

Training loss for different optimizer, learning rate = 0.5

