

$|\beta|_2^2$: 100, Trainable: last layer

