In [1]:

```
import numpy as np
import edf
from time import time
import sys
%matplotlib inline
import matplotlib.pyplot as plt
```

In [2]:

```
traindata = './mnist_data/train.npz'
valdata = './mnist_data/test.npz'

data = np.load(traindata)
t_imgs = np.float32(data['imgs'])/255.
t_labels = np.float32(data['labels'])

data = np.load(valdata)
v_imgs = np.float32(data['imgs'])/255.
v_labels = np.float32(data['labels'])
```

In [3]:

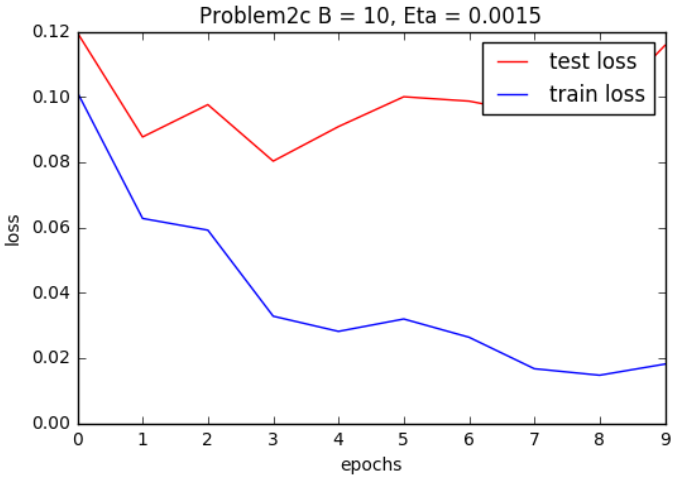
```
# Optimization functions with Adam optimization algo
rithm.
# For details, please see: https://arxiv.org/abs/141
2.6980,
# Please implement this function

# for repeatability
np.random.seed(0)

# some constant used in Adam
_a_b1t=edf.DT(1.0)
_a_b2t=edf.DT(1.0)




batch_array = [10,100]
eta_array = [0.0015]

for i in range(0,2):
    batch = batch_array[i]
    for j in range(0,11): #0 is 0.0015. 1-10 other l
earning rate
        eta = eta_array[0]
        if (j != 0):
            eta += (10**
(-4))*np.random.uniform(-2,2)

        edf.params = []
        edf.components = []

        # Inputs and parameters
        inp = edf.Value()
        lab = edf.Value()

        W1 = edf.Param(edf.xavier((28*28,128)))
        B1 = edf.Param(np.zeros((128)))
        W2 = edf.Param(edf.xavier((128,10)))
        B2 = edf.Param(np.zeros((10)))

        # models
        hidden = edf.RELU(edf.Add(edf.VDot(inp,W1),B
```

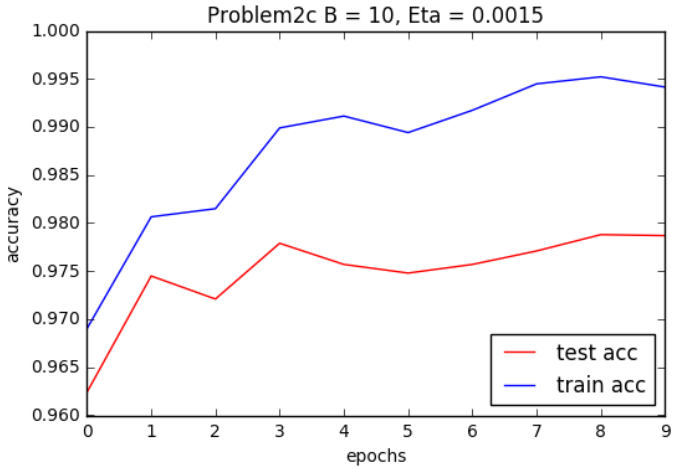
```
1))
        pred = edf.SoftMax(edf.Add(edf.VDot(hidden,W
2),B2))
        loss = edf.LogLoss(edf.Aref(pred,lab))
        acc = edf.Accuracy(pred,lab)


        # evaluate the random performance
        def eval(imgs, labels):

            batches = range(0, len(labels), batch)
            objective = 0
            accuracy = 0
            for k in batches:
                inp.set(imgs[k:k+batch])
                lab.set(labels[k:k+batch])
                edf.Forward()
                objective += np.mean(loss.value)
                accuracy += acc.value

            return accuracy/len(batches),
objective/len(batches)


        accuracy, objective = eval(t_imgs, t_labels)
        print("Random accuracy = %.4f" % accuracy)

        # train loop
        train_loss = []
        train_acc = []
        test_loss = []
        test_acc = []
        ep = 0
        stime = time()
        epoch = 10
        batches = range(0, len(t_labels), batch)

        while ep < epoch:

            # randon shuffle the train data in each
 epoch
            perm = np.random.permutation(len(t_label
```

```
s))
            for k in batches:
                inp.set(t_imgs[perm[k:k+batch]])
                lab.set(t_labels[perm[k:k+batch]])
                edf.Forward()
                edf.Backward(loss)

                b1 = 0.9
                b2 = 0.999
                epVal = 1e-8
                # here, we use Adam algorithm to opt
imize as in problem 2.c
                if 'grad_hist' not in
edf.params[0].__dict__.keys():
                    for p in edf.params:
                        p.grad_hist = edf.DT(0)
                        p.grad_h2 = edf.DT(0)

                # please finish this function
                for p in edf.params:
                    p.grad_hist = b1*p.grad_hist +
(_a_b1t-b1)*p.grad
                    p.grad_h2 = b2*p.grad_h2 + (_a_b
2t-b2)*p.grad**2
                    #_a_b1t =
                    p.value = p.value - (eta*p.grad_
hist)/(np.sqrt(p.grad_h2)+epVal)
                    p.grad = edf.DT(0)

            # evaluate on trainset
            t_acc, t_loss = eval(t_imgs, t_labels)
            print("Epoch %d: train loss = %.4f [%.3f
secs]" % (ep, t_loss,time()-stime))
            train_loss.append(t_loss)
            train_acc.append(t_acc)

            # evaluate on testset
            v_acc, v_loss = eval(v_imgs, v_labels)
            print("test accuracy=%.4f" % v_acc)
            test_loss.append(v_loss)
            test_acc.append(v_acc)
            stime = time()
```

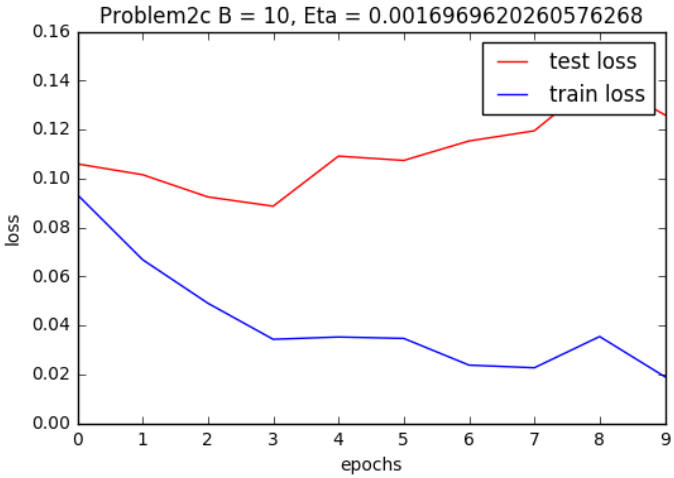
```
            ep += 1

        # plot
        plt.figure(1)
        plt.xlabel("epochs")
        plt.ylabel("loss")
        plt.plot(np.arange(len(test_loss)), test_los
s, color='red')
        plt.plot(np.arange(len(train_loss)), train_l
oss, color='blue')
        plt.legend(['test loss', 'train loss'],
loc='upper right')
        plt.title("Problem2c B = {}, Eta = {}".forma
t(batch,eta))
        plt.show()

        plt.figure(2)
        plt.xlabel("epochs")
        plt.ylabel("accuracy")
        plt.plot(np.arange(len(test_acc)), test_acc,
 color='red')
        plt.plot(np.arange(len(train_acc)), train_ac
c, color='blue')
        plt.legend(['test acc', 'train acc'], loc='l
ower right')
        plt.title("Problem2c B = {}, Eta = {}".forma
t(batch,eta))
        plt.show()
```

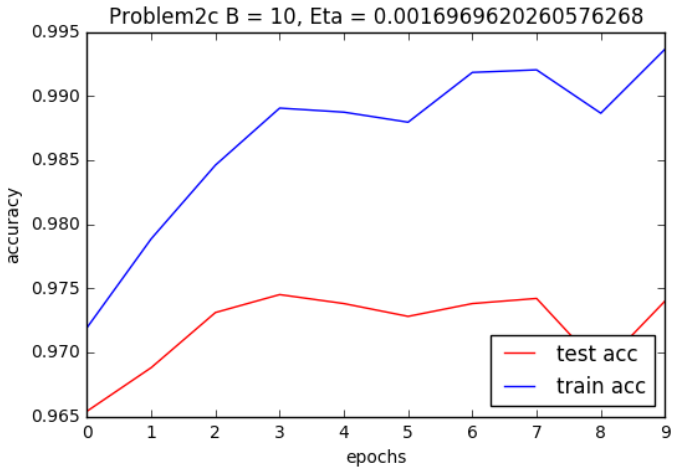
```
Random accuracy = 0.1329
Epoch 0: train loss = 0.1014 [27.351 sec
s]
test accuracy=0.9624
Epoch 1: train loss = 0.0627 [30.399 sec
s]
test accuracy=0.9745
Epoch 2: train loss = 0.0591 [30.318 sec
s]
test accuracy=0.9721
Epoch 3: train loss = 0.0328 [30.874 sec
s]
test accuracy=0.9779
Epoch 4: train loss = 0.0281 [38.596 sec
s]
test accuracy=0.9757
Epoch 5: train loss = 0.0319 [41.681 sec
s]
test accuracy=0.9748
Epoch 6: train loss = 0.0263 [42.578 sec
s]
test accuracy=0.9757
Epoch 7: train loss = 0.0167 [41.875 sec
s]
test accuracy=0.9771
Epoch 8: train loss = 0.0147 [42.419 sec
s]
test accuracy=0.9788
Epoch 9: train loss = 0.0181 [42.325 sec
s]
test accuracy=0.9787
```


