CS224n Assignment3 Solution

shawn

August 2020

1 Machine Learning & Neural networks

(a)

(i) Based on the equation of gradient m for each step, we can see that m is weighted sum of previous accumulated gradient m and current step gradient. So, for the gradients with different directions, the summation operation will reduce the magnitude of component vectors in different directions, then make m not to vary much.
(ii) Since v is the accumulated magnitude of gradients and m is divided bu $\sqrt{v}$, the model parameters with small gradient magnitude would get larger updates. Benefited from this adaptive learning, the parameters with large gradient magnitude would only get small update, the model will not over shoot; the large updates for parameters with small gradient magnitude will also accelerate the convergence.

(b)

(i) $\gamma = \frac{1}{1-p_{drop}}$. According to the properties of expectation, $E(h_{drop}) = E(\gamma d \odot h) = \gamma E(d \odot h) = \gamma(1 - p_{drop})E(h)$, so $\gamma = \frac{1}{1-p_{drop}}$ would not change the expectation.
(ii)We apply dropout as a regularization method during training to promote the generalization performance of the model. As we have kept the expectation of h to be unchanged with the application of dropout during training, then we cannot apply dropout during test, otherwise the expectation would not equal to training, then the model will predict wrong output.

2 Neural transition-based dependency parsing

(a)

All the remaining steps of parsing is listed in the table 1.

(b)

A sentence containing n words will be parsed in $2n$ steps. Because for each word, it needs one step to shift into the stack, then another step to be parsed and removed from the stack.

(e)

The best UAS on dev set is 87.77, the best USA on test set is 87.78.

Stack	Buffer	New dependency	Transition
[ROOT, parsed, this]	[sentence, correctly]		SHIFT
[ROOT, parsed, this, sentence]	[correctly]		SHIFT
[ROOT, parsed, sentence]	[correctly]	sentence → this	LEFT-ARC
[ROOT, parsed]	[correctly]	parsed → sentence	RIGHT-ARC
[ROOT, parsed, correctly]	[]		SHIFT
[ROOT, parsed]	[]	parsed → correctly	RIGHT-ARC
[ROOT]	[]	ROOT → parsed	RIGHT-ARC

Table 1: Remaining steps for parsing the sentence

(f)

(i) Error Type: Verb Phrase Attachment Error
Incorrect dependency: wedding → fearing
correct dependency: heading → fearing

(ii) Error Type: Coordination Attachment Error
Incorrect dependency: makes → rescue
Correct dependency: rush → rescue

(iii) Error Type: Prepositional Phrase Attachment Error
Incorrect dependency: named → Midland
Correct dependency: guy → Midland

(iv) Error Type: Verb Phrase Attachment Error
Incorrect dependency: root → one
Correct dependency: root → been