

Layer	Hyper-parameter	Value
CNN	window size	3
	number of filters	50
LSTM	number of layers	2
	state size	256
	initial state	0.0
	peepholes	Hadamard
MLP	number of layers	1
	dimension	100
Dropout	embeddings	0.15
	LSTM hidden states	0.25
	LSTM layers	0.33
Learning	optimizer	Adam
	initial learning rate	0.002
	decay rate	0.5
	gradient clipping	5.0