

Method	η	Other params
Grad OLS	0.15	-
Grad Ridge	0.15	$\lambda=0.0001$
Momentum	0.05	$\beta=0.9$
Stochastic	0.2	Batch size=100
AdaGrad	0.05	$\rho=0.99$
RMSProp	0.001	-
ADAM	0.005	$\beta_1=0.9, \beta_2=0.999$