Word embedding size	$m = 25$
Distance embedding size	$m_d = 25$
Initial learning rate	$\rho = 0.1$
Margin loss discount	$\kappa = 2.0$
Regularization	$\lambda = 10^{-4}$
k-best	$k = 64$