Problem2c B = 10, Eta = 0.0015
test loss
train loss
loss
epochs
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
1
2
3
4
5
6
7
8
9

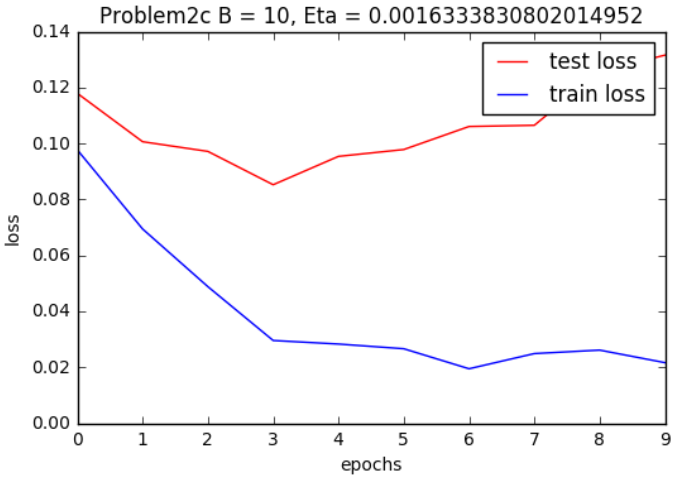

Problem2c B = 10, Eta = 0.0015
test acc
train acc
accuracy
epochs
1.000
0.995
0.990
0.985
0.980
0.975
0.970
0.965
0.960
0
1
2
3
4
5
6
7
8
9

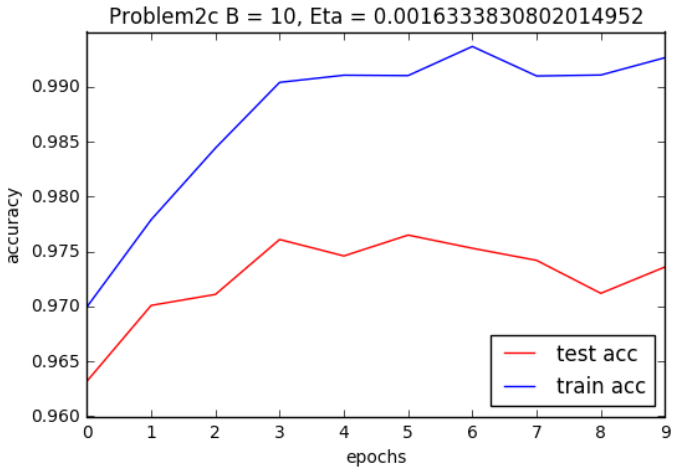
```
Random accuracy = 0.1098
Epoch 0: train loss = 0.0934 [39.491 sec
s]
test accuracy=0.9654
Epoch 1: train loss = 0.0668 [42.088 sec
s]
test accuracy=0.9688
Epoch 2: train loss = 0.0490 [43.247 sec
s]
test accuracy=0.9731
Epoch 3: train loss = 0.0343 [43.229 sec
s]
test accuracy=0.9745
Epoch 4: train loss = 0.0352 [43.785 sec
s]
test accuracy=0.9738
Epoch 5: train loss = 0.0346 [43.701 sec
s]
test accuracy=0.9728
Epoch 6: train loss = 0.0237 [44.364 sec
s]
test accuracy=0.9738
Epoch 7: train loss = 0.0226 [34.875 sec
s]
test accuracy=0.9742
Epoch 8: train loss = 0.0354 [32.348 sec
s]
test accuracy=0.9688
Epoch 9: train loss = 0.0190 [32.879 sec
s]
test accuracy=0.9740
```


Problem2c B = 10, Eta = 0.0016969620260576268
test loss
train loss
loss
epochs
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
1
2
3
4
5
6
7
8
9

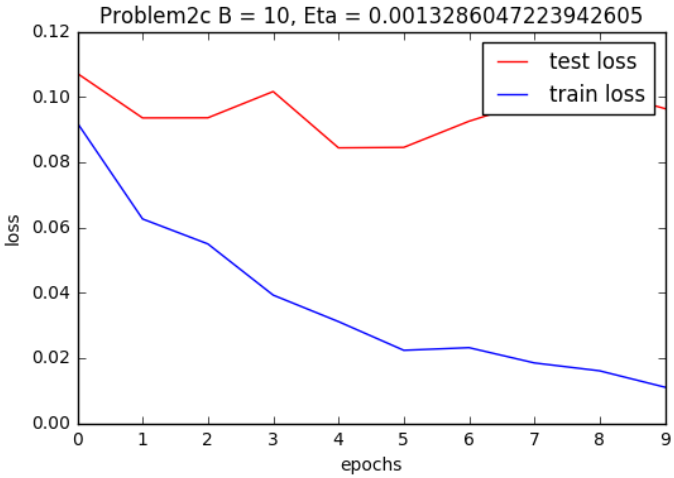

Problem2c B = 10, Eta = 0.0016969620260576268
test acc
train acc
accuracy
epochs
0.995
0.990
0.985
0.980
0.975
0.970
0.965
0
1
2
3
4
5
6
7
8
9

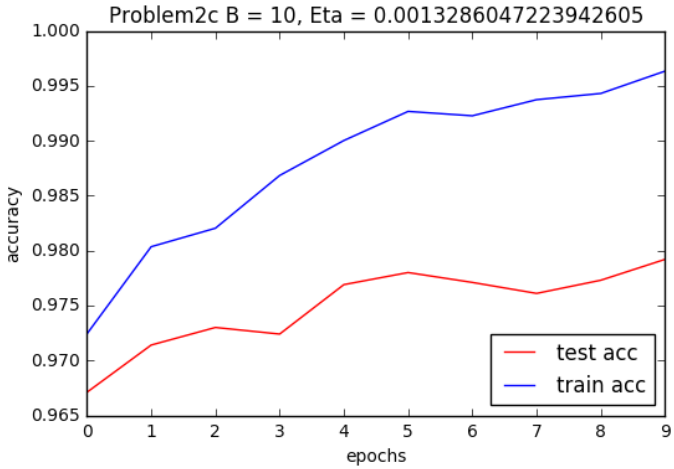
```
Random accuracy = 0.1056
Epoch 0: train loss = 0.0977 [27.908 sec
s]
test accuracy=0.9632
Epoch 1: train loss = 0.0694 [37.068 sec
s]
test accuracy=0.9701
Epoch 2: train loss = 0.0488 [42.517 sec
s]
test accuracy=0.9711
Epoch 3: train loss = 0.0296 [43.090 sec
s]
test accuracy=0.9761
Epoch 4: train loss = 0.0283 [42.438 sec
s]
test accuracy=0.9746
Epoch 5: train loss = 0.0266 [42.467 sec
s]
test accuracy=0.9765
Epoch 6: train loss = 0.0195 [43.345 sec
s]
test accuracy=0.9753
Epoch 7: train loss = 0.0249 [42.987 sec
s]
test accuracy=0.9742
Epoch 8: train loss = 0.0261 [43.905 sec
s]
test accuracy=0.9712
Epoch 9: train loss = 0.0216 [44.516 sec
s]
test accuracy=0.9736
```


Problem2c B = 10, Eta = 0.0016333830802014952
test loss
train loss
loss
epochs
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
1
2
3
4
5
6
7
8
9

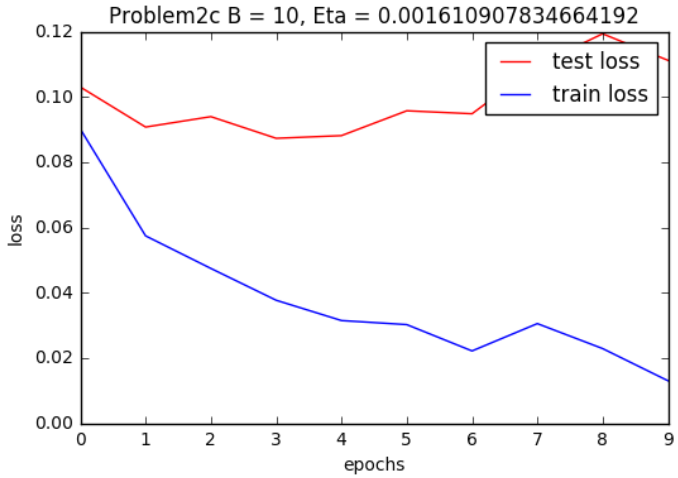

Problem2c B = 10, Eta = 0.0016333830802014952
test acc
train acc
accuracy
epochs
0.990
0.985
0.980
0.975
0.970
0.965
0.960
0
1
2
3
4
5
6
7
8
9

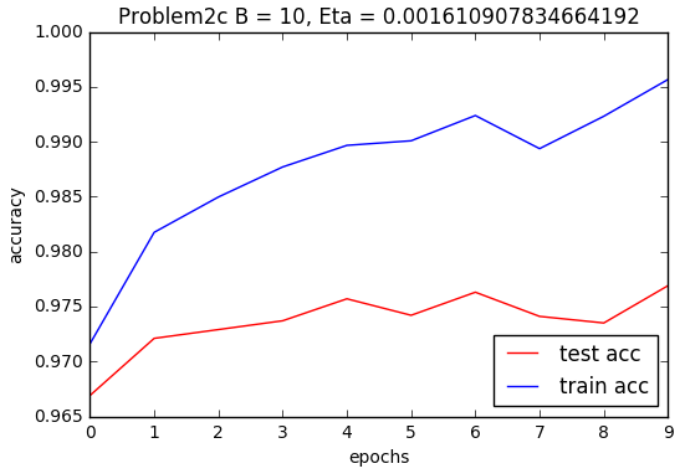
```
Random accuracy = 0.1500
Epoch 0: train loss = 0.0920 [36.548 sec
s]
test accuracy=0.9671
Epoch 1: train loss = 0.0626 [39.754 sec
s]
test accuracy=0.9714
Epoch 2: train loss = 0.0549 [40.735 sec
s]
test accuracy=0.9730
Epoch 3: train loss = 0.0392 [41.343 sec
s]
test accuracy=0.9724
Epoch 4: train loss = 0.0311 [42.593 sec
s]
test accuracy=0.9769
Epoch 5: train loss = 0.0223 [42.886 sec
s]
test accuracy=0.9780
Epoch 6: train loss = 0.0231 [42.992 sec
s]
test accuracy=0.9771
Epoch 7: train loss = 0.0185 [43.396 sec
s]
test accuracy=0.9761
Epoch 8: train loss = 0.0160 [43.933 sec
s]
test accuracy=0.9773
Epoch 9: train loss = 0.0110 [44.570 sec
s]
test accuracy=0.9792
```


