Hyperparameter	Value
Number of input variables	28
Number of hidden layers	8
Number of nodes [per layer]	[78, 66, 57, 48, 36, 24, 12, 6]
Number of Maxout layers [position]	3 [1, 2, 6]
Number of parallel layers per Maxout layer	25
Number of training epochs	240
Learning rate	0.0005
Training minibatch size	500