

In [9]:

```
weight0, weight1, weight2, bias0 ,bias1, bias2, loss_training_set ,loss_test_set = DNN()
```

0.0 % proceeding

.....  
.....  
.....  
.....  
.....

training loss is 2.5149347659670824

test loss is 1.6118094750958207

0.9 accuracy!

2.0 % proceeding

.....  
.....  
.....  
.....  
.....

training loss is 1.3611264849016464

test loss is 1.4140900466854691

0.912 accuracy!

4.0 % proceeding

.....  
.....  
.....  
.....  
.....

training loss is 1.0605172344904537

test loss is 1.2572120929347177

0.922 accuracy!

6.0 % proceeding

.....  
.....  
.....  
.....  
.....

training loss is 0.8422102388057583

test loss is 1.2231850932067416

0.923 accuracy!

8.0 % proceeding

.....  
.....  
.....  
.....  
.....

training loss is 0.726239135430163

test loss is 1.1626161206843304

0.927 accuracy!

10.0 % proceeding

.....  
.....  
.....  
.....  
.....

training loss is 0.6198121717872271

test loss is 1.160502795365448

0.928 accuracy!

12.0 % proceeding

.....  
.....  
.....  
.....  
.....

training loss is 0.5324898632013291

test loss is 1.0965966972596262

0.9319999999999999 accuracy!

14.000000000000002 % proceeding

.....  
.....

-----  
-----  
-----  
training loss is 0.46029361476551617  
test loss is 1.0960243550167224  
0.9319999999999999 accuracy!  
16.0 % proceeding  
-----  
-----  
-----  
-----  
-----

training loss is 0.4008714877695912  
test loss is 1.0152999944723609  
0.937 accuracy!  
18.0 % proceeding  
-----  
-----  
-----  
-----  
-----

training loss is 0.33506932505201975  
test loss is 1.035664666375785  
0.9339999999999999 accuracy!  
20.0 % proceeding  
-----  
-----  
-----  
-----  
-----

training loss is 0.3074991732074506  
test loss is 1.0307298039992197  
0.9359999999999999 accuracy!  
22.0 % proceeding  
-----  
-----  
-----  
-----  
-----

training loss is 0.26766254580996374  
test loss is 1.0211479334231863  
0.9359999999999999 accuracy!  
24.0 % proceeding  
-----  
-----  
-----  
-----  
-----

training loss is 0.22295709006632428  
test loss is 1.0679271217540682  
0.933 accuracy!  
26.0 % proceeding  
-----  
-----  
-----  
-----  
-----

training loss is 0.19941956402879477  
test loss is 1.0503457755780874  
0.9339999999999999 accuracy!  
28.000000000000004 % proceeding  
-----  
-----  
-----  
-----  
-----

training loss is 0.1780044230141076

test loss is 1.0155299987611037  
0.937 accuracy!  
30.0 % proceeding

-----  
-----  
-----  
-----  
-----

training loss is 0.14832513173285466  
test loss is 1.0154413234487647  
0.937 accuracy!  
32.0 % proceeding

-----  
-----  
-----  
-----  
-----

training loss is 0.13286747904725585  
test loss is 1.0533943647530484  
0.9339999999999999 accuracy!  
34.0 % proceeding

-----  
-----  
-----  
-----  
-----

training loss is 0.1144092205087674  
test loss is 1.0462395127176412  
0.9339999999999999 accuracy!  
36.0 % proceeding

-----  
-----  
-----  
-----  
-----

training loss is 0.1073018462836586  
test loss is 1.0631116470344022  
0.9339999999999999 accuracy!  
38.0 % proceeding

-----  
-----  
-----  
-----  
-----

training loss is 0.09659691666200783  
test loss is 1.061376457379356  
0.9339999999999999 accuracy!  
40.0 % proceeding

-----  
-----  
-----  
-----  
-----

training loss is 0.08137558567803722  
test loss is 1.0654026015723381  
0.9319999999999999 accuracy!  
42.0 % proceeding

-----  
-----  
-----  
-----  
-----

training loss is 0.06804135623211276  
test loss is 1.031719379157748  
0.9359999999999999 accuracy!  
44.0 % proceeding

-----  
-----  
-----

-----  
-----  
training loss is 0.057715049082359256  
test loss is 0.9699213937703438  
0.9390000000000001 accuracy!  
46.0 % proceeding  
-----  
-----  
-----  
-----  
-----