Problem2c B = 10, Eta = 0.0013286047223942605
test loss
train loss
loss
epochs
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
1
2
3
4
5
6
7
8
9

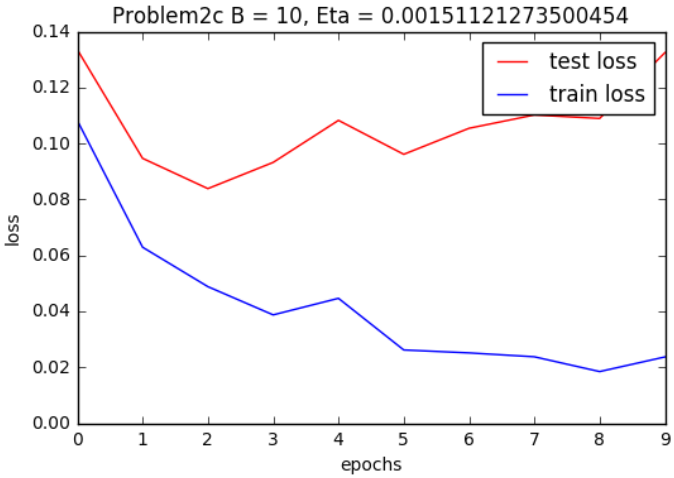

Problem2c B = 10, Eta = 0.0013286047223942605
test acc
train acc
accuracy
epochs
1.000
0.995
0.990
0.985
0.980
0.975
0.970
0.965
0
1
2
3
4
5
6
7
8
9

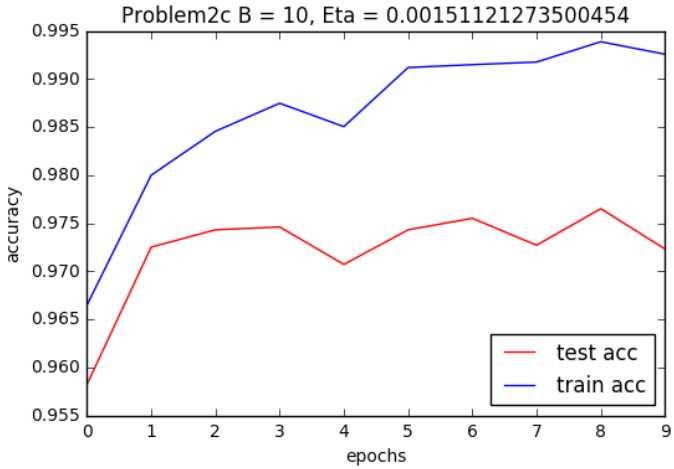
```
Random accuracy = 0.1061
Epoch 0: train loss = 0.0901 [42.478 sec
s]
test accuracy=0.9669
Epoch 1: train loss = 0.0574 [45.805 sec
s]
test accuracy=0.9721
Epoch 2: train loss = 0.0474 [45.826 sec
s]
test accuracy=0.9729
Epoch 3: train loss = 0.0377 [46.097 sec
s]
test accuracy=0.9737
Epoch 4: train loss = 0.0314 [44.226 sec
s]
test accuracy=0.9757
Epoch 5: train loss = 0.0302 [43.199 sec
s]
test accuracy=0.9742
Epoch 6: train loss = 0.0221 [43.441 sec
s]
test accuracy=0.9763
Epoch 7: train loss = 0.0305 [44.067 sec
s]
test accuracy=0.9741
Epoch 8: train loss = 0.0229 [44.760 sec
s]
test accuracy=0.9735
Epoch 9: train loss = 0.0130 [43.559 sec
s]
test accuracy=0.9769
```


Problem2c B = 10, Eta = 0.001610907834664192
test loss
train loss
loss
epochs
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
1
2
3
4
5
6
7
8
9

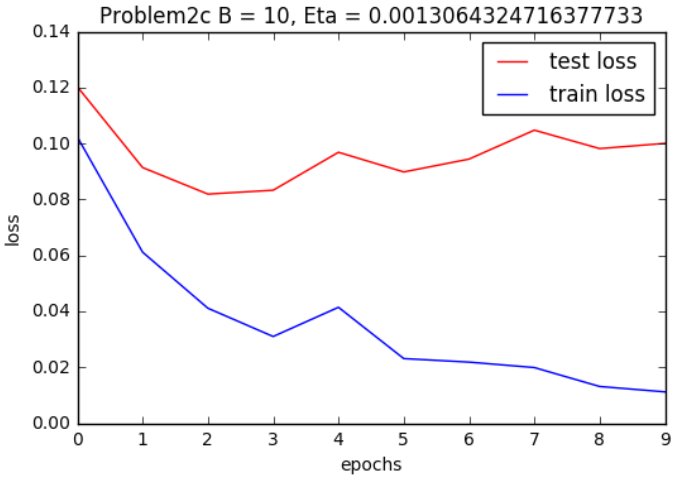

Problem2c B = 10, Eta = 0.001610907834664192
test acc
train acc
accuracy
epochs
1.000
0.995
0.990
0.985
0.980
0.975
0.970
0.965
0
1
2
3
4
5
6
7
8
9

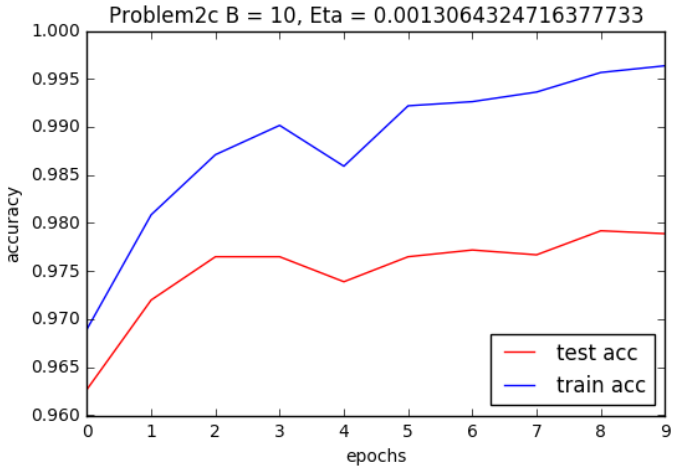
```
Random accuracy = 0.0829
Epoch 0: train loss = 0.1082 [38.642 sec
s]
test accuracy=0.9582
Epoch 1: train loss = 0.0629 [41.661 sec
s]
test accuracy=0.9725
Epoch 2: train loss = 0.0488 [41.477 sec
s]
test accuracy=0.9743
Epoch 3: train loss = 0.0387 [41.767 sec
s]
test accuracy=0.9746
Epoch 4: train loss = 0.0446 [42.111 sec
s]
test accuracy=0.9707
Epoch 5: train loss = 0.0262 [41.472 sec
s]
test accuracy=0.9743
Epoch 6: train loss = 0.0251 [31.508 sec
s]
test accuracy=0.9755
Epoch 7: train loss = 0.0237 [31.880 sec
s]
test accuracy=0.9727
Epoch 8: train loss = 0.0185 [31.854 sec
s]
test accuracy=0.9765
Epoch 9: train loss = 0.0237 [32.484 sec
s]
test accuracy=0.9723
```


Problem2c B = 10, Eta = 0.00151121273500454
test loss
train loss
loss
epochs
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
1
2
3
4
5
6
7
8
9

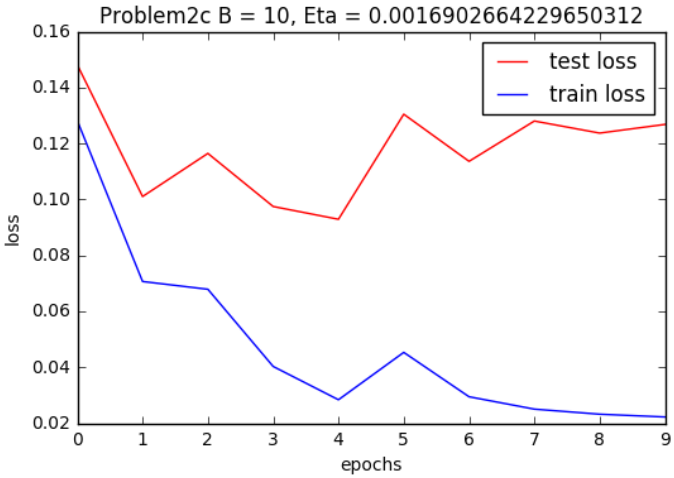

Problem2c B = 10, Eta = 0.00151121273500454
test acc
train acc
accuracy
epochs
0.995
0.990
0.985
0.980
0.975
0.970
0.965
0.960
0.955
0
1
2
3
4
5
6
7
8
9

