gaoxqiecx

January 23, 2025

[ ]:
```
import numpy as np
import torch
import torch.nn as nn
import matplotlib.pyplot as plt
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score, confusion_matrix,␣
↪ConfusionMatrixDisplay
```

[ ]:
```
#load Datasets
iris=load_iris()
X=iris.data
y=iris.target #documentation from iris dataset
#removing all features that belong to category 2
X=X[y!=2] #conditional slicing
y=y[y!=2]
```

[ ]:
```
X.shape
y.shape #reshaping
# y=y.reshape(100,1)
#y = y[:, np.newaxis] # adding new axis
```

[ ]:
```
(100,)
```

[ ]:
```
#spliting test and train using library train_test_split
X_train,X_test,y_train,y_test=train_test_split(X,y, test_size=0.2)
```

[ ]:
```
X_train_tensor=torch.tensor(X_train, dtype=torch.float32) #is float cause
X_test_tensor=torch.tensor(X_test, dtype=torch.float32) #is int cause

y_train_tensor=torch.tensor(y_train, dtype=int)
y_test_tensor=torch.tensor(y_test, dtype=int)
```

[ ]:
```
'''class ClassificationModel(nn.Module):
  def __init__(self):
    super(ClassificationModel, self).__init__()
    self.linear=nn.Linear(4,2) #4 features 1 target
```

```
    #self.linear=nn.Linear(4,2)-> auto  uses sigmod activation

    def forward(self, x):
      return (self.linear(x))
      # return torch.sigmod()
      #here we need to use sigmod activation to change any continuous value␣
 ↪between 0 and 1'''
```

[ ]:
```
'class ClassificationModel(nn.Module):\n  def __init__(self):\n
super(ClassificationModel, self).__init__()\n    self.linear=nn.Linear(4,2) #4
features 1 target\n    #self.linear=nn.Linear(4,2)-> auto  uses sigmod
activation\n\n    def forward(self, x):\n      return (self.linear(x))\n      #
return torch.sigmod()\n      #here we need to use sigmod activation to change
any continuous value between 0 and 1'
```

[ ]:
```
class ClassificationModel(nn.Module):
  def __init__(self):
    super(ClassificationModel,self).__init__()
    self.linear=nn.Linear(4,2)  #(target,features)

  def forward(self,x):
    return (self.linear(x))
```

[ ]:
```
#loss function -> cross entropy loss
model = ClassificationModel() #obect creation
loss=nn.CrossEntropyLoss()
criteria=torch.optim.SGD(model.parameters(),lr=0.01)
```

[ ]:
```
num_epochs=1000
train_loss=[]
test_loss=[]
train_accuracy=[]
test_accuracy=[]
```

