

---

# Assignment 3 writeup

---

Dima Lipin

January 22, 2019

## 1 THEORY

a) Adam

1.  $m$  helps to gain a *momentum* to update needed dimensions as raw gradient may be too wobbly around noisy ones and have a little impact on others. It also helps to average gradient on noisy directions to zeros, so we don't step in their directions.
2. We may want to update weights that had large gradients with smaller learning rates to prevent overshooting minimums

b) Dropout

1.  $\gamma = \frac{1}{1-p_{drop}}$  as  $\mathbb{E}[\gamma d \circ h]_i = h_i$
2. Because during inference time we would like to take the most certain prediction, without stochasticity that may hurt the final score, no dropout corresponds to that case as we take the mean value of each  $h$ .

c) Dependency parsing

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]		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

For a sequence with  $n$  words we need  $n$  shifts. Then we need to push each word out of stack once, so the total number of operations will be  $2n$ .

Errors in dependency diagrams

1. Verb Phrase Attachment Error, incorrect: wedding  $\rightarrow$  fearing, should be: heading  $\rightarrow$  fearing
2. Coordination Attachment Error, incorrect: makes  $\rightarrow$  rescue, should be: rush  $\rightarrow$  rescue
3. Prepositional Phrase Attachment Error, incorrect: named  $\rightarrow$  Midland, should be: loan  $\rightarrow$  Midland
4. Modifier Attachment Error, incorrect: elements  $\rightarrow$  most, should be: be crucial  $\rightarrow$  most