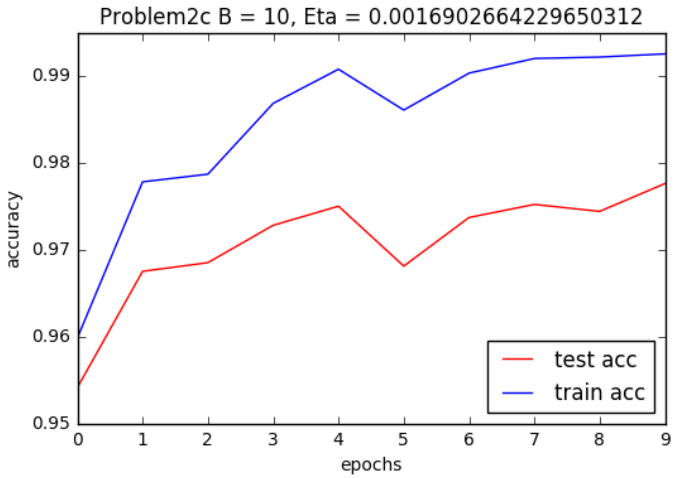
```
Random accuracy = 0.0844
Epoch 0: train loss = 0.1023 [26.609 sec
s]
test accuracy=0.9627
Epoch 1: train loss = 0.0611 [29.570 sec
s]
test accuracy=0.9720
Epoch 2: train loss = 0.0411 [28.733 sec
s]
test accuracy=0.9765
Epoch 3: train loss = 0.0310 [28.603 sec
s]
test accuracy=0.9765
Epoch 4: train loss = 0.0414 [28.833 sec
s]
test accuracy=0.9739
Epoch 5: train loss = 0.0231 [28.949 sec
s]
test accuracy=0.9765
Epoch 6: train loss = 0.0218 [28.803 sec
s]
test accuracy=0.9772
Epoch 7: train loss = 0.0199 [28.680 sec
s]
test accuracy=0.9767
Epoch 8: train loss = 0.0131 [28.637 sec
s]
test accuracy=0.9792
Epoch 9: train loss = 0.0112 [29.156 sec
s]
test accuracy=0.9789
```


Problem2c B = 10, Eta = 0.0013064324716377733
test loss
train loss
loss
epochs
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
1
2
3
4
5
6
7
8
9

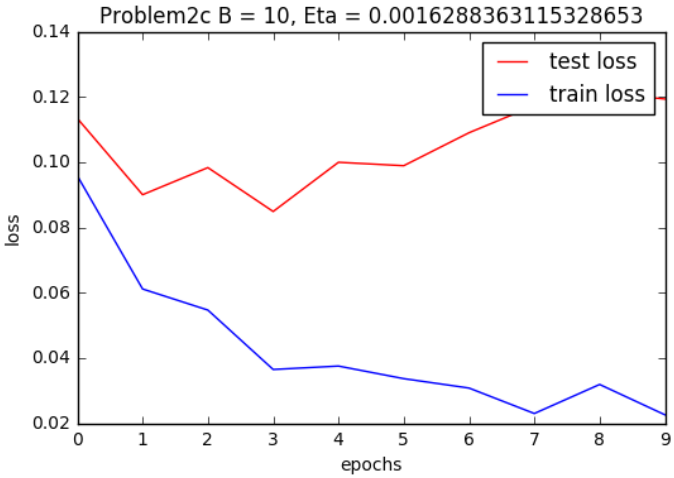

Problem2c B = 10, Eta = 0.0013064324716377733
test acc
train acc
accuracy
epochs
1.000
0.995
0.990
0.985
0.980
0.975
0.970
0.965
0.960
0
1
2
3
4
5
6
7
8
9

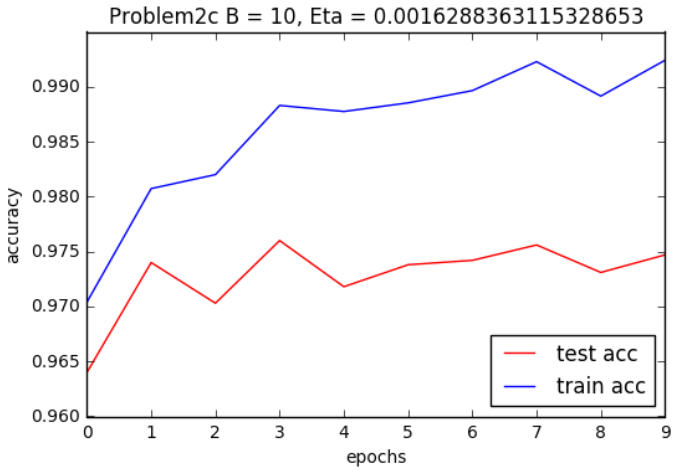
```
Random accuracy = 0.0823
Epoch 0: train loss = 0.1281 [28.671 sec
s]
test accuracy=0.9541
Epoch 1: train loss = 0.0706 [33.199 sec
s]
test accuracy=0.9675
Epoch 2: train loss = 0.0679 [33.401 sec
s]
test accuracy=0.9685
Epoch 3: train loss = 0.0403 [33.953 sec
s]
test accuracy=0.9728
Epoch 4: train loss = 0.0284 [33.856 sec
s]
test accuracy=0.9750
Epoch 5: train loss = 0.0453 [34.102 sec
s]
test accuracy=0.9681
Epoch 6: train loss = 0.0295 [33.319 sec
s]
test accuracy=0.9737
Epoch 7: train loss = 0.0250 [33.320 sec
s]
test accuracy=0.9752
Epoch 8: train loss = 0.0232 [32.697 sec
s]
test accuracy=0.9744
Epoch 9: train loss = 0.0222 [34.861 sec
s]
test accuracy=0.9776
```


Problem2c B = 10, Eta = 0.0016902664229650312
test loss
train loss
loss
epochs
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0
1
2
3
4
5
6
7
8
9

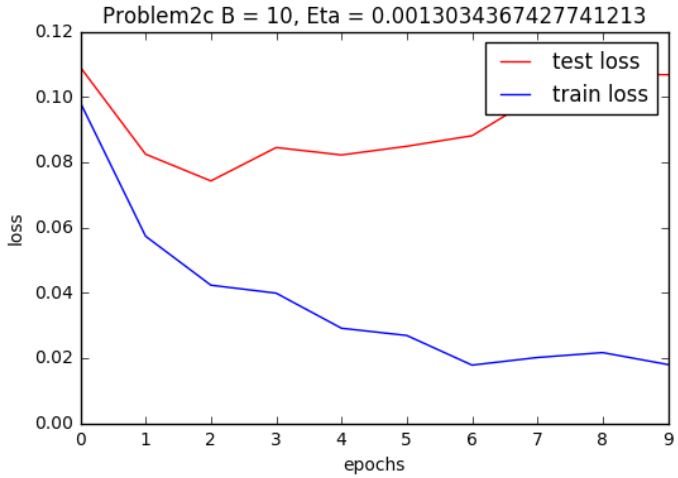

Problem2c B = 10, Eta = 0.0016902664229650312
test acc
train acc
accuracy
epochs
0.99
0.98
0.97
0.96
0.95
0
1
2
3
4
5
6
7
8
9

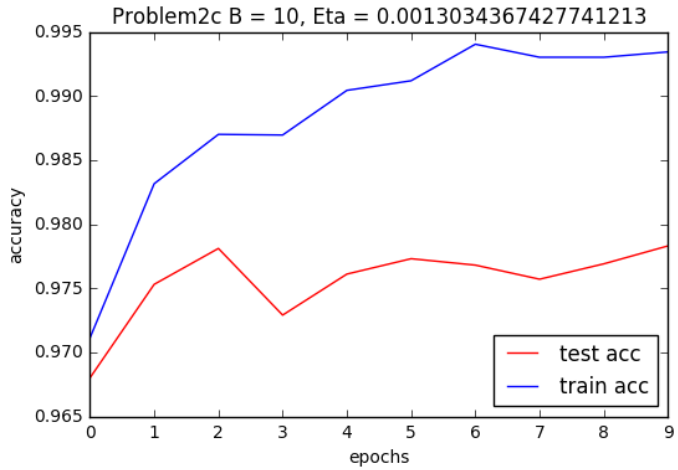
```
Random accuracy = 0.0953
Epoch 0: train loss = 0.0958 [29.977 sec
s]
test accuracy=0.9640
Epoch 1: train loss = 0.0611 [33.451 sec
s]
test accuracy=0.9740
Epoch 2: train loss = 0.0547 [34.087 sec
s]
test accuracy=0.9703
Epoch 3: train loss = 0.0365 [34.869 sec
s]
test accuracy=0.9760
Epoch 4: train loss = 0.0375 [34.544 sec
s]
test accuracy=0.9718
Epoch 5: train loss = 0.0336 [34.152 sec
s]
test accuracy=0.9738
Epoch 6: train loss = 0.0308 [34.216 sec
s]
test accuracy=0.9742
Epoch 7: train loss = 0.0230 [34.684 sec
s]
test accuracy=0.9756
Epoch 8: train loss = 0.0319 [34.802 sec
s]
test accuracy=0.9731
Epoch 9: train loss = 0.0226 [35.314 sec
s]
test accuracy=0.9747
```


Problem2c B = 10, Eta = 0.0016288363115328653
test loss
train loss
loss
epochs
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0
1
2
3
4
5
6
7
8
9

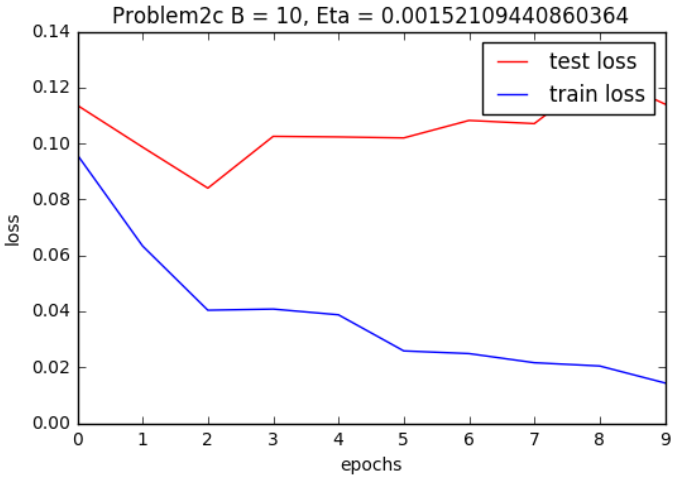

Problem2c B = 10, Eta = 0.0016288363115328653
test acc
train acc
accuracy
epochs
0.990
0.985
0.980
0.975
0.970
0.965
0.960
0
1
2
3
4
5
6
7
8
9

