

Optimizer (Plain SGD, Momentum, Nesterov Momentum, Adam)
Learning Rate (0.001, 0.01, 0.1, 0.5, 1.0, 2.0, 10.0, 50.0, 100.0)
Momentum (Off, 0.7, 0.8, 0.9)
Learning rate Scheduler (On, Off)
Batch Normalization (On, Off)
Dropout (Off, 0.1, 0.3, 0.5)
 L_1 Penalty (Off, 1e-07, 1e-08, 1e-09)
 L_2 Penalty (Off, 1e-07, 1e-08, 1e-09)