Youssef Kishk 3812

NLP - A3

1) mL & NN

a) Adam optimizer

i) Using momentum m, speed up the optimizer by improving the horizontal step towards the global minimum, and decreasing the vertical step that's unnecessary and prevent us from increasing the learning rate

ii) weights that have large gradients at certain point of the training will start learning with smaller learning rates, in contrast weights that have small gradients would have larger learning rates

b) Dropout

i) P_{drop} is the prob. of dropping this unit, so the constant γ is the value of $1/1-P_{drop}$ as $(1-P_{Prob})$ is the prob. of not dropping this unit, so γ will represt a constant equivilant to $(1-Prob)$

ii) bec, dropout add noise to the data, and we don't want to have that noise during evaluating the predictions so we can end up with true not random predictions

2) Dependency parsing

a)

Stack	Buffer	new dep.	transition
[root]	[I, parsed, this, sentence, cor]		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, pared]	[]	parsed → correctly	right_arc
[root]	[]	root → parsed	right_arc

b) Sentence containing n words need:
- (n) shift operations
- (n) arc operation

So total of 2n steps

c) i) verb phrase error
- wedding → Pearing (incorrect) ✗
- I → Pearing (correct) ✓

ii) Coordination error

iii) Prepositional phrase Error
- named → Midland ✗
- loan → Midland ✓

iv) Modifier error
- elements → most ✗
- crucial → most ✓

e) Best UAS:
- on dev set: 88.59
- on test set: 88.93