Ray Li
CSC 578
October 9th, 2016
Project One

Checklist

[X] Early stopping
[X] Cost functions
  [X] Quadratic
  [X] Cross-entropy
  [X] log-likelihood
  [X] allow choice of cost function with a parameter
[X] Momentum
[X] L2 Regularization
[X] Better initial weights
[X] Transfer functions
  [X] tanh
  [X] softmax
  [X] ReLU
[X] Minibatch shuffling
[X] Learning rate schedule description
[X] Returning learned network
[X] Returning accuracy and costs for plotting
[X] Did NOT include the MNIST data with my submission

Incomplete, Details _____
Not sure, Details _____

Pseudocode

Function MyNetwork

Training, Testing & Validation Split

Check if using previously trained network, if no
    Weight and bias initialization
Else
    use previous network settings and weights and biases

Batching

Ask user to input a transfer function for output layer if the parameter specifies ReLu
Ask user to input a transfer function for hidden layers if the parameter specifies Softmax

Start epoch
    For each batch
        Batch shuffling by shuffling the batch indexes
        Calculate activations for each layer according to user specified transfer function
        [Normalize tanh() output if transfer function is tanh()]
        Calculate error according to user specified cost function with L2
        Calculate delta according to user specified transfer function
        Backpropagation
        Gradient descent with momentum
    End batch iteration

    Use updated weight to predict on training, testing and validation set
    Calculate cost and accuracy
    Print out results for each set for this epoch
    Save cost and accuracy for each set
    Early stopping criteria (after 75% of epochs, if validation cost increases)
        Graph accuracy and cost
    Graph accuracy and cost
End epoch

Function act

Check if input matches a certain type of activation function
    Execute that activation function

Function dact

Check if input matches a certain type of the derivative activation function
    Execute that derivative of activation function

Instruction

```
function [ weights, bias, accuracy, cost ] =
        MyNetwork(inputs, targets, nodeLayers, numEpochs, batchSize, eta, split,    momentum, ActFunc,
        CostFunc, L2lambda, previousNetwork)
```

Input variable definitions [Data type specified in parentheses]

inputs: (matrix) input values; rows represent features and columns represent sample
targets: (matrix) target values
nodeLayers: (array) number of neurons on each layer
numEpochs: (double) number of epochs
batchSize: (double) batch size for each mini batch
eta: (double) learning rate for gradient descent
split: (array of three doubles) each number represent the percentage of training, testing and validation dataset (e.x. [80 10 10])
momentum: (double) momentum coeffcient
ActFunc: (string) activation or transfer function
CostFunc: (string) cost or error function
L2lambda: (double) lambda, the coefficient for L2 regularization
previousNetwork: (cell or 'None') if with to train a previously trained network, save the 1) nodeLayers in the first cell 2) weights in the second cell 3) bias in the third cell

Output variable definitions [Data type specified in parentheses]

weights: (cell) weights for each layer
bias: (cell) bias for each layer
accuracy: (cell) first cell represent the training accuracy, and the second and the third represent testing and validation accuracies respectively
cost: (cell) first cell represent the training cost, and the second and the third represent testing and validation cost respectively

Note

To save all the output, you should have an array of four elements when calling the function. For example,

```
>> [a b c d] = MyNetwork(trn, trnAns, [784,30,10], 30, 10, 3, [80, 10, 10], 0.3, 'sigmoid', 'quad', 5, 'None');
```

Description

This neural network is implemented in Matlab. This system allows user to try different network settings including number of layers, number of neurons in each hidden layer, cost function, activation function, etc. Here is a high level overview of how the program works.

Step One: Weight and Bias Initialization

If the network is used to train a previously trained network, existing weights and bias will be used to start. Otherwise, the weights and bias will be initialized with a mean of zero and a standard deviation of (1/total number of training samples).

Step Two: Batching

Step Three: Start learning batch by batch and epoch by epoch

In this part, user can test out different cost and activation function combinations. For gradient descent, user can test different learning rate as well. Since momentum is added to the gradient descent, user can try different momentum coefficients as well.

Step Four: Calculate the final prediction and evaluation

Analysis

My program works for all datasets with multiple configurations, including changes in epochs, hidden layer neurons, different batch size, eta, different activation and cost functions, momentum coefficient and L2 regularization coefficient. One thing that did not work quite well was the swapping of cost function within each batch with L2 regularization. However, it has been resolved by calculating the difference between activation and target values.

Ideas for enhancement

1. Learning rate schedule description. This will be accomplished like momentum. By checking the accuracies and costs of training, testing and validation, if any of them starts to drift away from the others, the learning rate will adjust, by increasing or decreasing a certain value, to try to optimize the learning rate.
2. Drop out methods for neurons. This can avoid overfitting network or saturated neurons.
3. Implement a grid search algorithm to find the best optimized combination of parameters.

Outputs

Configurations:

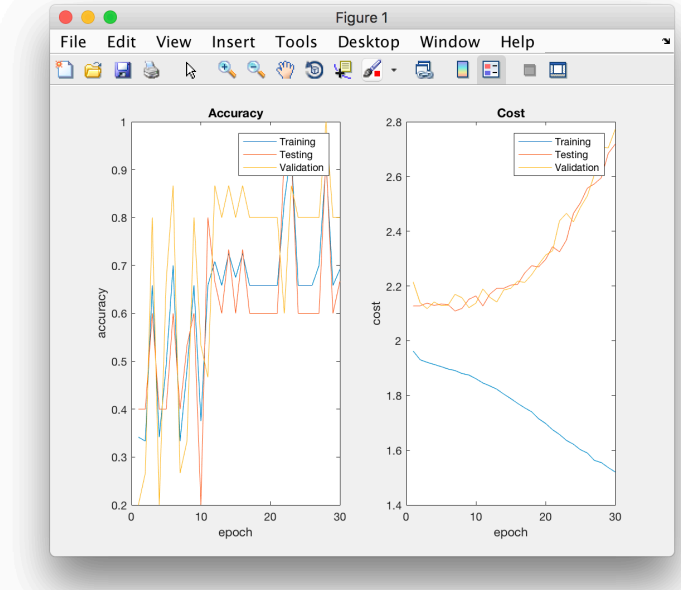
data set	epochs	hids	batch	eta	trans.	cost	mom.	reg.
iris.csv	40	20	10	0.1	sigmoid	cross	.3	5

Matlab Code:

>> MyNetwork(iris_inputs, iris_targets, [4,20,3], 40, 10, 0.1, [80, 10, 10], 0.3, 'sigmoid', 'cross', 5, 'None');

