Hyper-parameter	Value
Learning rate	0.1
Size of word vectors	100
Size of the context window	5
Number of epochs	100
Minimum number of word occurences	1
Number of negatives sampled	5
Loss function	Softmax
Minimum length of char n-gram	0
Maximum length of char n-gram	0
Maximum length of word n-gram	1
Sampling threshold	10^{-4}
Rate of updates for the learning rate	100
Use of pre-trained word vectors for supervised learning	Yes