[ ]:
```
for ep in range(num_epochs):
  model.train()
  predicted_y= model(X_train_tensor)
  losses=loss(predicted_y, y_train_tensor)
  print(losses)

  criteria.zero_grad()
  losses.backward()
  criteria.step()

  train_loss.append(losses.item())
  model.eval()
  with torch.no_grad():
```

```
    predicted_test_y=model(X_test_tensor)
    loss_test=loss(predicted_test_y,y_test_tensor)
    test_loss.append(loss_test.item())
```

```
tensor(1.4888, grad_fn=<NllLossBackward0>)
tensor(1.2511, grad_fn=<NllLossBackward0>)
tensor(1.0353, grad_fn=<NllLossBackward0>)
tensor(0.8492, grad_fn=<NllLossBackward0>)
tensor(0.6994, grad_fn=<NllLossBackward0>)
tensor(0.5874, grad_fn=<NllLossBackward0>)
tensor(0.5091, grad_fn=<NllLossBackward0>)
tensor(0.4568, grad_fn=<NllLossBackward0>)
tensor(0.4227, grad_fn=<NllLossBackward0>)
tensor(0.4004, grad_fn=<NllLossBackward0>)
tensor(0.3854, grad_fn=<NllLossBackward0>)
tensor(0.3750, grad_fn=<NllLossBackward0>)
tensor(0.3674, grad_fn=<NllLossBackward0>)
tensor(0.3615, grad_fn=<NllLossBackward0>)
tensor(0.3567, grad_fn=<NllLossBackward0>)
tensor(0.3526, grad_fn=<NllLossBackward0>)
tensor(0.3489, grad_fn=<NllLossBackward0>)
tensor(0.3455, grad_fn=<NllLossBackward0>)
tensor(0.3423, grad_fn=<NllLossBackward0>)
tensor(0.3392, grad_fn=<NllLossBackward0>)
tensor(0.3362, grad_fn=<NllLossBackward0>)
tensor(0.3333, grad_fn=<NllLossBackward0>)
tensor(0.3305, grad_fn=<NllLossBackward0>)
tensor(0.3277, grad_fn=<NllLossBackward0>)
tensor(0.3250, grad_fn=<NllLossBackward0>)
tensor(0.3223, grad_fn=<NllLossBackward0>)
tensor(0.3197, grad_fn=<NllLossBackward0>)
tensor(0.3171, grad_fn=<NllLossBackward0>)
tensor(0.3145, grad_fn=<NllLossBackward0>)
tensor(0.3120, grad_fn=<NllLossBackward0>)
tensor(0.3095, grad_fn=<NllLossBackward0>)
tensor(0.3071, grad_fn=<NllLossBackward0>)
tensor(0.3047, grad_fn=<NllLossBackward0>)
tensor(0.3023, grad_fn=<NllLossBackward0>)
tensor(0.3000, grad_fn=<NllLossBackward0>)
tensor(0.2977, grad_fn=<NllLossBackward0>)
tensor(0.2954, grad_fn=<NllLossBackward0>)
tensor(0.2931, grad_fn=<NllLossBackward0>)
tensor(0.2909, grad_fn=<NllLossBackward0>)
tensor(0.2887, grad_fn=<NllLossBackward0>)
tensor(0.2866, grad_fn=<NllLossBackward0>)
tensor(0.2844, grad_fn=<NllLossBackward0>)
tensor(0.2824, grad_fn=<NllLossBackward0>)
```

```
tensor(0.2803, grad_fn=<NllLossBackward0>)
tensor(0.2782, grad_fn=<NllLossBackward0>)
tensor(0.2762, grad_fn=<NllLossBackward0>)
tensor(0.2742, grad_fn=<NllLossBackward0>)
tensor(0.2723, grad_fn=<NllLossBackward0>)
tensor(0.2703, grad_fn=<NllLossBackward0>)
tensor(0.2684, grad_fn=<NllLossBackward0>)
tensor(0.2665, grad_fn=<NllLossBackward0>)
tensor(0.2647, grad_fn=<NllLossBackward0>)
tensor(0.2628, grad_fn=<NllLossBackward0>)
tensor(0.2610, grad_fn=<NllLossBackward0>)
tensor(0.2592, grad_fn=<NllLossBackward0>)
tensor(0.2575, grad_fn=<NllLossBackward0>)
tensor(0.2557, grad_fn=<NllLossBackward0>)
tensor(0.2540, grad_fn=<NllLossBackward0>)
tensor(0.2523, grad_fn=<NllLossBackward0>)
tensor(0.2506, grad_fn=<NllLossBackward0>)
tensor(0.2490, grad_fn=<NllLossBackward0>)
tensor(0.2473, grad_fn=<NllLossBackward0>)
tensor(0.2457, grad_fn=<NllLossBackward0>)
tensor(0.2441, grad_fn=<NllLossBackward0>)
tensor(0.2425, grad_fn=<NllLossBackward0>)
tensor(0.2410, grad_fn=<NllLossBackward0>)
tensor(0.2394, grad_fn=<NllLossBackward0>)
tensor(0.2379, grad_fn=<NllLossBackward0>)
tensor(0.2364, grad_fn=<NllLossBackward0>)
tensor(0.2349, grad_fn=<NllLossBackward0>)
tensor(0.2334, grad_fn=<NllLossBackward0>)
tensor(0.2320, grad_fn=<NllLossBackward0>)
tensor(0.2305, grad_fn=<NllLossBackward0>)
tensor(0.2291, grad_fn=<NllLossBackward0>)
tensor(0.2277, grad_fn=<NllLossBackward0>)
tensor(0.2263, grad_fn=<NllLossBackward0>)
tensor(0.2250, grad_fn=<NllLossBackward0>)
tensor(0.2236, grad_fn=<NllLossBackward0>)
tensor(0.2223, grad_fn=<NllLossBackward0>)
tensor(0.2209, grad_fn=<NllLossBackward0>)
tensor(0.2196, grad_fn=<NllLossBackward0>)
tensor(0.2183, grad_fn=<NllLossBackward0>)
tensor(0.2170, grad_fn=<NllLossBackward0>)
tensor(0.2158, grad_fn=<NllLossBackward0>)
tensor(0.2145, grad_fn=<NllLossBackward0>)
tensor(0.2133, grad_fn=<NllLossBackward0>)
tensor(0.2121, grad_fn=<NllLossBackward0>)
tensor(0.2108, grad_fn=<NllLossBackward0>)
tensor(0.2096, grad_fn=<NllLossBackward0>)
tensor(0.2085, grad_fn=<NllLossBackward0>)
tensor(0.2073, grad_fn=<NllLossBackward0>)
```

```
tensor(0.2061, grad_fn=<NllLossBackward0>)
tensor(0.2050, grad_fn=<NllLossBackward0>)
tensor(0.2038, grad_fn=<NllLossBackward0>)
tensor(0.2027, grad_fn=<NllLossBackward0>)
tensor(0.2016, grad_fn=<NllLossBackward0>)
tensor(0.2005, grad_fn=<NllLossBackward0>)
tensor(0.1994, grad_fn=<NllLossBackward0>)
tensor(0.1983, grad_fn=<NllLossBackward0>)
tensor(0.1973, grad_fn=<NllLossBackward0>)
tensor(0.1962, grad_fn=<NllLossBackward0>)
tensor(0.1952, grad_fn=<NllLossBackward0>)
tensor(0.1941, grad_fn=<NllLossBackward0>)
tensor(0.1931, grad_fn=<NllLossBackward0>)
tensor(0.1921, grad_fn=<NllLossBackward0>)
tensor(0.1911, grad_fn=<NllLossBackward0>)
tensor(0.1901, grad_fn=<NllLossBackward0>)
tensor(0.1891, grad_fn=<NllLossBackward0>)
tensor(0.1881, grad_fn=<NllLossBackward0>)
tensor(0.1872, grad_fn=<NllLossBackward0>)
tensor(0.1862, grad_fn=<NllLossBackward0>)
tensor(0.1853, grad_fn=<NllLossBackward0>)
tensor(0.1843, grad_fn=<NllLossBackward0>)
tensor(0.1834, grad_fn=<NllLossBackward0>)
tensor(0.1825, grad_fn=<NllLossBackward0>)
tensor(0.1816, grad_fn=<NllLossBackward0>)
tensor(0.1807, grad_fn=<NllLossBackward0>)
tensor(0.1798, grad_fn=<NllLossBackward0>)
tensor(0.1789, grad_fn=<NllLossBackward0>)
tensor(0.1780, grad_fn=<NllLossBackward0>)
tensor(0.1772, grad_fn=<NllLossBackward0>)
tensor(0.1763, grad_fn=<NllLossBackward0>)
tensor(0.1754, grad_fn=<NllLossBackward0>)
tensor(0.1746, grad_fn=<NllLossBackward0>)
tensor(0.1738, grad_fn=<NllLossBackward0>)
tensor(0.1729, grad_fn=<NllLossBackward0>)
tensor(0.1721, grad_fn=<NllLossBackward0>)
