

Assignment 3 (Written)

1. Machine Learning & Neural Networks

- a)(i) m acts as a weighted moving average and thus prevents noisy updates by considering previous values of gradient as well. Random spikes in gradient do not affect as much and our learning is smoother.
- (ii) Model parameters which till now have not changed much will get higher updates.

It acts as a normaliser for learning rate for each parameter, ensuring we learn all the parameters of the model fairly equally.

b)(i) $h_{\text{drop}} = \gamma d o h$
 $E_{p_{\text{drop}}} [h_{\text{drop}}]_i = h_i \Rightarrow E_{p_{\text{drop}}} [\gamma d_i h_i] = h_i$

$\Rightarrow \gamma E_{p_{\text{drop}}} [d_i h_i] = h_i \Rightarrow \gamma p_{\text{drop}} \cdot h_i = h_i$

$\Rightarrow \boxed{\gamma = \frac{1}{p_{\text{drop}}}}$

- (ii) By applying dropout, we aim to reduce overfitting which is not a concern during testing.

2. Neural Transition-Based Dependency Parsing

a)

Stack	Buffer	Transition
[Root]	[I, passed, this, sentence, correctly]	Initial
[Root, I]	[passed, this, sentence, correctly]	SHIFT
[Root, I, passed]	[this, sentence, correctly]	SHIFT
[Root, passed]	[this, sentence, correctly]	LEFT-ARC
[Root, passed, this]	[sentence, correctly]	SHIFT
[Root, passed, this, sentence]	[correctly]	SHIFT
[Root, passed, sentence]	[correctly]	LEFT-ARC
[Root, passed]	[correctly]	RIGHT-ARC
[Root, passed, correctly]	-	SHIFT
[Root, passed]	-	RIGHT-ARC
[Root]	-	RIGHT-ARC

b) For a sentence containing n words (+ ROOT), $2n$ steps are required.

- n SHIFT operations
- n RELATION operations

i. Error Type: Verb Phrase Attachment Error
 Incorrect Dependency: wedding \rightarrow fearing
 Correct Dependency: heading \rightarrow fearing

ii. Error Type: Coordination Attachment Error
 Incorrect Dependency: makes \rightarrow rescue
 Correct Dependency: rush \rightarrow rescue

iii) Error Type : Prepositional Phrase Attachment Error
Incorrect Dependency : named → ~~mid~~ Midland
Correct Dependency : guy → Midland

iv) Error Type : Modifier Attachment Error
Incorrect Dependency : elements → most
Correct Dependency : crucial → most