-----  
-----  
training loss is 0.0492960273555089  
test loss is 0.98439075728883  
0.938 accuracy!  
48.0 % proceeding  
-----  
-----  
-----  
-----  
-----

-----  
-----  
training loss is 0.04890186231384991  
test loss is 1.035957936108084  
0.9339999999999999 accuracy!  
50.0 % proceeding  
-----  
-----  
-----  
-----  
-----

-----  
-----  
training loss is 0.041416323493454206  
test loss is 0.9523526259382074  
0.94 accuracy!  
52.0 % proceeding  
-----  
-----  
-----  
-----  
-----

-----  
-----  
training loss is 0.03893742072213069  
test loss is 0.9836329667270168  
0.938 accuracy!  
54.0 % proceeding  
-----  
-----  
-----  
-----  
-----

-----  
-----  
training loss is 0.03216594083755486  
test loss is 1.0476570312153932  
0.935 accuracy!  
56.00000000000001 % proceeding  
-----  
-----  
-----  
-----  
-----

-----  
-----  
training loss is 0.028319811179531116  
test loss is 0.9348595398347663  
0.942 accuracy!  
57.99999999999999 % proceeding  
-----  
-----  
-----  
-----  
-----

-----  
-----  
training loss is 0.024044030420369077  
test loss is 0.9383687714946148  
-----  
-----  
-----

0.9410000000000001 accuracy!  
60.0 % proceeding

training loss is 0.01795469955199769  
test loss is 0.9659495655732327  
0.9390000000000001 accuracy!  
62.0 % proceeding

training loss is 0.016448282781528788  
test loss is 0.9846731557335892  
0.938 accuracy!  
64.0 % proceeding

training loss is 0.012488898017275946  
test loss is 1.0424003746871906  
0.935 accuracy!  
66.0 % proceeding

training loss is 0.016868747519025708  
test loss is 1.020335244472253  
0.9359999999999999 accuracy!  
68.0 % proceeding

training loss is 0.012179407548810741  
test loss is 0.9993177767089149  
0.938 accuracy!  
70.0 % proceeding

training loss is 0.013001181330945898  
test loss is 0.9997127270323134  
0.938 accuracy!  
72.0 % proceeding

training loss is 0.010252817346090556  
test loss is 0.9670726722741555  
0.94 accuracy!  
74.0 % proceeding

-----  
-----  
training loss is 0.012148169288423466  
test loss is 0.9619179961236524  
0.94 accuracy!  
76.0 % proceeding  
-----  
-----  
-----  
-----  
-----  
-----

-----  
-----  
training loss is 0.010745262998210338  
test loss is 1.003380833457053  
0.937 accuracy!  
78.0 % proceeding  
-----  
-----  
-----  
-----  
-----  
-----

-----  
-----  
training loss is 0.009490179483876504  
test loss is 0.9358499302645163  
0.9410000000000001 accuracy!  
80.0 % proceeding  
-----  
-----  
-----  
-----  
-----  
-----

-----  
-----  
training loss is 0.006396593196803525  
test loss is 0.9991055675244346  
0.938 accuracy!  
82.0 % proceeding  
-----  
-----  
-----  
-----  
-----  
-----

-----  
-----  
training loss is 0.007056705645462451  
test loss is 0.9347761989165964  
0.942 accuracy!  
84.0 % proceeding  
-----  
-----  
-----  
-----  
-----  
-----

-----  
-----  
training loss is 0.007406090936302602  
test loss is 0.9049490159101365  
0.943 accuracy!  
86.0 % proceeding  
-----  
-----  
-----  
-----  
-----  
-----

-----  
-----  
training loss is 0.005892544544452425  
test loss is 0.9348808416467967  
0.942 accuracy!  
88.0 % proceeding  
-----  
-----  
-----  
-----  
-----  
-----

-----  
-----  
training loss is 0.0024030067138877333  
test loss is 0.9340942348970782  
0.942 accuracy!  
-----  
-----  
-----  
-----  
-----  
-----

90.0 % proceeding

training loss is 0.0010467341585904887  
test loss is 0.9348460279372178  
0.942 accuracy!  
92.0 % proceeding

training loss is 0.005186301831370774  
test loss is 0.9031487629179922  
0.944 accuracy!  
94.0 % proceeding

training loss is 0.00695910540839891  
test loss is 0.9805674060457873  
0.9390000000000001 accuracy!  
96.0 % proceeding

training loss is 0.0033000310162809806  
test loss is 0.9509675600956976  
0.9410000000000001 accuracy!  
98.0 % proceeding

training loss is 0.002779154419935798  
test loss is 0.9683039815842942  
0.938 accuracy!

```
In [13]: confusionmatrix, top3matrix, top3matrix_index = ConfusionMatrix_n_top3(weight0, weight1, weight2, bias0 ,bias1, b
```

