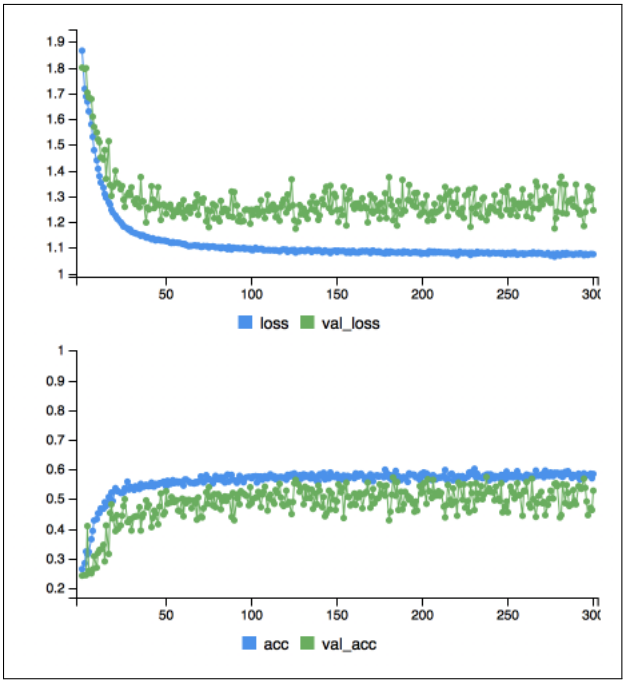
Homework 2 ECS 171

Atharva Chalke

July 31, 2018

Problem 1-4.



Learning rate: 0.1
Hidden layer activation: RELU

layers	Hidden units	function	generalization error	epoch	Cost Function	Optimizer
3	3	softmax	42.7%	300	Cross entropy	SGD
3	3 FULL DATA SET	softmax	45.68%	300	Cross entropy	SGD
4	3	softmax	63.22%	300	Cross entropy	SGD
5	3	softmax	59.23%	300	Cross entropy	SGD
5	6	softmax	39.01%	300	Cross entropy	SGD
5	9	softmax	38.08%	300	Cross entropy	SGD
5	12	softmax	42.22%	300	Cross entropy	SGD

Proof. .
Question: What is the optimal configuration ? What do you find the relationship between number of layers and number of nodes to generalization error ?

Answer:

Optimal configuration:
5 layer 9 units.

As the number of layers increased with the number of units to be constant(3), the generalization error went down. I strongly suspect this behavior of the neural network, however, since the learning rate was held constant, and so were the number of iterations, maybe the reason for this was vanishing gradients due to deeper network. When the number of iterations were increased on a deeper network, the ANN was able to generalize better, as seen by running it for 500 iterations with 0.5.

Conclusion:
1. Increasing the number of layers with 3 units increased the generalization error.
2. Adding units increased error for 6 units, and then decreased error substantially.

□

Problem 5. Which class does the following sample belong to ?

NUC

Problem 6.