

Method	Dev	Standard
Two-layer LSTM	62.7	63.1
Region selection	62.4	62.4
DPPNet	62.5	62.7
MCB	65.4	—
Co-Attention	65.9	66.1
MRN	66.3	66.3
MLB	—	68.9
MLP + ResNet	67.4	—
MLP + ResNet +Atten.	68.9	—
MLP + Attri.	68.4	—
MLP + Attri. + Atten.	<b>69.5</b>	<b>69.8</b>
10 ensemble models	<b>70.5</b>	<b>70.5</b>