

## Homework 5 Minji Kim

(a). For Linear regression model, we use the model loss  $l(y, x) = \log(1 + e^{-yw^Tx})$ . For Linear SVM model, we use  $l(y, x) = \max(0, 1 - yw^Tx)$  without  $L_2$  regularization. We define the class of each loss to use it.

```
class Logistic_Loss(nn.modules.Module):
    def __init__(self):
        super(Logistic_Loss, self).__init__()
    def forward(self, outputs, labels):
        batch_size = outputs.size()[0]
        return torch.sum(torch.log(1+ torch.exp(-(outputs.t()*labels))))/batch_size

class SVM_Loss(nn.modules.Module):
    def __init__(self):
        super(SVM_Loss, self).__init__()
    def forward(self, outputs, labels):
        return torch.sum( torch.clamp(1- outputs.t()*labels, min=0)/batch_size )
```

The general torch optimization steps are as follows.

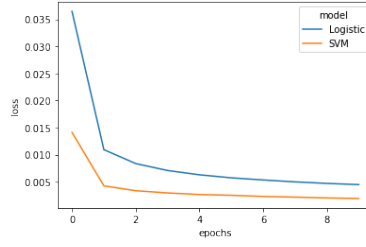
```
for input, target in dataset:
    optimizer.zero_grad()
    output = model(input)
    loss = loss_fn(output, target)
    loss.backward()
    optimizer.step()
```

To implement our own SGD and SGD moment optimizer, I used the code

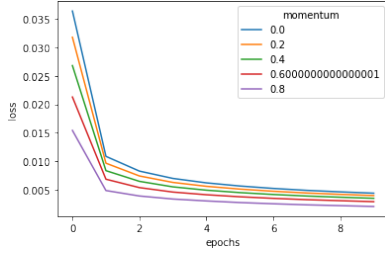
```
sgd_optimizer= torch.optim.SGD(model.parameters(), lr = learning_rate, momentum = momentum).
```

Setting `momentum = 0.0` corresponds to the SGD optimization, while setting momentum parameters between 0 and 1 can do the momentum methods. For each epoch, we report the average of batch losses. All printings are attached at the appendix. We can observe decreasing losses for each epoch after plotting it. For example, below figure shows the loss changes after Logistic regression (Left) and Linear SVM (right) fitting using SGD optimization with learning rate 0.01. Both reported small losses after 2 epochs, indicating that the learning stabilizes quite fast.

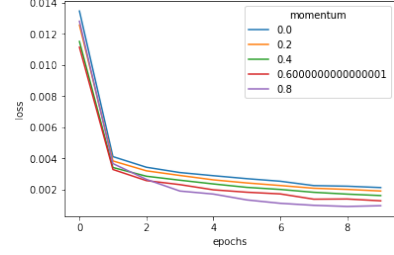
(b). Here we use the learning rate 0.01. The final testing accuracy of trained models are both 99.905434% for Logistic Regression and for Linear SVM model.



(c). We run the code with the following set of momentum:  $\{0.0, 0.2, 0.4, 0.6, 0.8\}$ . We can see from the figure that larger the momentum, slightly lower the loss values. For Logistic Regression model, the accuracy stays the same, while SVM model showed the best accuracy of 95.95% when momentum was 0.4 or 0.6. Overall, the momentum made fast convergence in the optimization process.

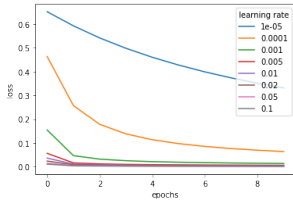


(a) Logistic Regression

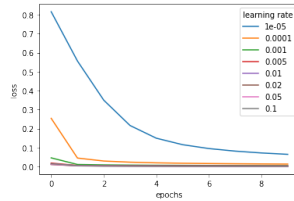


(b) Linear SVM

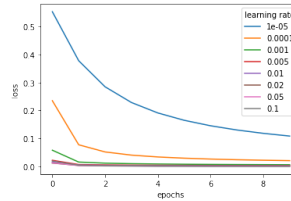
(d). We run the code with following set of learning rate:  $\{1e-5, 1e-4, 1e-3, 0.005, 0.01, 0.02, 0.05, 0.1\}$ . SGD optimization and Momentum with 0.8 optimization were used. Logistic regression result showed the best test accuracy of 99.95% with learning rate (lr) 0.1 using SGD method, and with lr larger than 0.02 using momentum. Linear SVM result showed the best accuracy of 99.95% with lr 0.02 using SGD, and with lr 0.005 using momentum. Below figure shows that with very small learning rate like 0.00001, the learning progress very slowly, while learning rate larger than 0.001 showed sufficiently fast and efficient learning. Overall, the momentum made fast convergence in the optimization process.



