Model	Parameters	Validation	Test
- LSTM (medium)	10M	86.2	82.7
- LSTM (large)	24M	82.2	78.4
- Variational LSTM (medium)	20M	81.9	79.7
- Variational LSTM (large)	66M	77.9	75.2
- Variational LSTM	51M	71.1	68.5
- Variational RHN	24M	68.1	66.0
- Variational RHN	23M	67.9	65.4
- 5-layer RHN	24M	64.8	62.2
- 4-layer skip connection LSTM	24M	60.9	58.3
- AWD-LSTM 3-layer (baseline)	24M	60.0	57.3
Fraternal dropout + AWD-LSTM 3-layer	24M	**58.9**	**56.8**