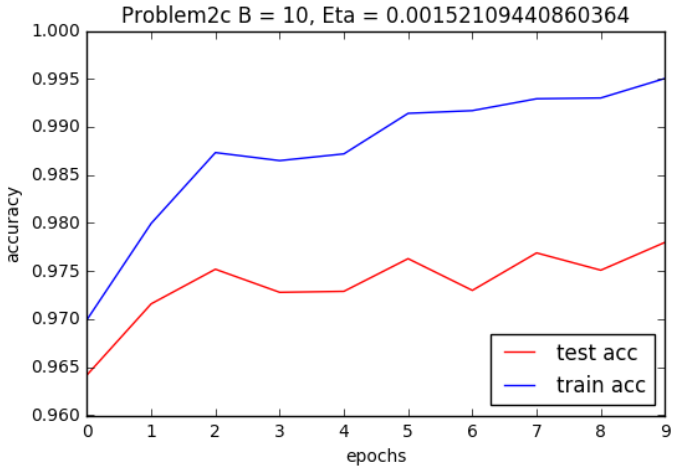
```
Random accuracy = 0.0928
Epoch 0: train loss = 0.0983 [29.133 sec
s]
test accuracy=0.9680
Epoch 1: train loss = 0.0573 [31.562 sec
s]
test accuracy=0.9753
Epoch 2: train loss = 0.0423 [30.520 sec
s]
test accuracy=0.9781
Epoch 3: train loss = 0.0398 [30.277 sec
s]
test accuracy=0.9729
Epoch 4: train loss = 0.0291 [30.224 sec
s]
test accuracy=0.9761
Epoch 5: train loss = 0.0269 [30.400 sec
s]
test accuracy=0.9773
Epoch 6: train loss = 0.0178 [30.650 sec
s]
test accuracy=0.9768
Epoch 7: train loss = 0.0201 [30.829 sec
s]
test accuracy=0.9757
Epoch 8: train loss = 0.0216 [30.996 sec
s]
test accuracy=0.9769
Epoch 9: train loss = 0.0180 [31.309 sec
s]
test accuracy=0.9783
```


Problem2c B = 10, Eta = 0.0013034367427741213
test loss
train loss
loss
epochs
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
1
2
3
4
5
6
7
8
9

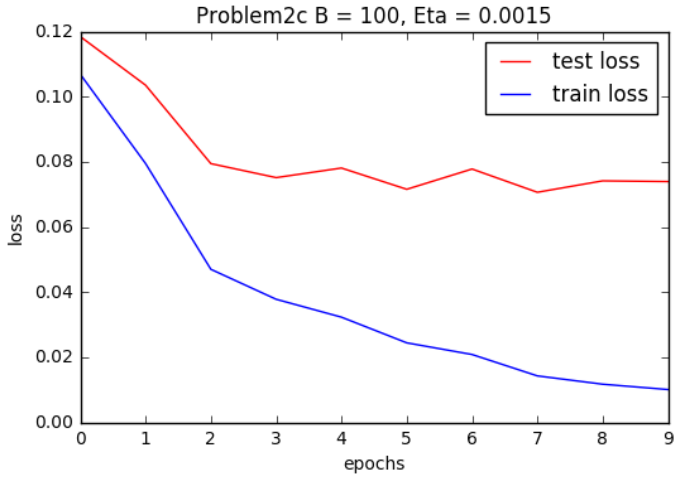

Problem2c B = 10, Eta = 0.0013034367427741213
test acc
train acc
accuracy
epochs
0.995
0.990
0.985
0.980
0.975
0.970
0.965
0
1
2
3
4
5
6
7
8
9

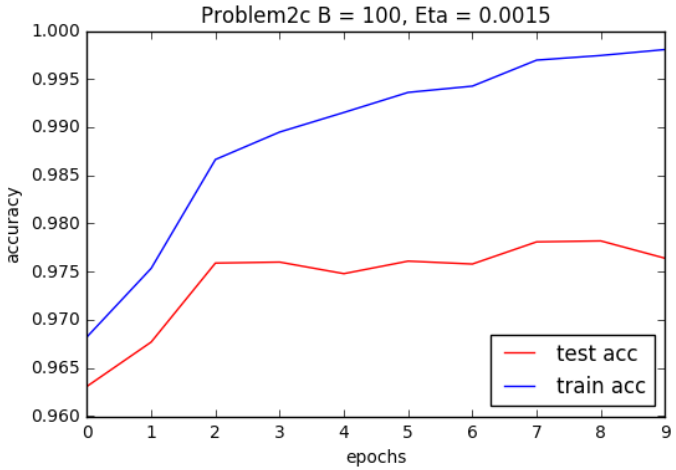
```
Random accuracy = 0.0910
Epoch 0: train loss = 0.0959 [26.854 sec
s]
test accuracy=0.9642
Epoch 1: train loss = 0.0633 [30.017 sec
s]
test accuracy=0.9716
Epoch 2: train loss = 0.0404 [30.350 sec
s]
test accuracy=0.9752
Epoch 3: train loss = 0.0408 [30.445 sec
s]
test accuracy=0.9728
Epoch 4: train loss = 0.0387 [30.655 sec
s]
test accuracy=0.9729
Epoch 5: train loss = 0.0258 [30.733 sec
s]
test accuracy=0.9763
Epoch 6: train loss = 0.0249 [30.846 sec
s]
test accuracy=0.9730
Epoch 7: train loss = 0.0216 [31.062 sec
s]
test accuracy=0.9769
Epoch 8: train loss = 0.0205 [31.243 sec
s]
test accuracy=0.9751
Epoch 9: train loss = 0.0144 [31.699 sec
s]
test accuracy=0.9780
```


Problem2c B = 10, Eta = 0.00152109440860364
test loss
train loss
loss
epochs
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
1
2
3
4
5
6
7
8
9

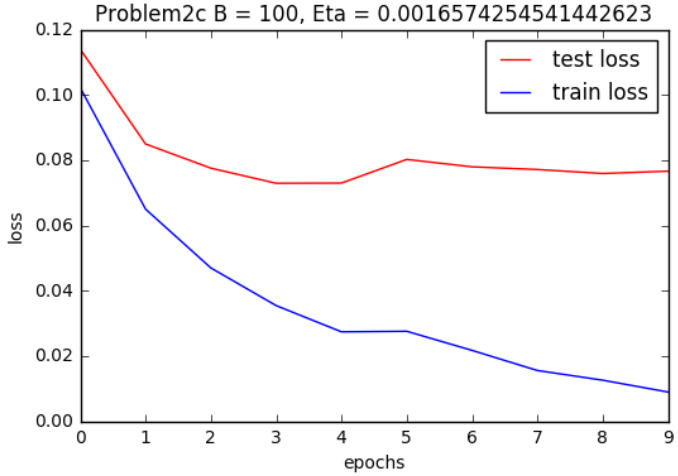

Problem2c B = 10, Eta = 0.00152109440860364
test acc
train acc
accuracy
epochs
1.000
0.995
0.990
0.985
0.980
0.975
0.970
0.965
0.960
0
1
2
3
4
5
6
7
8
9

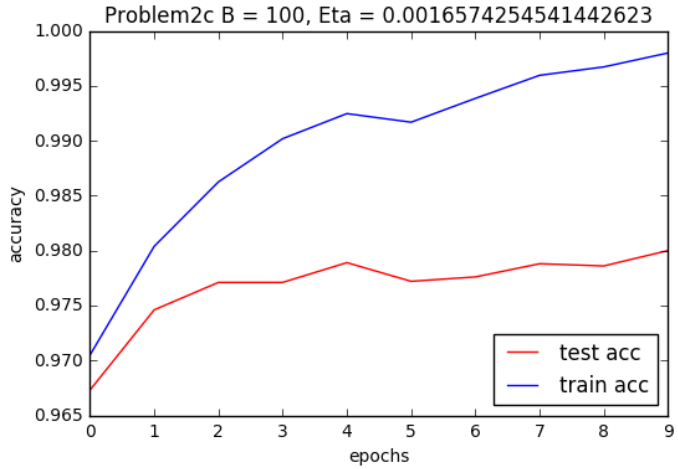
```
Random accuracy = 0.1142
Epoch 0: train loss = 0.1068 [1.794 secs]
test accuracy=0.9631
Epoch 1: train loss = 0.0794 [2.525 secs]
test accuracy=0.9677
Epoch 2: train loss = 0.0470 [2.609 secs]
test accuracy=0.9759
Epoch 3: train loss = 0.0378 [2.625 secs]
test accuracy=0.9760
Epoch 4: train loss = 0.0323 [2.616 secs]
test accuracy=0.9748
Epoch 5: train loss = 0.0244 [2.683 secs]
test accuracy=0.9761
Epoch 6: train loss = 0.0208 [2.664 secs]
test accuracy=0.9758
Epoch 7: train loss = 0.0143 [2.642 secs]
test accuracy=0.9781
Epoch 8: train loss = 0.0117 [2.669 secs]
test accuracy=0.9782
Epoch 9: train loss = 0.0101 [2.632 secs]
test accuracy=0.9764
```



