

Character embedding size	$d = 50$
Hidden unit number	$H = 50$
Initial learning rate	$\alpha = 0.2$
Margin loss discount	$\mu = 0.2$
Regularization	$\lambda = 10^{-6}$
Dropout rate on input layer	$p = 0.2$
Maximum word length	$w = 4$