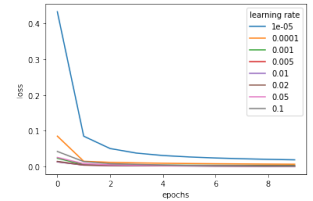
(a) LR SGD



(b) SVM SGD



(c) LR Momentum



(d) SVM Momentum

## APPENDIX

### Output Result of (b)

Learning rate : 0.01, momentum : 0.0

Epoch [1/10], Logistic Regression SGD:[average loss 0.0365]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.0109]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.0083]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.0071]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.0063]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0057]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0053]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0050]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.0047]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.0045]  
Accuracy on the test images: 99.905434 %

Learning rate : 0.01, momentum : 0.0

Epoch [1/10], Linear SVM SGD:[average loss 0.0141]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0042]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0033]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0029]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0026]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0025]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0023]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0021]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0020]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0019]  
Accuracy on the test images: 99.905434 %

### Output Result of (c):Logistic Regression

Learning rate : 0.01, momentum : 0.0

Epoch [1/10], Logistic Regression SGD:[average loss 0.0363]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.0109]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.0083]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.0070]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.0063]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0057]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0053]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0049]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.0047]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.0044]

Accuracy on the test images: 99.905434 %

Learning rate : 0.01, momentum : 0.2

Epoch [1/10], Logistic Regression SGD:[average loss 0.0317]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.0097]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.0075]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.0063]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.0056]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0052]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0048]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0045]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.0042]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.0040]

Accuracy on the test images: 99.905434 %

Learning rate : 0.01, momentum : 0.4

Epoch [1/10], Logistic Regression SGD:[average loss 0.0268]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.0084]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.0065]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.0056]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.0050]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0045]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0042]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0039]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.0037]

Epoch [10/10], Logistic Regression SGD:[average loss 0.0035]

Accuracy on the test images: 99.905434 %

Learning rate : 0.01, momentum : 0.6

Epoch [1/10], Logistic Regression SGD:[average loss 0.0213]

Epoch [2/10], Logistic Regression SGD:[average loss 0.0069]

Epoch [3/10], Logistic Regression SGD:[average loss 0.0054]

Epoch [4/10], Logistic Regression SGD:[average loss 0.0046]

Epoch [5/10], Logistic Regression SGD:[average loss 0.0042]

Epoch [6/10], Logistic Regression SGD:[average loss 0.0038]

Epoch [7/10], Logistic Regression SGD:[average loss 0.0035]

Epoch [8/10], Logistic Regression SGD:[average loss 0.0033]

Epoch [9/10], Logistic Regression SGD:[average loss 0.0031]

Epoch [10/10], Logistic Regression SGD:[average loss 0.0030]

Accuracy on the test images: 99.905434 %

Learning rate : 0.01, momentum : 0.8

Epoch [1/10], Logistic Regression SGD:[average loss 0.0154]

Epoch [2/10], Logistic Regression SGD:[average loss 0.0049]

Epoch [3/10], Logistic Regression SGD:[average loss 0.0040]

Epoch [4/10], Logistic Regression SGD:[average loss 0.0034]

Epoch [5/10], Logistic Regression SGD:[average loss 0.0031]

Epoch [6/10], Logistic Regression SGD:[average loss 0.0028]

Epoch [7/10], Logistic Regression SGD:[average loss 0.0026]

Epoch [8/10], Logistic Regression SGD:[average loss 0.0024]

Epoch [9/10], Logistic Regression SGD:[average loss 0.0023]

Epoch [10/10], Logistic Regression SGD:[average loss 0.0021]

Accuracy on the test images: 99.905434 %

### Output Result of (c):Linear SVM

Learning rate : 0.01, momentum : 0.0

Epoch [1/10], Linear SVM SGD:[average loss 0.0135]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0041]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0034]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0031]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0029]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0027]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0025]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0022]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0022]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0021]

