Training Loss Progression 2.2 2.0 1.8 Cross-Entropy Loss 1.2 1.0 25 75 100 200 50 125 0 150 175 **Epochs**