## Module

```
In [3]: import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
```

## Loader

```
In [4]: import gzip
import numpy as np
from pathlib import Path
import math
import random

class Dataloader():
    def __init__(self, path, is_train=True, shuffle=True, batch_size=8):
```

```

path = Path(path)
imagePath = Path(path/'train-images-idx3-ubyte.gz') if is_train else Path(path/'t10k-images-idx3-ubyte.gz')
labelPath = Path(path/'train-labels-idx1-ubyte.gz') if is_train else Path(path/'t10k-labels-idx1-ubyte.gz')

self.batch_size = batch_size
self.images = self.loadImages(imagePath)
self.labels = self.loadLabels(labelPath)
self.index = 0
self.idx = np.arange(0, self.images.shape[0])
if shuffle: np.random.shuffle(self.idx) # shuffle images

def __len__(self):
    n_images, _, _, _ = self.images.shape
    n_images = math.ceil(n_images / self.batch_size)
    return n_images

def __iter__(self):
    return datasetIterator(self)

def __getitem__(self, index):
    image = self.images[self.idx[index * self.batch_size:(index + 1) * self.batch_size]]
    label = self.labels[self.idx[index * self.batch_size:(index + 1) * self.batch_size]]
    image = image/255.0
    return image, label

def loadImages(self, path):
    with gzip.open(path) as f:
        images = np.frombuffer(f.read(), 'B', offset=16)
        images = images.reshape(-1, 1, 28, 28).astype(np.float32)
        return images

def loadLabels(self, path):
    with gzip.open(path) as f:
        labels = np.frombuffer(f.read(), 'B', offset=8)
        rows = len(labels)
        cols = labels.max() + 1
        one_hot = np.zeros((rows, cols)).astype(np.uint8)
        one_hot[np.arange(rows), labels] = 1
        one_hot = one_hot.astype(np.float64)
        return one_hot

# for enumerate magic python function returns Iterator
class datasetIterator():
    def __init__(self, dataloader):
        self.index = 0
        self.dataloader = dataloader

    def __next__(self):
        if self.index < len(self.dataloader):
            item = self.dataloader[self.index]
            self.index += 1
            return item
        # end of iteration
        raise StopIteration

```

## Function

```

In [5]: def ReLU(value):
        return max(0, value)

def converter_ReLU(array):
    return np.array([ReLU(x) for x in array])

def zeroorone(value):
    if value > 0:
        return 1
    else:
        return 0

def converter_zeroorone_10(array):
    return np.array([zeroorone(x) for x in array.reshape(10)])

def converter_zeroorone_784(array):
    return np.array([zeroorone(x) for x in array.reshape(784)])

def SoftMax(z):
    c = np.max(z)
    exp_z = np.exp(z-c)
    sum_exp_z = np.sum(exp_z)
    y = exp_z / sum_exp_z
    return y

def Cross_entropy_loss(y_label, y_prediction):
    return -np.sum(y_label*np.log(y_prediction+1e-7))