Accuracy on the test images: 99.905434 %

Learning rate : 0.01, momentum : 0.2

Epoch [1/10], Linear SVM SGD:[average loss 0.0126]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0038]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0032]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0029]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0026]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0024]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0023]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0021]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0020]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0019]

Accuracy on the test images: 99.905434 %

Learning rate : 0.01, momentum : 0.4

Epoch [1/10], Linear SVM SGD:[average loss 0.0115]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0034]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0029]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0026]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0024]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0021]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0020]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0018]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0017]

Epoch [10/10], Linear SVM SGD:[average loss 0.0016]

Accuracy on the test images: 99.952721 %

Learning rate : 0.01, momentum : 0.6

Epoch [1/10], Linear SVM SGD:[average loss 0.0111]

Epoch [2/10], Linear SVM SGD:[average loss 0.0033]

Epoch [3/10], Linear SVM SGD:[average loss 0.0026]

Epoch [4/10], Linear SVM SGD:[average loss 0.0023]

Epoch [5/10], Linear SVM SGD:[average loss 0.0020]

Epoch [6/10], Linear SVM SGD:[average loss 0.0018]

Epoch [7/10], Linear SVM SGD:[average loss 0.0017]

Epoch [8/10], Linear SVM SGD:[average loss 0.0014]

Epoch [9/10], Linear SVM SGD:[average loss 0.0014]

Epoch [10/10], Linear SVM SGD:[average loss 0.0013]

Accuracy on the test images: 99.952721 %

Learning rate : 0.01, momentum : 0.8

Epoch [1/10], Linear SVM SGD:[average loss 0.0128]

Epoch [2/10], Linear SVM SGD:[average loss 0.0037]

Epoch [3/10], Linear SVM SGD:[average loss 0.0026]

Epoch [4/10], Linear SVM SGD:[average loss 0.0019]

Epoch [5/10], Linear SVM SGD:[average loss 0.0017]

Epoch [6/10], Linear SVM SGD:[average loss 0.0013]

Epoch [7/10], Linear SVM SGD:[average loss 0.0011]

Epoch [8/10], Linear SVM SGD:[average loss 0.0010]

Epoch [9/10], Linear SVM SGD:[average loss 0.0009]

Epoch [10/10], Linear SVM SGD:[average loss 0.0010]

Accuracy on the test images: 99.905434 %

## Output Result of (d):Logistic Regression

Learning rate : 1e-05, momentum : 0.0

Epoch [1/10], Logistic Regression SGD:[average loss 0.6512]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.5918]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.5410]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.4975]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.4598]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.4272]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.3985]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.3733]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.3510]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.3311]

Accuracy on the test images: 98.345154 %

Learning rate : 0.0001, momentum : 0.0

Epoch [1/10], Logistic Regression SGD:[average loss 0.4629]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.2563]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.1781]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.1379]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.1135]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0971]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0853]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0764]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.0694]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.0637]

Accuracy on the test images: 99.716309 %

Learning rate : 0.001, momentum : 0.0

Epoch [1/10], Logistic Regression SGD:[average loss 0.1544]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.0465]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.0318]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.0253]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.0215]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0189]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0171]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0157]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.0146]



Epoch [10/10], Logistic Regression SGD:[average loss 0.0137]  
 Accuracy on the test images: 99.858162 %  
 Learning rate : 0.005, momentum : 0.0

Epoch [1/10], Logistic Regression SGD:[average loss 0.0562]  
 Epoch [2/10], Logistic Regression SGD:[average loss 0.0160]  
 Epoch [3/10], Logistic Regression SGD:[average loss 0.0119]  
 Epoch [4/10], Logistic Regression SGD:[average loss 0.0099]  
 Epoch [5/10], Logistic Regression SGD:[average loss 0.0087]  
 Epoch [6/10], Logistic Regression SGD:[average loss 0.0079]  
 Epoch [7/10], Logistic Regression SGD:[average loss 0.0073]  
 Epoch [8/10], Logistic Regression SGD:[average loss 0.0068]  
 Epoch [9/10], Logistic Regression SGD:[average loss 0.0064]  
 Epoch [10/10], Logistic Regression SGD:[average loss 0.0061]  
 Accuracy on the test images: 99.905434 %  
 Learning rate : 0.01, momentum : 0.0

