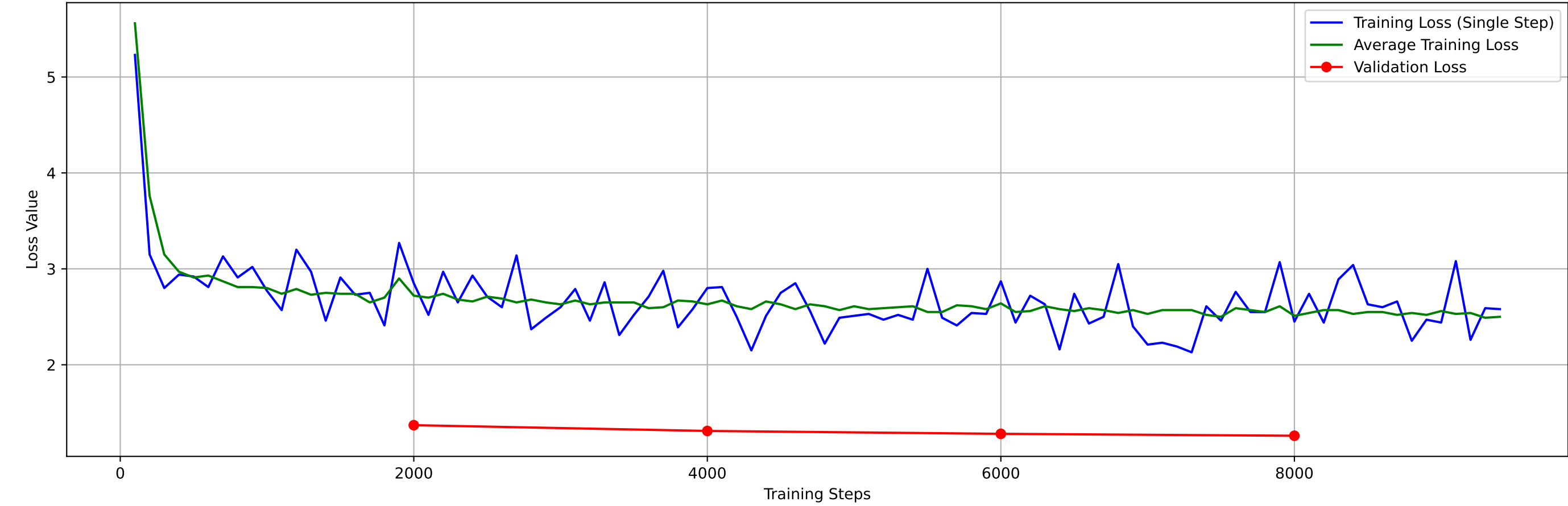
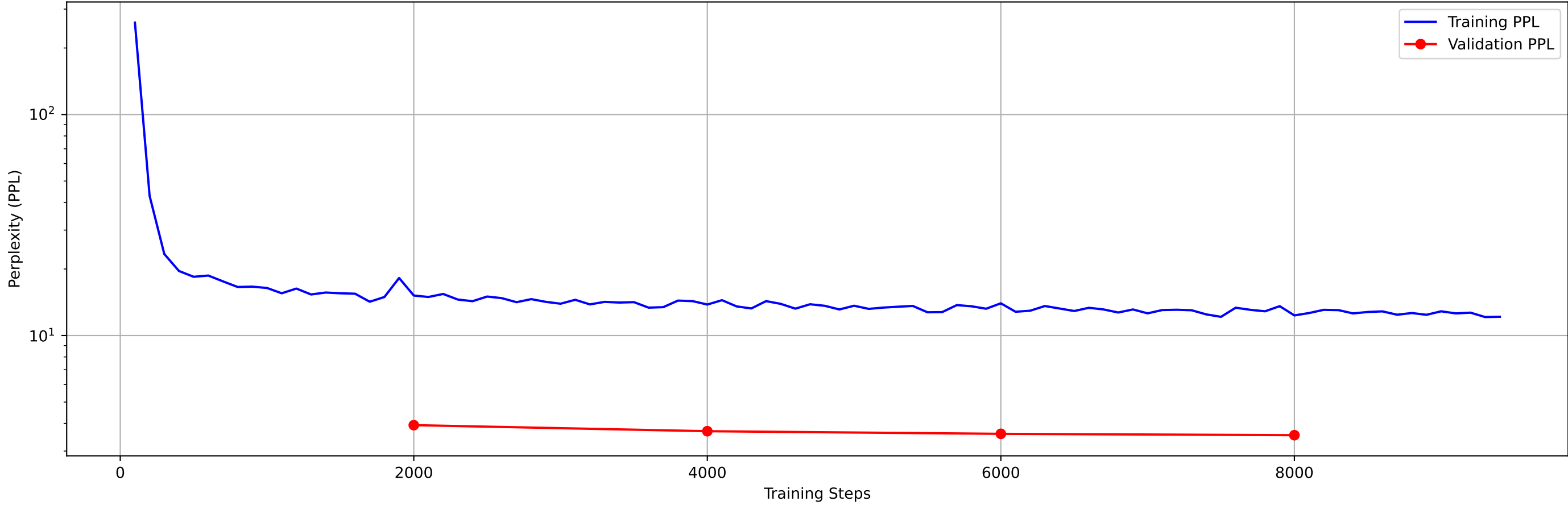