tensor(0.1713, grad_fn=<NllLossBackward0>)
tensor(0.1705, grad_fn=<NllLossBackward0>)
tensor(0.1697, grad_fn=<NllLossBackward0>)
tensor(0.1689, grad_fn=<NllLossBackward0>)
tensor(0.1681, grad_fn=<NllLossBackward0>)
tensor(0.1674, grad_fn=<NllLossBackward0>)
tensor(0.1666, grad_fn=<NllLossBackward0>)
tensor(0.1658, grad_fn=<NllLossBackward0>)
tensor(0.1651, grad_fn=<NllLossBackward0>)
tensor(0.1643, grad_fn=<NllLossBackward0>)
tensor(0.1636, grad_fn=<NllLossBackward0>)
tensor(0.1628, grad_fn=<NllLossBackward0>)
```

```
tensor(0.1621, grad_fn=<NllLossBackward0>)
tensor(0.1614, grad_fn=<NllLossBackward0>)
tensor(0.1607, grad_fn=<NllLossBackward0>)
tensor(0.1600, grad_fn=<NllLossBackward0>)
tensor(0.1593, grad_fn=<NllLossBackward0>)
tensor(0.1586, grad_fn=<NllLossBackward0>)
tensor(0.1579, grad_fn=<NllLossBackward0>)
tensor(0.1572, grad_fn=<NllLossBackward0>)
tensor(0.1565, grad_fn=<NllLossBackward0>)
tensor(0.1558, grad_fn=<NllLossBackward0>)
tensor(0.1552, grad_fn=<NllLossBackward0>)
tensor(0.1545, grad_fn=<NllLossBackward0>)
tensor(0.1538, grad_fn=<NllLossBackward0>)
tensor(0.1532, grad_fn=<NllLossBackward0>)
tensor(0.1525, grad_fn=<NllLossBackward0>)
tensor(0.1519, grad_fn=<NllLossBackward0>)
tensor(0.1513, grad_fn=<NllLossBackward0>)
tensor(0.1506, grad_fn=<NllLossBackward0>)
tensor(0.1500, grad_fn=<NllLossBackward0>)
tensor(0.1494, grad_fn=<NllLossBackward0>)
tensor(0.1488, grad_fn=<NllLossBackward0>)
tensor(0.1481, grad_fn=<NllLossBackward0>)
tensor(0.1475, grad_fn=<NllLossBackward0>)
tensor(0.1469, grad_fn=<NllLossBackward0>)
tensor(0.1463, grad_fn=<NllLossBackward0>)
tensor(0.1457, grad_fn=<NllLossBackward0>)
tensor(0.1452, grad_fn=<NllLossBackward0>)
tensor(0.1446, grad_fn=<NllLossBackward0>)
tensor(0.1440, grad_fn=<NllLossBackward0>)
tensor(0.1434, grad_fn=<NllLossBackward0>)
tensor(0.1429, grad_fn=<NllLossBackward0>)
tensor(0.1423, grad_fn=<NllLossBackward0>)
tensor(0.1417, grad_fn=<NllLossBackward0>)
tensor(0.1412, grad_fn=<NllLossBackward0>)
tensor(0.1406, grad_fn=<NllLossBackward0>)
tensor(0.1401, grad_fn=<NllLossBackward0>)
tensor(0.1395, grad_fn=<NllLossBackward0>)
tensor(0.1390, grad_fn=<NllLossBackward0>)
tensor(0.1385, grad_fn=<NllLossBackward0>)
tensor(0.1379, grad_fn=<NllLossBackward0>)
tensor(0.1374, grad_fn=<NllLossBackward0>)
tensor(0.1369, grad_fn=<NllLossBackward0>)
tensor(0.1364, grad_fn=<NllLossBackward0>)
tensor(0.1358, grad_fn=<NllLossBackward0>)
tensor(0.1353, grad_fn=<NllLossBackward0>)
tensor(0.1348, grad_fn=<NllLossBackward0>)
tensor(0.1343, grad_fn=<NllLossBackward0>)
tensor(0.1338, grad_fn=<NllLossBackward0>)
```

```
tensor(0.1333, grad_fn=<NllLossBackward0>)
tensor(0.1328, grad_fn=<NllLossBackward0>)
tensor(0.1323, grad_fn=<NllLossBackward0>)
tensor(0.1319, grad_fn=<NllLossBackward0>)
tensor(0.1314, grad_fn=<NllLossBackward0>)
tensor(0.1309, grad_fn=<NllLossBackward0>)
tensor(0.1304, grad_fn=<NllLossBackward0>)
tensor(0.1300, grad_fn=<NllLossBackward0>)
tensor(0.1295, grad_fn=<NllLossBackward0>)
tensor(0.1290, grad_fn=<NllLossBackward0>)
tensor(0.1286, grad_fn=<NllLossBackward0>)
tensor(0.1281, grad_fn=<NllLossBackward0>)
tensor(0.1276, grad_fn=<NllLossBackward0>)
tensor(0.1272, grad_fn=<NllLossBackward0>)
tensor(0.1267, grad_fn=<NllLossBackward0>)
tensor(0.1263, grad_fn=<NllLossBackward0>)
tensor(0.1259, grad_fn=<NllLossBackward0>)
tensor(0.1254, grad_fn=<NllLossBackward0>)
tensor(0.1250, grad_fn=<NllLossBackward0>)
tensor(0.1246, grad_fn=<NllLossBackward0>)
tensor(0.1241, grad_fn=<NllLossBackward0>)
tensor(0.1237, grad_fn=<NllLossBackward0>)
tensor(0.1233, grad_fn=<NllLossBackward0>)
tensor(0.1228, grad_fn=<NllLossBackward0>)
tensor(0.1224, grad_fn=<NllLossBackward0>)
tensor(0.1220, grad_fn=<NllLossBackward0>)
tensor(0.1216, grad_fn=<NllLossBackward0>)
tensor(0.1212, grad_fn=<NllLossBackward0>)
tensor(0.1208, grad_fn=<NllLossBackward0>)
tensor(0.1204, grad_fn=<NllLossBackward0>)
tensor(0.1200, grad_fn=<NllLossBackward0>)
tensor(0.1196, grad_fn=<NllLossBackward0>)
tensor(0.1192, grad_fn=<NllLossBackward0>)
tensor(0.1188, grad_fn=<NllLossBackward0>)
tensor(0.1184, grad_fn=<NllLossBackward0>)
tensor(0.1180, grad_fn=<NllLossBackward0>)
tensor(0.1176, grad_fn=<NllLossBackward0>)
tensor(0.1172, grad_fn=<NllLossBackward0>)
tensor(0.1169, grad_fn=<NllLossBackward0>)
tensor(0.1165, grad_fn=<NllLossBackward0>)
tensor(0.1161, grad_fn=<NllLossBackward0>)
tensor(0.1157, grad_fn=<NllLossBackward0>)
tensor(0.1154, grad_fn=<NllLossBackward0>)
tensor(0.1150, grad_fn=<NllLossBackward0>)
tensor(0.1146, grad_fn=<NllLossBackward0>)
tensor(0.1142, grad_fn=<NllLossBackward0>)
tensor(0.1139, grad_fn=<NllLossBackward0>)
tensor(0.1135, grad_fn=<NllLossBackward0>)
```

```
tensor(0.1132, grad_fn=<NllLossBackward0>)
tensor(0.1128, grad_fn=<NllLossBackward0>)
tensor(0.1125, grad_fn=<NllLossBackward0>)
tensor(0.1121, grad_fn=<NllLossBackward0>)
tensor(0.1118, grad_fn=<NllLossBackward0>)
tensor(0.1114, grad_fn=<NllLossBackward0>)
tensor(0.1111, grad_fn=<NllLossBackward0>)
tensor(0.1107, grad_fn=<NllLossBackward0>)
tensor(0.1104, grad_fn=<NllLossBackward0>)
tensor(0.1100, grad_fn=<NllLossBackward0>)
tensor(0.1097, grad_fn=<NllLossBackward0>)
tensor(0.1094, grad_fn=<NllLossBackward0>)
tensor(0.1090, grad_fn=<NllLossBackward0>)
tensor(0.1087, grad_fn=<NllLossBackward0>)
tensor(0.1084, grad_fn=<NllLossBackward0>)
tensor(0.1081, grad_fn=<NllLossBackward0>)
tensor(0.1077, grad_fn=<NllLossBackward0>)
tensor(0.1074, grad_fn=<NllLossBackward0>)
tensor(0.1071, grad_fn=<NllLossBackward0>)
tensor(0.1068, grad_fn=<NllLossBackward0>)
tensor(0.1065, grad_fn=<NllLossBackward0>)
tensor(0.1061, grad_fn=<NllLossBackward0>)
tensor(0.1058, grad_fn=<NllLossBackward0>)
tensor(0.1055, grad_fn=<NllLossBackward0>)
tensor(0.1052, grad_fn=<NllLossBackward0>)
tensor(0.1049, grad_fn=<NllLossBackward0>)
tensor(0.1046, grad_fn=<NllLossBackward0>)
tensor(0.1043, grad_fn=<NllLossBackward0>)
tensor(0.1040, grad_fn=<NllLossBackward0>)