```



In [6]:

```
def DNN(batchsize=100, epoch=50, testing = 1000):
    #ready for dataset
    learning_rate = batchSize/60000
    iteration = math.ceil(60000/batchsize)
    loss_training_set = []
    loss_test_set = []

    training_data = Dataloader(
        path="./",
        shuffle=True,
        batch_size=batchsize
    )
    test_data = Dataloader(
        path="./",
        shuffle=True,
        is_train = False,
        batch_size = 1
    )
    #initialize function
    weight0 = np.random.randn(784,784)
    weight1 = np.random.randn(784,784)
    weight2 = np.random.randn(784,10)
    bias0 = np.random.randn(784)
    bias1 = np.random.randn(784)
    bias2 = np.random.randn(10)

    for k in range(0,epoch):
        print(100*(k/epoch), "% proceeding")
        training_loss = 0
        error = 0
        test_loss = 0
        for i in range(0,iteration):
            #forward propagation
            print("-",end='')
            delta_3 = np.zeros(10).reshape(1,10)
            delta_2 = np.zeros(784).reshape(1,784)
            delta_1 = np.zeros(784).reshape(1,784)
            chain_delta_3 = np.zeros(7840).reshape(784,10)
            chain_delta_2 = np.zeros(614656).reshape(784,784)
            chain_delta_1 = np.zeros(614656).reshape(784,784)
            for j in range(0, batchSize):
                #forward propagation
                y_label = training_data.__getitem__(i)[1][j]

                layer_input = converter_zeroorone_784(training_data.__getitem__(i)[0][j])
                layer_1 = np.dot(layer_input, weight0)+bias0
                layer_1_ReLU = converter_ReLU(layer_1)

                layer_2 = np.dot(layer_1_ReLU,weight1)+bias1
                layer_2_ReLU = converter_ReLU(layer_2)

                layer_3 = np.dot(layer_2_ReLU, weight2)+bias2
                y_prediction = SoftMax(layer_3)
                training_loss += Cross_entropy_loss(y_label,y_prediction)
                #Backwardpropagation weight2

                delta_3_b = ((y_prediction - y_label)/batchsize)
                chain_delta_3_b = np.dot(layer_2_ReLU.reshape(784,1), delta_3_b.reshape(1,10))

                #Backwardpropagation weight1

                delta_2_b = np.dot(delta_3_b,weight2.T)*converter_zeroorone_784(layer_2)
                chain_delta_2_b = np.dot(layer_1_ReLU.reshape(784,1), delta_2_b.reshape(1,784))

                #Backwardpropagation weight0

                delta_1_b = np.dot(delta_2_b,weight1.T)*converter_zeroorone_784(layer_1)
                chain_delta_1_b = np.dot(layer_input.reshape(784,1), delta_1_b.reshape(1,784))

                delta_3 += delta_3_b
                delta_2 += delta_2_b
                delta_1 += delta_1_b
                chain_delta_3 += chain_delta_3_b
                chain_delta_2 += chain_delta_2_b
                chain_delta_1 += chain_delta_1_b
            weight2 -= (learning_rate * chain_delta_3)
            weight1 -= (learning_rate * chain_delta_2)
            weight0 -= (learning_rate * chain_delta_1)
            bias2 -= delta_3.reshape(10)*learning_rate
            bias1 -= delta_2.reshape(784)*learning_rate
            bias0 -= delta_1.reshape(784)*learning_rate
            print("\n")
            print("training loss is",training_loss/60000)
            loss_training_set.append(training_loss/60000)

        for i in range(0,testing):
```

```

        y_label = test_data.__getitem__(i)[1]
        layer_input = converter_zeroorone_784(test_data.__getitem__(i)[0])
        layer_1 = np.dot(layer_input, weight0)+bias0
        layer_1_ReLU = converter_ReLU(layer_1)
        layer_2 = np.dot(layer_1_ReLU,weight1)+bias1
        layer_2_ReLU = converter_ReLU(layer_2)
        layer_3 = np.dot(layer_2_ReLU, weight2)+bias2
        y_prediction = SoftMax(layer_3)
        test_loss += Cross_entropy_loss(y_label,y_prediction)
        if(np.argmax(y_prediction)!=np.argmax(test_data.__getitem__(i)[1])):
            error += 1
    print("test loss is",test_loss/testing)
    print(1-(error/testing),"accuracy!")
    loss_test_set.append(test_loss/testing)
return weight0, weight1, weight2, bias0 ,bias1, bias2, loss_training_set ,loss_test_set

```

```

In [12]: test_data = Dataloader(
    path=".",
    shuffle=True,
    is_train = False,
    batch_size = 1
)

```

```

In [11]: def ConfusionMatrix_n_top3(weight0, weight1, weight2, bias0 ,bias1, bias2, test_data):
    confusionmatrix = np.zeros(100).reshape(10,10)
    top3matrix = np.zeros(30).reshape(10,3)
    top3matrix_index = np.zeros(30).reshape(10,3)
    for i in range(0,10000):
        y_label = test_data.__getitem__(i)[1]
        layer_input = converter_zeroorone_784(test_data.__getitem__(i)[0])
        layer_1 = np.dot(layer_input, weight0)+bias0
        layer_1_ReLU = converter_ReLU(layer_1)
        layer_2 = np.dot(layer_1_ReLU,weight1)+bias1
        layer_2_ReLU = converter_ReLU(layer_2)
        layer_3 = np.dot(layer_2_ReLU, weight2)+bias2
        y_prediction = SoftMax(layer_3)

        #confusionmatrix
        confusionmatrix[np.argmax(y_label)][np.argmax(y_prediction)]+=1
        index = np.argmin(top3matrix[np.argmax(y_prediction)])
        if top3matrix[np.argmax(y_prediction)][index] <= y_prediction[np.argmax(y_prediction)]:
            top3matrix[np.argmax(y_prediction)][index] = y_prediction[np.argmax(y_prediction)]
            top3matrix_index[np.argmax(y_prediction)][index] = i

    return confusionmatrix , top3matrix, top3matrix_index

