Different untuned estimators' performance $GD_{\lambda}=0.10, \eta=0.05$ $NAG, \lambda = 0.10, \eta = 0.05, \gamma = 0.90$ Adadelta, $\lambda = 0.10$, $\gamma = 0.90$ RMSprop, $\lambda = 0.10, \eta = 0.05, \gamma = 0.90$ Adam, $\lambda = 0.10, \eta = 0.05, \beta_1 = 0.90, \beta_2 = 0.999$