tensor(0.1037, grad_fn=<NllLossBackward0>)
tensor(0.1034, grad_fn=<NllLossBackward0>)
tensor(0.1031, grad_fn=<NllLossBackward0>)
tensor(0.1028, grad_fn=<NllLossBackward0>)
tensor(0.1025, grad_fn=<NllLossBackward0>)
tensor(0.1022, grad_fn=<NllLossBackward0>)
tensor(0.1019, grad_fn=<NllLossBackward0>)
tensor(0.1016, grad_fn=<NllLossBackward0>)
tensor(0.1014, grad_fn=<NllLossBackward0>)
tensor(0.1011, grad_fn=<NllLossBackward0>)
tensor(0.1008, grad_fn=<NllLossBackward0>)
tensor(0.1005, grad_fn=<NllLossBackward0>)
tensor(0.1002, grad_fn=<NllLossBackward0>)
tensor(0.0999, grad_fn=<NllLossBackward0>)
tensor(0.0997, grad_fn=<NllLossBackward0>)
tensor(0.0994, grad_fn=<NllLossBackward0>)
tensor(0.0991, grad_fn=<NllLossBackward0>)
tensor(0.0988, grad_fn=<NllLossBackward0>)
tensor(0.0986, grad_fn=<NllLossBackward0>)
```

```
tensor(0.0983, grad_fn=<NllLossBackward0>)
tensor(0.0980, grad_fn=<NllLossBackward0>)
tensor(0.0978, grad_fn=<NllLossBackward0>)
tensor(0.0975, grad_fn=<NllLossBackward0>)
tensor(0.0972, grad_fn=<NllLossBackward0>)
tensor(0.0970, grad_fn=<NllLossBackward0>)
tensor(0.0967, grad_fn=<NllLossBackward0>)
tensor(0.0965, grad_fn=<NllLossBackward0>)
tensor(0.0962, grad_fn=<NllLossBackward0>)
tensor(0.0959, grad_fn=<NllLossBackward0>)
tensor(0.0957, grad_fn=<NllLossBackward0>)
tensor(0.0954, grad_fn=<NllLossBackward0>)
tensor(0.0952, grad_fn=<NllLossBackward0>)
tensor(0.0949, grad_fn=<NllLossBackward0>)
tensor(0.0947, grad_fn=<NllLossBackward0>)
tensor(0.0944, grad_fn=<NllLossBackward0>)
tensor(0.0942, grad_fn=<NllLossBackward0>)
tensor(0.0939, grad_fn=<NllLossBackward0>)
tensor(0.0937, grad_fn=<NllLossBackward0>)
tensor(0.0935, grad_fn=<NllLossBackward0>)
tensor(0.0932, grad_fn=<NllLossBackward0>)
tensor(0.0930, grad_fn=<NllLossBackward0>)
tensor(0.0927, grad_fn=<NllLossBackward0>)
tensor(0.0925, grad_fn=<NllLossBackward0>)
tensor(0.0923, grad_fn=<NllLossBackward0>)
tensor(0.0920, grad_fn=<NllLossBackward0>)
tensor(0.0918, grad_fn=<NllLossBackward0>)
tensor(0.0915, grad_fn=<NllLossBackward0>)
tensor(0.0913, grad_fn=<NllLossBackward0>)
tensor(0.0911, grad_fn=<NllLossBackward0>)
tensor(0.0909, grad_fn=<NllLossBackward0>)
tensor(0.0906, grad_fn=<NllLossBackward0>)
tensor(0.0904, grad_fn=<NllLossBackward0>)
tensor(0.0902, grad_fn=<NllLossBackward0>)
tensor(0.0899, grad_fn=<NllLossBackward0>)
tensor(0.0897, grad_fn=<NllLossBackward0>)
tensor(0.0895, grad_fn=<NllLossBackward0>)
tensor(0.0893, grad_fn=<NllLossBackward0>)
tensor(0.0891, grad_fn=<NllLossBackward0>)
tensor(0.0888, grad_fn=<NllLossBackward0>)
tensor(0.0886, grad_fn=<NllLossBackward0>)
tensor(0.0884, grad_fn=<NllLossBackward0>)
tensor(0.0882, grad_fn=<NllLossBackward0>)
tensor(0.0880, grad_fn=<NllLossBackward0>)
tensor(0.0878, grad_fn=<NllLossBackward0>)
tensor(0.0875, grad_fn=<NllLossBackward0>)
tensor(0.0873, grad_fn=<NllLossBackward0>)
tensor(0.0871, grad_fn=<NllLossBackward0>)
```

```
tensor(0.0869, grad_fn=<NllLossBackward0>)
tensor(0.0867, grad_fn=<NllLossBackward0>)
tensor(0.0865, grad_fn=<NllLossBackward0>)
tensor(0.0863, grad_fn=<NllLossBackward0>)
tensor(0.0861, grad_fn=<NllLossBackward0>)
tensor(0.0859, grad_fn=<NllLossBackward0>)
tensor(0.0857, grad_fn=<NllLossBackward0>)
tensor(0.0855, grad_fn=<NllLossBackward0>)
tensor(0.0853, grad_fn=<NllLossBackward0>)
tensor(0.0851, grad_fn=<NllLossBackward0>)
tensor(0.0849, grad_fn=<NllLossBackward0>)
tensor(0.0847, grad_fn=<NllLossBackward0>)
tensor(0.0845, grad_fn=<NllLossBackward0>)
tensor(0.0843, grad_fn=<NllLossBackward0>)
tensor(0.0841, grad_fn=<NllLossBackward0>)
tensor(0.0839, grad_fn=<NllLossBackward0>)
tensor(0.0837, grad_fn=<NllLossBackward0>)
tensor(0.0835, grad_fn=<NllLossBackward0>)
tensor(0.0833, grad_fn=<NllLossBackward0>)
tensor(0.0831, grad_fn=<NllLossBackward0>)
tensor(0.0829, grad_fn=<NllLossBackward0>)
tensor(0.0827, grad_fn=<NllLossBackward0>)
tensor(0.0825, grad_fn=<NllLossBackward0>)
tensor(0.0823, grad_fn=<NllLossBackward0>)
tensor(0.0822, grad_fn=<NllLossBackward0>)
tensor(0.0820, grad_fn=<NllLossBackward0>)
tensor(0.0818, grad_fn=<NllLossBackward0>)
tensor(0.0816, grad_fn=<NllLossBackward0>)
tensor(0.0814, grad_fn=<NllLossBackward0>)
tensor(0.0812, grad_fn=<NllLossBackward0>)
tensor(0.0811, grad_fn=<NllLossBackward0>)
tensor(0.0809, grad_fn=<NllLossBackward0>)
tensor(0.0807, grad_fn=<NllLossBackward0>)
tensor(0.0805, grad_fn=<NllLossBackward0>)
tensor(0.0803, grad_fn=<NllLossBackward0>)
tensor(0.0802, grad_fn=<NllLossBackward0>)
tensor(0.0800, grad_fn=<NllLossBackward0>)
tensor(0.0798, grad_fn=<NllLossBackward0>)
tensor(0.0796, grad_fn=<NllLossBackward0>)
tensor(0.0795, grad_fn=<NllLossBackward0>)
tensor(0.0793, grad_fn=<NllLossBackward0>)
tensor(0.0791, grad_fn=<NllLossBackward0>)
tensor(0.0789, grad_fn=<NllLossBackward0>)
tensor(0.0788, grad_fn=<NllLossBackward0>)
tensor(0.0786, grad_fn=<NllLossBackward0>)
tensor(0.0784, grad_fn=<NllLossBackward0>)
tensor(0.0782, grad_fn=<NllLossBackward0>)
tensor(0.0781, grad_fn=<NllLossBackward0>)
```

```
tensor(0.0779, grad_fn=<NllLossBackward0>)
tensor(0.0777, grad_fn=<NllLossBackward0>)
tensor(0.0776, grad_fn=<NllLossBackward0>)
tensor(0.0774, grad_fn=<NllLossBackward0>)
tensor(0.0772, grad_fn=<NllLossBackward0>)
tensor(0.0771, grad_fn=<NllLossBackward0>)
tensor(0.0769, grad_fn=<NllLossBackward0>)
tensor(0.0768, grad_fn=<NllLossBackward0>)
tensor(0.0766, grad_fn=<NllLossBackward0>)
tensor(0.0764, grad_fn=<NllLossBackward0>)
tensor(0.0763, grad_fn=<NllLossBackward0>)
tensor(0.0761, grad_fn=<NllLossBackward0>)
tensor(0.0759, grad_fn=<NllLossBackward0>)
tensor(0.0758, grad_fn=<NllLossBackward0>)
tensor(0.0756, grad_fn=<NllLossBackward0>)
tensor(0.0755, grad_fn=<NllLossBackward0>)
tensor(0.0753, grad_fn=<NllLossBackward0>)
tensor(0.0752, grad_fn=<NllLossBackward0>)
tensor(0.0750, grad_fn=<NllLossBackward0>)
tensor(0.0749, grad_fn=<NllLossBackward0>)
tensor(0.0747, grad_fn=<NllLossBackward0>)
