ML Assignment-3 Report
MT18117

**1. MNIST Handwritten image classification using Neural Network:**

I used Multi layer perceptron to train and test the MNIST images. In this, implemented using 3 layers and 5 layers Neural Network using Sigmoid and ReLu activation function in hidden layer and Softmax function in output layer.

In output layer, we used softmax function which gives probability of each image by all 10 neuron units and all sums to 1. So ,I used Cross-Entroy loss along with One-Hot-Encoding of actual label(Y).

Our input layer is of 784 neuron units and output layer is of 10 neuron units.

a) Sigmoid 3 Layers:

Implemented Sigmoid function as activation function for hidden layer(size=100). Used Gradient Descent algorithm for weight vector updation in backpropagation.
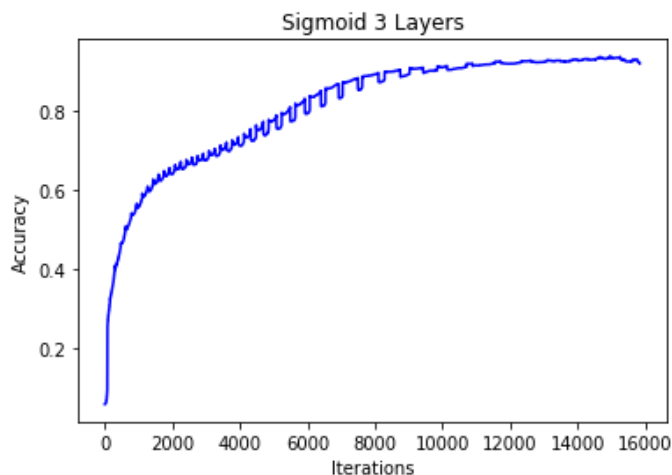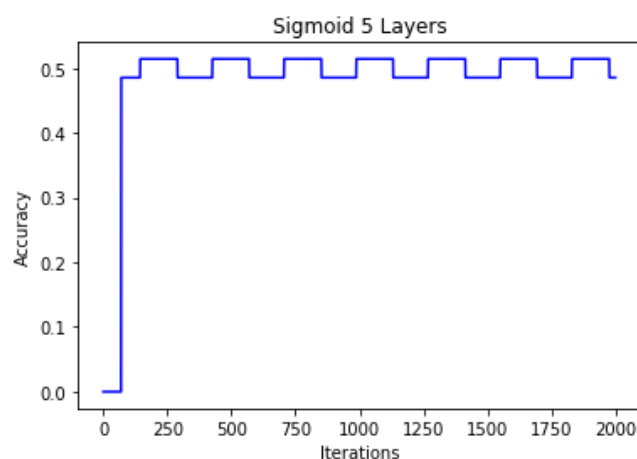

Here for every iteration accuracy is increasing and sometimes decreasing but by increasing iterations it is Overfitting.
Used learning rate (alpha)=0.01 and maximum iterations=8000

Sigmoid 5 Layers:

Implemented Sigmoid function as activation function for hidden layer(size=100,50,50) i.e., 3 Hidden layers. Used Gradient Descent algorithm for weight vector updation in backpropagation.

Here for every iteration accuracy is fluctuating and after some iterations it is constant.
Used learning rate (alpha)=0.01 and maximum iterations=

ReLu 3 Layers:

Implemented ReLu function as activation function for hidden layer(size=100) i.e., 1 Hidden layer.
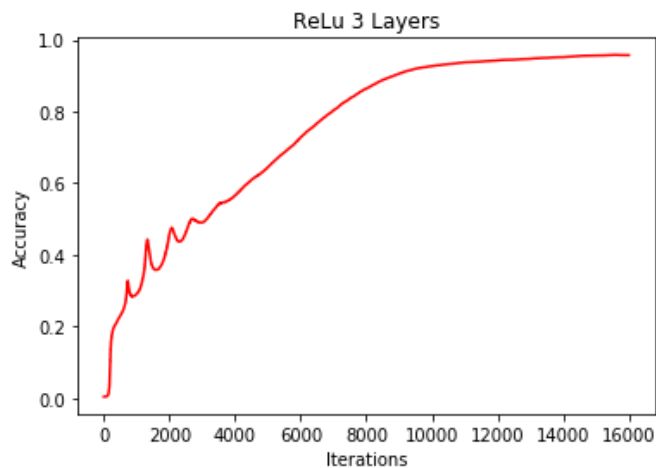

