

Bowman + CNN: trained with AdaDelta (learning rate = 1.0) and L2 regularization on all the weights except for the bias with $\gamma = 5 \cdot 10^{-6}$

