

Training with 20000 samples, $L_1 = 0.0001$