Epoch [1/10], Logistic Regression SGD:[average loss 0.0363]  
 Epoch [2/10], Logistic Regression SGD:[average loss 0.0109]  
 Epoch [3/10], Logistic Regression SGD:[average loss 0.0083]  
 Epoch [4/10], Logistic Regression SGD:[average loss 0.0070]  
 Epoch [5/10], Logistic Regression SGD:[average loss 0.0063]  
 Epoch [6/10], Logistic Regression SGD:[average loss 0.0057]  
 Epoch [7/10], Logistic Regression SGD:[average loss 0.0053]  
 Epoch [8/10], Logistic Regression SGD:[average loss 0.0049]  
 Epoch [9/10], Logistic Regression SGD:[average loss 0.0047]  
 Epoch [10/10], Logistic Regression SGD:[average loss 0.0044]  
 Accuracy on the test images: 99.905434 %  
 Learning rate : 0.02, momentum : 0.0

Epoch [1/10], Logistic Regression SGD:[average loss 0.0239]  
 Epoch [2/10], Logistic Regression SGD:[average loss 0.0077]  
 Epoch [3/10], Logistic Regression SGD:[average loss 0.0060]  
 Epoch [4/10], Logistic Regression SGD:[average loss 0.0051]  
 Epoch [5/10], Logistic Regression SGD:[average loss 0.0046]  
 Epoch [6/10], Logistic Regression SGD:[average loss 0.0042]  
 Epoch [7/10], Logistic Regression SGD:[average loss 0.0039]  
 Epoch [8/10], Logistic Regression SGD:[average loss 0.0036]

Epoch [9/10], Logistic Regression SGD:[average loss 0.0034]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.0033]  
Accuracy on the test images: 99.905434 %  
Learning rate : 0.05, momentum : 0.0  
Epoch [1/10], Logistic Regression SGD:[average loss 0.0144]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.0051]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.0040]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.0034]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.0031]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0028]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0026]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0024]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.0023]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.0021]  
Accuracy on the test images: 99.905434 %  
Learning rate : 0.1, momentum : 0.0  
Epoch [1/10], Logistic Regression SGD:[average loss 0.0107]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.0037]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.0029]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.0025]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.0022]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0021]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0019]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0017]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.0016]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.0015]  
Accuracy on the test images: 99.952721 %

### Output Result of (d):Linear SVM

Learning rate : 1e-05, momentum : 0.0

Epoch [1/10], Linear SVM SGD:[average loss 0.8160]  
Epoch [2/10], Linear SVM SGD:[average loss 0.5560]  
Epoch [3/10], Linear SVM SGD:[average loss 0.3497]  
Epoch [4/10], Linear SVM SGD:[average loss 0.2161]  
Epoch [5/10], Linear SVM SGD:[average loss 0.1491]  
Epoch [6/10], Linear SVM SGD:[average loss 0.1150]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0945]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0810]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0713]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0641]  
Accuracy on the test images: 99.101654 %

Learning rate : 0.0001, momentum : 0.0

Epoch [1/10], Linear SVM SGD:[average loss 0.2540]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0444]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0288]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0227]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0193]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0171]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0156]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0144]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0135]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0128]  
Accuracy on the test images: 99.763596 %

Learning rate : 0.001, momentum : 0.0

Epoch [1/10], Linear SVM SGD:[average loss 0.0454]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0105]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0084]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0072]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0064]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0058]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0054]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0050]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0047]

Epoch [10/10], Linear SVM SGD:[average loss 0.0045]  
Accuracy on the test images: 99.905434 %  
Learning rate : 0.005, momentum : 0.0  
Epoch [1/10], Linear SVM SGD:[average loss 0.0180]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0053]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0041]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0036]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0034]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0033]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0031]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0029]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0028]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0027]  
Accuracy on the test images: 99.905434 %  
Learning rate : 0.01, momentum : 0.0  
Epoch [1/10], Linear SVM SGD:[average loss 0.0135]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0041]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0034]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0031]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0029]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0027]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0025]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0022]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0022]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0021]  
Accuracy on the test images: 99.905434 %  
Learning rate : 0.02, momentum : 0.0  
Epoch [1/10], Linear SVM SGD:[average loss 0.0104]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0033]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0027]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0025]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0022]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0020]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0018]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0016]