Matlab Output:

```
Trial>> MyNetwork(iris_inputs, iris_targets, [4,20,3], 40, 10, 0.1, [80, 10, 10], 0.3, 'sigmoid', 'cross', 5, 'None');
        |        TRAIN          ||             TEST     ||     VALIDATION
-------------------------------------------------------------------------------------
Ep      | Cost |   Corr  |  Acc  ||  Cost |  Corr |  Acc ||  Cost | Corr |  Acc
-------------------------------------------------------------------------------------
1       | 1.962 | 41/120 | 0.342 || 2.127 | 6/15 | 0.400 || 2.215 | 3/15 | 0.200
2       | 1.930 | 40/120 | 0.333 || 2.127 | 6/15 | 0.400 || 2.139 | 4/15 | 0.267
3       | 1.921 | 79/120 | 0.658 || 2.138 | 9/15 | 0.600 || 2.117 | 12/15 | 0.800
4       | 1.913 | 41/120 | 0.342 || 2.129 | 6/15 | 0.400 || 2.142 | 3/15 | 0.200
5       | 1.905 | 59/120 | 0.492 || 2.134 | 6/15 | 0.400 || 2.128 | 10/15 | 0.667
6       | 1.897 | 84/120 | 0.700 || 2.132 | 9/15 | 0.600 || 2.129 | 13/15 | 0.867
7       | 1.891 | 40/120 | 0.333 || 2.108 | 6/15 | 0.400 || 2.170 | 4/15 | 0.267
8       | 1.880 | 58/120 | 0.483 || 2.118 | 8/15 | 0.533 || 2.157 | 5/15 | 0.333
9       | 1.875 | 79/120 | 0.658 || 2.152 | 9/15 | 0.600 || 2.121 | 12/15 | 0.800
10      | 1.861 | 45/120 | 0.375 || 2.164 | 3/15 | 0.200 || 2.137 | 8/15 | 0.533
11      | 1.846 | 79/120 | 0.658 || 2.127 | 12/15 | 0.800 || 2.189 | 7/15 | 0.467
12      | 1.835 | 85/120 | 0.708 || 2.170 | 10/15 | 0.667 || 2.159 | 13/15 | 0.867
13      | 1.823 | 79/120 | 0.658 || 2.192 | 9/15 | 0.600 || 2.141 | 12/15 | 0.800
14      | 1.805 | 87/120 | 0.725 || 2.192 | 11/15 | 0.733 || 2.186 | 13/15 | 0.867
15      | 1.789 | 81/120 | 0.675 || 2.203 | 9/15 | 0.600 || 2.191 | 12/15 | 0.800
16      | 1.771 | 87/120 | 0.725 || 2.206 | 11/15 | 0.733 || 2.218 | 13/15 | 0.867
17      | 1.755 | 79/120 | 0.658 || 2.247 | 9/15 | 0.600 || 2.213 | 12/15 | 0.800
18      | 1.740 | 79/120 | 0.658 || 2.274 | 9/15 | 0.600 || 2.241 | 12/15 | 0.800
19      | 1.715 | 79/120 | 0.658 || 2.270 | 9/15 | 0.600 || 2.278 | 12/15 | 0.800
20      | 1.698 | 79/120 | 0.658 || 2.297 | 9/15 | 0.600 || 2.313 | 12/15 | 0.800
21      | 1.675 | 79/120 | 0.658 || 2.344 | 9/15 | 0.600 || 2.328 | 12/15 | 0.800
22      | 1.658 | 100/120 | 0.833 || 2.325 | 14/15 | 0.933 || 2.438 | 9/15 | 0.600
23      | 1.636 | 113/120 | 0.942 || 2.368 | 14/15 | 0.933 || 2.465 | 13/15 | 0.867
24      | 1.622 | 79/120 | 0.658 || 2.465 | 9/15 | 0.600 || 2.434 | 12/15 | 0.800
25      | 1.602 | 79/120 | 0.658 || 2.503 | 9/15 | 0.600 || 2.487 | 12/15 | 0.800
26      | 1.590 | 79/120 | 0.658 || 2.557 | 9/15 | 0.600 || 2.529 | 12/15 | 0.800
27      | 1.563 | 84/120 | 0.700 || 2.573 | 9/15 | 0.600 || 2.606 | 12/15 | 0.800
28      | 1.555 | 110/120 | 0.917 || 2.595 | 14/15 | 0.933 || 2.706 | 15/15 | 1.000
29      | 1.537 | 79/120 | 0.658 || 2.684 | 9/15 | 0.600 || 2.704 | 12/15 | 0.800
30      | 1.520 | 83/120 | 0.692 || 2.718 | 10/15 | 0.667 || 2.772 | 12/15 | 0.800
Early Stopping: Validation error increased after 30 epochs.
```



Configurations:

data set	epochs	hids	batch	eta	trans.	cost	mom.	reg.
iris.csv	40	20	10	0.1	relu	cross	.3	5

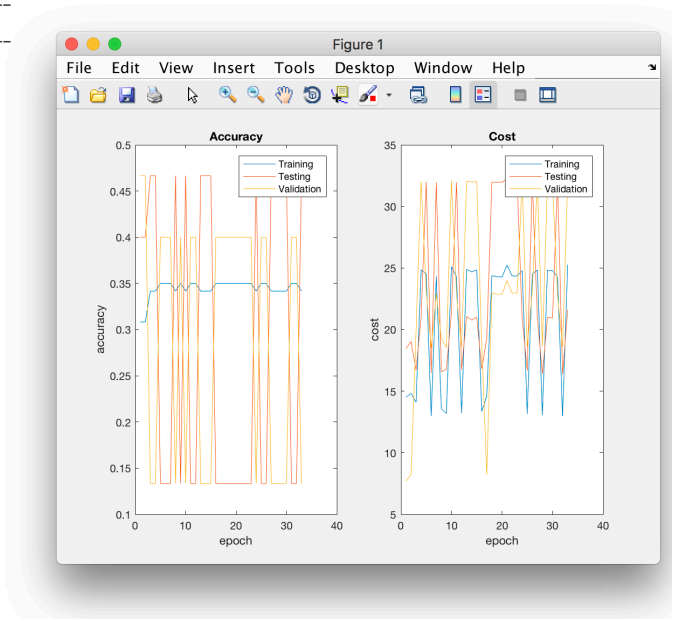
Matlab Code:

>> MyNetwork(iris_inputs, iris_targets, [4,20,3], 40, 10, 0.1, [80, 10, 10], 0.3, 'relu', 'cross', 5, 'None');

note since the input transfer function is ‘relu’, the application will prompt the user to input a transfer function for the output layer (as shown below)