Here for some iteration accuracy is increasing and decreasing, and after some iterations it is
constantly increasing. By increase in iterations, it is overfitting.
Used learning rate (alpha)=0.01 and maximum iterations =

ReLu 5 Layers:

Implemented ReLu function as activation function for hidden layer(size=100,50,50) i.e., 3 Hidden
layers.

Here this some iteration accuracy is increasing and suddenly decreased, and after some iterations it
is constantly decreasing. By increase in iterations, it is Underfitting.
Used learning rate (alpha)=0.01 and maximum iterations =
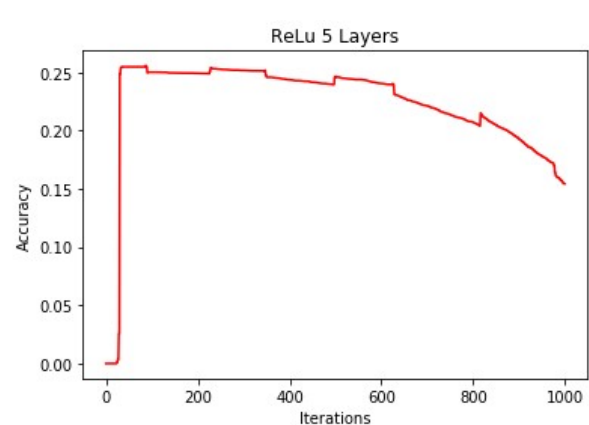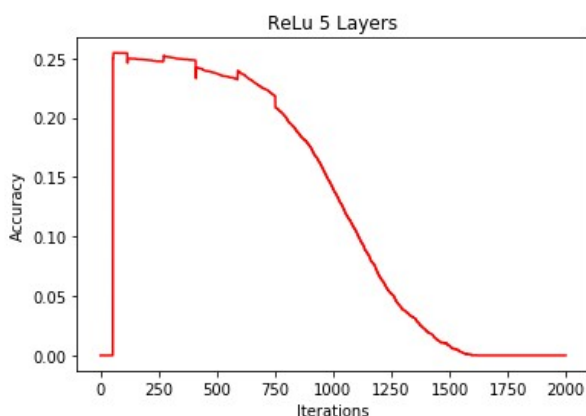
Implemented using Sklearn:

By using Sklearn, It is giving very good accuracy i.e., overfitting the training data and also giving very good accuracy for test data. (taken parameter solver= 'lbfgs' and alpha=0.00001).

I used activation function as 'logistic' for Sigmoid activation function for hidden layers and 'relu' for ReLu activation function. It uses Softmax function at output for multiclass labels(MNIST dataset has 10 classes).

		3-Layers	5-Layers
Sigmoid	Train accuracy	0.936	0.511
	Test Accuracy		0.4991231146965977
	Iterations	7933	1000
ReLu	Train accuracy	0.966	0.251
	Test Accuracy	0.9540512101017187	0.1648544370396352
	Iterations	8000	500
Sklearn sigmoid and ReLu	Train accuracy(Sigmoid)	1	0.9997368421052631
	Test Accuracy	0.9800070150824273	0.9793055068397054
	Train accuracy(Relu)	1	1
	Test Accuracy	0.9891266222378113	.9880743598737285

2. AlexNet using Sklearn:

In this, I implemented by extracting the features of CIFAR images from pertained AlexNet model from layer 8. By using these features I have trained the model using Linear SVM and got an accuracy of 0.4(because I have taken less no.of images because training is taking time) and constructed the Confusion matrix and ROC Curve.