

Adam Optimizer vs SGD with Learning Rate $\eta = 0.001$