tensor(0.0745, grad_fn=<NllLossBackward0>)
tensor(0.0744, grad_fn=<NllLossBackward0>)
tensor(0.0742, grad_fn=<NllLossBackward0>)
tensor(0.0741, grad_fn=<NllLossBackward0>)
tensor(0.0739, grad_fn=<NllLossBackward0>)
tensor(0.0738, grad_fn=<NllLossBackward0>)
tensor(0.0736, grad_fn=<NllLossBackward0>)
tensor(0.0735, grad_fn=<NllLossBackward0>)
tensor(0.0733, grad_fn=<NllLossBackward0>)
tensor(0.0732, grad_fn=<NllLossBackward0>)
tensor(0.0730, grad_fn=<NllLossBackward0>)
tensor(0.0729, grad_fn=<NllLossBackward0>)
tensor(0.0727, grad_fn=<NllLossBackward0>)
tensor(0.0726, grad_fn=<NllLossBackward0>)
tensor(0.0725, grad_fn=<NllLossBackward0>)
tensor(0.0723, grad_fn=<NllLossBackward0>)
tensor(0.0722, grad_fn=<NllLossBackward0>)
tensor(0.0720, grad_fn=<NllLossBackward0>)
tensor(0.0719, grad_fn=<NllLossBackward0>)
tensor(0.0717, grad_fn=<NllLossBackward0>)
tensor(0.0716, grad_fn=<NllLossBackward0>)
tensor(0.0715, grad_fn=<NllLossBackward0>)
tensor(0.0713, grad_fn=<NllLossBackward0>)
tensor(0.0712, grad_fn=<NllLossBackward0>)
tensor(0.0710, grad_fn=<NllLossBackward0>)
tensor(0.0709, grad_fn=<NllLossBackward0>)
tensor(0.0708, grad_fn=<NllLossBackward0>)
```

```
tensor(0.0706, grad_fn=<NllLossBackward0>)
tensor(0.0705, grad_fn=<NllLossBackward0>)
tensor(0.0704, grad_fn=<NllLossBackward0>)
tensor(0.0702, grad_fn=<NllLossBackward0>)
tensor(0.0701, grad_fn=<NllLossBackward0>)
tensor(0.0699, grad_fn=<NllLossBackward0>)
tensor(0.0698, grad_fn=<NllLossBackward0>)
tensor(0.0697, grad_fn=<NllLossBackward0>)
tensor(0.0695, grad_fn=<NllLossBackward0>)
tensor(0.0694, grad_fn=<NllLossBackward0>)
tensor(0.0693, grad_fn=<NllLossBackward0>)
tensor(0.0691, grad_fn=<NllLossBackward0>)
tensor(0.0690, grad_fn=<NllLossBackward0>)
tensor(0.0689, grad_fn=<NllLossBackward0>)
tensor(0.0688, grad_fn=<NllLossBackward0>)
tensor(0.0686, grad_fn=<NllLossBackward0>)
tensor(0.0685, grad_fn=<NllLossBackward0>)
tensor(0.0684, grad_fn=<NllLossBackward0>)
tensor(0.0682, grad_fn=<NllLossBackward0>)
tensor(0.0681, grad_fn=<NllLossBackward0>)
tensor(0.0680, grad_fn=<NllLossBackward0>)
tensor(0.0679, grad_fn=<NllLossBackward0>)
tensor(0.0677, grad_fn=<NllLossBackward0>)
tensor(0.0676, grad_fn=<NllLossBackward0>)
tensor(0.0675, grad_fn=<NllLossBackward0>)
tensor(0.0674, grad_fn=<NllLossBackward0>)
tensor(0.0672, grad_fn=<NllLossBackward0>)
tensor(0.0671, grad_fn=<NllLossBackward0>)
tensor(0.0670, grad_fn=<NllLossBackward0>)
tensor(0.0669, grad_fn=<NllLossBackward0>)
tensor(0.0667, grad_fn=<NllLossBackward0>)
tensor(0.0666, grad_fn=<NllLossBackward0>)
tensor(0.0665, grad_fn=<NllLossBackward0>)
tensor(0.0664, grad_fn=<NllLossBackward0>)
tensor(0.0663, grad_fn=<NllLossBackward0>)
tensor(0.0661, grad_fn=<NllLossBackward0>)
tensor(0.0660, grad_fn=<NllLossBackward0>)
tensor(0.0659, grad_fn=<NllLossBackward0>)
tensor(0.0658, grad_fn=<NllLossBackward0>)
tensor(0.0657, grad_fn=<NllLossBackward0>)
tensor(0.0655, grad_fn=<NllLossBackward0>)
tensor(0.0654, grad_fn=<NllLossBackward0>)
tensor(0.0653, grad_fn=<NllLossBackward0>)
tensor(0.0652, grad_fn=<NllLossBackward0>)
tensor(0.0651, grad_fn=<NllLossBackward0>)
tensor(0.0650, grad_fn=<NllLossBackward0>)
tensor(0.0648, grad_fn=<NllLossBackward0>)
tensor(0.0647, grad_fn=<NllLossBackward0>)
```

```
tensor(0.0646, grad_fn=<NllLossBackward0>)
tensor(0.0645, grad_fn=<NllLossBackward0>)
tensor(0.0644, grad_fn=<NllLossBackward0>)
tensor(0.0643, grad_fn=<NllLossBackward0>)
tensor(0.0642, grad_fn=<NllLossBackward0>)
tensor(0.0640, grad_fn=<NllLossBackward0>)
tensor(0.0639, grad_fn=<NllLossBackward0>)
tensor(0.0638, grad_fn=<NllLossBackward0>)
tensor(0.0637, grad_fn=<NllLossBackward0>)
tensor(0.0636, grad_fn=<NllLossBackward0>)
tensor(0.0635, grad_fn=<NllLossBackward0>)
tensor(0.0634, grad_fn=<NllLossBackward0>)
tensor(0.0633, grad_fn=<NllLossBackward0>)
tensor(0.0632, grad_fn=<NllLossBackward0>)
tensor(0.0630, grad_fn=<NllLossBackward0>)
tensor(0.0629, grad_fn=<NllLossBackward0>)
tensor(0.0628, grad_fn=<NllLossBackward0>)
tensor(0.0627, grad_fn=<NllLossBackward0>)
tensor(0.0626, grad_fn=<NllLossBackward0>)
tensor(0.0625, grad_fn=<NllLossBackward0>)
tensor(0.0624, grad_fn=<NllLossBackward0>)
tensor(0.0623, grad_fn=<NllLossBackward0>)
tensor(0.0622, grad_fn=<NllLossBackward0>)
tensor(0.0621, grad_fn=<NllLossBackward0>)
tensor(0.0620, grad_fn=<NllLossBackward0>)
tensor(0.0619, grad_fn=<NllLossBackward0>)
tensor(0.0618, grad_fn=<NllLossBackward0>)
tensor(0.0617, grad_fn=<NllLossBackward0>)
tensor(0.0616, grad_fn=<NllLossBackward0>)
tensor(0.0615, grad_fn=<NllLossBackward0>)
tensor(0.0614, grad_fn=<NllLossBackward0>)
tensor(0.0612, grad_fn=<NllLossBackward0>)
tensor(0.0611, grad_fn=<NllLossBackward0>)
tensor(0.0610, grad_fn=<NllLossBackward0>)
tensor(0.0609, grad_fn=<NllLossBackward0>)
tensor(0.0608, grad_fn=<NllLossBackward0>)
tensor(0.0607, grad_fn=<NllLossBackward0>)
tensor(0.0606, grad_fn=<NllLossBackward0>)
tensor(0.0605, grad_fn=<NllLossBackward0>)
tensor(0.0604, grad_fn=<NllLossBackward0>)
tensor(0.0603, grad_fn=<NllLossBackward0>)
tensor(0.0602, grad_fn=<NllLossBackward0>)
tensor(0.0601, grad_fn=<NllLossBackward0>)
tensor(0.0600, grad_fn=<NllLossBackward0>)
tensor(0.0599, grad_fn=<NllLossBackward0>)
tensor(0.0598, grad_fn=<NllLossBackward0>)
tensor(0.0597, grad_fn=<NllLossBackward0>)
tensor(0.0597, grad_fn=<NllLossBackward0>)
```

```
tensor(0.0596, grad_fn=<NllLossBackward0>)
tensor(0.0595, grad_fn=<NllLossBackward0>)
tensor(0.0594, grad_fn=<NllLossBackward0>)
tensor(0.0593, grad_fn=<NllLossBackward0>)
tensor(0.0592, grad_fn=<NllLossBackward0>)
tensor(0.0591, grad_fn=<NllLossBackward0>)
tensor(0.0590, grad_fn=<NllLossBackward0>)
tensor(0.0589, grad_fn=<NllLossBackward0>)
tensor(0.0588, grad_fn=<NllLossBackward0>)
tensor(0.0587, grad_fn=<NllLossBackward0>)
tensor(0.0586, grad_fn=<NllLossBackward0>)
tensor(0.0585, grad_fn=<NllLossBackward0>)
tensor(0.0584, grad_fn=<NllLossBackward0>)
tensor(0.0583, grad_fn=<NllLossBackward0>)
