Parameter	Final value	Experiment range
Word embedding dimensionality	200	100-300
Embedding dimensionality for part-of-speech tags	30	10-50
Embedding dimensionality for relative positions	20	10-50
Number of CNN filters	192	64-384
Sizes of CNN filters	2 to 7	2-4 to 5-9
Norm regularization parameter (λ)	0.01	0.0-1.0
Number of LSTM units (RNN)	600	0-2400
Dropout probability (CNN and RNN)	0.5	0.0-0.7
Initial learning rate	0.01	0.001-0.1
Number of epochs (Subtask 1)	200	20-400
Number of epochs (Subtask 2)	10	5-40
Ensemble size	20	1-30
Training batch size	64	32-192
Upsampling ratio (only Subtask 2)	1.0	0.0-5.0
Max. sentence length (only subtask 2)	19	7-23