CS 224n Assignment 3: Dependency Parsing

Chrysa Dikonimaki

August 2020

1 Machine Learning and Neural Networks

1.a

1.a.i

Using m stops the updates from varying as much because the parameters' updates take into account the previous values gradient.
This low variance may be helpful to learning because we will have milder changes.

1.a.ii

The parameters with larger gradients will get smaller updates and vice versa.
This solve the vanishing/exploding gradient problem.

1.b

1.b.i

$h_i = 0 * p_{drop} + (1 - p_{drop}) * x \Leftrightarrow x = \frac{1}{(1-p_{drop})} * h_i$
Consequently , $\gamma = \frac{1}{(1-p_{drop})}$

1.b.ii

Because during training we want to learn the parameters and avoid overfitting, but during evaluation we want to see what our model learned. So, we want to take into account all parameters, we can't arbitrary remove some of them.

2 Neural Transition-Based Dependency Parsing

2.a

Table 1.

Stack	Buffer	New dependency	Transition
ROOT	[I, parsed, this, sentence, correctly]		Initial Configuration
ROOT, I	[parsed, this, sentence, correctly]		SHIFT
ROOT, I, parsed	[this, sentence, correctly]	parsed → I	LEFT-ARC
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: 2.a

2.b

2n = $O(n)$, because the parser parses all the words one time(one SHIFT for each word) and each word is deleted after one ARC, so n ARCs so as to conclude with one element in the stack in the end.

2.e

Less than half an hour.

```
Epoch 10 out of 10
100%|                                                                                | 184
Average Train Loss: 0.06602585252406665
Evaluating on dev set
1445850it [00:00, 31515314.74it/s]
- dev UAS: 88.52

================================================================================
TESTING
================================================================================
Restoring the best model weights found on the dev set
Final evaluation on test set
2919736it [00:00, 45743305.00it/s]
- test UAS: 88.51
Done!
```

2.f

2.f.i

Error type : Verb Phrase Attachment Error
Incorrect dependency : wedding → fearing
Correct dependency : disembarked → fearing

2.f.ii

Error type : Coordination Attachment Error
Incorrect dependency : makes → rescue
Correct dependency : rush → rescue

2.f.iii

Error type : Prepositional Phrase Attachment Error
Incorrect dependency : named → Midland
Correct dependency : guy → Midland

2.f.iv

Error type : Modifier Attachment Error
Incorrect dependency : elements → most
Correct dependency : crucial → most