**David R Mohler**

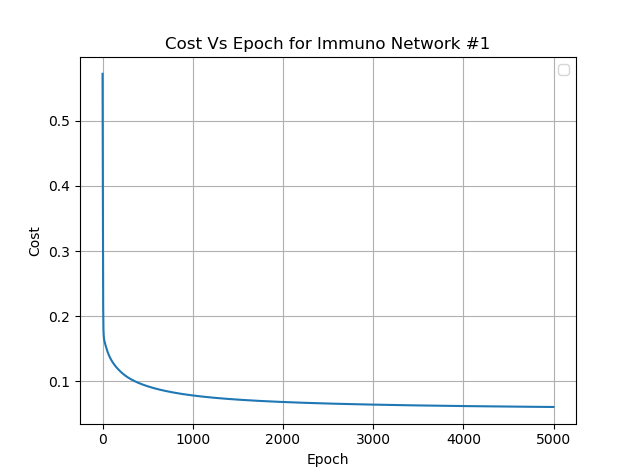
**EE-5410: Neural Networks**

**Assignment #4**

**04-13-2018**

For this assignment I have implemented six unique neural networks for the purpose of classifying the data sets given that describe the success rate of two different wart treatments for patients in conjunction with other patient data. The two methods being classified are that of Immunotherapy and Cryotherapy, with the added information of age, time, number of warts, type, area, and in the case of immunotherapy, induration diameter. With each network I will show its ability to classify the data associated with its own data set as well as cross validation of the network against its counterpart data set. Each network is tested under a series of ten trials for testing under its own data and against the cross performance testing data.

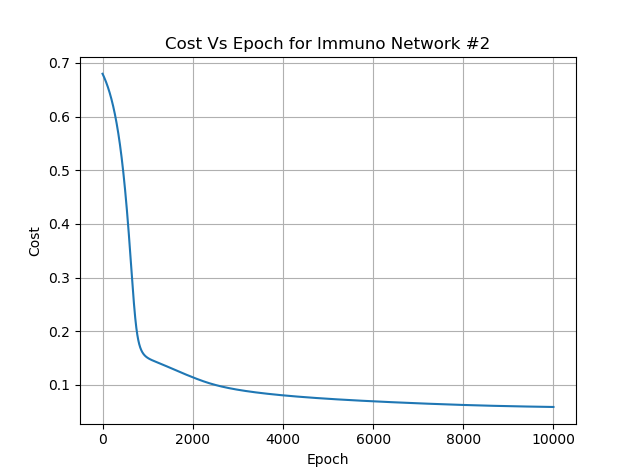
* **Immunology Network #1:**

This is a basic implementation of the back propagation algorithm with only a single hidden layer consisting of six inputs, 36 neurons in the hidden layer, and a single output neuron. All of the networks are tested using a global learning rate of in order to view the effect of different topologies and optimizations independently of the learning rate. Using the momentum optimizer, the single hidden layer network was trained and tested 10 times with the immunology data set. The average across all trials for this network was 78.89%, with a maximum achieved accuracy of 83.33%. When generalizing the network (i.e. training the network with immunology treatment data and attempting classification of the cryotherapy data set), the network achieved an average of 72.22% classification.

* + **Topology:** 
    - 6 inputs (ignoring the sex of the patient)
    - 1 Hidden Layer: 36 Neurons
    - Single output
  + **Optimizer:**
    - Momentum Optimizer
  + **Network Accuracy: (10 Trial Average | 5000 Epochs)** 
    - Immunology Data: 78.89%
    - Cryotherapy Data: 68.33%
* **Immunology Network #2:**

The second network is an expansion to three hidden layers with a change in optimizer to the AdaDelta optimizer. Through experimentation, the AdaDelta was able to provide a slight increase in the average performance of the network on the immunology data (79.44%), with no loss of generality such that the network was able to provide the same average performance when presented with the cryotherapy test set (68.33%).

* + **Topology:**

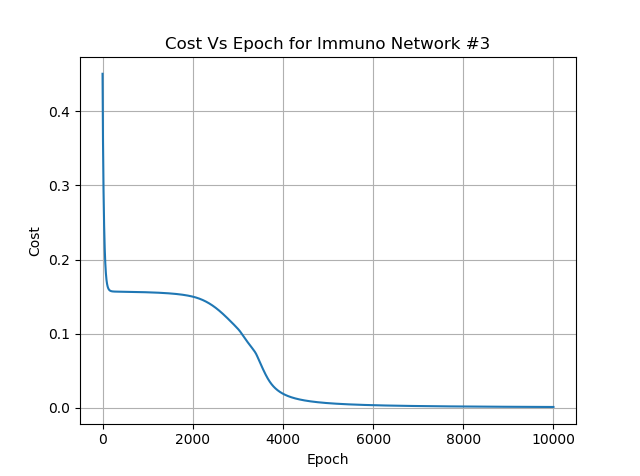


* + - 6 inputs (ignoring the sex of the patient)
    - 3 Hidden Layers:

|  |  |  |  |
| --- | --- | --- | --- |
| Layer | 1 | 2 | 3 |
| Neurons | 18 | 24 | 18 |

