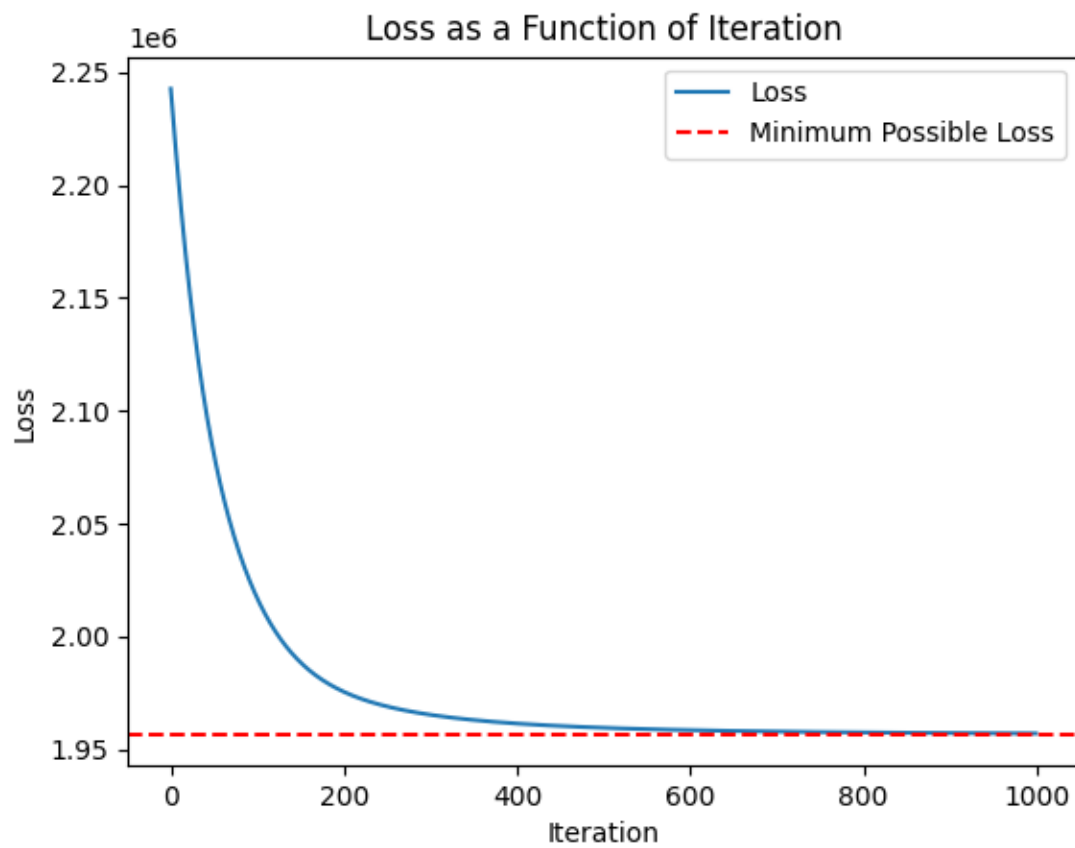NLP - Assignment – 2
Jiechen Li

Visualizations of:
1). the loss as a function of time/iteration - also include the (known) minimum possible loss

Final Token Probabilities vs. Optimal Probabilities