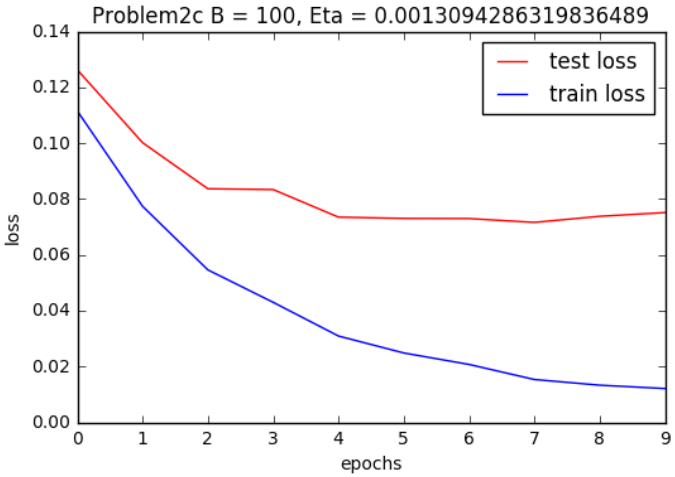


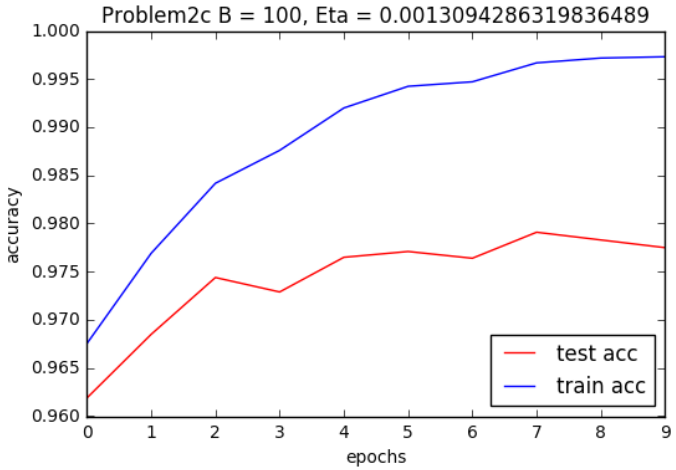

```
Random accuracy = 0.0875
Epoch 0: train loss = 0.1021 [1.782 secs]
test accuracy=0.9673
Epoch 1: train loss = 0.0650 [2.465 secs]
test accuracy=0.9746
Epoch 2: train loss = 0.0470 [2.573 secs]
test accuracy=0.9771
Epoch 3: train loss = 0.0354 [2.589 secs]
test accuracy=0.9771
Epoch 4: train loss = 0.0274 [2.556 secs]
test accuracy=0.9789
Epoch 5: train loss = 0.0275 [2.602 secs]
test accuracy=0.9772
Epoch 6: train loss = 0.0217 [2.543 secs]
test accuracy=0.9776
Epoch 7: train loss = 0.0155 [2.558 secs]
test accuracy=0.9788
Epoch 8: train loss = 0.0126 [2.551 secs]
test accuracy=0.9786
Epoch 9: train loss = 0.0089 [2.552 secs]
test accuracy=0.9800
```


Problem2c B = 100, Eta = 0.0016574254541442623
test loss
train loss
loss
epochs
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
1
2
3
4
5
6
7
8
9

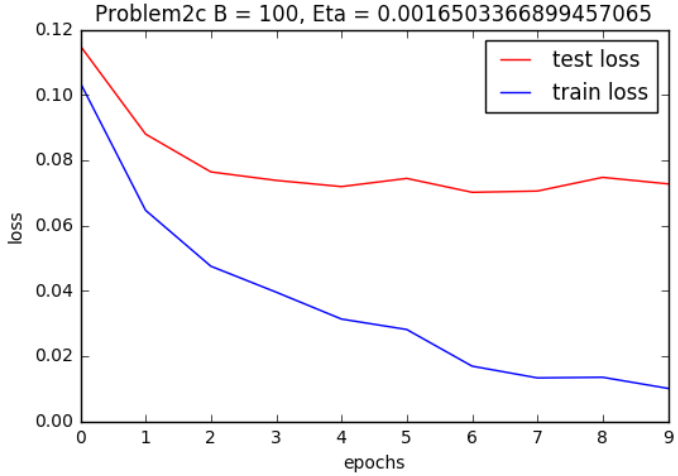

Problem2c B = 100, Eta = 0.0016574254541442623
test acc
train acc
accuracy
epochs
1.000
0.995
0.990
0.985
0.980
0.975
0.970
0.965
0
1
2
3
4
5
6
7
8
9

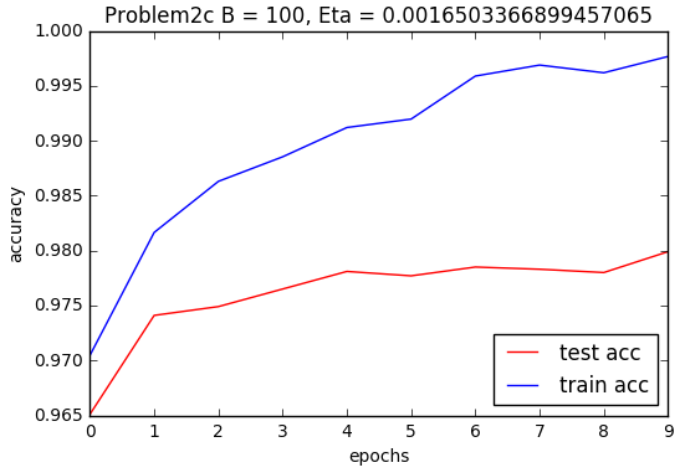
```
Random accuracy = 0.1188
Epoch 0: train loss = 0.1115 [1.787 secs]
test accuracy=0.9619
Epoch 1: train loss = 0.0774 [2.346 secs]
test accuracy=0.9685
Epoch 2: train loss = 0.0546 [2.413 secs]
test accuracy=0.9744
Epoch 3: train loss = 0.0430 [2.441 secs]
test accuracy=0.9729
Epoch 4: train loss = 0.0309 [2.468 secs]
test accuracy=0.9765
Epoch 5: train loss = 0.0248 [2.475 secs]
test accuracy=0.9771
Epoch 6: train loss = 0.0207 [2.450 secs]
test accuracy=0.9764
Epoch 7: train loss = 0.0153 [2.448 secs]
test accuracy=0.9791
Epoch 8: train loss = 0.0133 [2.471 secs]
test accuracy=0.9783
Epoch 9: train loss = 0.0121 [2.466 secs]
test accuracy=0.9775
```



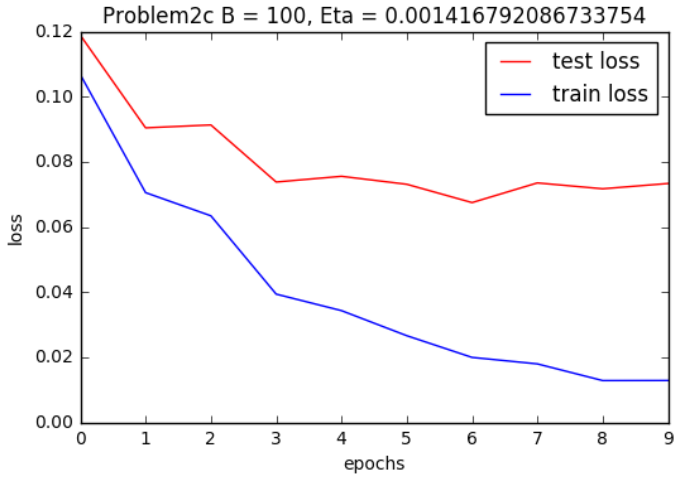


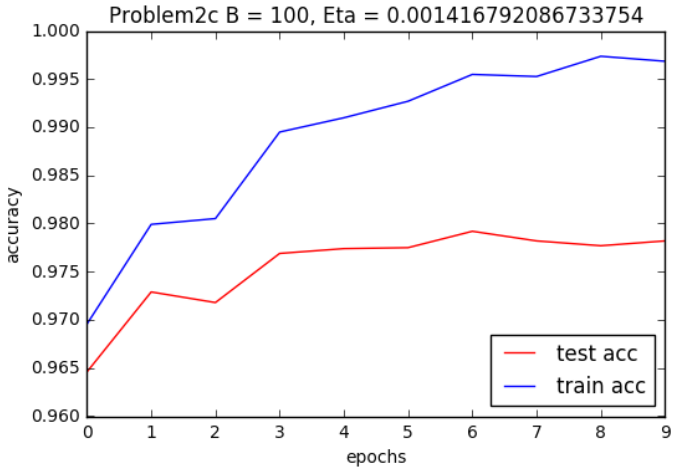

```
Random accuracy = 0.1039
Epoch 0: train loss = 0.1037 [1.811 secs]
test accuracy=0.9651
Epoch 1: train loss = 0.0646 [2.662 secs]
test accuracy=0.9741
Epoch 2: train loss = 0.0475 [2.759 secs]
test accuracy=0.9749
Epoch 3: train loss = 0.0395 [2.800 secs]
test accuracy=0.9765
Epoch 4: train loss = 0.0313 [2.824 secs]
test accuracy=0.9781
Epoch 5: train loss = 0.0281 [2.836 secs]
test accuracy=0.9777
Epoch 6: train loss = 0.0169 [2.839 secs]
test accuracy=0.9785
Epoch 7: train loss = 0.0133 [2.826 secs]
test accuracy=0.9783
Epoch 8: train loss = 0.0135 [2.830 secs]
test accuracy=0.9780
Epoch 9: train loss = 0.0101 [2.797 secs]
test accuracy=0.9799
```


Problem2c B = 100, Eta = 0.0016503366899457065
test loss
train loss
loss
epochs
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
1
2
3
4
5
6
7
8
9

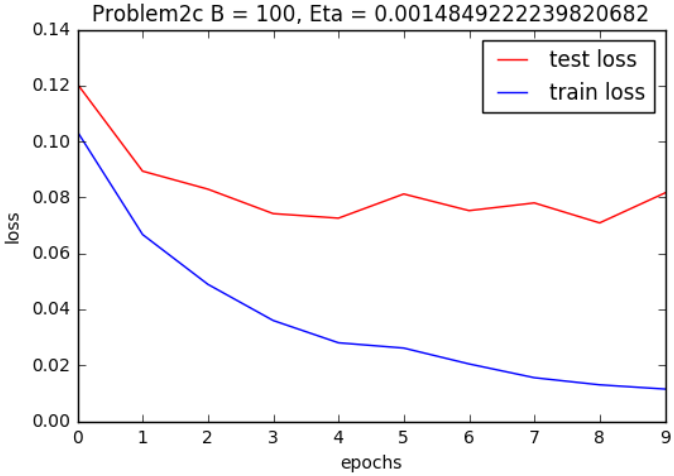

Problem2c B = 100, Eta = 0.0016503366899457065
test acc
train acc
accuracy
epochs
1.000
0.995
0.990
0.985
0.980
0.975
0.970
0.965
0
1
2
3
4
5
6
7
8
9