Epoch [9/10], Linear SVM SGD:[average loss 0.0015]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0015]  
Accuracy on the test images: 99.952721 %  
Learning rate : 0.05, momentum : 0.0  
Epoch [1/10], Linear SVM SGD:[average loss 0.0112]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0033]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0023]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0018]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0016]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0013]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0012]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0009]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0009]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0008]  
Accuracy on the test images: 99.905434 %  
Learning rate : 0.1, momentum : 0.0  
Epoch [1/10], Linear SVM SGD:[average loss 0.0144]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0042]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0025]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0019]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0015]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0012]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0010]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0007]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0007]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0006]  
Accuracy on the test images: 99.905434 %

### Output Result of (d):Logistic Regression with Momentum

Learning rate : 1e-05, momentum : 0.8

Epoch [1/10], Logistic Regression SGD:[average loss 0.5527]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.3776]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.2845]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.2284]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.1911]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.1648]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.1453]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.1302]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.1182]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.1085]  
Accuracy on the test images: 99.716309 %

Learning rate : 0.0001, momentum : 0.8

Epoch [1/10], Logistic Regression SGD:[average loss 0.2348]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.0773]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.0518]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.0404]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.0338]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0295]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0264]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0240]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.0221]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.0206]  
Accuracy on the test images: 99.810875 %

Learning rate : 0.001, momentum : 0.8

Epoch [1/10], Logistic Regression SGD:[average loss 0.0580]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.0156]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.0117]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.0098]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.0087]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0079]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0072]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0068]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.0064]

Epoch [10/10], Logistic Regression SGD:[average loss 0.0061]  
Accuracy on the test images: 99.905434 %  
Learning rate : 0.005, momentum : 0.8  
Epoch [1/10], Logistic Regression SGD:[average loss 0.0220]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.0067]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.0054]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.0046]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.0042]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0038]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0035]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0033]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.0031]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.0030]  
Accuracy on the test images: 99.905434 %  
Learning rate : 0.01, momentum : 0.8  
Epoch [1/10], Logistic Regression SGD:[average loss 0.0154]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.0049]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.0040]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.0034]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.0031]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0028]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0026]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0024]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.0023]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.0021]  
Accuracy on the test images: 99.905434 %  
Learning rate : 0.02, momentum : 0.8  
Epoch [1/10], Logistic Regression SGD:[average loss 0.0121]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.0036]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.0029]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.0025]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.0022]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0020]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0018]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0017]

Epoch [9/10], Logistic Regression SGD:[average loss 0.0015]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.0014]  
Accuracy on the test images: 99.952721 %  
Learning rate : 0.05, momentum : 0.8  
Epoch [1/10], Logistic Regression SGD:[average loss 0.0128]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.0031]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.0019]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.0015]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.0011]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0011]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0009]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0008]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.0007]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.0006]  
Accuracy on the test images: 99.952721 %  
Learning rate : 0.1, momentum : 0.8  
Epoch [1/10], Logistic Regression SGD:[average loss 0.0193]  
Epoch [2/10], Logistic Regression SGD:[average loss 0.0043]  
Epoch [3/10], Logistic Regression SGD:[average loss 0.0021]  
Epoch [4/10], Logistic Regression SGD:[average loss 0.0016]  
Epoch [5/10], Logistic Regression SGD:[average loss 0.0008]  
Epoch [6/10], Logistic Regression SGD:[average loss 0.0006]  
Epoch [7/10], Logistic Regression SGD:[average loss 0.0003]  
Epoch [8/10], Logistic Regression SGD:[average loss 0.0003]  
Epoch [9/10], Logistic Regression SGD:[average loss 0.0002]  
Epoch [10/10], Logistic Regression SGD:[average loss 0.0002]  
Accuracy on the test images: 99.952721 %



### Output Result of (d):Linear SVM with Momentum

Learning rate : 1e-05, momentum : 0.8

Epoch [1/10], Linear SVM SGD:[average loss 0.4331]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0846]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0501]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0373]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0306]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0265]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0236]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0215]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0198]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0185]  
Accuracy on the test images: 99.763596 %

