

## A Model Sizes

Here we give more detail on the trained models from Table 2.

<b>model</b>	<b># HL</b>	<b>Dim of HL</b>	<b>Dim of enc</b>
trn-NC-skip	2	256, 64	-
val-NC-skip	2	256, 64	-
trn-FC-skip	2	256, 64	4800
val-FC-skip	2	256, 64	4800
trn-LS-skip	3	2400, 1200, 600	-
val-LS-skip	3	2400, 1200, 600	-
val-LS-GloVe	3	2400, 1200, 600	4800

Table 3: Description of Models trained (HL: Hidden Layer, enc: GRU or LSTM encoder used to encode inputs)