cs224n-2019 Assignment 3

Falco Winkler

June 2020

1 Neural Networks

(a) (i)

By adding the rolling average we take larger steps initially when the gradient is big and don't fall into tiny local minima. This article explains momentum in-depth: https://distill.pub/2017/momentum/.

(ii)

TODO

(b)

TODO

2 Neural Dependency Parser

(a)

stack	buffer	new dependency	transition
[ROOT]	[I, parsed, this, sentence, correctly]		Initial
[ROOT, I]	[parsed, this, sentence, correctly]		SHIFT
[ROOT, I, parsed]	[this, sentence, correctly]		SHIFT
[ROOT, 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

(b)

We can observe that, in order to process a word it needs to be $SHIFT$'ed onto the stack first. Each LEFT-ARC and RIGHT-ARC command deletes one item from the stack. So, you need $2n$ operations, for every word a shift and an arc.

(f)

(i)

- Error type: Verb Phrase Attachment Error
- Incorrect: wedding → fearing
- Correct: I → fearing

(ii)

- Error type: Coordination attachment error
- Incorrect: makes → rescue
- Correct: rush → rescue

(iii)

- Error type: Prepositional Phrase Attachment Error

- Incorrect: named → Midland
- Correct: guy → Midland

(iv)

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