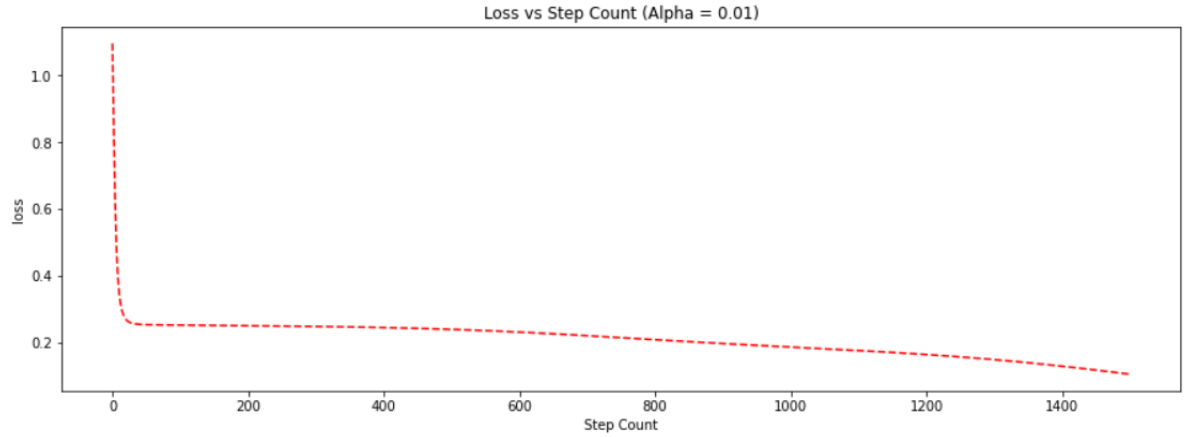
Assignment – 02
Report file

Shivani Panchiwala
UTA ID: 1001982478
CSE 6363 – 001
Prof.: Alex Dillhoff

XOR Problem

The model uses one hidden layer and Hyperbolic tangent Function with a learning rate (alpha) of 0.01, given step count as 1500.




```
Accuracy : 100.0 %
```

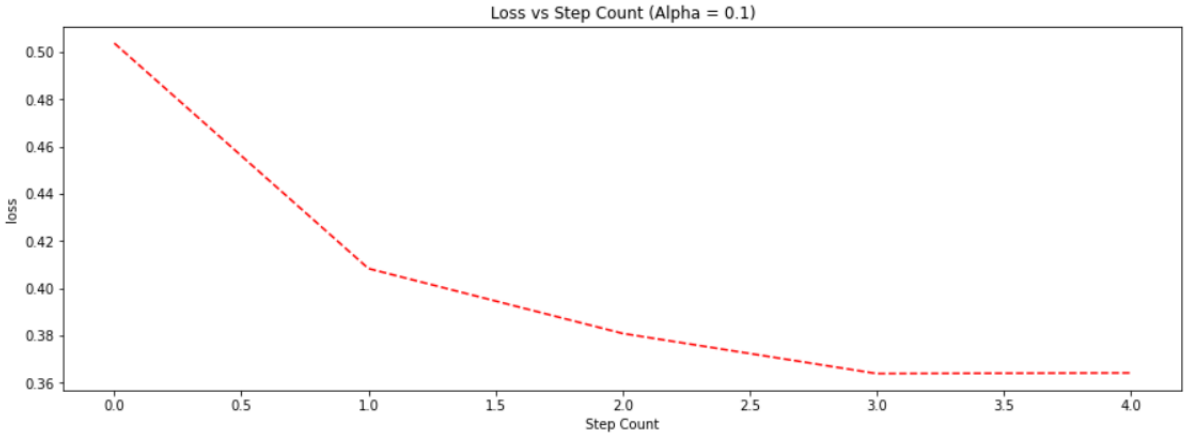
To solve this problem with the architecture using sigmoid activations and hyperbolic tangent activations. But hyperbolic tangent functions is more easier than sigmoid function to train this problem.

Handwritten Digit Recognition

I built three different models with different number of hidden layers and learning rate parameter. Each model when trained over 60000 samples of MNIST dataset and tested over 10000 samples. Randomly selected 10% of the training images to use as a validation set.