Matlab Output:

```
Trial>> MyNetwork(iris_inputs, iris_targets, [4,20,3], 40, 10, 0.1, [80, 10, 10], 0.3, 'relu', 'cross', 5, 'None');
ReLu should only be used for the hidden layers. Which activation function do you want to use for the output layer?
'softmax'
        |        TRAIN          ||              TEST     ||     VALIDATION
-------------------------------------------------------------------------------------------
Ep      | Cost |   Corr  |  Acc  ||  Cost |  Corr |  Acc ||  Cost | Corr |  Acc
-------------------------------------------------------------------------------------------
1       | 14.509 | 37/120 | 0.308 || 18.445 | 6/15 | 0.400 || 7.695 | 7/15 | 0.467
2       | 14.823 | 37/120 | 0.308 || 19.038 | 6/15 | 0.400 || 8.210 | 7/15 | 0.467
3       | 14.099 | 41/120 | 0.342 || 16.688 | 7/15 | 0.467 || 19.932 | 2/15 | 0.133
4       | 24.847 | 41/120 | 0.342 || 21.011 | 7/15 | 0.467 || 32.022 | 2/15 | 0.133
5       | 24.533 | 42/120 | 0.350 || 32.011 | 2/15 | 0.133 || 23.167 | 6/15 | 0.400
6       | 12.980 | 42/120 | 0.350 || 16.468 | 2/15 | 0.133 || 18.521 | 6/15 | 0.400
7       | 24.312 | 42/120 | 0.350 || 31.967 | 2/15 | 0.133 || 22.906 | 6/15 | 0.400
8       | 13.588 | 41/120 | 0.342 || 16.568 | 7/15 | 0.467 || 19.287 | 2/15 | 0.133
9       | 13.176 | 42/120 | 0.350 || 16.798 | 2/15 | 0.133 || 18.546 | 6/15 | 0.400
10      | 25.105 | 41/120 | 0.342 || 21.373 | 7/15 | 0.467 || 32.106 | 2/15 | 0.133
11      | 24.289 | 42/120 | 0.350 || 31.966 | 2/15 | 0.133 || 22.877 | 6/15 | 0.400
12      | 13.225 | 42/120 | 0.350 || 16.750 | 2/15 | 0.133 || 18.640 | 6/15 | 0.400
13      | 24.895 | 41/120 | 0.342 || 21.079 | 7/15 | 0.467 || 32.036 | 2/15 | 0.133
14      | 24.681 | 41/120 | 0.342 || 20.770 | 7/15 | 0.467 || 31.980 | 2/15 | 0.133
15      | 24.849 | 41/120 | 0.342 || 21.013 | 7/15 | 0.467 || 32.022 | 2/15 | 0.133
16      | 13.354 | 42/120 | 0.350 || 16.817 | 2/15 | 0.133 || 18.764 | 6/15 | 0.400
17      | 14.624 | 42/120 | 0.350 || 19.307 | 2/15 | 0.133 || 8.212 | 6/15 | 0.400
18      | 24.377 | 42/120 | 0.350 || 31.976 | 2/15 | 0.133 || 22.983 | 6/15 | 0.400
19      | 24.286 | 42/120 | 0.350 || 31.965 | 2/15 | 0.133 || 22.874 | 6/15 | 0.400
20      | 24.275 | 42/120 | 0.350 || 31.965 | 2/15 | 0.133 || 22.860 | 6/15 | 0.400
21      | 25.235 | 42/120 | 0.350 || 32.249 | 2/15 | 0.133 || 23.977 | 6/15 | 0.400
22      | 24.360 | 42/120 | 0.350 || 31.973 | 2/15 | 0.133 || 22.963 | 6/15 | 0.400
23      | 24.360 | 42/120 | 0.350 || 31.973 | 2/15 | 0.133 || 22.963 | 6/15 | 0.400
24      | 24.787 | 41/120 | 0.342 || 20.925 | 7/15 | 0.467 || 32.005 | 2/15 | 0.133
25      | 13.159 | 42/120 | 0.350 || 16.661 | 2/15 | 0.133 || 18.616 | 6/15 | 0.400
26      | 24.507 | 42/120 | 0.350 || 32.004 | 2/15 | 0.133 || 23.137 | 6/15 | 0.400
27      | 24.846 | 41/120 | 0.342 || 21.009 | 7/15 | 0.467 || 32.022 | 2/15 | 0.133
28      | 13.048 | 41/120 | 0.342 || 16.410 | 7/15 | 0.467 || 18.653 | 2/15 | 0.133
29      | 24.827 | 41/120 | 0.342 || 20.981 | 7/15 | 0.467 || 32.016 | 2/15 | 0.133
30      | 24.798 | 41/120 | 0.342 || 20.940 | 7/15 | 0.467 || 32.008 | 2/15 | 0.133
31      | 24.229 | 42/120 | 0.350 || 31.967 | 2/15 | 0.133 || 22.804 | 6/15 | 0.400
32      | 12.967 | 42/120 | 0.350 || 16.378 | 2/15 | 0.133 || 18.570 | 6/15 | 0.400
33      | 25.271 | 41/120 | 0.342 || 21.603 | 7/15 | 0.467 || 32.166 | 2/15 | 0.133
Early Stopping: Validation error increased after 30 epochs.
```



Configurations:

data set	epochs	hids	batch	eta	trans.	cost	mom.	reg.
iris.csv	40	20	10	0.1	relu	cross	0	5

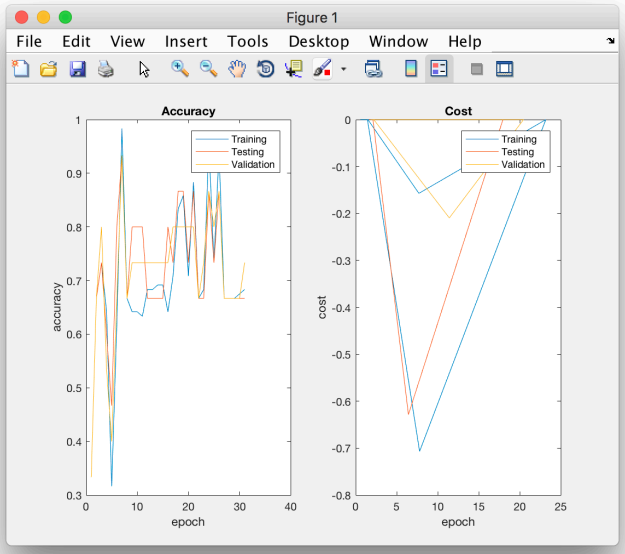
Matlab Code:

>> MyNetwork(iris_inputs, iris_targets, [4,20,3], 40, 10, 0.1, [80, 10, 10], 0, 'relu', 'cross', 5, 'None');

note since the input transfer function is ‘relu’, the application will prompt the user to input a transfer function for the output layer (as shown below)

