ML Assignment-2 : Neural Network

(for SPAM classification)

NOTE :-

1. Error was calculated at 5, 10, 20, 30, 50, 100, 200, 300, 500, 1000, 2000 till 15000 iterations.
2. Error was not calculated after each iteration to reduce time consumption.
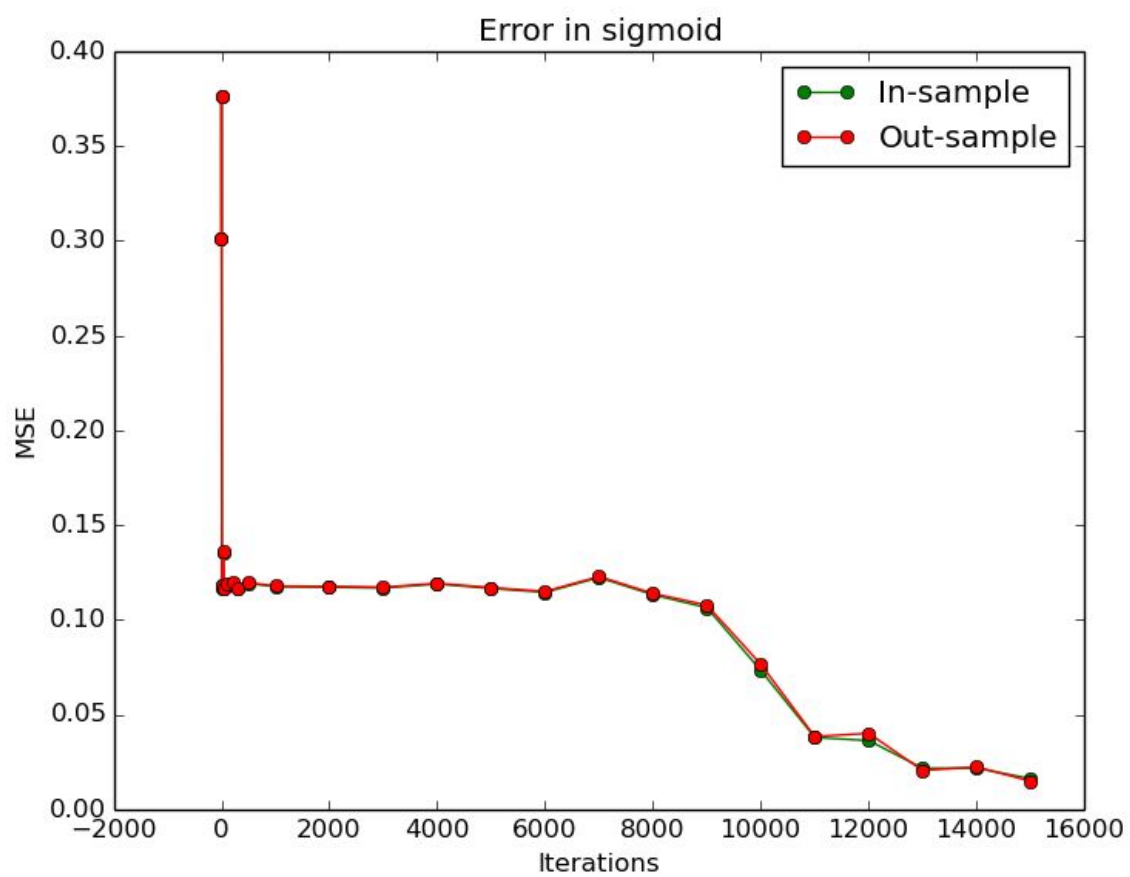
PART A1: Sigmoid function as activation function
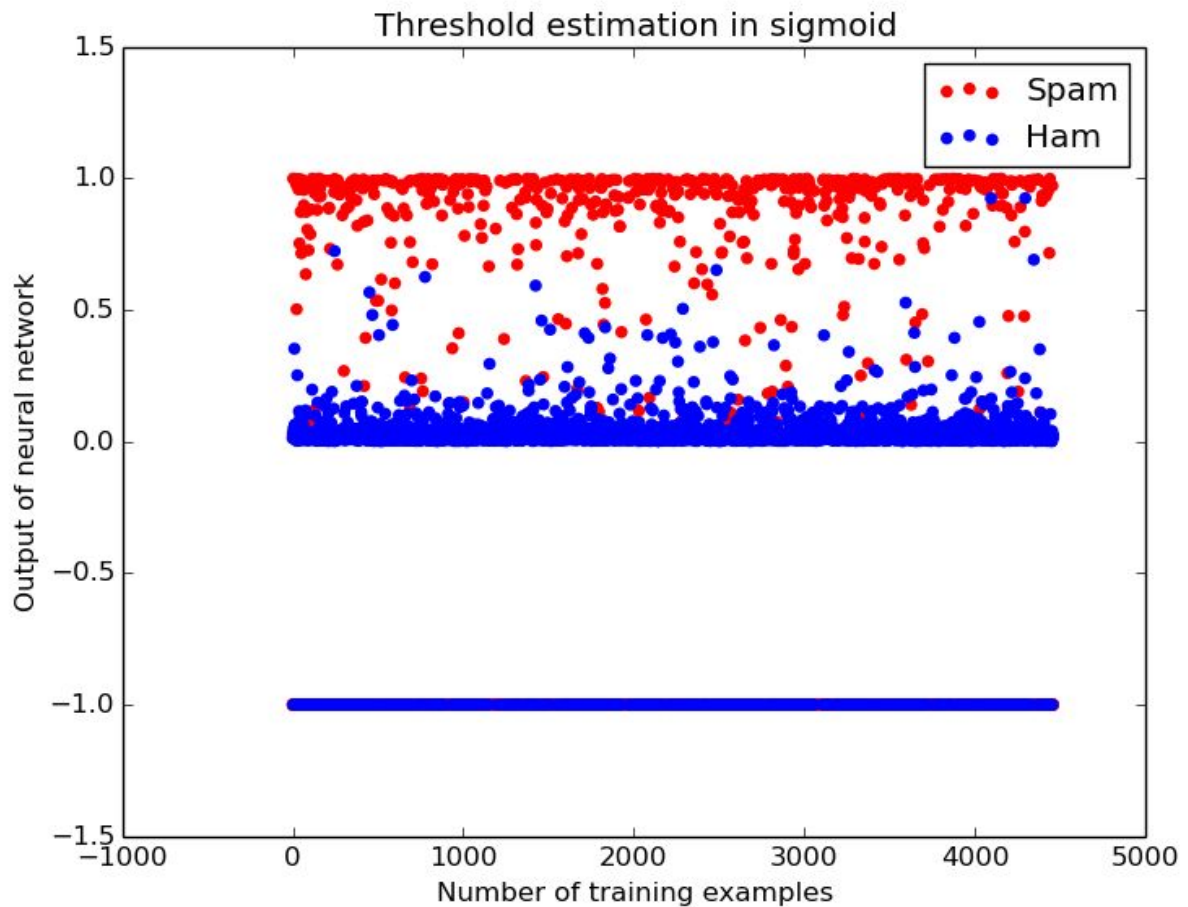
Details :-