Model 1:

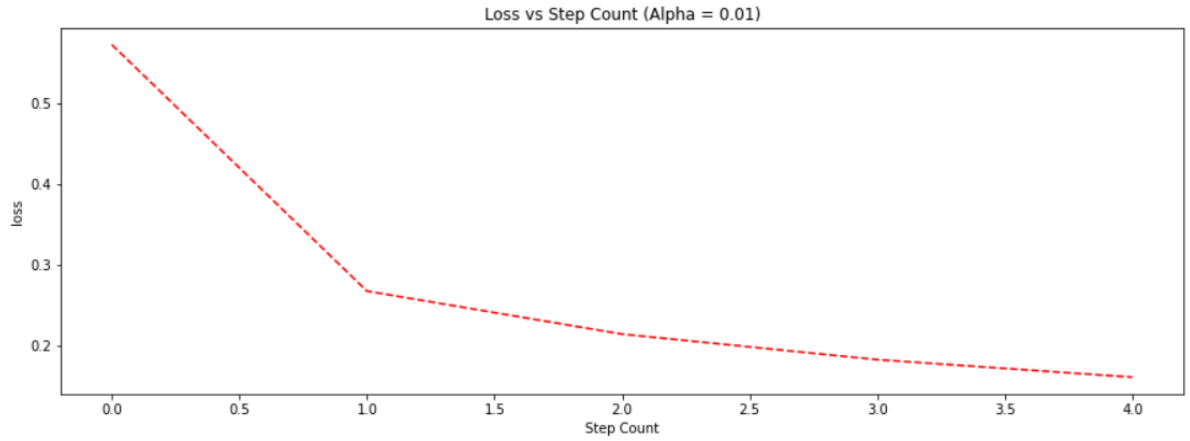
Number of Hidden Layers = 1 (Using Hyperbolic Tangent function at Hidden Layer and SoftMax Function at Output layer)
Learning Rate (Alpha) = 0.1
Accuracy Observed = 94.99 %
Step Count = 5




```
Accuracy For Model 1: (94.99) %
```

Model 2:

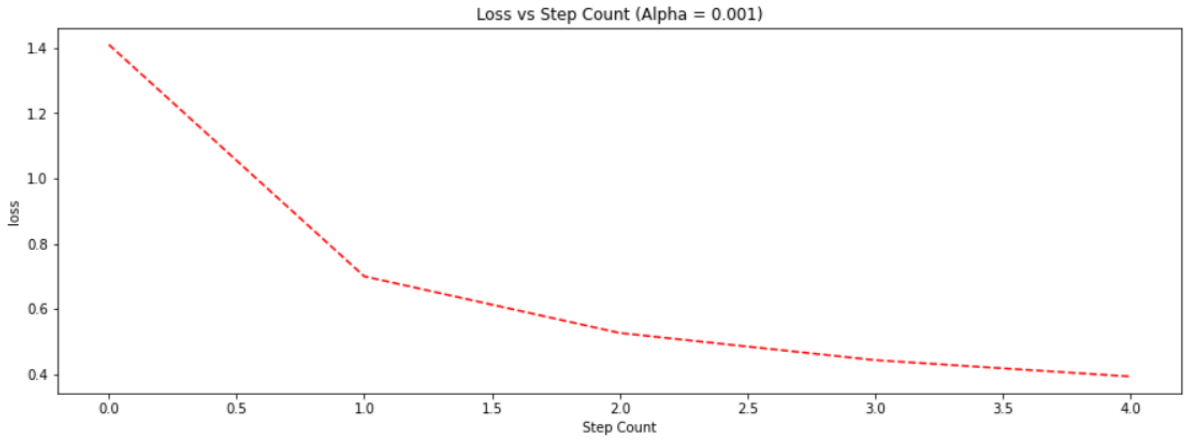
Number of Hidden Layers = 2 (Using Hyperbolic Tangent function at Hidden Layer and SoftMax Function at Output layer)
Learning Rate (Alpha) = 0.01
Accuracy Observed = 94.73 %
Step Count = 5




```
Accuracy For Model 2: (94.73) %
```

Model 3:

Number of Hidden Layers = 3 (Using Hyperbolic Tangent function at Hidden Layer and SoftMax Function at Output layer)
Learning Rate (Alpha) = 0.001
Accuracy Observed = 89.32 %
Step Count = 5