Matlab Output:

```
Trial>> MyNetwork(iris_inputs, iris_targets, [4,20,3], 40, 10, 0.1, [80, 10, 10], 0, 'relu', 'cross', 5, 'None');
ReLu should only be used for the hidden layers. Which activation function do you want to use for the output layer?
'softmax'
        |        TRAIN          ||              TEST     ||     VALIDATION
-------------------------------------------------------------------------------------------
Ep      | Cost |   Corr  |  Acc  ||  Cost |  Corr |  Acc ||  Cost | Corr |  Acc
-------------------------------------------------------------------------------------------
1       | 12.493 | 40/120 | 0.333 || 18.258 | 5/15 | 0.333 || 16.063 | 5/15 | 0.333
2       | 12.716 | 80/120 | 0.667 || 13.058 | 10/15 | 0.667 || 13.033 | 10/15 | 0.667
3       | 1.490 | 88/120 | 0.733 || 1.623 | 11/15 | 0.733 || 1.801 | 12/15 | 0.800
4       | 7.705 | 77/120 | 0.642 || 9.933 | 9/15 | 0.600 || 11.753 | 8/15 | 0.533
5       | 23.179 | 38/120 | 0.317 || 17.954 | 7/15 | 0.467 || 20.460 | 6/15 | 0.400
6       | 7.792 | 74/120 | 0.617 || 6.430 | 12/15 | 0.800 || 11.416 | 10/15 | 0.667
7       | 1.431 | 118/120 | 0.983 || 2.108 | 14/15 | 0.933 || 2.206 | 14/15 | 0.933
8       | 12.131 | 80/120 | 0.667 || 13.076 | 10/15 | 0.667 || 13.203 | 10/15 | 0.667
9       | 1.410 | 77/120 | 0.642 || 2.283 | 12/15 | 0.800 || 2.495 | 11/15 | 0.733
10      | 1.066 | 77/120 | 0.642 || 1.958 | 12/15 | 0.800 || 2.200 | 11/15 | 0.733
11      | 0.884 | 76/120 | 0.633 || 1.904 | 12/15 | 0.800 || 2.106 | 11/15 | 0.733
12      | 0.838 | 82/120 | 0.683 || 1.924 | 10/15 | 0.667 || 2.098 | 11/15 | 0.733
13      | 0.809 | 82/120 | 0.683 || 1.931 | 10/15 | 0.667 || 2.098 | 11/15 | 0.733
14      | 0.805 | 83/120 | 0.692 || 1.929 | 10/15 | 0.667 || 2.089 | 11/15 | 0.733
15      | 0.791 | 83/120 | 0.692 || 1.937 | 10/15 | 0.667 || 2.095 | 11/15 | 0.733
16      | 0.777 | 77/120 | 0.642 || 1.906 | 12/15 | 0.800 || 2.085 | 11/15 | 0.733
17      | 0.787 | 85/120 | 0.708 || 1.954 | 11/15 | 0.733 || 2.121 | 12/15 | 0.800
18      | 0.744 | 100/120 | 0.833 || 1.914 | 13/15 | 0.867 || 2.124 | 12/15 | 0.800
19      | 0.731 | 103/120 | 0.858 || 1.917 | 13/15 | 0.867 || 2.130 | 12/15 | 0.800
20      | 0.998 | 85/120 | 0.708 || 2.226 | 11/15 | 0.733 || 2.320 | 12/15 | 0.800
21      | 0.712 | 106/120 | 0.883 || 1.943 | 13/15 | 0.867 || 2.179 | 12/15 | 0.800
22      | 1.470 | 80/120 | 0.667 || 2.918 | 10/15 | 0.667 || 3.019 | 10/15 | 0.667
23      | 1.101 | 82/120 | 0.683 || 2.464 | 10/15 | 0.667 || 2.534 | 11/15 | 0.733
24      | 0.574 | 115/120 | 0.958 || 2.008 | 13/15 | 0.867 || 2.172 | 13/15 | 0.867
25      | 0.971 | 88/120 | 0.733 || 2.478 | 11/15 | 0.733 || 2.630 | 12/15 | 0.800
26      | 0.690 | 113/120 | 0.942 || 2.257 | 13/15 | 0.867 || 2.570 | 13/15 | 0.867
27      | 1.745 | 80/120 | 0.667 || 3.481 | 10/15 | 0.667 || 3.615 | 10/15 | 0.667
28      | 1.457 | 80/120 | 0.667 || 3.112 | 10/15 | 0.667 || 3.303 | 10/15 | 0.667
29      | 1.645 | 80/120 | 0.667 || 3.316 | 10/15 | 0.667 || 3.517 | 10/15 | 0.667
30      | 1.548 | 81/120 | 0.675 || 3.188 | 10/15 | 0.667 || 3.412 | 10/15 | 0.667
31      | 1.924 | 82/120 | 0.683 || 3.617 | 10/15 | 0.667 || 3.620 | 11/15 | 0.733
Early Stopping: Validation error increased after 30 epochs.
```



Configurations:

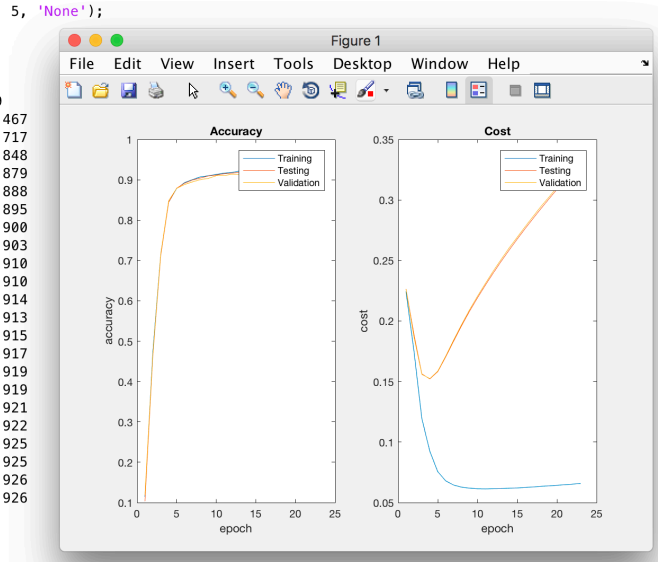
data set	epochs	hids	batch	eta	trans.	cost	mom.	reg.
MNIST	30	30	10	3.0	sigmoid	quad	.3	5

Matlab Code:

>> MyNetwork(trn, trnAns, [784,30,10], 30, 10, 3, [80, 10, 10], 0.3, 'sigmoid', 'quad', 5, 'None');

