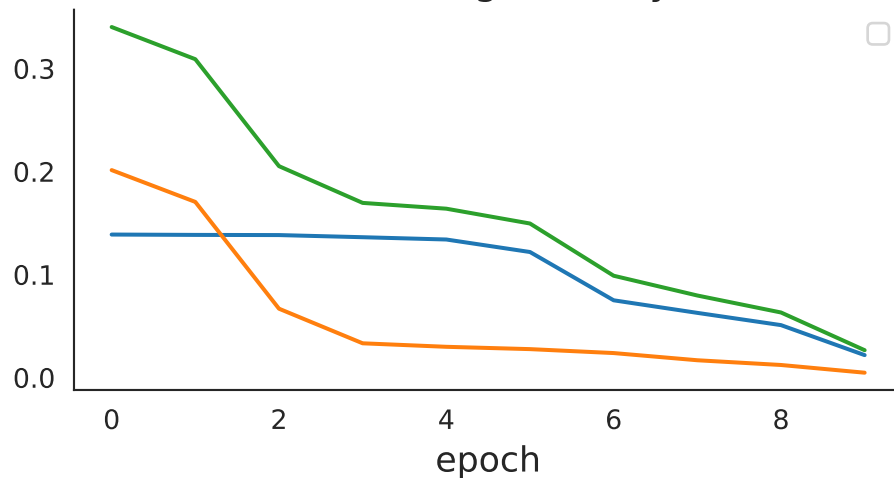**Pretraining loss when increasing weight decay**

0.001 weight decay · 0.05 weight decay · 0.01 weight decay

Legend:
- class loss
- cont loss
- (total) loss

Axis labels: loss (y-axis), epoch (x-axis)