```

## Loss graph

```

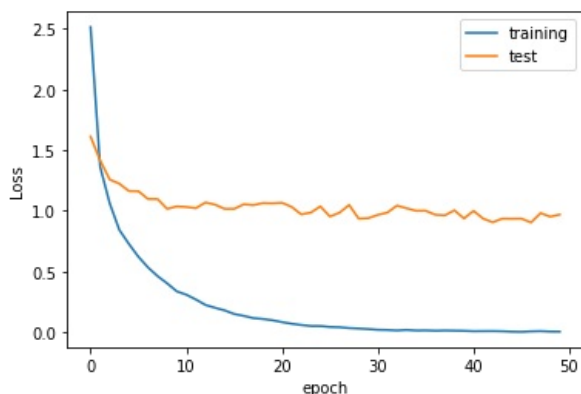
In [14]: index = [x for x in range(50)]

```

```

In [15]: plt.xlabel('epoch')
plt.ylabel("Loss")
plt.plot(index,loss_training_set)
plt.plot(index,loss_test_set)
plt.legend(['training', 'test'])
plt.show()

```



## Confusion Matrix

```
In [16]: confusionmatrix_visualization = np.array([(100*x)/np.sum(x) for x in confusionmatrix])
```

```
In [17]: confusionmatrix_visualization = np.array([ np.fix(x) for x in confusionmatrix])
```

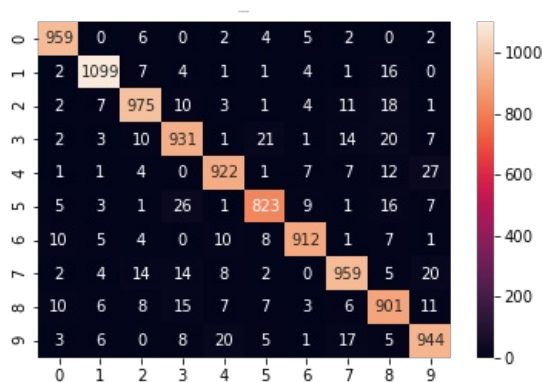
```
In [18]: confusionmatrix_visualization.astype(int)
```

```
Out[18]: array([[ 959,    0,    6,    0,    2,    4,    5,    2,    0,    2],
 [    2, 1099,    7,    4,    1,    1,    4,    1,   16,    0],
 [    2,    7,   975,   10,    3,    1,    4,   11,   18,    1],
 [    2,    3,   10,   931,    1,   21,    1,   14,   20,    7],
 [    1,    1,    4,    0,   922,    1,    7,    7,   12,   27],
 [    5,    3,    1,   26,    1,   823,    9,    1,   16,    7],
 [   10,    5,    4,    0,   10,    8,   912,    1,    7,    1],
 [    2,    4,   14,   14,    8,    2,    0,   959,    5,   20],
 [   10,    6,    8,   15,    7,    7,    3,    6,   901,   11],
 [    3,    6,    0,    8,   20,    5,    1,   17,    5,   944]])
```

```
In [19]: sns.heatmap(confusionmatrix_visualization.astype(int), annot=True, fmt='d')

plt.title('confusionmatrix', fontsize=1)

plt.show()
```



## Top-3 images with probability

```
In [20]: print(top3matrix)
```

```
[[1. 1. 1.]
 [1. 1. 1.]
 [1. 1. 1.]
 [1. 1. 1.]
 [1. 1. 1.]
 [1. 1. 1.]
 [1. 1. 1.]
 [1. 1. 1.]
 [1. 1. 1.]
 [1. 1. 1.]]
```

```
In [22]: fig, axes = plt.subplots(10,3, figsize=(3,10))
j=0
for i,ax in enumerate(axes.flat):
    index = math.ceil(top3matrix_index.reshape(30)[j])
    ax.imshow(test_data.__getitem__(index)[0].reshape(28,28))
    j+=1
```