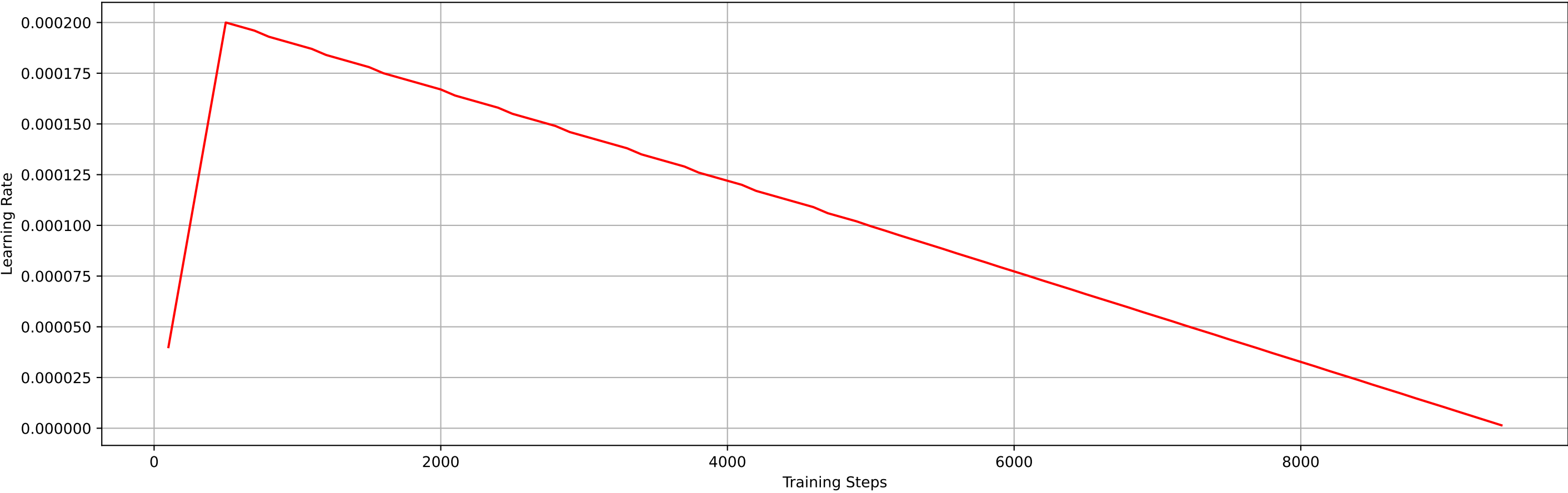
Loss Changes During Training



Perplexity Changes During Training



Learning Rate Adjustment Curve



Smoothed Training Loss Curve (Window Size=5)

