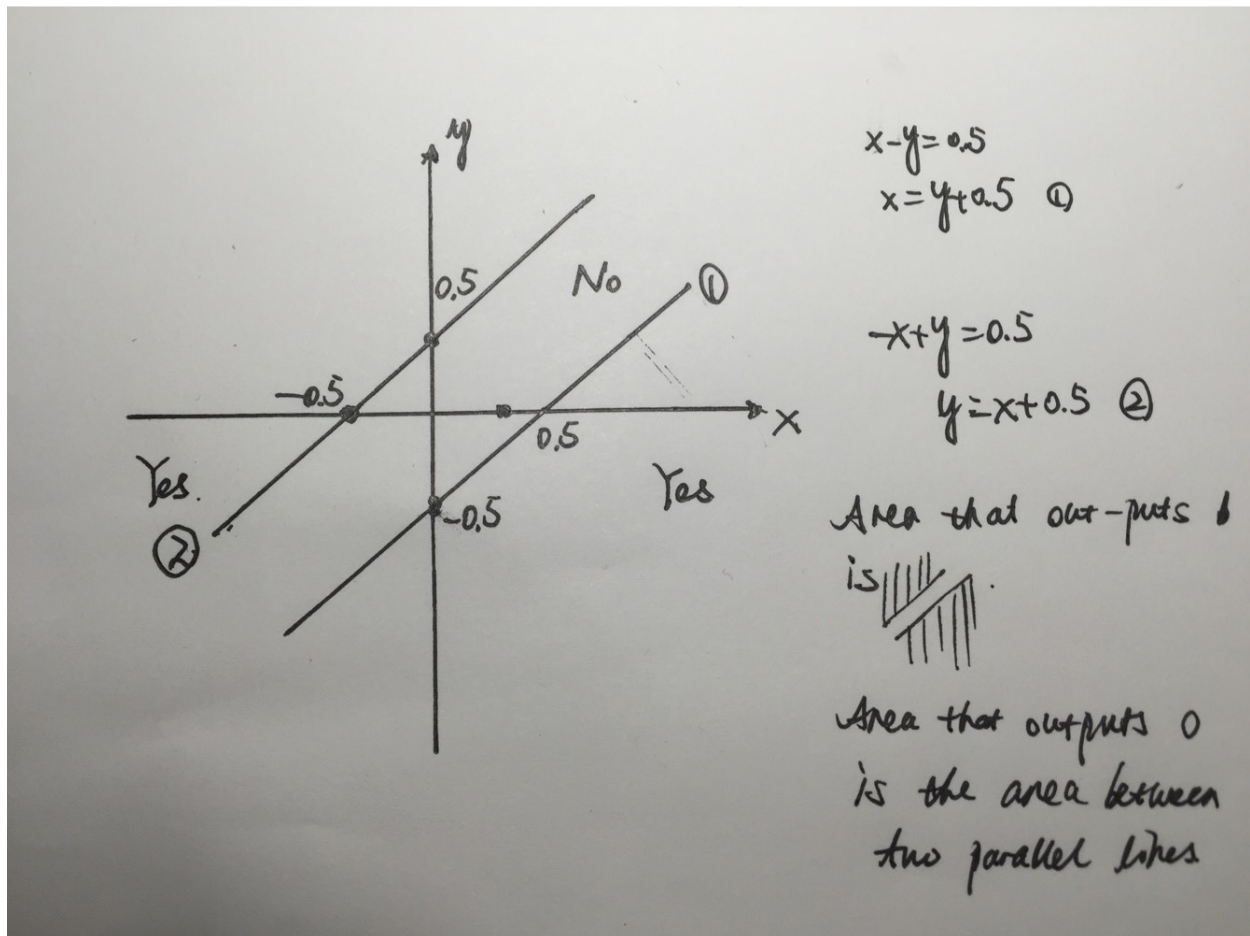


CSCI 360  
Mingzhe Fang  
5301084138  
Oct.14th

## Project 2

### Part 2



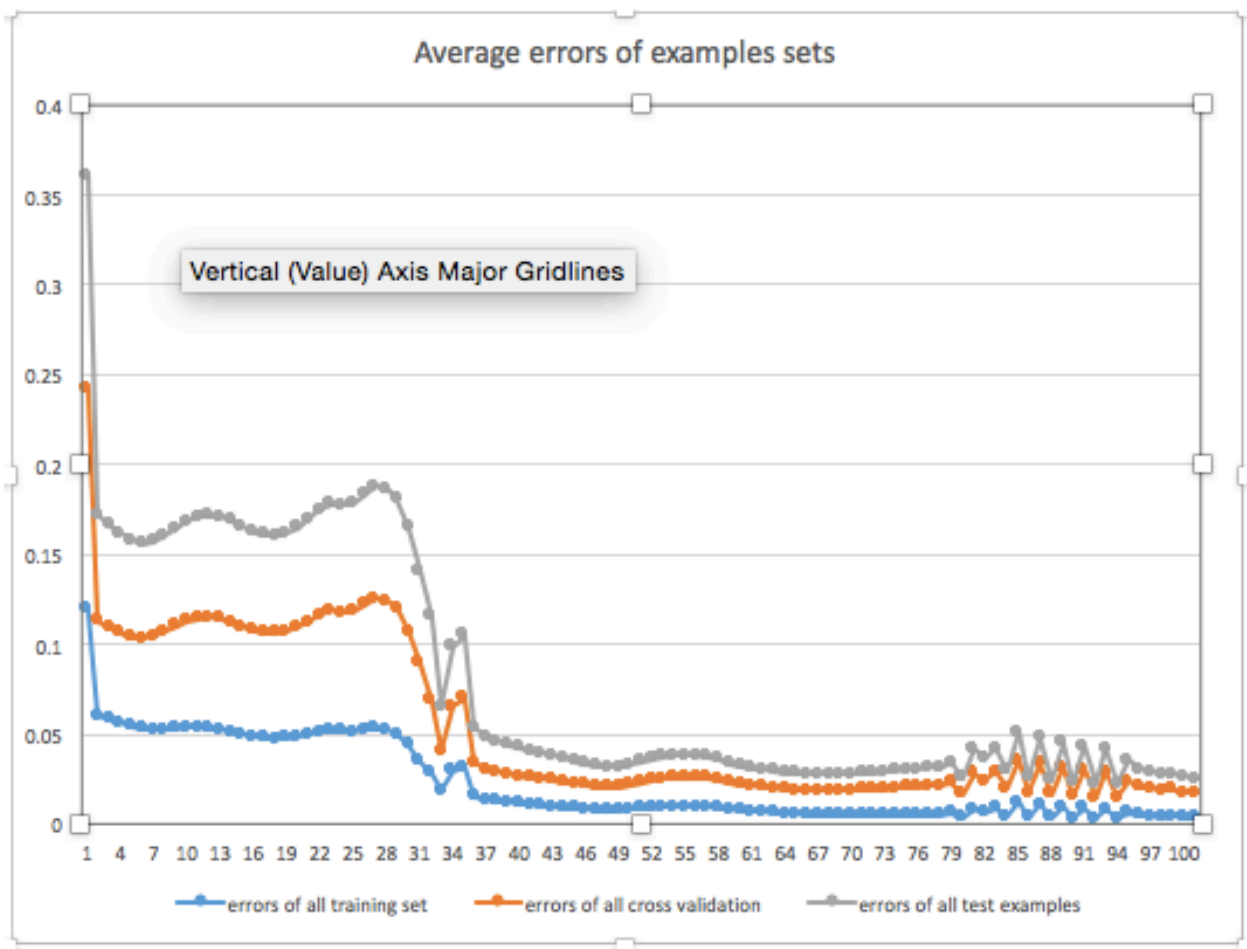
As shown in the above drawing, the area between two parallel lines output 0 while the area on the both sides of the parallel lines output one. It is deducted from the existing formula.

### Part3

#### Question1:

*Explain why the backpropagation algorithm uses a training and a cross validation set and terminates once the error over all cross validation examples (instead of, say, all training examples) increases. Read the section "Output of gesturetrain" carefully. Then, graph the error on the training set, cross validation set and test set as a function of the number of epochs.*

*Explain your experimental results (including your observations, detailed explanations of the observations and your overall conclusions).*



#### Answer1:

It uses a cross validation set to make sure the model isn't overtrained. Overtraining will exist because criterion used for training the model isn't necessarily the same the criterion used to judge the efficacy of a model. Cross validation is a technique generally being used to avoid overtraining.

