Loss of baseline model 2.4 2.2 Cross-Entropy Loss 1.8 1.6 50 100 150 200 250 Ò Epochs