tensor(0.0582, grad_fn=<NllLossBackward0>)
tensor(0.0581, grad_fn=<NllLossBackward0>)
tensor(0.0580, grad_fn=<NllLossBackward0>)
tensor(0.0580, grad_fn=<NllLossBackward0>)
tensor(0.0579, grad_fn=<NllLossBackward0>)
tensor(0.0578, grad_fn=<NllLossBackward0>)
tensor(0.0577, grad_fn=<NllLossBackward0>)
tensor(0.0576, grad_fn=<NllLossBackward0>)
tensor(0.0575, grad_fn=<NllLossBackward0>)
tensor(0.0574, grad_fn=<NllLossBackward0>)
tensor(0.0573, grad_fn=<NllLossBackward0>)
tensor(0.0572, grad_fn=<NllLossBackward0>)
tensor(0.0571, grad_fn=<NllLossBackward0>)
tensor(0.0571, grad_fn=<NllLossBackward0>)
tensor(0.0570, grad_fn=<NllLossBackward0>)
tensor(0.0569, grad_fn=<NllLossBackward0>)
tensor(0.0568, grad_fn=<NllLossBackward0>)
tensor(0.0567, grad_fn=<NllLossBackward0>)
tensor(0.0566, grad_fn=<NllLossBackward0>)
tensor(0.0565, grad_fn=<NllLossBackward0>)
tensor(0.0564, grad_fn=<NllLossBackward0>)
tensor(0.0564, grad_fn=<NllLossBackward0>)
tensor(0.0563, grad_fn=<NllLossBackward0>)
tensor(0.0562, grad_fn=<NllLossBackward0>)
tensor(0.0561, grad_fn=<NllLossBackward0>)
tensor(0.0560, grad_fn=<NllLossBackward0>)
tensor(0.0559, grad_fn=<NllLossBackward0>)
tensor(0.0558, grad_fn=<NllLossBackward0>)
tensor(0.0558, grad_fn=<NllLossBackward0>)
tensor(0.0557, grad_fn=<NllLossBackward0>)
tensor(0.0556, grad_fn=<NllLossBackward0>)
tensor(0.0555, grad_fn=<NllLossBackward0>)
tensor(0.0554, grad_fn=<NllLossBackward0>)
tensor(0.0553, grad_fn=<NllLossBackward0>)
```

```
tensor(0.0553, grad_fn=<NllLossBackward0>)
tensor(0.0552, grad_fn=<NllLossBackward0>)
tensor(0.0551, grad_fn=<NllLossBackward0>)
tensor(0.0550, grad_fn=<NllLossBackward0>)
tensor(0.0549, grad_fn=<NllLossBackward0>)
tensor(0.0548, grad_fn=<NllLossBackward0>)
tensor(0.0548, grad_fn=<NllLossBackward0>)
tensor(0.0547, grad_fn=<NllLossBackward0>)
tensor(0.0546, grad_fn=<NllLossBackward0>)
tensor(0.0545, grad_fn=<NllLossBackward0>)
tensor(0.0544, grad_fn=<NllLossBackward0>)
tensor(0.0544, grad_fn=<NllLossBackward0>)
tensor(0.0543, grad_fn=<NllLossBackward0>)
tensor(0.0542, grad_fn=<NllLossBackward0>)
tensor(0.0541, grad_fn=<NllLossBackward0>)
tensor(0.0540, grad_fn=<NllLossBackward0>)
tensor(0.0540, grad_fn=<NllLossBackward0>)
tensor(0.0539, grad_fn=<NllLossBackward0>)
tensor(0.0538, grad_fn=<NllLossBackward0>)
tensor(0.0537, grad_fn=<NllLossBackward0>)
tensor(0.0536, grad_fn=<NllLossBackward0>)
tensor(0.0536, grad_fn=<NllLossBackward0>)
tensor(0.0535, grad_fn=<NllLossBackward0>)
tensor(0.0534, grad_fn=<NllLossBackward0>)
tensor(0.0533, grad_fn=<NllLossBackward0>)
tensor(0.0533, grad_fn=<NllLossBackward0>)
tensor(0.0532, grad_fn=<NllLossBackward0>)
tensor(0.0531, grad_fn=<NllLossBackward0>)
tensor(0.0530, grad_fn=<NllLossBackward0>)
tensor(0.0530, grad_fn=<NllLossBackward0>)
tensor(0.0529, grad_fn=<NllLossBackward0>)
tensor(0.0528, grad_fn=<NllLossBackward0>)
tensor(0.0527, grad_fn=<NllLossBackward0>)
tensor(0.0526, grad_fn=<NllLossBackward0>)
tensor(0.0526, grad_fn=<NllLossBackward0>)
tensor(0.0525, grad_fn=<NllLossBackward0>)
tensor(0.0524, grad_fn=<NllLossBackward0>)
tensor(0.0523, grad_fn=<NllLossBackward0>)
tensor(0.0523, grad_fn=<NllLossBackward0>)
tensor(0.0522, grad_fn=<NllLossBackward0>)
tensor(0.0521, grad_fn=<NllLossBackward0>)
tensor(0.0521, grad_fn=<NllLossBackward0>)
tensor(0.0520, grad_fn=<NllLossBackward0>)
tensor(0.0519, grad_fn=<NllLossBackward0>)
tensor(0.0518, grad_fn=<NllLossBackward0>)
tensor(0.0518, grad_fn=<NllLossBackward0>)
tensor(0.0517, grad_fn=<NllLossBackward0>)
tensor(0.0516, grad_fn=<NllLossBackward0>)
```

```
tensor(0.0515, grad_fn=<NllLossBackward0>)
tensor(0.0515, grad_fn=<NllLossBackward0>)
tensor(0.0514, grad_fn=<NllLossBackward0>)
tensor(0.0513, grad_fn=<NllLossBackward0>)
tensor(0.0513, grad_fn=<NllLossBackward0>)
tensor(0.0512, grad_fn=<NllLossBackward0>)
tensor(0.0511, grad_fn=<NllLossBackward0>)
tensor(0.0510, grad_fn=<NllLossBackward0>)
tensor(0.0510, grad_fn=<NllLossBackward0>)
tensor(0.0509, grad_fn=<NllLossBackward0>)
tensor(0.0508, grad_fn=<NllLossBackward0>)
tensor(0.0508, grad_fn=<NllLossBackward0>)
tensor(0.0507, grad_fn=<NllLossBackward0>)
tensor(0.0506, grad_fn=<NllLossBackward0>)
tensor(0.0506, grad_fn=<NllLossBackward0>)
tensor(0.0505, grad_fn=<NllLossBackward0>)
tensor(0.0504, grad_fn=<NllLossBackward0>)
tensor(0.0504, grad_fn=<NllLossBackward0>)
tensor(0.0503, grad_fn=<NllLossBackward0>)
tensor(0.0502, grad_fn=<NllLossBackward0>)
tensor(0.0501, grad_fn=<NllLossBackward0>)
tensor(0.0501, grad_fn=<NllLossBackward0>)
tensor(0.0500, grad_fn=<NllLossBackward0>)
tensor(0.0499, grad_fn=<NllLossBackward0>)
tensor(0.0499, grad_fn=<NllLossBackward0>)
tensor(0.0498, grad_fn=<NllLossBackward0>)
tensor(0.0497, grad_fn=<NllLossBackward0>)
tensor(0.0497, grad_fn=<NllLossBackward0>)
tensor(0.0496, grad_fn=<NllLossBackward0>)
tensor(0.0495, grad_fn=<NllLossBackward0>)
tensor(0.0495, grad_fn=<NllLossBackward0>)
tensor(0.0494, grad_fn=<NllLossBackward0>)
tensor(0.0493, grad_fn=<NllLossBackward0>)
tensor(0.0493, grad_fn=<NllLossBackward0>)
tensor(0.0492, grad_fn=<NllLossBackward0>)
tensor(0.0492, grad_fn=<NllLossBackward0>)
tensor(0.0491, grad_fn=<NllLossBackward0>)
tensor(0.0490, grad_fn=<NllLossBackward0>)
tensor(0.0490, grad_fn=<NllLossBackward0>)
tensor(0.0489, grad_fn=<NllLossBackward0>)
tensor(0.0488, grad_fn=<NllLossBackward0>)
tensor(0.0488, grad_fn=<NllLossBackward0>)
tensor(0.0487, grad_fn=<NllLossBackward0>)
tensor(0.0486, grad_fn=<NllLossBackward0>)
tensor(0.0486, grad_fn=<NllLossBackward0>)
tensor(0.0485, grad_fn=<NllLossBackward0>)
tensor(0.0484, grad_fn=<NllLossBackward0>)
tensor(0.0484, grad_fn=<NllLossBackward0>)
```

```
tensor(0.0483, grad_fn=<NllLossBackward0>)
tensor(0.0483, grad_fn=<NllLossBackward0>)
tensor(0.0482, grad_fn=<NllLossBackward0>)
tensor(0.0481, grad_fn=<NllLossBackward0>)
tensor(0.0481, grad_fn=<NllLossBackward0>)
tensor(0.0480, grad_fn=<NllLossBackward0>)
tensor(0.0479, grad_fn=<NllLossBackward0>)
