Different optimizers at 0.01 learning rate