1. Training set contains randomly chosen 80% of training examples from the dataset and test set contains remaining 20% of training examples from the dataset.
2. Training examples were distributed to training set and test set in a way which will ensure that ratio of ham examples to spam examples in both training and test set is equal.
3. Values of some parameters were taken as follows :-
    a. Learning rate = 0.1
    b. Threshold = 0.3
4. There are 2 hidden layers and an output layer in this neural network ;
    a. 1st hidden layer has 100 neurons.
    b. 2nd hidden layer has 50 neurons.
    c. Output layer has 1 neuron.

Results :-

1. Optimal number of iterations = 15000
2. Performance on test set :-
    a. Accuracy = 98.74%
    b. Recall = 94%
    c. Precision = 96.57%
    d. F1 score = 95.27%

Plots :-

Threshold estimation in sigmoid

PART A2: Hyperbolic tangent function as activation function
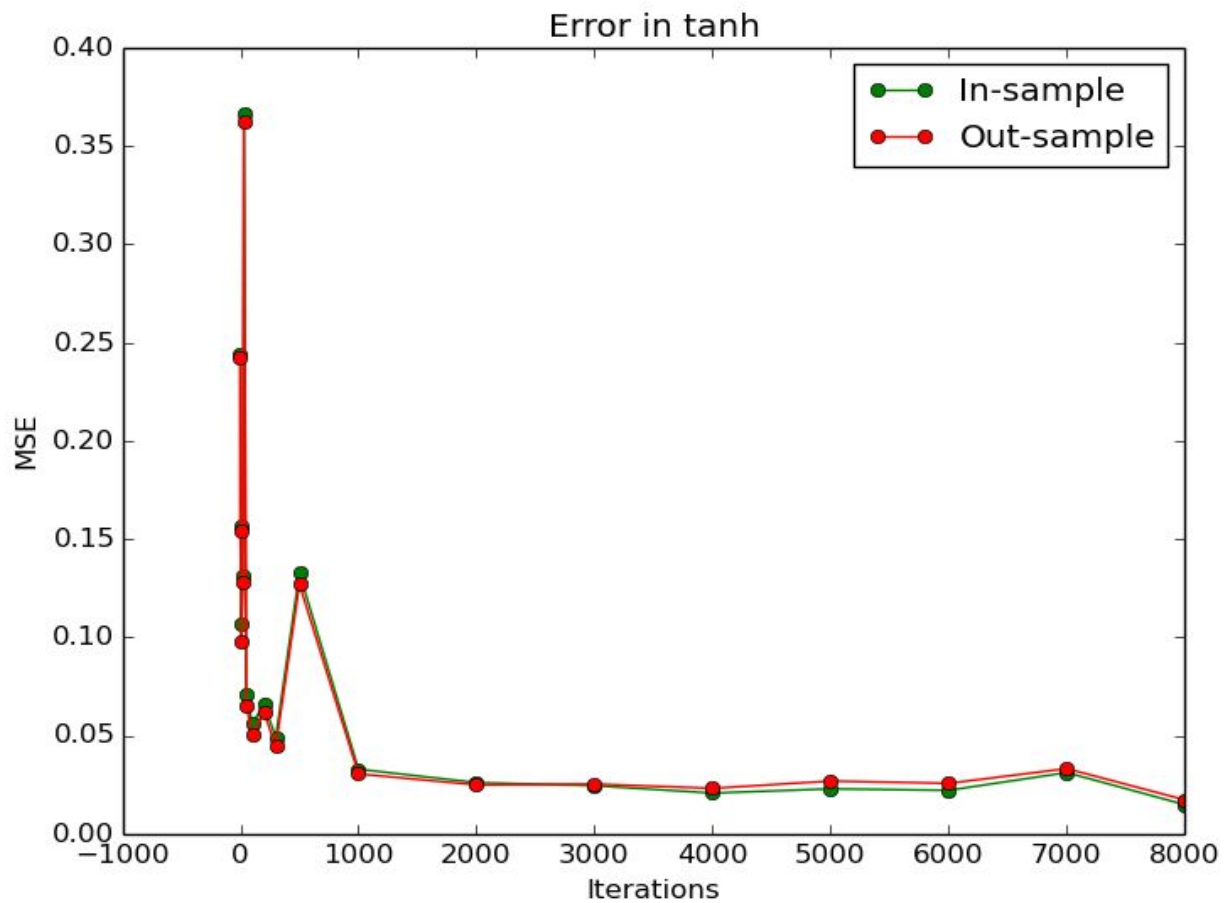
Details :-