tensor(0.0479, grad_fn=<NllLossBackward0>)
tensor(0.0478, grad_fn=<NllLossBackward0>)
tensor(0.0478, grad_fn=<NllLossBackward0>)
tensor(0.0477, grad_fn=<NllLossBackward0>)
tensor(0.0476, grad_fn=<NllLossBackward0>)
tensor(0.0476, grad_fn=<NllLossBackward0>)
tensor(0.0475, grad_fn=<NllLossBackward0>)
tensor(0.0475, grad_fn=<NllLossBackward0>)
tensor(0.0474, grad_fn=<NllLossBackward0>)
tensor(0.0473, grad_fn=<NllLossBackward0>)
tensor(0.0473, grad_fn=<NllLossBackward0>)
tensor(0.0472, grad_fn=<NllLossBackward0>)
tensor(0.0472, grad_fn=<NllLossBackward0>)
tensor(0.0471, grad_fn=<NllLossBackward0>)
tensor(0.0470, grad_fn=<NllLossBackward0>)
tensor(0.0470, grad_fn=<NllLossBackward0>)
tensor(0.0469, grad_fn=<NllLossBackward0>)
tensor(0.0469, grad_fn=<NllLossBackward0>)
tensor(0.0468, grad_fn=<NllLossBackward0>)
tensor(0.0467, grad_fn=<NllLossBackward0>)
tensor(0.0467, grad_fn=<NllLossBackward0>)
tensor(0.0466, grad_fn=<NllLossBackward0>)
tensor(0.0466, grad_fn=<NllLossBackward0>)
tensor(0.0465, grad_fn=<NllLossBackward0>)
tensor(0.0464, grad_fn=<NllLossBackward0>)
tensor(0.0464, grad_fn=<NllLossBackward0>)
tensor(0.0463, grad_fn=<NllLossBackward0>)
tensor(0.0463, grad_fn=<NllLossBackward0>)
tensor(0.0462, grad_fn=<NllLossBackward0>)
tensor(0.0462, grad_fn=<NllLossBackward0>)
tensor(0.0461, grad_fn=<NllLossBackward0>)
tensor(0.0460, grad_fn=<NllLossBackward0>)
tensor(0.0460, grad_fn=<NllLossBackward0>)
tensor(0.0459, grad_fn=<NllLossBackward0>)
tensor(0.0459, grad_fn=<NllLossBackward0>)
tensor(0.0458, grad_fn=<NllLossBackward0>)
tensor(0.0458, grad_fn=<NllLossBackward0>)
tensor(0.0457, grad_fn=<NllLossBackward0>)
tensor(0.0456, grad_fn=<NllLossBackward0>)
tensor(0.0456, grad_fn=<NllLossBackward0>)
tensor(0.0455, grad_fn=<NllLossBackward0>)
```

```
tensor(0.0455, grad_fn=<NllLossBackward0>)
tensor(0.0454, grad_fn=<NllLossBackward0>)
tensor(0.0454, grad_fn=<NllLossBackward0>)
tensor(0.0453, grad_fn=<NllLossBackward0>)
tensor(0.0453, grad_fn=<NllLossBackward0>)
tensor(0.0452, grad_fn=<NllLossBackward0>)
tensor(0.0452, grad_fn=<NllLossBackward0>)
tensor(0.0451, grad_fn=<NllLossBackward0>)
tensor(0.0450, grad_fn=<NllLossBackward0>)
tensor(0.0450, grad_fn=<NllLossBackward0>)
tensor(0.0449, grad_fn=<NllLossBackward0>)
tensor(0.0449, grad_fn=<NllLossBackward0>)
tensor(0.0448, grad_fn=<NllLossBackward0>)
tensor(0.0448, grad_fn=<NllLossBackward0>)
tensor(0.0447, grad_fn=<NllLossBackward0>)
tensor(0.0447, grad_fn=<NllLossBackward0>)
tensor(0.0446, grad_fn=<NllLossBackward0>)
tensor(0.0446, grad_fn=<NllLossBackward0>)
tensor(0.0445, grad_fn=<NllLossBackward0>)
tensor(0.0445, grad_fn=<NllLossBackward0>)
tensor(0.0444, grad_fn=<NllLossBackward0>)
tensor(0.0443, grad_fn=<NllLossBackward0>)
tensor(0.0443, grad_fn=<NllLossBackward0>)
tensor(0.0442, grad_fn=<NllLossBackward0>)
tensor(0.0442, grad_fn=<NllLossBackward0>)
tensor(0.0441, grad_fn=<NllLossBackward0>)
tensor(0.0441, grad_fn=<NllLossBackward0>)
tensor(0.0440, grad_fn=<NllLossBackward0>)
tensor(0.0440, grad_fn=<NllLossBackward0>)
tensor(0.0439, grad_fn=<NllLossBackward0>)
tensor(0.0439, grad_fn=<NllLossBackward0>)
tensor(0.0438, grad_fn=<NllLossBackward0>)
tensor(0.0438, grad_fn=<NllLossBackward0>)
tensor(0.0437, grad_fn=<NllLossBackward0>)
tensor(0.0437, grad_fn=<NllLossBackward0>)
tensor(0.0436, grad_fn=<NllLossBackward0>)
tensor(0.0436, grad_fn=<NllLossBackward0>)
tensor(0.0435, grad_fn=<NllLossBackward0>)
tensor(0.0435, grad_fn=<NllLossBackward0>)
tensor(0.0434, grad_fn=<NllLossBackward0>)
tensor(0.0434, grad_fn=<NllLossBackward0>)
tensor(0.0433, grad_fn=<NllLossBackward0>)
tensor(0.0433, grad_fn=<NllLossBackward0>)
tensor(0.0432, grad_fn=<NllLossBackward0>)
tensor(0.0432, grad_fn=<NllLossBackward0>)
tensor(0.0431, grad_fn=<NllLossBackward0>)
tensor(0.0431, grad_fn=<NllLossBackward0>)
tensor(0.0430, grad_fn=<NllLossBackward0>)
```

```
tensor(0.0430, grad_fn=<NllLossBackward0>)
tensor(0.0429, grad_fn=<NllLossBackward0>)
tensor(0.0429, grad_fn=<NllLossBackward0>)
tensor(0.0428, grad_fn=<NllLossBackward0>)
tensor(0.0428, grad_fn=<NllLossBackward0>)
tensor(0.0427, grad_fn=<NllLossBackward0>)
tensor(0.0427, grad_fn=<NllLossBackward0>)
tensor(0.0426, grad_fn=<NllLossBackward0>)
tensor(0.0426, grad_fn=<NllLossBackward0>)
tensor(0.0425, grad_fn=<NllLossBackward0>)
tensor(0.0425, grad_fn=<NllLossBackward0>)
tensor(0.0424, grad_fn=<NllLossBackward0>)
tensor(0.0424, grad_fn=<NllLossBackward0>)
tensor(0.0423, grad_fn=<NllLossBackward0>)
tensor(0.0423, grad_fn=<NllLossBackward0>)
tensor(0.0422, grad_fn=<NllLossBackward0>)
tensor(0.0422, grad_fn=<NllLossBackward0>)
tensor(0.0421, grad_fn=<NllLossBackward0>)
tensor(0.0421, grad_fn=<NllLossBackward0>)
tensor(0.0421, grad_fn=<NllLossBackward0>)
tensor(0.0420, grad_fn=<NllLossBackward0>)
tensor(0.0420, grad_fn=<NllLossBackward0>)
tensor(0.0419, grad_fn=<NllLossBackward0>)
tensor(0.0419, grad_fn=<NllLossBackward0>)
tensor(0.0418, grad_fn=<NllLossBackward0>)
tensor(0.0418, grad_fn=<NllLossBackward0>)
tensor(0.0417, grad_fn=<NllLossBackward0>)
tensor(0.0417, grad_fn=<NllLossBackward0>)
tensor(0.0416, grad_fn=<NllLossBackward0>)
tensor(0.0416, grad_fn=<NllLossBackward0>)
tensor(0.0415, grad_fn=<NllLossBackward0>)
tensor(0.0415, grad_fn=<NllLossBackward0>)
tensor(0.0415, grad_fn=<NllLossBackward0>)
tensor(0.0414, grad_fn=<NllLossBackward0>)
tensor(0.0414, grad_fn=<NllLossBackward0>)
tensor(0.0413, grad_fn=<NllLossBackward0>)
tensor(0.0413, grad_fn=<NllLossBackward0>)
tensor(0.0412, grad_fn=<NllLossBackward0>)
tensor(0.0412, grad_fn=<NllLossBackward0>)
tensor(0.0411, grad_fn=<NllLossBackward0>)
tensor(0.0411, grad_fn=<NllLossBackward0>)
tensor(0.0410, grad_fn=<NllLossBackward0>)
tensor(0.0410, grad_fn=<NllLossBackward0>)
tensor(0.0410, grad_fn=<NllLossBackward0>)
tensor(0.0409, grad_fn=<NllLossBackward0>)
tensor(0.0409, grad_fn=<NllLossBackward0>)
tensor(0.0408, grad_fn=<NllLossBackward0>)
tensor(0.0408, grad_fn=<NllLossBackward0>)
```

