

CS 224n Problem Set #3 Solutions: Dependency Parsing

LA DUC CHINH

Due Wednesday, Nov 14 at 11:59 pm on Gradescope.

Written: Dependency Parsing

1. Machine Learning & Neural Networks

(a) Adam Optimizer

$$\theta \leftarrow \theta - \alpha \nabla_{\theta} J_{minibatch}(\theta) \quad (1)$$

- i. Momentum stops updates from varying because the momentum is slowly update with β_1
- ii. I don't know :(

(b) Drop out

i.

$$E_{p_{drop}}[\mathbf{h}_{drop}]_i = \gamma h_i (1 \times (1 - p_{drop}) + 0 \times p_{drop}) = h_i \iff \gamma h_i (1 - p_{drop}) = h_i \iff \gamma = \frac{1}{1 - p_{drop}}$$

- ii. Because when we want to evaluate we need our model to work at full capacity.

2. Neural Transition - Based Dependency Parsing

(a) Transition step

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 \rightarrow I	LEFT-ARC
[ROOT, parsed, this]	[sentence, correctly]		SHIFT
[ROOT, parsed, this, sentence]	[correctly]		SHIFT
[ROOT, parsed, sentence]	[correctly]	sentence \rightarrow this	LEFT-ARROW
[ROOT, parsed]	[correctly]	parsed \rightarrow sentence	RIGHT-ARROW
[ROOT, parsed, correctly]	[]		SHIFT
[ROOT, parsed]	[]	parsed \rightarrow correctly	RIGHT-ARROW
[ROOT]	[]	ROOT \rightarrow parsed	RIGHT-ARROW

(b) n words will need n SHIFT steps and n step(s) of X-ARROW so in total we need $2n$ steps