1. Training set contains randomly chosen 80% of training examples from the dataset and test set contains remaining 20% of training examples from the dataset.
2. Training examples were distributed to training set and test set in a way which will ensure that ratio of ham examples to spam examples in both training and test set is equal.
3. Values of some parameters were taken as follows :-
    a. Learning rate = 0.1
    b. Threshold = 0.5
4. There are 2 hidden layers and an output layer in this neural network ;
    a. 1st hidden layer has 100 neurons.
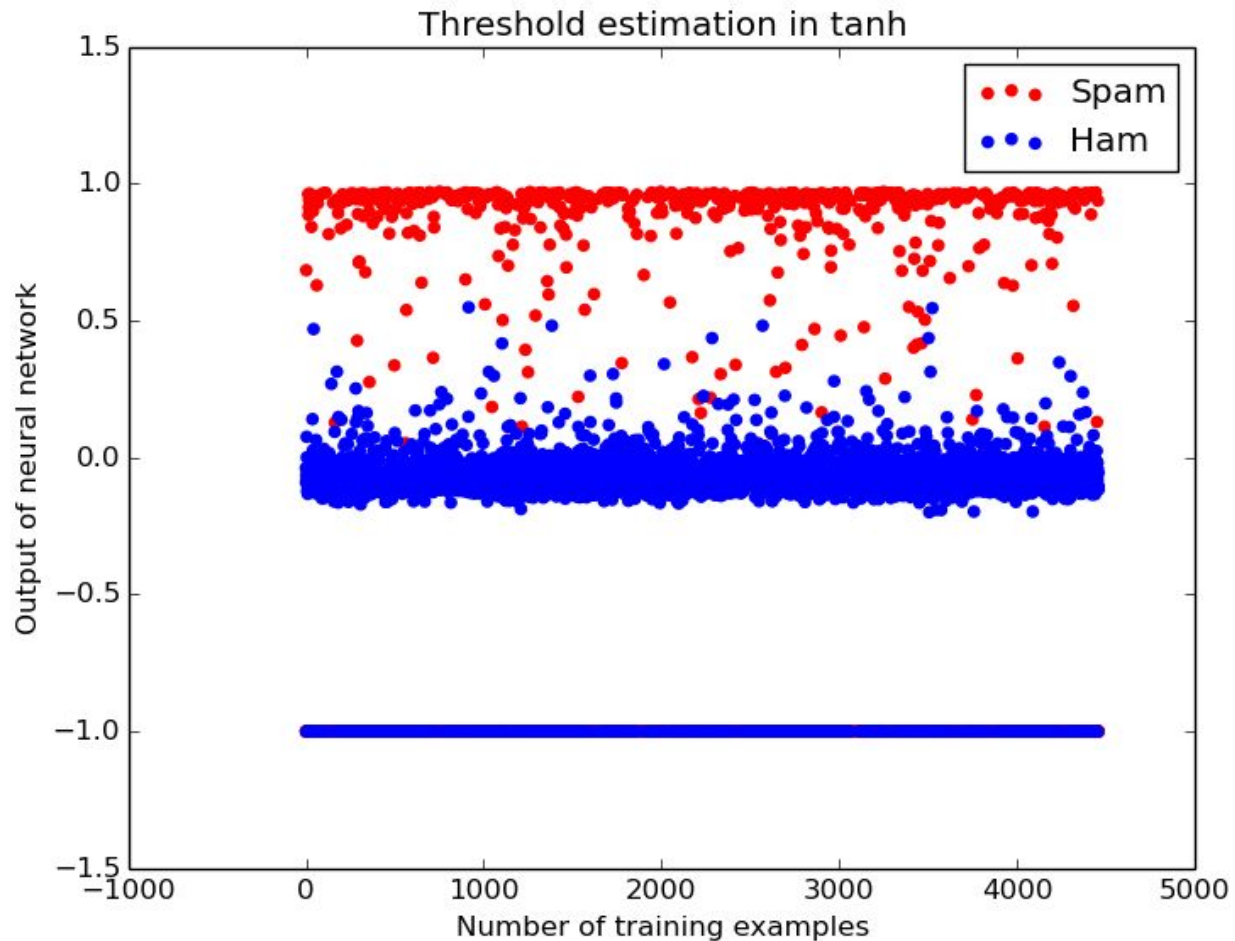   b. 2nd hidden layer has 50 neurons.

     c. Output layer has 1 neuron.