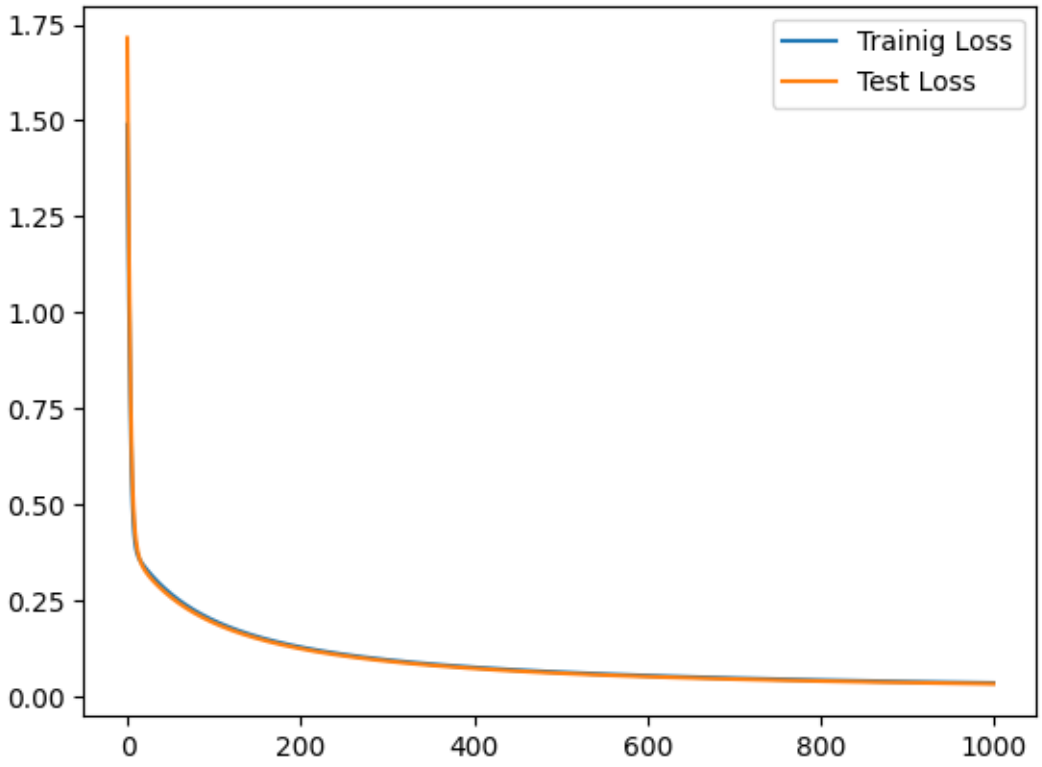
```
tensor(0.0407, grad_fn=<NllLossBackward0>)
tensor(0.0407, grad_fn=<NllLossBackward0>)
tensor(0.0406, grad_fn=<NllLossBackward0>)
tensor(0.0406, grad_fn=<NllLossBackward0>)
tensor(0.0406, grad_fn=<NllLossBackward0>)
tensor(0.0405, grad_fn=<NllLossBackward0>)
tensor(0.0405, grad_fn=<NllLossBackward0>)
tensor(0.0404, grad_fn=<NllLossBackward0>)
tensor(0.0404, grad_fn=<NllLossBackward0>)
tensor(0.0403, grad_fn=<NllLossBackward0>)
tensor(0.0403, grad_fn=<NllLossBackward0>)
tensor(0.0403, grad_fn=<NllLossBackward0>)
tensor(0.0402, grad_fn=<NllLossBackward0>)
tensor(0.0402, grad_fn=<NllLossBackward0>)
tensor(0.0401, grad_fn=<NllLossBackward0>)
tensor(0.0401, grad_fn=<NllLossBackward0>)
tensor(0.0400, grad_fn=<NllLossBackward0>)
tensor(0.0400, grad_fn=<NllLossBackward0>)
tensor(0.0400, grad_fn=<NllLossBackward0>)
tensor(0.0399, grad_fn=<NllLossBackward0>)
tensor(0.0399, grad_fn=<NllLossBackward0>)
tensor(0.0398, grad_fn=<NllLossBackward0>)
tensor(0.0398, grad_fn=<NllLossBackward0>)
tensor(0.0397, grad_fn=<NllLossBackward0>)
tensor(0.0397, grad_fn=<NllLossBackward0>)
tensor(0.0397, grad_fn=<NllLossBackward0>)
tensor(0.0396, grad_fn=<NllLossBackward0>)
tensor(0.0396, grad_fn=<NllLossBackward0>)
tensor(0.0395, grad_fn=<NllLossBackward0>)
tensor(0.0395, grad_fn=<NllLossBackward0>)
tensor(0.0394, grad_fn=<NllLossBackward0>)
tensor(0.0394, grad_fn=<NllLossBackward0>)
tensor(0.0394, grad_fn=<NllLossBackward0>)
tensor(0.0393, grad_fn=<NllLossBackward0>)
tensor(0.0393, grad_fn=<NllLossBackward0>)
tensor(0.0392, grad_fn=<NllLossBackward0>)
tensor(0.0392, grad_fn=<NllLossBackward0>)
tensor(0.0392, grad_fn=<NllLossBackward0>)
tensor(0.0391, grad_fn=<NllLossBackward0>)
tensor(0.0391, grad_fn=<NllLossBackward0>)
tensor(0.0390, grad_fn=<NllLossBackward0>)
tensor(0.0390, grad_fn=<NllLossBackward0>)
tensor(0.0390, grad_fn=<NllLossBackward0>)
tensor(0.0389, grad_fn=<NllLossBackward0>)
tensor(0.0389, grad_fn=<NllLossBackward0>)
tensor(0.0388, grad_fn=<NllLossBackward0>)
tensor(0.0388, grad_fn=<NllLossBackward0>)
tensor(0.0388, grad_fn=<NllLossBackward0>)
```

```
tensor(0.0387, grad_fn=<NllLossBackward0>)
tensor(0.0387, grad_fn=<NllLossBackward0>)
tensor(0.0386, grad_fn=<NllLossBackward0>)
tensor(0.0386, grad_fn=<NllLossBackward0>)
tensor(0.0386, grad_fn=<NllLossBackward0>)
tensor(0.0385, grad_fn=<NllLossBackward0>)
tensor(0.0385, grad_fn=<NllLossBackward0>)
tensor(0.0384, grad_fn=<NllLossBackward0>)
tensor(0.0384, grad_fn=<NllLossBackward0>)
tensor(0.0384, grad_fn=<NllLossBackward0>)
tensor(0.0383, grad_fn=<NllLossBackward0>)
tensor(0.0383, grad_fn=<NllLossBackward0>)
tensor(0.0382, grad_fn=<NllLossBackward0>)
tensor(0.0382, grad_fn=<NllLossBackward0>)
tensor(0.0382, grad_fn=<NllLossBackward0>)
tensor(0.0381, grad_fn=<NllLossBackward0>)
tensor(0.0381, grad_fn=<NllLossBackward0>)
tensor(0.0381, grad_fn=<NllLossBackward0>)
tensor(0.0380, grad_fn=<NllLossBackward0>)
tensor(0.0380, grad_fn=<NllLossBackward0>)
tensor(0.0379, grad_fn=<NllLossBackward0>)
tensor(0.0379, grad_fn=<NllLossBackward0>)
tensor(0.0379, grad_fn=<NllLossBackward0>)
tensor(0.0378, grad_fn=<NllLossBackward0>)
tensor(0.0378, grad_fn=<NllLossBackward0>)
tensor(0.0378, grad_fn=<NllLossBackward0>)
tensor(0.0377, grad_fn=<NllLossBackward0>)
tensor(0.0377, grad_fn=<NllLossBackward0>)
tensor(0.0376, grad_fn=<NllLossBackward0>)
tensor(0.0376, grad_fn=<NllLossBackward0>)
tensor(0.0376, grad_fn=<NllLossBackward0>)
tensor(0.0375, grad_fn=<NllLossBackward0>)
tensor(0.0375, grad_fn=<NllLossBackward0>)
tensor(0.0375, grad_fn=<NllLossBackward0>)
tensor(0.0374, grad_fn=<NllLossBackward0>)
tensor(0.0374, grad_fn=<NllLossBackward0>)
tensor(0.0373, grad_fn=<NllLossBackward0>)
tensor(0.0373, grad_fn=<NllLossBackward0>)
tensor(0.0373, grad_fn=<NllLossBackward0>)
tensor(0.0372, grad_fn=<NllLossBackward0>)
tensor(0.0372, grad_fn=<NllLossBackward0>)
tensor(0.0372, grad_fn=<NllLossBackward0>)
tensor(0.0371, grad_fn=<NllLossBackward0>)
tensor(0.0371, grad_fn=<NllLossBackward0>)
tensor(0.0371, grad_fn=<NllLossBackward0>)
tensor(0.0370, grad_fn=<NllLossBackward0>)
tensor(0.0370, grad_fn=<NllLossBackward0>)
tensor(0.0369, grad_fn=<NllLossBackward0>)
```

```
tensor(0.0369, grad_fn=<NllLossBackward0>)
tensor(0.0369, grad_fn=<NllLossBackward0>)
tensor(0.0368, grad_fn=<NllLossBackward0>)
tensor(0.0368, grad_fn=<NllLossBackward0>)
tensor(0.0368, grad_fn=<NllLossBackward0>)
tensor(0.0367, grad_fn=<NllLossBackward0>)
tensor(0.0367, grad_fn=<NllLossBackward0>)
tensor(0.0367, grad_fn=<NllLossBackward0>)
tensor(0.0366, grad_fn=<NllLossBackward0>)
tensor(0.0366, grad_fn=<NllLossBackward0>)
tensor(0.0366, grad_fn=<NllLossBackward0>)
tensor(0.0365, grad_fn=<NllLossBackward0>)
tensor(0.0365, grad_fn=<NllLossBackward0>)
tensor(0.0364, grad_fn=<NllLossBackward0>)
tensor(0.0364, grad_fn=<NllLossBackward0>)
tensor(0.0364, grad_fn=<NllLossBackward0>)
tensor(0.0363, grad_fn=<NllLossBackward0>)
tensor(0.0363, grad_fn=<NllLossBackward0>)
tensor(0.0363, grad_fn=<NllLossBackward0>)
tensor(0.0362, grad_fn=<NllLossBackward0>)
tensor(0.0362, grad_fn=<NllLossBackward0>)
tensor(0.0362, grad_fn=<NllLossBackward0>)
tensor(0.0361, grad_fn=<NllLossBackward0>)
tensor(0.0361, grad_fn=<NllLossBackward0>)
tensor(0.0361, grad_fn=<NllLossBackward0>)
tensor(0.0360, grad_fn=<NllLossBackward0>)
tensor(0.0360, grad_fn=<NllLossBackward0>)
tensor(0.0360, grad_fn=<NllLossBackward0>)
tensor(0.0359, grad_fn=<NllLossBackward0>)
tensor(0.0359, grad_fn=<NllLossBackward0>)
tensor(0.0359, grad_fn=<NllLossBackward0>)
tensor(0.0358, grad_fn=<NllLossBackward0>)
tensor(0.0358, grad_fn=<NllLossBackward0>)
tensor(0.0358, grad_fn=<NllLossBackward0>)
tensor(0.0357, grad_fn=<NllLossBackward0>)
tensor(0.0357, grad_fn=<NllLossBackward0>)
tensor(0.0357, grad_fn=<NllLossBackward0>)
tensor(0.0356, grad_fn=<NllLossBackward0>)
tensor(0.0356, grad_fn=<NllLossBackward0>)
tensor(0.0356, grad_fn=<NllLossBackward0>)
tensor(0.0355, grad_fn=<NllLossBackward0>)
tensor(0.0355, grad_fn=<NllLossBackward0>)
tensor(0.0355, grad_fn=<NllLossBackward0>)
tensor(0.0354, grad_fn=<NllLossBackward0>)
tensor(0.0354, grad_fn=<NllLossBackward0>)
tensor(0.0354, grad_fn=<NllLossBackward0>)
tensor(0.0353, grad_fn=<NllLossBackward0>)
tensor(0.0353, grad_fn=<NllLossBackward0>)
```

```
tensor(0.0353, grad_fn=<NllLossBackward0>)
tensor(0.0352, grad_fn=<NllLossBackward0>)
tensor(0.0352, grad_fn=<NllLossBackward0>)
tensor(0.0352, grad_fn=<NllLossBackward0>)
tensor(0.0351, grad_fn=<NllLossBackward0>)
tensor(0.0351, grad_fn=<NllLossBackward0>)
tensor(0.0351, grad_fn=<NllLossBackward0>)
tensor(0.0350, grad_fn=<NllLossBackward0>)
tensor(0.0350, grad_fn=<NllLossBackward0>)
tensor(0.0350, grad_fn=<NllLossBackward0>)
tensor(0.0349, grad_fn=<NllLossBackward0>)
tensor(0.0349, grad_fn=<NllLossBackward0>)
tensor(0.0349, grad_fn=<NllLossBackward0>)
tensor(0.0348, grad_fn=<NllLossBackward0>)
tensor(0.0348, grad_fn=<NllLossBackward0>)
tensor(0.0348, grad_fn=<NllLossBackward0>)
tensor(0.0347, grad_fn=<NllLossBackward0>)
tensor(0.0347, grad_fn=<NllLossBackward0>)
tensor(0.0347, grad_fn=<NllLossBackward0>)
tensor(0.0347, grad_fn=<NllLossBackward0>)
tensor(0.0346, grad_fn=<NllLossBackward0>)
tensor(0.0346, grad_fn=<NllLossBackward0>)
tensor(0.0346, grad_fn=<NllLossBackward0>)
tensor(0.0345, grad_fn=<NllLossBackward0>)
tensor(0.0345, grad_fn=<NllLossBackward0>)
tensor(0.0345, grad_fn=<NllLossBackward0>)
tensor(0.0344, grad_fn=<NllLossBackward0>)
tensor(0.0344, grad_fn=<NllLossBackward0>)
tensor(0.0344, grad_fn=<NllLossBackward0>)
tensor(0.0343, grad_fn=<NllLossBackward0>)
tensor(0.0343, grad_fn=<NllLossBackward0>)
tensor(0.0343, grad_fn=<NllLossBackward0>)
tensor(0.0342, grad_fn=<NllLossBackward0>)
tensor(0.0342, grad_fn=<NllLossBackward0>)
tensor(0.0342, grad_fn=<NllLossBackward0>)
tensor(0.0342, grad_fn=<NllLossBackward0>)
tensor(0.0341, grad_fn=<NllLossBackward0>)
tensor(0.0341, grad_fn=<NllLossBackward0>)
tensor(0.0341, grad_fn=<NllLossBackward0>)
tensor(0.0340, grad_fn=<NllLossBackward0>)
tensor(0.0340, grad_fn=<NllLossBackward0>)
tensor(0.0340, grad_fn=<NllLossBackward0>)
tensor(0.0339, grad_fn=<NllLossBackward0>)
tensor(0.0339, grad_fn=<NllLossBackward0>)
tensor(0.0339, grad_fn=<NllLossBackward0>)
```

[ ]:
```
plt.plot(train_loss, label="Trainig Loss")
plt.plot(test_loss, label='Test Loss')
plt.legend()
plt.show()
```



[ ]:
```
# confusion matrix
y_pred=model(X_test_tensor)
conf_matirx=confusion_matrix(y_test_tensor.numpy(),y_pred.numpy())
disp=ConfusionMatrixDisplay(confusion_matrix=conf_matrix, display_labels=[0,1])

disp.plot()
```