**Homework 2**

**COSC 6342: Machine Learning**

**Submitted by**

**S M Salah Uddin Kadir (1800503)**

**Rubayat Jinnah (1891217)**

**Dataset:**

Training Dataset URL: <https://storage.googleapis.com/download.tensorflow.org/data/iris_training.csv>

Testing Dataset URL:

<https://storage.googleapis.com/download.tensorflow.org/data/iris_test.csv>

**Feature types:**

There are 4 types of features in the dataset:

1. Sepal length
2. Sepal width
3. Petal length
4. Petal width

**Ranges of values:???**

**Classes:**

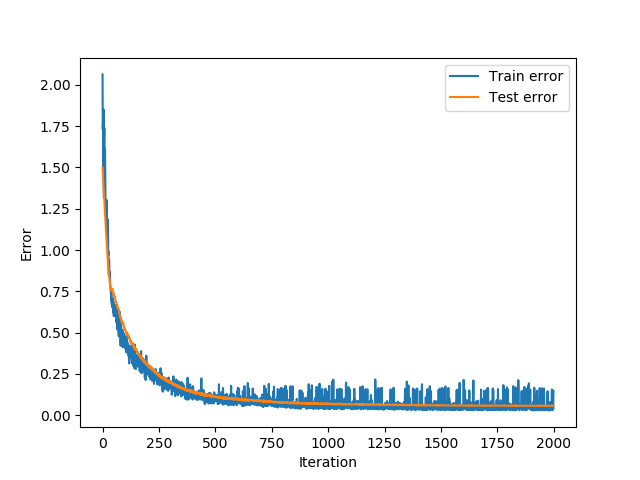
The dataset will be divided into 3 classes:

1. Iris setosa
2. Iris versicolor
3. Iris virginica

**Figures:**

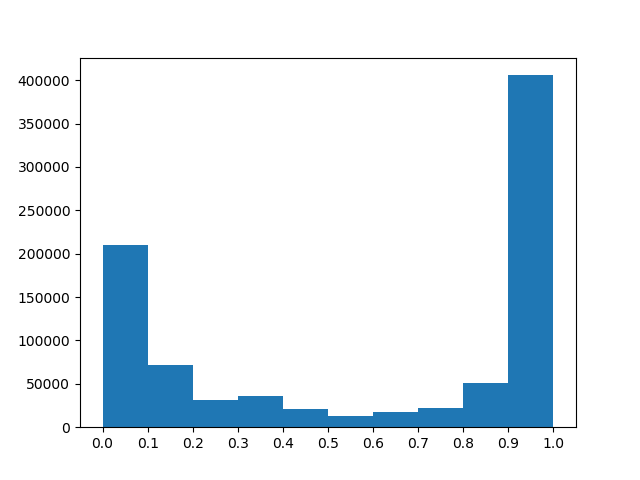
**Training and Test error changing rate along iterations:**

The figure below shows an example of training versus test error along iterations. The X axis corresponds to the iteration and the y axis corresponds to the error.

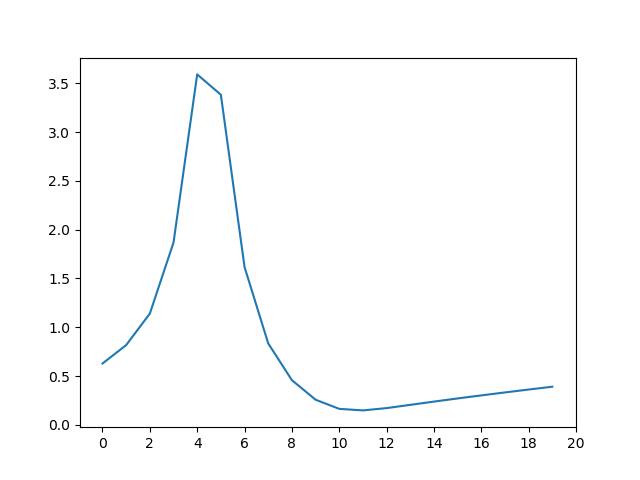
****

**Histogram of activation of a hidden unit**

The figure below illustrates a sample output for histogram of activation of a hidden unit. The X axis corresponds to the activation value and the y axis corresponds to the frequency of samples with that activation value.



**Weight change:**



**Question Answer**

A) **Weight changes** - By plotting the weight changes for hidden units along iterations, answer the following questions:

1. Do the weights of units in the same layer become stable almost at the same time? Does it depend on the layer?

**Ans**: No, the weights of units in the same layer does not become stable almost at the same time. It depends on the layer.

1. Do the weights in different layers become stable around the same time? Explain.

**Ans**: Yes. The weights in different layers become stable around the same time. Because, in neural network the error rate reduces from each layer by calibrating values of weights in the layers. Thus, all the layers in the neural network changes weights in begging stage of forward and backward propagations. And the weights become stable for different layers at the same time when the error rate becomes significantly low.

1. In which iteration do you think the model overfits the data? Do you see any relation between the weight change and overfitting?

**Ans**:

B) **Activation changes** – By plotting the histogram of activations for hidden units along iterations, answer the following questions:

1. Do different neurons have similar distribution of activations in the same layer? Explain.

**Ans:**

1. Do different neurons have similar distribution of activations in different layers? Explain.

**Ans:**

C) **Parameter changes** – By changing the input parameters and training the network, answer the following questions:

1. How does nonlinearity (sigmoid, tanh or ReLU) affect overfitting?

In overfitting you are interested in error on your training and test set. If your training set error is dropping but your test set is increasing (as on the picture) then you are overfitting. So this is what you should be looking at. You can be looking at on which point ooverfitting appears, or how severe it can be. But number of neurons and regularization will have way higher impact than activation function

1. How does the number of hidden layers and number of units in hidden layers affect the training?