Results :-
1. Optimal number of iterations = 8000
2. Performance on test set :-
     a. Accuracy = 98.74%
     b. Recall = 92%
     c. Precision = 98.57%
     d. F1 score = 95.17%

Plots :-

Threshold estimation in tanh

PART B: Sigmoid with softmax function as activation function

Details :-

1. Training set contains randomly chosen 80% of training examples from the dataset and test set contains remaining 20% of training examples from the dataset.
2. Training examples were distributed to training set and test set in a way which will ensure that ratio of ham examples to spam examples in both training and test set is equal.
3. Values of some parameters were taken as follows :-
   a. Learning = 0.1
   b. Threshold = 0.4
4. There are 2 hidden layers and an output layer in this neural network ;
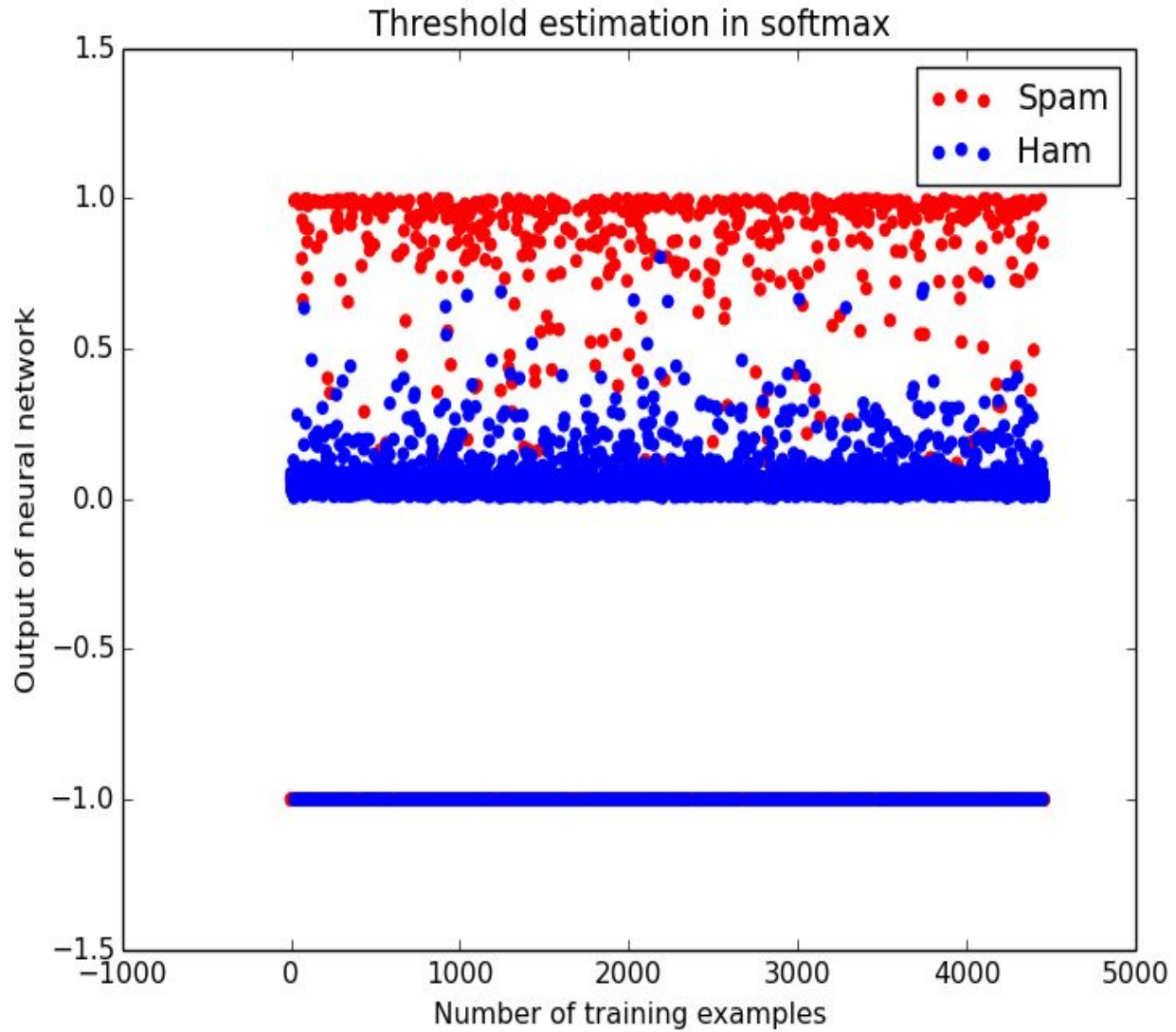   a. 1st hidden layer has 100 neurons.
   b. 2nd hidden layer has 50 neurons.

c. Output layer has 2 neurons with softmax as activation function.