Matlab Output:

```
Early Stopping: Validation error increased after 30 epochs.
Trial>> MyNetwork(trn, trnAns, [784,30,10], 30, 10, 3, [80, 10, 10], 0.3, 'sigmoid', 'quad', 5, 'None');
        |        TRAIN          ||             TEST     ||     VALIDATION
-------------------------------------------------------------------------------------------
Ep      | Cost |   Corr  |  Acc  ||  Cost |  Corr |  Acc ||  Cost | Corr |  Acc
-------------------------------------------------------------------------------------------
1       | 0.224 | 4567/40000 | 0.114 || 0.227 | 517/5000 | 0.103 || 0.226 | 594/5000 | 0.119
2       | 0.175 | 18986/40000 | 0.475 || 0.189 | 2321/5000 | 0.464 || 0.187 | 2335/5000 | 0.467
3       | 0.119 | 28715/40000 | 0.718 || 0.156 | 3567/5000 | 0.713 || 0.156 | 3584/5000 | 0.717
4       | 0.092 | 33772/40000 | 0.844 || 0.152 | 4223/5000 | 0.845 || 0.152 | 4241/5000 | 0.848
5       | 0.076 | 35129/40000 | 0.878 || 0.158 | 4395/5000 | 0.879 || 0.159 | 4393/5000 | 0.879
6       | 0.068 | 35716/40000 | 0.893 || 0.170 | 4458/5000 | 0.892 || 0.171 | 4441/5000 | 0.888
7       | 0.064 | 36012/40000 | 0.900 || 0.183 | 4497/5000 | 0.899 || 0.184 | 4474/5000 | 0.895
8       | 0.063 | 36284/40000 | 0.907 || 0.196 | 4521/5000 | 0.904 || 0.197 | 4502/5000 | 0.900
9       | 0.062 | 36399/40000 | 0.910 || 0.208 | 4547/5000 | 0.909 || 0.209 | 4517/5000 | 0.903
10      | 0.061 | 36527/40000 | 0.913 || 0.219 | 4559/5000 | 0.912 || 0.220 | 4550/5000 | 0.910
11      | 0.061 | 36658/40000 | 0.916 || 0.230 | 4575/5000 | 0.915 || 0.231 | 4552/5000 | 0.910
12      | 0.062 | 36724/40000 | 0.918 || 0.240 | 4583/5000 | 0.917 || 0.241 | 4569/5000 | 0.914
13      | 0.062 | 36875/40000 | 0.922 || 0.249 | 4593/5000 | 0.919 || 0.251 | 4564/5000 | 0.913
14      | 0.062 | 36942/40000 | 0.924 || 0.259 | 4612/5000 | 0.922 || 0.260 | 4576/5000 | 0.915
15      | 0.062 | 37026/40000 | 0.926 || 0.268 | 4618/5000 | 0.924 || 0.269 | 4585/5000 | 0.917
16      | 0.063 | 37079/40000 | 0.927 || 0.276 | 4620/5000 | 0.924 || 0.278 | 4595/5000 | 0.919
17      | 0.063 | 37145/40000 | 0.929 || 0.285 | 4625/5000 | 0.925 || 0.286 | 4594/5000 | 0.919
18      | 0.063 | 37185/40000 | 0.930 || 0.293 | 4625/5000 | 0.925 || 0.295 | 4605/5000 | 0.921
19      | 0.064 | 37254/40000 | 0.931 || 0.301 | 4634/5000 | 0.927 || 0.302 | 4610/5000 | 0.922
20      | 0.064 | 37288/40000 | 0.932 || 0.309 | 4635/5000 | 0.927 || 0.310 | 4623/5000 | 0.925
21      | 0.065 | 37337/40000 | 0.933 || 0.316 | 4639/5000 | 0.928 || 0.318 | 4623/5000 | 0.925
22      | 0.065 | 37393/40000 | 0.935 || 0.324 | 4641/5000 | 0.928 || 0.325 | 4632/5000 | 0.926
23      | 0.066 | 37410/40000 | 0.935 || 0.331 | 4648/5000 | 0.930 || 0.333 | 4629/5000 | 0.926
Early Stopping: Validation error increased after 23 epochs.
Trial>>
Trial>>
```



Configurations:

data set	epochs	hids	batch	eta	trans.	cost	mom.	reg.
MNIST	30	30	10	3.0	softmax	log	.3	0

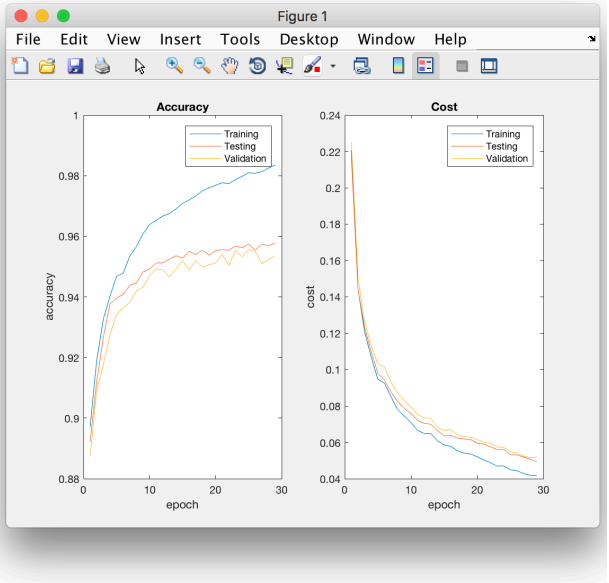
Matlab Code:

>> MyNetwork(trn, trnAns, [784,30,10], 30, 10, 3, [80, 10, 10], 0.3, 'softmax', 'log', 0, 'None');

note since the input transfer function is ‘soft’, the application will prompt the user to input a transfer function for the hidden layer (as shown below)