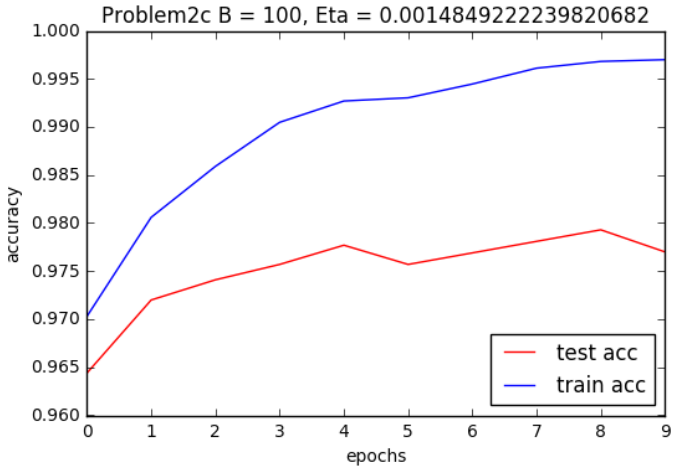
```
Random accuracy = 0.0974
Epoch 0: train loss = 0.1068 [1.785 secs]
test accuracy=0.9646
Epoch 1: train loss = 0.0705 [2.520 secs]
test accuracy=0.9729
Epoch 2: train loss = 0.0634 [2.636 secs]
test accuracy=0.9718
Epoch 3: train loss = 0.0394 [2.618 secs]
test accuracy=0.9769
Epoch 4: train loss = 0.0343 [2.660 secs]
test accuracy=0.9774
Epoch 5: train loss = 0.0266 [2.654 secs]
test accuracy=0.9775
Epoch 6: train loss = 0.0199 [2.719 secs]
test accuracy=0.9792
Epoch 7: train loss = 0.0179 [2.693 secs]
test accuracy=0.9782
Epoch 8: train loss = 0.0128 [2.676 secs]
test accuracy=0.9777
Epoch 9: train loss = 0.0128 [2.667 secs]
test accuracy=0.9782
```



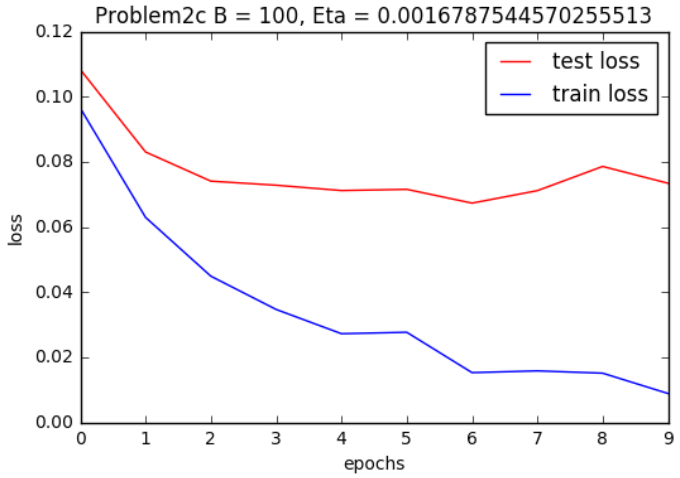


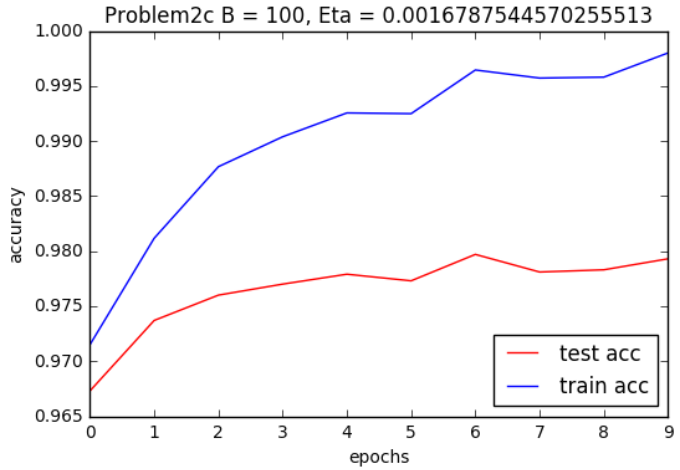

```
Random accuracy = 0.1377
Epoch 0: train loss = 0.1035 [1.828 secs]
test accuracy=0.9644
Epoch 1: train loss = 0.0667 [2.380 secs]
test accuracy=0.9720
Epoch 2: train loss = 0.0489 [2.445 secs]
test accuracy=0.9741
Epoch 3: train loss = 0.0360 [2.437 secs]
test accuracy=0.9757
Epoch 4: train loss = 0.0280 [2.480 secs]
test accuracy=0.9777
Epoch 5: train loss = 0.0261 [2.501 secs]
test accuracy=0.9757
Epoch 6: train loss = 0.0205 [2.505 secs]
test accuracy=0.9769
Epoch 7: train loss = 0.0155 [2.504 secs]
test accuracy=0.9781
Epoch 8: train loss = 0.0130 [2.523 secs]
test accuracy=0.9793
Epoch 9: train loss = 0.0115 [2.495 secs]
test accuracy=0.9770
```


Problem2c B = 100, Eta = 0.0014849222239820682
test loss
train loss
loss
epochs
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0 1 2 3 4 5 6 7 8 9

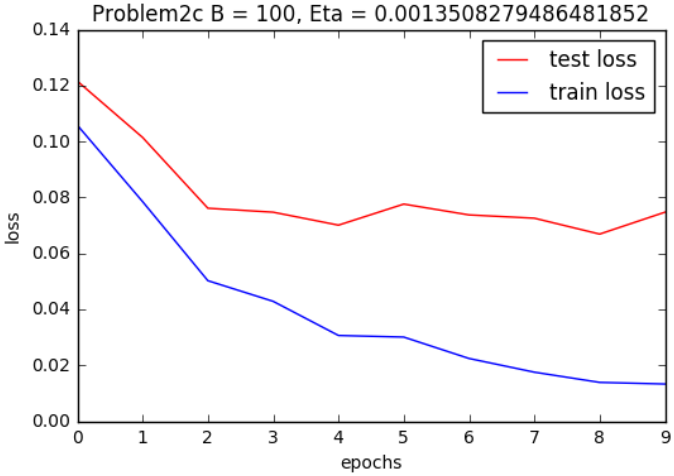

Problem2c B = 100, Eta = 0.0014849222239820682
test acc
train acc
accuracy
epochs
1.000
0.995
0.990
0.985
0.980
0.975
0.970
0.965
0.960
0 1 2 3 4 5 6 7 8 9

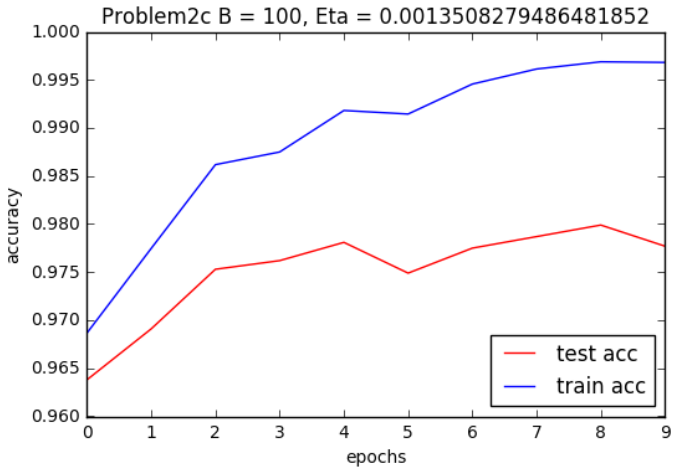
```
Random accuracy = 0.1194
Epoch 0: train loss = 0.0965 [1.783 secs]
test accuracy=0.9673
Epoch 1: train loss = 0.0629 [2.491 secs]
test accuracy=0.9737
Epoch 2: train loss = 0.0449 [2.615 secs]
test accuracy=0.9760
Epoch 3: train loss = 0.0347 [2.645 secs]
test accuracy=0.9770
Epoch 4: train loss = 0.0272 [2.640 secs]
test accuracy=0.9779
Epoch 5: train loss = 0.0276 [2.648 secs]
test accuracy=0.9773
Epoch 6: train loss = 0.0153 [2.664 secs]
test accuracy=0.9797
Epoch 7: train loss = 0.0158 [2.654 secs]
test accuracy=0.9781
Epoch 8: train loss = 0.0151 [2.661 secs]
test accuracy=0.9783
Epoch 9: train loss = 0.0089 [2.678 secs]
test accuracy=0.9793
```



