

Training with 40000 samples,  $L_1 = 0.0001$