Matlab Output:

```
Trial>> MyNetwork(trn, trnAns, [784,30,10], 30, 10, 3, [80, 10, 10], 0.3, 'softmax', 'log', 0, 'None');
Softmax should only be used for the output layer. Which activation function do you want to use for hidden layers?
'sigmoid'
        |        TRAIN          ||             TEST     ||     VALIDATION
-------------------------------------------------------------------------------------------
Ep      | Cost |   Corr  |  Acc  ||  Cost |  Corr |  Acc ||  Cost | Corr |  Acc
-------------------------------------------------------------------------------------------
1       | 0.221 | 35884/40000 | 0.897 || 0.217 | 4461/5000 | 0.892 || 0.225 | 4437/5000 | 0.887
2       | 0.144 | 36765/40000 | 0.919 || 0.144 | 4561/5000 | 0.912 || 0.151 | 4546/5000 | 0.909
3       | 0.121 | 37308/40000 | 0.933 || 0.123 | 4633/5000 | 0.927 || 0.127 | 4588/5000 | 0.918
4       | 0.107 | 37606/40000 | 0.940 || 0.110 | 4688/5000 | 0.938 || 0.113 | 4637/5000 | 0.927
5       | 0.095 | 37870/40000 | 0.947 || 0.098 | 4698/5000 | 0.940 || 0.103 | 4671/5000 | 0.934
6       | 0.093 | 37912/40000 | 0.948 || 0.094 | 4704/5000 | 0.941 || 0.101 | 4682/5000 | 0.936
7       | 0.085 | 38136/40000 | 0.953 || 0.088 | 4719/5000 | 0.944 || 0.093 | 4691/5000 | 0.938
8       | 0.078 | 38261/40000 | 0.957 || 0.082 | 4723/5000 | 0.945 || 0.087 | 4710/5000 | 0.942
9       | 0.074 | 38431/40000 | 0.961 || 0.079 | 4741/5000 | 0.948 || 0.083 | 4716/5000 | 0.943
10      | 0.071 | 38555/40000 | 0.964 || 0.076 | 4746/5000 | 0.949 || 0.079 | 4735/5000 | 0.947
11      | 0.067 | 38611/40000 | 0.965 || 0.072 | 4756/5000 | 0.951 || 0.075 | 4746/5000 | 0.949
12      | 0.065 | 38667/40000 | 0.967 || 0.070 | 4756/5000 | 0.951 || 0.073 | 4745/5000 | 0.949
13      | 0.065 | 38699/40000 | 0.967 || 0.070 | 4762/5000 | 0.952 || 0.073 | 4733/5000 | 0.947
14      | 0.061 | 38762/40000 | 0.969 || 0.066 | 4768/5000 | 0.954 || 0.068 | 4746/5000 | 0.949
15      | 0.059 | 38836/40000 | 0.971 || 0.064 | 4764/5000 | 0.953 || 0.067 | 4760/5000 | 0.952
16      | 0.058 | 38879/40000 | 0.972 || 0.064 | 4775/5000 | 0.955 || 0.067 | 4744/5000 | 0.949
17      | 0.055 | 38932/40000 | 0.973 || 0.062 | 4770/5000 | 0.954 || 0.064 | 4760/5000 | 0.952
18      | 0.054 | 39000/40000 | 0.975 || 0.062 | 4777/5000 | 0.955 || 0.063 | 4749/5000 | 0.950
19      | 0.054 | 39040/40000 | 0.976 || 0.062 | 4769/5000 | 0.954 || 0.063 | 4753/5000 | 0.951
20      | 0.052 | 39073/40000 | 0.977 || 0.059 | 4776/5000 | 0.955 || 0.062 | 4756/5000 | 0.951
21      | 0.050 | 39106/40000 | 0.978 || 0.059 | 4778/5000 | 0.956 || 0.060 | 4770/5000 | 0.954
22      | 0.049 | 39096/40000 | 0.977 || 0.058 | 4777/5000 | 0.955 || 0.060 | 4752/5000 | 0.950
23      | 0.047 | 39149/40000 | 0.979 || 0.056 | 4784/5000 | 0.957 || 0.057 | 4777/5000 | 0.955
24      | 0.047 | 39191/40000 | 0.980 || 0.056 | 4781/5000 | 0.956 || 0.057 | 4766/5000 | 0.953
25      | 0.045 | 39238/40000 | 0.981 || 0.053 | 4787/5000 | 0.957 || 0.055 | 4778/5000 | 0.956
26      | 0.044 | 39232/40000 | 0.981 || 0.053 | 4777/5000 | 0.955 || 0.054 | 4775/5000 | 0.955
27      | 0.043 | 39252/40000 | 0.981 || 0.052 | 4787/5000 | 0.957 || 0.052 | 4755/5000 | 0.951
28      | 0.042 | 39300/40000 | 0.983 || 0.051 | 4784/5000 | 0.957 || 0.052 | 4762/5000 | 0.952
29      | 0.042 | 39340/40000 | 0.984 || 0.049 | 4789/5000 | 0.958 || 0.052 | 4767/5000 | 0.953
Early Stopping: Validation error increased after 23 epochs.
```



Configurations:

data set	epochs	hids	batch	eta	trans.	cost	mom.	reg.
MNIST	30	30	10	1.0	softmax	log	.3	5

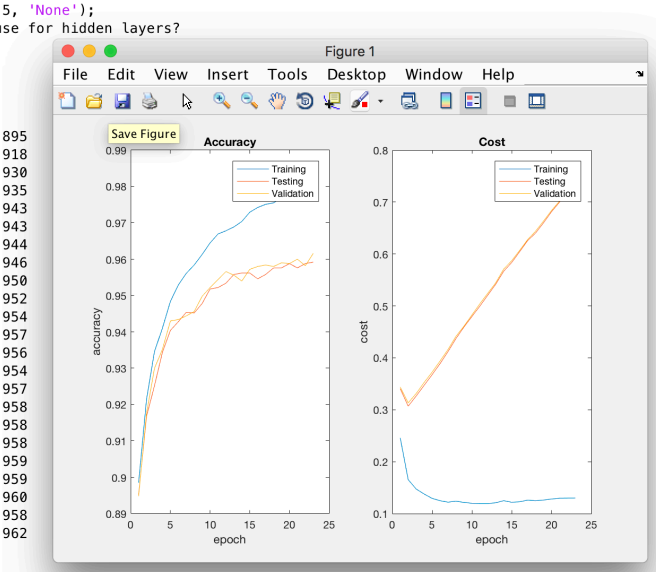
Matlab Code:

>> MyNetwork(trn, trnAns, [784,30,10], 30, 10, 3, [80, 10, 10], 0.3, 'softmax', 'log', 5, 'None');

note since the input transfer function is ‘soft’, the application will prompt the user to input a transfer function for the hidden layer (as shown below)

Matlab Output:

```
Trial>> MyNetwork(trn, trnAns, [784,30,10], 30, 10, 3, [80, 10, 10], 0.3, 'softmax', 'log', 5, 'None');
Softmax should only be used for the output layer. Which activation function do you want to use for hidden layers?
'sigmoid'
        |        TRAIN          ||             TEST     ||     VALIDATION
-------------------------------------------------------------------------------------------
Ep      | Cost |   Corr  |  Acc  ||  Cost |  Corr |  Acc ||  Cost | Corr |  Acc
-------------------------------------------------------------------------------------------
1       | 0.246 | 35941/40000 | 0.899 || 0.340 | 4476/5000 | 0.895 || 0.344 | 4474/5000 | 0.895
2       | 0.165 | 36859/40000 | 0.921 || 0.307 | 4584/5000 | 0.917 || 0.313 | 4588/5000 | 0.918
3       | 0.147 | 37385/40000 | 0.935 || 0.326 | 4626/5000 | 0.925 || 0.331 | 4650/5000 | 0.930
4       | 0.138 | 37641/40000 | 0.941 || 0.347 | 4672/5000 | 0.934 || 0.353 | 4676/5000 | 0.935
5       | 0.129 | 37935/40000 | 0.948 || 0.368 | 4702/5000 | 0.940 || 0.372 | 4715/5000 | 0.943
6       | 0.125 | 38116/40000 | 0.953 || 0.389 | 4714/5000 | 0.943 || 0.394 | 4717/5000 | 0.943
7       | 0.122 | 38244/40000 | 0.956 || 0.412 | 4727/5000 | 0.945 || 0.416 | 4722/5000 | 0.944
8       | 0.124 | 38336/40000 | 0.958 || 0.437 | 4726/5000 | 0.945 || 0.441 | 4728/5000 | 0.946
9       | 0.122 | 38452/40000 | 0.961 || 0.459 | 4739/5000 | 0.948 || 0.461 | 4749/5000 | 0.950
10      | 0.120 | 38579/40000 | 0.964 || 0.479 | 4759/5000 | 0.952 || 0.482 | 4761/5000 | 0.952
11      | 0.120 | 38677/40000 | 0.967 || 0.500 | 4761/5000 | 0.952 || 0.504 | 4772/5000 | 0.954
12      | 0.120 | 38713/40000 | 0.968 || 0.521 | 4767/5000 | 0.953 || 0.524 | 4783/5000 | 0.957
13      | 0.121 | 38755/40000 | 0.969 || 0.542 | 4779/5000 | 0.956 || 0.545 | 4778/5000 | 0.956
14      | 0.125 | 38815/40000 | 0.970 || 0.567 | 4781/5000 | 0.956 || 0.571 | 4770/5000 | 0.954
15      | 0.122 | 38917/40000 | 0.973 || 0.584 | 4781/5000 | 0.956 || 0.587 | 4786/5000 | 0.957
16      | 0.123 | 38969/40000 | 0.974 || 0.604 | 4773/5000 | 0.955 || 0.607 | 4790/5000 | 0.958
17      | 0.126 | 39004/40000 | 0.975 || 0.626 | 4779/5000 | 0.956 || 0.628 | 4792/5000 | 0.958
18      | 0.125 | 39020/40000 | 0.976 || 0.640 | 4788/5000 | 0.958 || 0.644 | 4790/5000 | 0.958
19      | 0.126 | 39071/40000 | 0.977 || 0.660 | 4788/5000 | 0.958 || 0.663 | 4795/5000 | 0.959
20      | 0.128 | 39125/40000 | 0.978 || 0.681 | 4794/5000 | 0.959 || 0.684 | 4794/5000 | 0.959
21      | 0.130 | 39148/40000 | 0.979 || 0.700 | 4788/5000 | 0.958 || 0.702 | 4800/5000 | 0.960
22      | 0.130 | 39134/40000 | 0.978 || 0.716 | 4794/5000 | 0.959 || 0.720 | 4791/5000 | 0.958
23      | 0.130 | 39228/40000 | 0.981 || 0.735 | 4796/5000 | 0.959 || 0.737 | 4808/5000 | 0.962
Early Stopping: Validation error increased after 23 epochs.
Trial>>
```