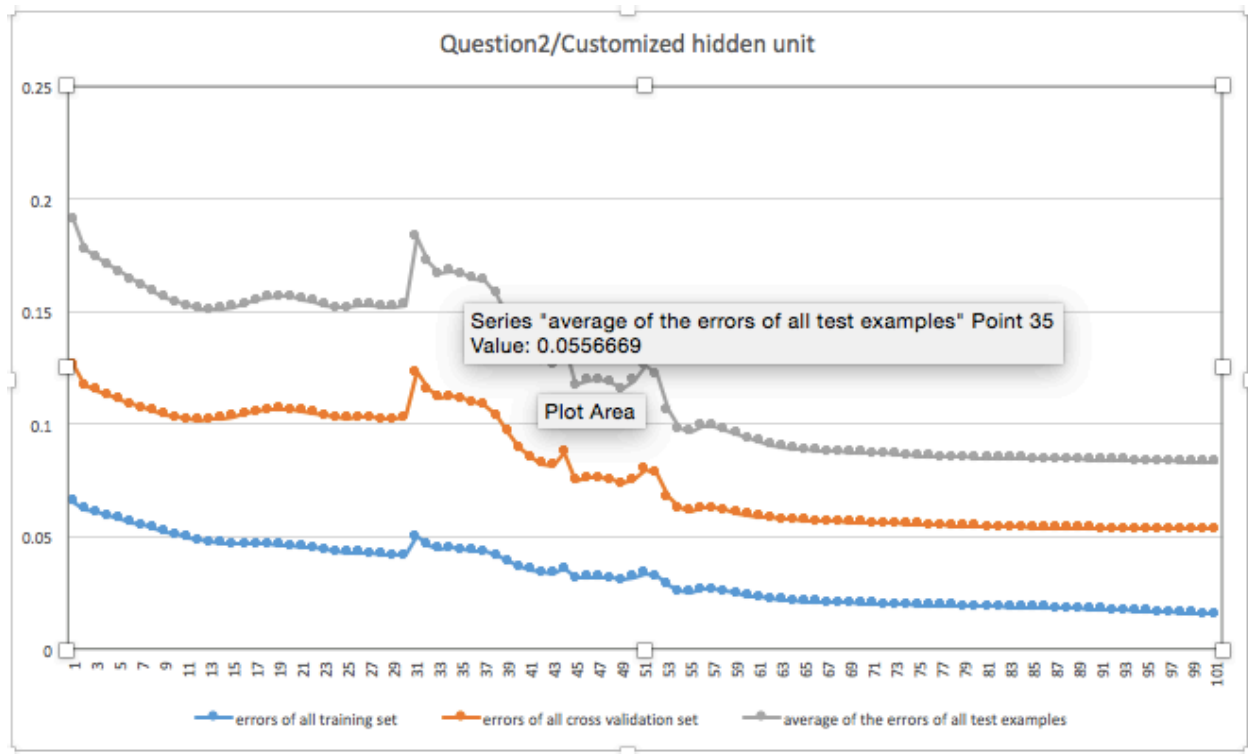
As shown in the graph above, errors over all training set, all cross validation and all test examples first decreases significant and then finally fluctuate at certain level as the number of epoch increases. In general, at same epoch, errors of all test examples are usually higher than errors over all cross validation and errors over all training set.

Question 2:

*Change the number of hidden units in the neural network to 2 and train it (make testgesturetrain). Compare the training time and the accuracy and error of the resulting network (100<sup>th</sup> epoch) to the network with the default values for the parameters.*

*Explain your experimental results (including your observations, detailed explanations of the observations and your overall conclusions).*

Answer2:



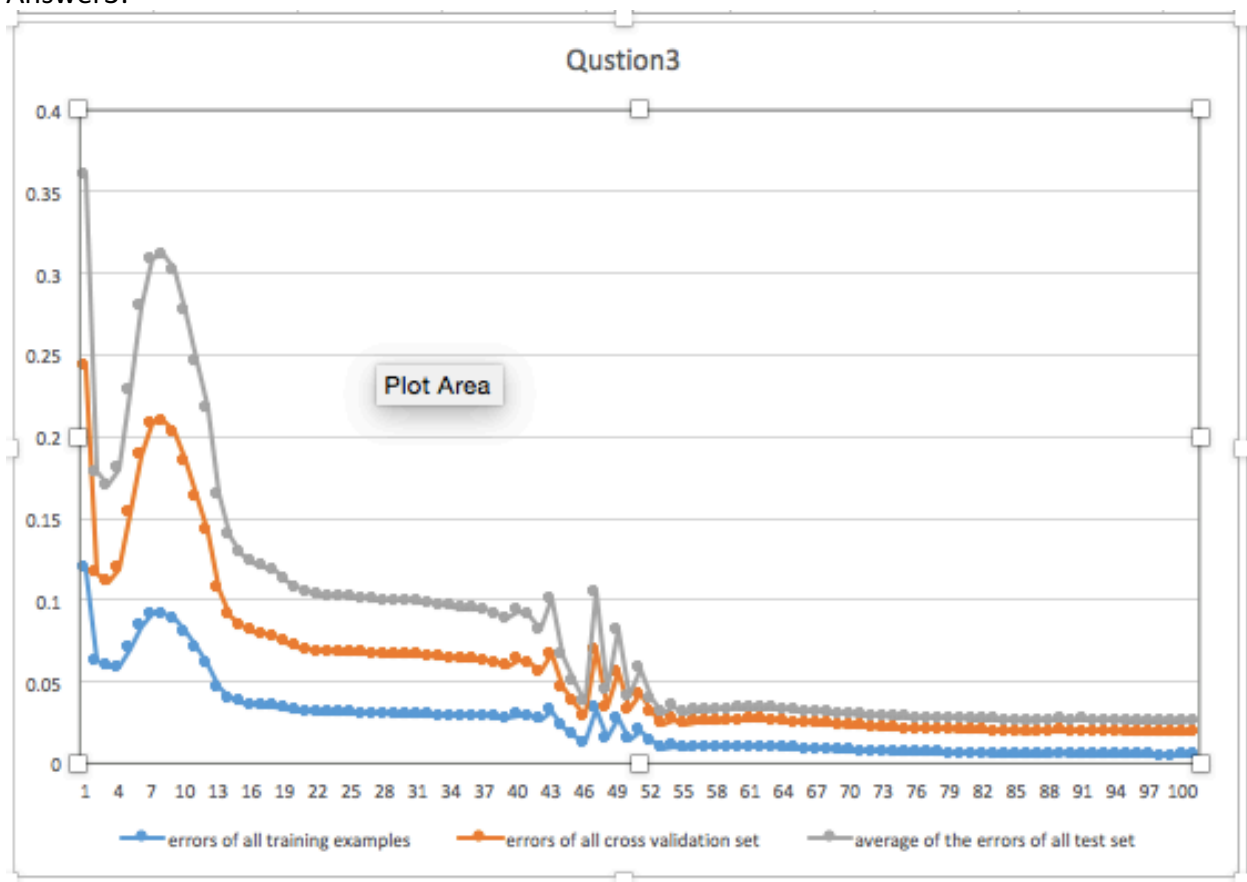
By comparing the 100<sup>th</sup> epoch value in the condition that hidden units is equal to 2 to the first graph, we get that the 100<sup>th</sup> epoch's error in graph 2 is bigger than that in graph 1, which indicates that the more hidden units being used, the more accurate it will be after certain training sets.

Overall, both three errors are decreasing with the increment of number of epoch. It fluctuates a little bit in the beginning, but finally become stable in the end.

Question 3

Change the learning rate,  $\eta$ , of the neural network to 0.8 and train it (make testgesturetrain). Graph the error on the training set, cross validation set and test set as a function of the number of epochs and compare it with the graph from Question 1. What can you say about the influence of  $\eta$  on the training process? Explain your experimental results (including your observations, detailed explanations of the observations and your overall conclusions).

Answer3:



After we change the learning rate from 0.3 to 0.8, graph 3 shifted a little bit to the right in compare with graph 1, which indicates that to reach certain accuracy level, it takes more epoch in graph 3 condition than that in graph 1. Also, the fluctuation level in the beginning stage in graph 3 is also bigger than that in graph 1, as condition 3 has a bigger learning rate. Even though, since they are using the same training set, finally in the stable condition, the error created by condition 3 and condition doesn't differ significantly.

In general, both three curves still decrease with the increment of number of epoch. They first fluctuate a little bit in the beginning of training, while become stable in the late stage.