* + - Single output
  + **Optimizer:**
    - AdaDelta Optimizer
  + **Network Accuracy: (10 Trial Average | 10,000 Epochs)** 
    - Immunology Data: 79.44%
    - Cryotherapy Data: 68.33%
* **Immunology Network #3:**

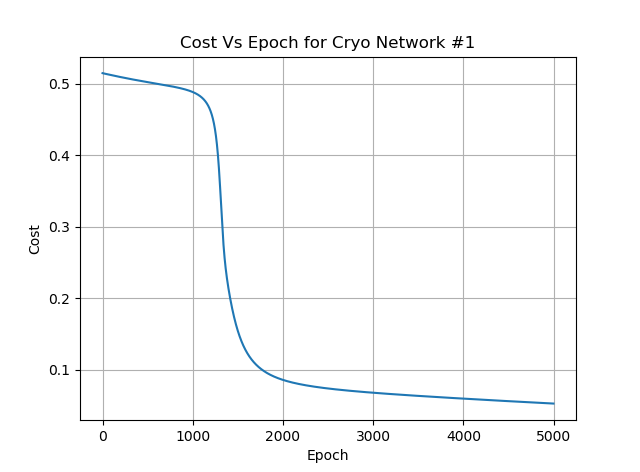
Again, we expand the number of hidden layers and the number of neurons per layer with a chosen “decrescendo” topology, such that each successive layer reduces in the number of neurons it contains, from 100 in the first hidden layer, down to the 7th layer which contains only two neurons. For this architecture, the Adagrad optimizer gave the highest average classification of the cryotherapy data. This architecture yielded interesting results. It was found that across 10 trials that the ability of the network to classify the testing patterns drawn from the Immunology data set was the worst across all six implemented networks at a total of 71.11%. In contrast, this network provided the highest accuracy during validation with the cryotherapy data set when compared against the other five network implementations with nearly 75% accuracy.

* + **Topology:** 
    - 6 inputs (ignoring the sex of the patient)
    - 7 Hidden Layers:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Layer | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Neurons | 100 | 75 | 50 | 25 | 10 | 5 | 2 |

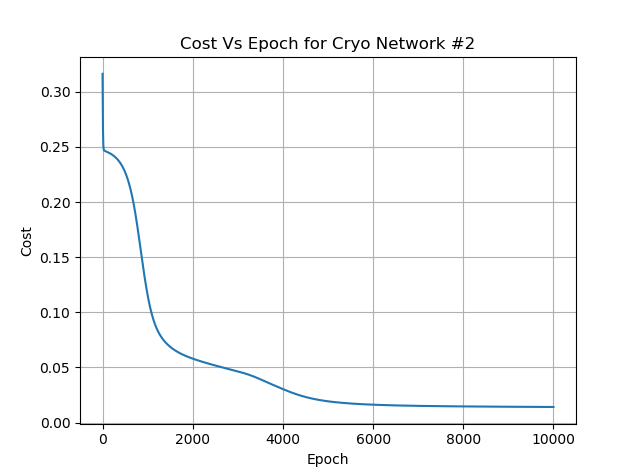
* Single output
  + **Optimizer:**
    - AdaGrad Optimizer
  + **Network Accuracy: (10 Trial Average | 10,000 Epochs)** 
    - Immunology Data: 71.11 %
    - Cryotherapy Data: 74.99%
* **Cryotherapy Network #1:**

For the cryotherapy data sets, the same general approach was used, beginning with a single layer network and successively growing in the following network architectures. For the initial Cryotherapy we began with classification using a single hidden layer comprised of 25 neurons and applying the AdaDelta optimizer. This network is able to tie for the best performance in terms of the classification of the testing set from the data it was trained with (i.e. when training with the cryotherapy data and testing with its associated test set) at an average of 83.33%. However, this network was the least generalized with the lowest average cross performance managing an average of only 54.45%, which is only slightly better than random chance for the given two class problem.



