

Model	Train Accuracy	Test Accuracy	#Parameters
Classifier(hand crafted features)	99.7	78.2	—
GRU encoders	98.8	81.4	15.0M
Tree-based CNN encoders	83.3	82.1	3.5M
SPINN-NP encoders	89.2	83.2	3.7M
LSTM with attention	85.3	83.5	252K
mLSTM	92.0	86.1	1.9M
LSTM Networks	88.5	86.3	3.4M
word-word attention and aggregation	90.5	86.8	582K
NTI with global attention	88.5	87.3	-
Our model with bi-LSTM encoders	89.8	86.4	6M
Our model with btree-LSTM encoders	88.6	87.6	2M