Learning rate : 0.0001, momentum : 0.8

Epoch [1/10], Linear SVM SGD:[average loss 0.0849]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0145]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0113]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0099]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0088]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0081]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0075]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0070]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0066]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0062]  
Accuracy on the test images: 99.858162 %

Learning rate : 0.001, momentum : 0.8

Epoch [1/10], Linear SVM SGD:[average loss 0.0229]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0052]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0041]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0037]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0034]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0032]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0031]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0029]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0028]

Epoch [10/10], Linear SVM SGD:[average loss 0.0027]  
 Accuracy on the test images: 99.905434 %  
 Learning rate : 0.005, momentum : 0.8

Epoch [1/10], Linear SVM SGD:[average loss 0.0130]  
 Epoch [2/10], Linear SVM SGD:[average loss 0.0036]  
 Epoch [3/10], Linear SVM SGD:[average loss 0.0029]  
 Epoch [4/10], Linear SVM SGD:[average loss 0.0024]  
 Epoch [5/10], Linear SVM SGD:[average loss 0.0021]  
 Epoch [6/10], Linear SVM SGD:[average loss 0.0019]  
 Epoch [7/10], Linear SVM SGD:[average loss 0.0017]  
 Epoch [8/10], Linear SVM SGD:[average loss 0.0014]  
 Epoch [9/10], Linear SVM SGD:[average loss 0.0014]  
 Epoch [10/10], Linear SVM SGD:[average loss 0.0013]  
 Accuracy on the test images: 99.952721 %  
 Learning rate : 0.01, momentum : 0.8

Epoch [1/10], Linear SVM SGD:[average loss 0.0128]  
 Epoch [2/10], Linear SVM SGD:[average loss 0.0037]  
 Epoch [3/10], Linear SVM SGD:[average loss 0.0026]  
 Epoch [4/10], Linear SVM SGD:[average loss 0.0019]  
 Epoch [5/10], Linear SVM SGD:[average loss 0.0017]  
 Epoch [6/10], Linear SVM SGD:[average loss 0.0013]  
 Epoch [7/10], Linear SVM SGD:[average loss 0.0011]  
 Epoch [8/10], Linear SVM SGD:[average loss 0.0010]  
 Epoch [9/10], Linear SVM SGD:[average loss 0.0009]  
 Epoch [10/10], Linear SVM SGD:[average loss 0.0010]  
 Accuracy on the test images: 99.905434 %  
 Learning rate : 0.02, momentum : 0.8

Epoch [1/10], Linear SVM SGD:[average loss 0.0141]  
 Epoch [2/10], Linear SVM SGD:[average loss 0.0040]  
 Epoch [3/10], Linear SVM SGD:[average loss 0.0024]  
 Epoch [4/10], Linear SVM SGD:[average loss 0.0020]  
 Epoch [5/10], Linear SVM SGD:[average loss 0.0015]  
 Epoch [6/10], Linear SVM SGD:[average loss 0.0014]  
 Epoch [7/10], Linear SVM SGD:[average loss 0.0009]  
 Epoch [8/10], Linear SVM SGD:[average loss 0.0007]

Epoch [9/10], Linear SVM SGD:[average loss 0.0008]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0005]  
Accuracy on the test images: 99.858162 %  
Learning rate : 0.05, momentum : 0.8  
Epoch [1/10], Linear SVM SGD:[average loss 0.0252]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0079]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0044]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0033]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0023]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0017]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0012]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0008]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0001]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0003]  
Accuracy on the test images: 99.905434 %  
Learning rate : 0.1, momentum : 0.8  
Epoch [1/10], Linear SVM SGD:[average loss 0.0419]  
Epoch [2/10], Linear SVM SGD:[average loss 0.0136]  
Epoch [3/10], Linear SVM SGD:[average loss 0.0077]  
Epoch [4/10], Linear SVM SGD:[average loss 0.0058]  
Epoch [5/10], Linear SVM SGD:[average loss 0.0040]  
Epoch [6/10], Linear SVM SGD:[average loss 0.0025]  
Epoch [7/10], Linear SVM SGD:[average loss 0.0017]  
Epoch [8/10], Linear SVM SGD:[average loss 0.0013]  
Epoch [9/10], Linear SVM SGD:[average loss 0.0003]  
Epoch [10/10], Linear SVM SGD:[average loss 0.0004]  
Accuracy on the test images: 99.905434 %