* + **Topology:** 
    - 6 inputs (ignoring the sex of the patient)
    - 1 Hidden Layer: 25 Neurons
    - Single output
  + **Optimizer:**
    - Momentum Optimizer
  + **Network Accuracy: (10 Trial Average | 5000 Epochs)** 
    - Immunology Data: 83.33%
    - Cryotherapy Data: 54.45%
* **Cryotherapy Network #2:**

The second Cryotherapy trained network consists of five hidden layers with an architecture such that the number of neurons increase by five per layer from 10 in the first to 20 in the middle (3rd) hidden layer and then decreases by increments of five to return to 10 neurons in the final hidden layer (5). Along with this, we return to the use of the momentum optimizer, which was yielding relatively strong results in selection trials. Using this topology, we were able to achieve an average accuracy of 83.33% with a maximum of 88.89% successful classification of the cryotherapy data. When tested with the immunology data we can also see that it provided the second best generalization for cross classification at 70%.

* + **Topology:** 
    - 6 inputs (ignoring the sex of the patient)
    - 5 Hidden Layers:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Layer | 1 | 2 | 3 | 4 | 5 |
| Neurons | 10 | 15 | 20 | 15 | 10 |

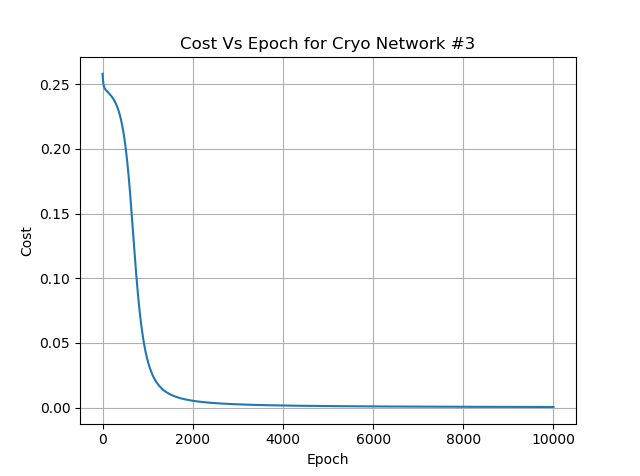
* + - Single output
  + **Optimizer:**
    - Momentum Optimizer
  + **Network Accuracy: (10 Trial Average | 10,000 Epochs)** 
    - Immunology Data: 83.33%
    - Cryotherapy Data: 70.00%
* **Cryotherapy Network #3:**

The final network was the largest tested of the six implementations. This network consisted of a total of 10 layers and following a general “decrescendo” structure as was seen in the third immunology implementation, as well as use of the momentum optimizer again. It can be noted that the major increase in neurons does not necessarily translate to an increase in network accuracy. While the network performed comparably to the top performing networks at 81.67%, and 66.11% (cross evaluation) it still does not provide the top performance, which makes the increase in computation necessary for the calculations hard to justify.

* + **Topology:** 
    - 6 inputs (ignoring the sex of the patient)
    - 10 Hidden Layers:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Layer | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Neurons | 300 | 200 | 100 | 50 | 25 | 20 | 15 | 10 | 5 | 2 |

* + - Single output
  + **Optimizer:**
    - AdaGrad Optimizer
  + **Network Accuracy: (10 Trial Average | 10,000 Epochs)** 
    - Immunology Data: 81.67 %
    - Cryotherapy Data: 66.11%



* **Summary of Results:**

By reviewing the results of the six implementations, the “Cryo2” network provides the best generalization, not necessarily in the sense that it had the maximum mean cross validation percentage, but rather in the sense that the average of the validation and cross validation percentages was largest. Cryo2 provides an average accuracy of approximately 76.65% when considering both data sets which is nearly 3% greater than the nearest competitor (Immuno2) with 73.89%.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Network** | **Hidden Layers** | **Hidden Neurons** | **Epochs** |  | **Optimizer** | **Validation Accuracy (%)** | **Cross Validation Accuracy (%)** |
| **Immuno1** | **1** | **36** | **5000** | **0.1** | **Momentum** | **78.89** | **68.33** |
| **Immuno2** | **3** | **18,24,18** | **10000** | **0.1** | **AdaDelta** | **79.44** | **68.33** |
| **Immuno3** | **7** | **100,75,50,25,10,5,2** | **10000** | **0.1** | **Adagrad** | **71.11** | **74.99** |
| **Cryo1** | **1** | **25** | **5000** | **0.1** | **AdaDelta** | **83.33** | **54.45** |
| **Cryo2** | **5** | **10,15,20,15,10** | **10000** | **0.1** | **Momentum** | **83.33** | **70** |
| **Cryo3** | **10** | **300,200,100,50,**  **25,20,15,10,5,2** | **10000** | **0.1** | **Momentum** | **81.67** | **66.11** |