Assignment 3
Zeyad Ezzat
4492

1. Machine Learning and Neural Networks

a. Adam
   i. what equations means that the update step value will have 0.9 from the previous value +0.1 from the gradient , so this means that the update step get updated every step to the same direction but with a less value to avoid overshooting the minimum
   ii. adaptive learning rate
   this means when you have a huge gradient you will get a smaller learning rate and vice versa

b. Dropout
   i. 1/(1-p) because the output of the equation must equal to the original hidden layer
   ii. Using dropout during training means that you drop some of your feature vector , so it becomes more general because it doesn't depend on a certain feature
   but in the test you want to use all of your model power or features to get maximum performance

2. Neural Translation

a. Table

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,	Correctly		Shift

this, sentence			
Root, parsed, sentence	Correctly	Sentence->this	Left-arc
Root, parsed	Correctly	Parsed->sentence	Right-arc
Root, parsed, correctly	Correctly		Shift
Root, parsed	[]	Parsed->correctly	Right-arc
Root	[]	Root->parsed	Right-arc

b. 2N
   in the case if all the words depend on the last word of the sentence
c. Code
d. Code
e. Code
f. Parsing errors
   i. Propositional phrase error
      incorrect :weeding->fearing
      correct: disembarked->fearing
   ii. Coordinate attachment error
      incorrect: makes->rescue
      correct: rush->rescue
   iii. Verb phrase attachment error
      incorrect: named->midland
      correct: loan->midland
   iv. Modifier attachment error
      incorrect: elements->most
      correct: crucial ->most