

Learning rate ( $\alpha$ )	0.02
Iterations	900
Momentum	0.05
L2 weight regularization	0.003
Dropout fraction	0.5