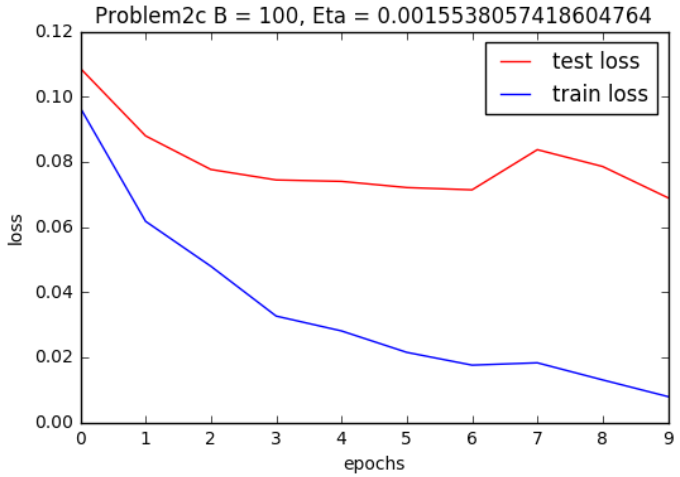


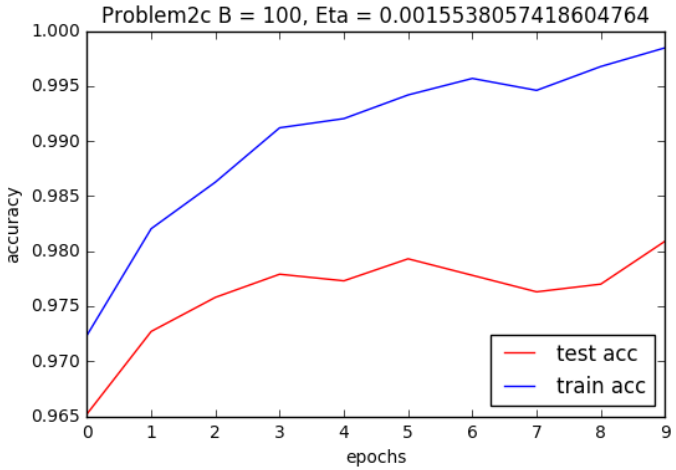

```
Random accuracy = 0.1205
Epoch 0: train loss = 0.1058 [1.797 secs]
test accuracy=0.9638
Epoch 1: train loss = 0.0785 [2.479 secs]
test accuracy=0.9691
Epoch 2: train loss = 0.0502 [2.635 secs]
test accuracy=0.9753
Epoch 3: train loss = 0.0428 [2.609 secs]
test accuracy=0.9762
Epoch 4: train loss = 0.0306 [2.641 secs]
test accuracy=0.9781
Epoch 5: train loss = 0.0301 [2.607 secs]
test accuracy=0.9749
Epoch 6: train loss = 0.0224 [2.636 secs]
test accuracy=0.9775
Epoch 7: train loss = 0.0175 [2.614 secs]
test accuracy=0.9787
Epoch 8: train loss = 0.0139 [2.613 secs]
test accuracy=0.9799
Epoch 9: train loss = 0.0133 [2.608 secs]
test accuracy=0.9777
```


Problem2c B = 100, Eta = 0.0013508279486481852
test loss
train loss
loss
epochs
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0 1 2 3 4 5 6 7 8 9

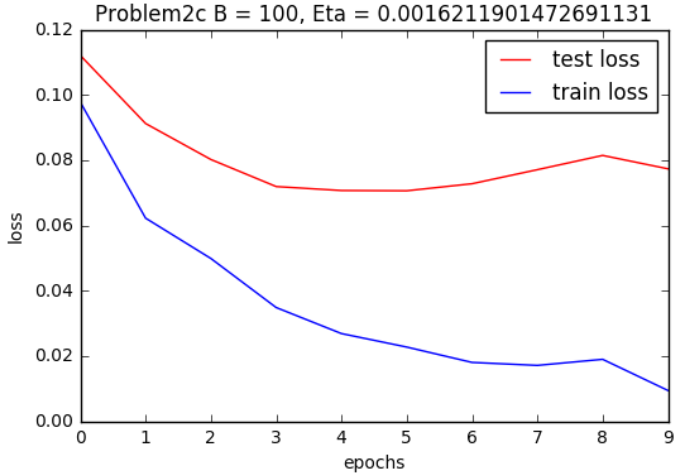

Problem2c B = 100, Eta = 0.0013508279486481852
test acc
train acc
accuracy
epochs
1.000
0.995
0.990
0.985
0.980
0.975
0.970
0.965
0.960
0 1 2 3 4 5 6 7 8 9

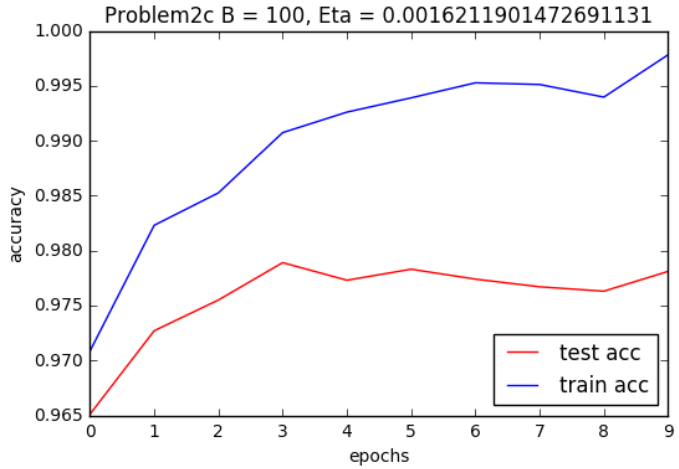
```
Random accuracy = 0.0850
Epoch 0: train loss = 0.0966 [1.797 secs]
test accuracy=0.9652
Epoch 1: train loss = 0.0617 [2.302 secs]
test accuracy=0.9727
Epoch 2: train loss = 0.0479 [2.392 secs]
test accuracy=0.9758
Epoch 3: train loss = 0.0326 [2.393 secs]
test accuracy=0.9779
Epoch 4: train loss = 0.0281 [2.394 secs]
test accuracy=0.9773
Epoch 5: train loss = 0.0215 [2.372 secs]
test accuracy=0.9793
Epoch 6: train loss = 0.0176 [2.383 secs]
test accuracy=0.9778
Epoch 7: train loss = 0.0183 [2.395 secs]
test accuracy=0.9763
Epoch 8: train loss = 0.0130 [2.392 secs]
test accuracy=0.9770
Epoch 9: train loss = 0.0079 [2.406 secs]
test accuracy=0.9809
```



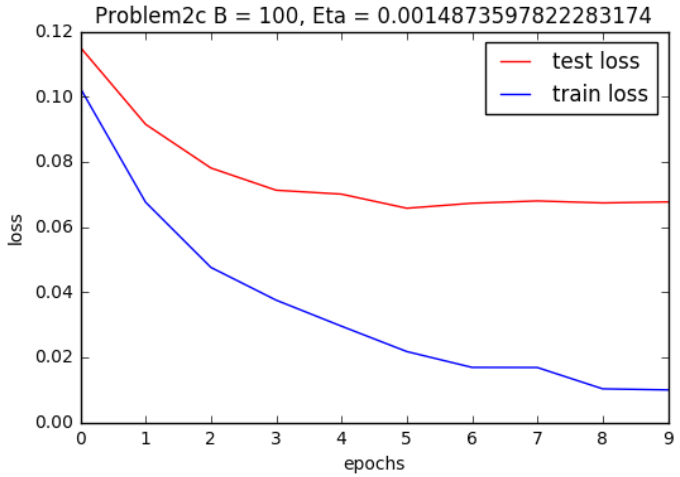


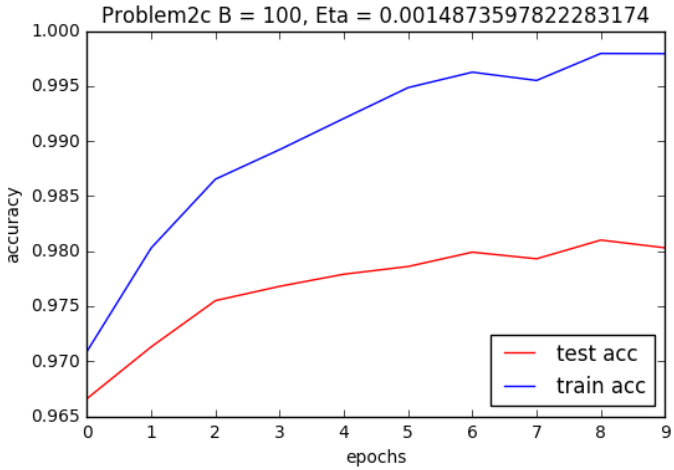

```
Random accuracy = 0.1307
Epoch 0: train loss = 0.0977 [1.786 secs]
test accuracy=0.9651
Epoch 1: train loss = 0.0622 [2.443 secs]
test accuracy=0.9727
Epoch 2: train loss = 0.0499 [2.569 secs]
test accuracy=0.9755
Epoch 3: train loss = 0.0348 [2.532 secs]
test accuracy=0.9789
Epoch 4: train loss = 0.0268 [2.557 secs]
test accuracy=0.9773
Epoch 5: train loss = 0.0227 [2.541 secs]
test accuracy=0.9783
Epoch 6: train loss = 0.0180 [2.562 secs]
test accuracy=0.9774
Epoch 7: train loss = 0.0171 [2.531 secs]
test accuracy=0.9767
Epoch 8: train loss = 0.0190 [2.572 secs]
test accuracy=0.9763
Epoch 9: train loss = 0.0094 [2.543 secs]
test accuracy=0.9781
```


Problem2c B = 100, Eta = 0.0016211901472691131
test loss
train loss
loss
epochs
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
1
2
3
4
5
6
7
8
9


Problem2c B = 100, Eta = 0.0016211901472691131
test acc
train acc
accuracy
epochs
1.000
0.995
0.990
0.985
0.980
0.975
0.970
0.965
0
1
2
3
4
5
6
7
8
9

```
Random accuracy = 0.1319
Epoch 0: train loss = 0.1026 [1.793 secs]
test accuracy=0.9666
Epoch 1: train loss = 0.0676 [2.265 secs]
test accuracy=0.9713
Epoch 2: train loss = 0.0476 [2.329 secs]
test accuracy=0.9755
Epoch 3: train loss = 0.0375 [2.364 secs]
test accuracy=0.9768
Epoch 4: train loss = 0.0295 [2.349 secs]
test accuracy=0.9779
Epoch 5: train loss = 0.0217 [2.378 secs]
test accuracy=0.9786
Epoch 6: train loss = 0.0169 [2.366 secs]
test accuracy=0.9799
Epoch 7: train loss = 0.0168 [2.374 secs]
test accuracy=0.9793
Epoch 8: train loss = 0.0103 [2.366 secs]
test accuracy=0.9810
Epoch 9: train loss = 0.0100 [2.352 secs]
test accuracy=0.9803
```




Problem2c B = 100, Eta = 0.0014873597822283174
1.000
0.995
0.990
0.985
0.980
0.975
0.970
0.965
accuracy
0
1
2
3
4
5
6
7
8
9
epochs
test acc
train acc