



**Iran University of
Science & Technology**

Natural Language Processing (CS22N)

Student Name:

Babak Behkamkia

Instructor Name:

Dr. Sauleh Eetemadi

Spring 2023

1. Machine Learning & Neural Networks

a. Adam Optimizer

- i. Imagine a ball falling from a hill. Despite the agronomy of the way, it goes in line with its momentum, and if there are small holes, it will pass them easily. The momentum algorithm works like this example. It forces the theta to change in the way of previously seen data rather than a new batch because it is more reliable and prevents dramatic changes in direction. It helps us to deal with local minimums.
- ii. The term help algorithm to normalize the gradients. For instance, the impact of a parameter with large gradients will decrease and vice-versa.

b. Dropout

- i. We know that d shows which values will be dropped and which aren't, and we want to determine γ somehow to force h_{drop} to be equal to h . Thus, we need to make $(1 - p_{drop})$ equal to one because this term determines the probability of 1s in d . So, the value of gamma should be $1/(1 - p_{drop})$.
- ii. Dropout helps the model prevent overfitting by eliminating unusual dependencies on some weights. In other words, the model learns not to compute its output according to some specific weights. In the testing phase, we need all of the weights, and it is not logical to use some proportion of them.

2. Neural Transition-Based Dependency Parsing

a.

Stack	Buffer	New dependency	Transition
[ROOT]	[I, attended, lectures, in, the, NLP, class]		Initial Configuration
[ROOT, I]	[attended, lectures, in, the, NLP, class]		shift
[ROOT, I, attended]	[lectures, in, the, NLP, class]		shift
[ROOT, attended]	[lectures, in, the, NLP, class]	attended → I	Left-arc
[ROOT, attended, lectures]	[in, the, NLP, class]		shift
[ROOT, attended]	[in, the, NLP, class]	attended → lectures	Right-arc
[ROOT, attended, in]	[the, NLP, class]		shift
[ROOT, attended, in, the]	[NLP, class]		shift
[ROOT, attended, in, the, NLP]	[class]		shift
[ROOT, attended, in, the, NLP, class]	[]		shift
[ROOT, attended, in, the, class]	[]	class → NLP	Left-arc
[ROOT, attended, in, class]	[]	class → the	Left-arc
[ROOT, attended, class]	[]	class → in	Left-arc
[ROOT, attended]	[]	attended → class	right-arc
[ROOT]	[]	ROOT → attended	right-arc

- b. $2n$ times. Because we should insert all words from the buffer to stack and also remove all of them from the stack, and these two stages take n steps each.
- e. Dev UAS score: 87.81 and Test UAS score: 88.46
- f. dependency parses
 - i.
 - **Error type:** Verb Phrase Attachment Error
 - **Incorrect dependency:** acquisition \rightarrow citing
 - **Correct dependency:** blocked \rightarrow citing
 - ii.
 - **Error type:** Modifier Attachment Error
 - **Incorrect dependency:** had \rightarrow already
 - **Correct dependency:** left \rightarrow already
 - iii.
 - **Error type:** Prepositional Phrase Attachment Error
 - **Incorrect dependency:** declined \rightarrow decision
 - **Correct dependency:** reasons \rightarrow decision
 - iv.
 - **Error type:** Coordination Attachment Error
 - **Incorrect dependency:** affects \rightarrow one
 - **Correct dependency:** plants \rightarrow one