Configurations:

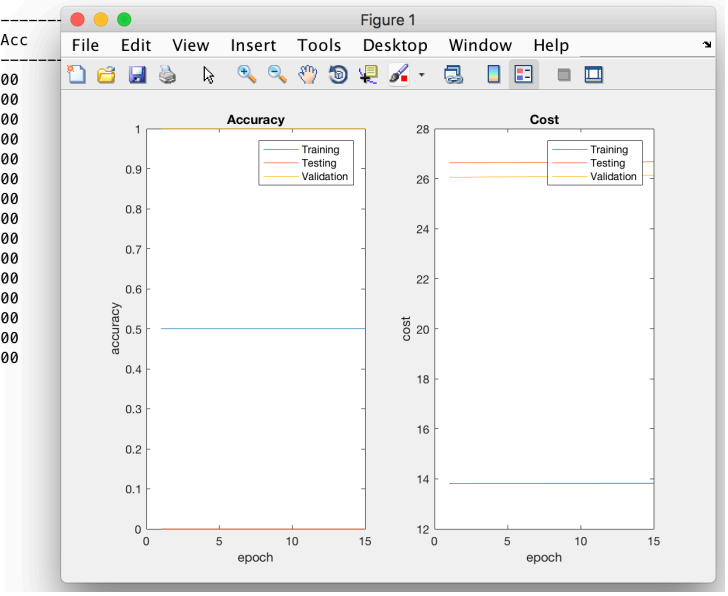
data set	epochs	hids	batch	eta	trans.	cost	mom.	reg.
xor.csv	20	[3 2]	1	0.1	sigmoid	cross	.3	5

Matlab Code:

>> MyNetwork(xor_input, xor_targets, [2,3,2,1], 20, 1, 0.1, [50, 25, 25], 0.3, 'sigmoid', 'cross', 5, 'None');

Matlab Output:

```
Trial>> MyNetwork(xor_input, xor_targets, [2,3,2,1], 20, 1, 0.1, [50, 25, 25], 0.3, 'sigmoid', 'cross', 5, 'None');
        |        TRAIN          ||             TEST     ||     VALIDATION
-------------------------------------------------------------------------------------------
Ep      | Cost |   Corr  |  Acc  ||  Cost |  Corr |  Acc ||  Cost | Corr |  Acc
-------------------------------------------------------------------------------------------
1       | 13.811 | 1/2 | 0.500 || 26.642 | 0/1 | 0.000 || 26.058 | 1/1 | 1.000
2       | 13.811 | 1/2 | 0.500 || 26.642 | 0/1 | 0.000 || 26.053 | 1/1 | 1.000
3       | 13.812 | 1/2 | 0.500 || 26.645 | 0/1 | 0.000 || 26.064 | 1/1 | 1.000
4       | 13.812 | 1/2 | 0.500 || 26.648 | 0/1 | 0.000 || 26.071 | 1/1 | 1.000
5       | 13.813 | 1/2 | 0.500 || 26.648 | 0/1 | 0.000 || 26.071 | 1/1 | 1.000
6       | 13.814 | 1/2 | 0.500 || 26.652 | 0/1 | 0.000 || 26.083 | 1/1 | 1.000
7       | 13.814 | 1/2 | 0.500 || 26.655 | 0/1 | 0.000 || 26.090 | 1/1 | 1.000
8       | 13.815 | 1/2 | 0.500 || 26.656 | 0/1 | 0.000 || 26.089 | 1/1 | 1.000
9       | 13.816 | 1/2 | 0.500 || 26.658 | 0/1 | 0.000 || 26.095 | 1/1 | 1.000
10      | 13.816 | 1/2 | 0.500 || 26.662 | 0/1 | 0.000 || 26.108 | 1/1 | 1.000
11      | 13.817 | 1/2 | 0.500 || 26.665 | 0/1 | 0.000 || 26.115 | 1/1 | 1.000
12      | 13.818 | 1/2 | 0.500 || 26.668 | 0/1 | 0.000 || 26.122 | 1/1 | 1.000
13      | 13.819 | 1/2 | 0.500 || 26.671 | 0/1 | 0.000 || 26.129 | 1/1 | 1.000
14      | 13.819 | 1/2 | 0.500 || 26.671 | 0/1 | 0.000 || 26.128 | 1/1 | 1.000
15      | 13.821 | 1/2 | 0.500 || 26.676 | 0/1 | 0.000 || 26.141 | 1/1 | 1.000
Early Stopping: Validation error increased after 15 epochs.
Trial>>
Trial>>
Trial>>
Trial>>
Trial>>
Trial>>
Trial>>
Trial>>
Trial>>
Trial>>
Trial>>
```



Configurations:

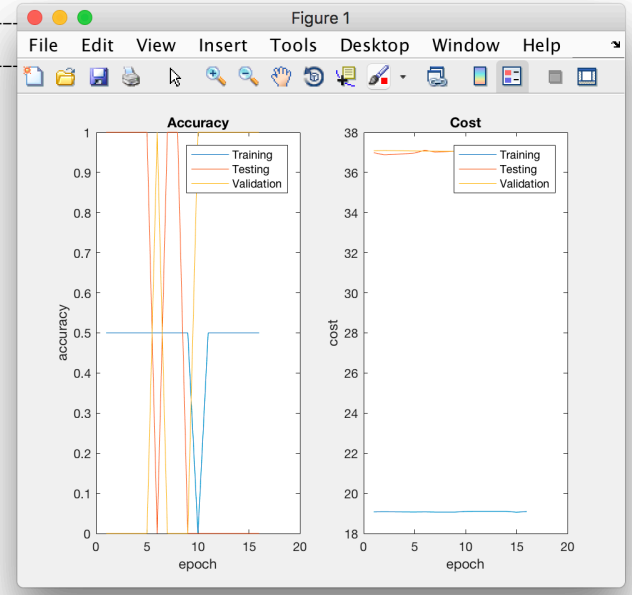
data set	epochs	hids	batch	eta	trans.	cost	mom.	reg.
xor.csv	20	[3 2]	1	0.1	tanh	cross	.3	5