Results :-
1. Optimal number of iterations = 17000
2. Performance on test set :-
    a. Accuracy = 97.49%
    b. Recall = 90%
    c. Precision = 91.22%
    d. F1 score = 90.6%

Plots :-

Threshold estimation in softmax
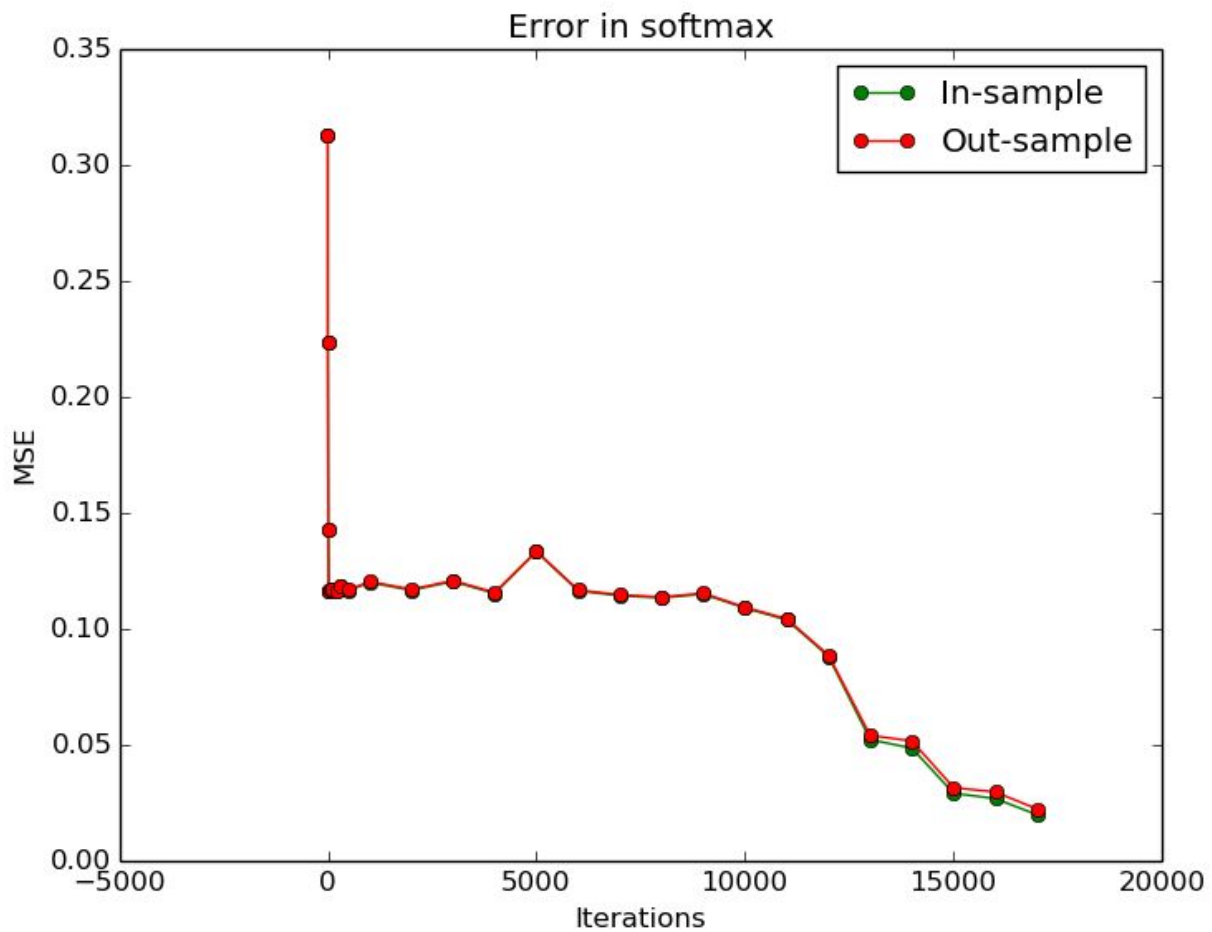
Conclusion :-

In my opinion the architecture with hyperbolic tangent as activation function performs the best because it requires less number of iterations to reach minimal error point and segregates spam from ham better than other architectures.