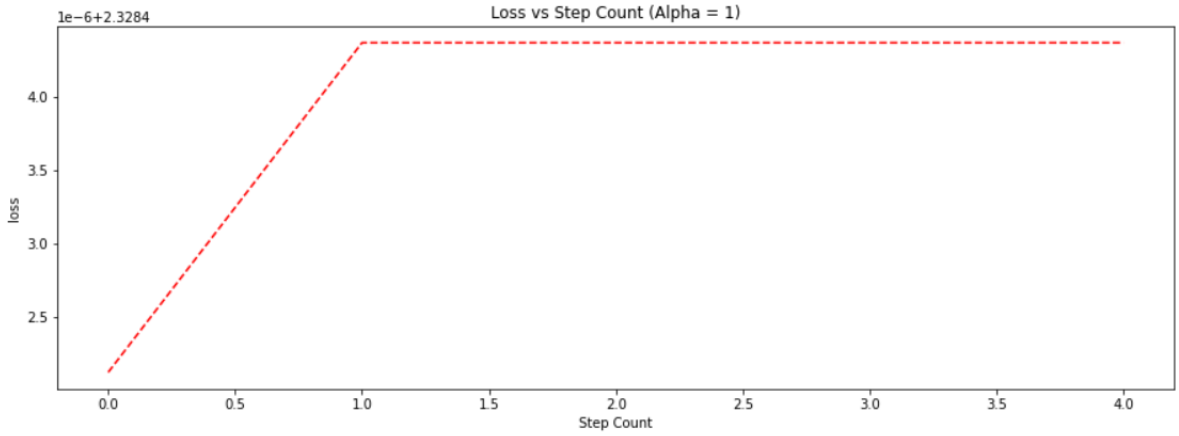



```
Accuracy For Model 3: (89.32) %
```

Playing with Hyperparameters

Experiment Model 1:

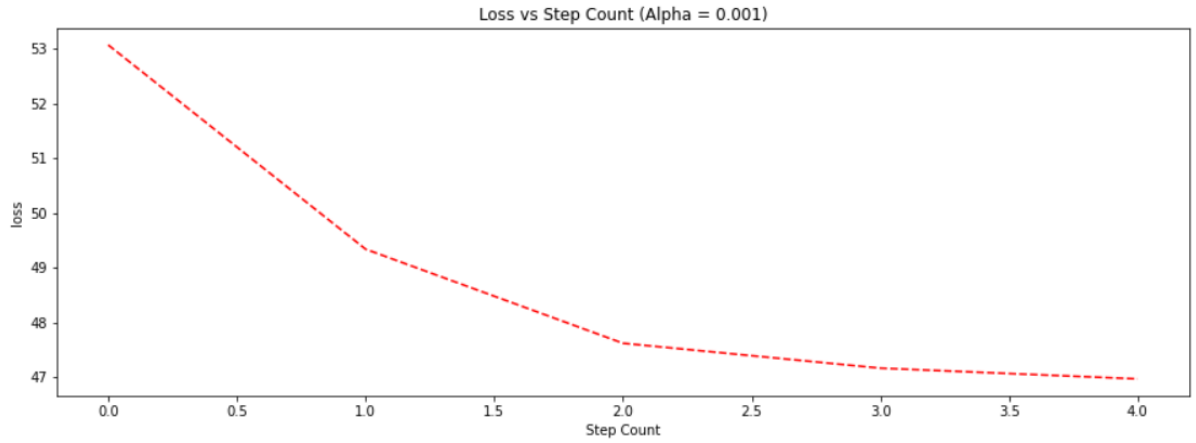
- Initializing the weights to zero
- Training the model and testing for alpha = 1
- Using Hyperbolic tangent function at Hidden Layer and Softmax Function at Output Layer
- Number of Hidden Layer: 1
- Accuracy For Experimental Model 1: 9.58 %




```
Accuracy For Experimental Model 1: (9.58) %
```

Experiment Model 2:

- Initializing the weights randomly between -10 and 10
- Training the model and testing for alpha = 0.001
- Using Sigmoid function at Hidden Layer and Softmax Function at Output Layer
- Number of Hidden Layer: 1
- Accuracy For Experimental Model 2: 16.72%




```
Accuracy For Experimental Model 2: (16.72) %
```

Conclusion:

In a neural network model, the hyperparameters play a major role in defining the performance of the model. The above experiment shows the change in accuracy in handwritten digit recognition problem changes as the number of hidden layers, learning rate, step count, hypermeter functions change.