Matlab Code:

>> MyNetwork(xor_input, xor_targets, [2,3,2,1], 20, 1, 0.1, [50, 25, 25], 0.3, 'tanh', 'cross', 5, 'None');

Matlab Output:

```
Trial>> MyNetwork(xor_input, xor_targets, [2,3,2,1], 20, 1, 0.1, [50, 25, 25], 0.3, 'tanh', 'cross', 5, 'None');
        |        TRAIN          ||             TEST     ||     VALIDATION
-------------------------------------------------------------------------------
Ep      | Cost |   Corr  | Acc  || Cost |  Corr |  Acc || Cost | Corr |  Acc
-------------------------------------------------------------------------------
1       | 19.075 | 1/2 | 0.500 || 36.984 | 1/1 | 1.000 || 37.082 | 0/1 | 0.000
2       | 19.085 | 1/2 | 0.500 || 36.878 | 1/1 | 1.000 || 37.096 | 0/1 | 0.000
3       | 19.079 | 1/2 | 0.500 || 36.898 | 1/1 | 1.000 || 37.088 | 0/1 | 0.000
4       | 19.072 | 1/2 | 0.500 || 36.932 | 1/1 | 1.000 || 37.078 | 0/1 | 0.000
5       | 19.068 | 1/2 | 0.500 || 36.966 | 1/1 | 1.000 || 37.071 | 0/1 | 0.000
6       | 19.076 | 1/2 | 0.500 || 37.107 | 0/1 | 0.000 || 37.066 | 1/1 | 1.000
7       | 19.065 | 1/2 | 0.500 || 37.019 | 1/1 | 1.000 || 37.060 | 0/1 | 0.000
8       | 19.064 | 1/2 | 0.500 || 37.035 | 1/1 | 1.000 || 37.057 | 0/1 | 0.000
9       | 19.066 | 1/2 | 0.500 || 37.060 | 0/1 | 0.000 || 37.055 | 0/1 | 0.000
10      | 19.097 | 0/2 | 0.000 || 37.224 | 0/1 | 0.000 || 37.067 | 1/1 | 1.000
11      | 19.104 | 1/2 | 0.500 || 37.253 | 0/1 | 0.000 || 37.069 | 1/1 | 1.000
12      | 19.104 | 1/2 | 0.500 || 37.256 | 0/1 | 0.000 || 37.067 | 1/1 | 1.000
13      | 19.103 | 1/2 | 0.500 || 37.254 | 0/1 | 0.000 || 37.064 | 1/1 | 1.000
14      | 19.101 | 1/2 | 0.500 || 37.252 | 0/1 | 0.000 || 37.061 | 1/1 | 1.000
15      | 19.060 | 1/2 | 0.500 || 37.079 | 0/1 | 0.000 || 37.037 | 1/1 | 1.000
16      | 19.091 | 1/2 | 0.500 || 37.221 | 0/1 | 0.000 || 37.050 | 1/1 | 1.000
Early Stopping: Validation error increased after 15 epochs.
Trial>>
Trial>>
Trial>>
Trial>>
Trial>>
Trial>>
Trial>>
```



Configurations:

data set	epochs	hids	batch	eta	trans.	cost	mom.	reg.
xor.csv	20	[3 2]	1	0.1	relu	cross	.3	5

Matlab Code:

>> MyNetwork(xor_input, xor_targets, [2,3,2,1], 20, 1, 0.1, [50, 25, 25], 0.3, 'relu', 'cross', 5, 'None');

note since the input transfer function is ‘relu’, the application will prompt the user to input a transfer function for the output layer (as shown below)

Matlab Output:

```
Trial>> MyNetwork(xor_input, xor_targets, [2,3,2,1], 20, 1, 0.1, [50, 25, 25], 0.3, 'relu', 'cross', 5, 'None');
ReLu should only be used for the hidden layers. Which activation function do you want to use for the output layer?
'sigmoid'
        |        TRAIN          ||             TEST     ||     VALIDATION
-------------------------------------------------------------------------------
Ep      | Cost |   Corr  | Acc  || Cost |  Corr |  Acc || Cost | Corr |  Acc
-------------------------------------------------------------------------------
1       | 11.467 | 2/2 | 1.000 || 23.270 | 0/1 | 0.000 || 23.024 | 0/1 | 0.000
2       | 11.468 | 2/2 | 1.000 || 23.284 | 0/1 | 0.000 || 23.036 | 1/1 | 1.000
3       | 11.475 | 2/2 | 1.000 || 23.308 | 0/1 | 0.000 || 23.060 | 1/1 | 1.000
4       | 11.473 | 2/2 | 1.000 || 23.329 | 0/1 | 0.000 || 23.073 | 1/1 | 1.000
5       | 11.489 | 2/2 | 1.000 || 23.357 | 0/1 | 0.000 || 23.111 | 1/1 | 1.000
6       | 11.497 | 2/2 | 1.000 || 23.382 | 0/1 | 0.000 || 23.138 | 1/1 | 1.000
7       | 11.501 | 2/2 | 1.000 || 23.404 | 0/1 | 0.000 || 23.152 | 1/1 | 1.000
8       | 11.515 | 2/2 | 1.000 || 23.433 | 0/1 | 0.000 || 23.191 | 1/1 | 1.000
9       | 11.529 | 2/2 | 1.000 || 23.458 | 0/1 | 0.000 || 23.219 | 1/1 | 1.000
10      | 11.542 | 2/2 | 1.000 || 23.484 | 0/1 | 0.000 || 23.246 | 1/1 | 1.000
11      | 11.554 | 2/2 | 1.000 || 23.506 | 0/1 | 0.000 || 23.261 | 1/1 | 1.000
12      | 11.569 | 2/2 | 1.000 || 23.536 | 0/1 | 0.000 || 23.300 | 1/1 | 1.000
13      | 11.600 | 2/2 | 1.000 || 23.596 | 0/1 | 0.000 || 23.360 | 1/1 | 1.000
14      | 11.615 | 2/2 | 1.000 || 23.625 | 0/1 | 0.000 || 23.391 | 1/1 | 1.000
15      | 11.628 | 2/2 | 1.000 || 23.648 | 0/1 | 0.000 || 23.407 | 1/1 | 1.000
Early Stopping: Validation error increased after 15 epochs.
Trial>>
Trial>>
Trial>>
Trial>>
Trial>>
Trial>>
Trial>>
Trial>>
```