

bolt_RNN

July 14, 2022

```
[ ]: import pandas as pd
import torch
import numpy as np
from sklearn.preprocessing import MinMaxScaler
from torch.utils.data import TensorDataset, DataLoader
import os
import cv2 as cv
from torch.utils.data import Dataset
from torchvision import transforms
import torch.nn.functional as F
from PIL import Image
import torch.nn as nn
from torch.optim import Adam
import matplotlib.pyplot as plt
import numpy as np
from torch.autograd import Variable
```

1 Define the data and label

```
[ ]: def getlabel(dir):
    labels=[]
    for name in os.listdir(dir):
        if os.path.splitext(name)[1] == '.png':
            fname=os.path.splitext(name)[0]
            labels.append(fname)
    labels = list(map(int, labels))
    labels = torch.tensor(labels)
    # Normalize labels
    #min = torch.min(labels)
    #max = torch.max(labels)
    #labels = (labels - min) / (max - min)
    labels = labels.type(torch.FloatTensor)
    return labels

def generate_dataset(dir):
    """
    set_label should be 'torch.tensor([1])' if two-catogory and positive sample
```

```

"""
train_data = []
for file_name in os.listdir(dir):
    if file_name != "Thumbs.db":
        img_dir = os.path.join(dir, file_name)
        img = cv.imread(img_dir)
        img = cv.resize(img, (769, 432))    # /5 resize img
        #img_gray = cv.cvtColor(img,cv.COLOR_RGB2GRAY)
        pimg = Image.fromarray(img)
        train_data.append(pimg)
return train_data

```

```

[ ]: trainindir= './train/'
      validdir = './test/'
      train_data0 = generate_dataset(trainindir)
      train_label0=getlabel(trainindir)
      valid_data0 = generate_dataset(validdir)
      valid_label0 = getlabel(validdir)

```

```

[ ]: train_label0

```

```

[ ]: tensor([ 0.,  1., 10., 100., 101., 102., 103., 104., 105., 106., 107., 108.,
            109., 11., 110., 111., 112., 113., 114., 115., 116., 117., 118., 119.,
            12., 120., 122., 123., 124., 125., 126., 129., 13., 130., 132., 133.,
            134., 135., 136., 137., 138., 139., 14., 140., 141., 142., 143., 144.,
            145., 146., 147., 148., 149., 15., 150., 151., 152., 153., 154., 155.,
            156., 157., 159., 16., 160., 162., 163., 164., 165., 166., 167., 168.,
            169., 17., 170., 171., 172., 173., 174., 175., 176., 177., 178., 179.,
            18., 180., 182., 183., 184., 185., 186., 189., 19., 190., 191., 192.,
            193., 194., 195., 196., 197., 198., 199.,  2., 20., 200., 201., 203.,
            204., 205., 21., 22., 23., 24., 25., 26., 27., 28., 29.,  3.,
            30., 32., 33., 34., 35., 36., 37., 38., 39.,  4., 43., 44.,
            45., 46., 47., 48., 49.,  5., 50., 51., 52., 53., 54., 55.,
            56., 57., 58.,  6., 60., 61., 62., 64., 65., 67., 68., 69.,
            7., 70., 71., 74., 75., 76., 77., 78., 79.,  8., 80., 81.,
            82., 83., 84., 85., 86., 87., 88., 89.,  9., 90., 91., 92.,
            93., 94., 95., 96., 97., 98., 99.])

```

2 dataset

```

[ ]: class MyDataset(Dataset):
      def __init__(self, data, labels, transform=None, target_transform=None):
          self.transform = transform
          self.target_transform = target_transform
          imgs = []
          for i in range(len(labels)):
              # print(type(data[i]))    # <class 'PIL.Image.Image'>

```

```

        im_tensor = transform(data[i]).to(torch.device("cpu"))
        imgs.append((im_tensor, labels[i]))
    self.imgs = imgs                                # DataLoader getitem
def __getitem__(self, index):
    fn, label = self.imgs[index]
    return fn, label
def __len__(self):
    return len(self.imgs)

```

```
[ ]:
```

3 MyDataset

```

[ ]: normMean = [0.5, 0.5, 0.5]
    normStd = [0.5, 0.5, 0.5]
    transform = transforms.Compose(
        [transforms.Grayscale(num_output_channels=1), # num_output_channels 1
         transforms.ToTensor(), # range [0, 255] -> [0.0,1.0]
         transforms.Normalize(normMean, normStd)]
    )

    # train_transform
    train_data = MyDataset(train_data0, train_label0, transform=transform)
    valid_data = MyDataset(valid_data0, valid_label0, transform=transform)
    train_loader = DataLoader(dataset=train_data, batch_size=1, shuffle=True)
    valid_loader = DataLoader(dataset=valid_data, batch_size=1, shuffle=True)
    dataiter=iter(train_loader)
    images, labels = dataiter.next()
    print(images.shape)
    print(labels.shape)

    torch.Size([1, 1, 432, 769])
    torch.Size([1])

```

```
[ ]: len(train_label0)
```

```
[ ]: 187
```

4

```

[ ]: class Net(torch.nn.Module):
    # Net
    def __init__(self, n_feature, n_hidden, n_output):
        #
        super(Net, self).__init__()

```

```

#
self.conv1 = nn.Conv2d(in_channels=1, out_channels=3, kernel_size=3,
↳ stride=1, padding=1)
self.bn1 = nn.BatchNorm2d(3)
self.conv2 = nn.Conv2d(in_channels=3, out_channels=6, kernel_size=3,
↳ stride=1, padding=1)
self.bn2 = nn.BatchNorm2d(6)
self.pool = nn.MaxPool2d(2,2)
self.fc1 = nn.Linear(6*108*192, n_feature)
self.hidden_layer = torch.nn.Linear(n_feature, n_hidden)
# ( )
self.predict_layer = torch.nn.Linear(n_hidden, n_output)

#
def forward(self, x):
    x = F.relu(self.bn1(self.conv1(x)))           #432-3+1=430, 430/2=215;
↳ 769-3+1=767, 767/2=384
    x = self.pool(x)
    x = F.relu(self.bn2(self.conv2(x)))           #215-3+1=213, 213/
↳ 2=108; 384-3+1=382, 382/2=192
    x = self.pool(x)
    x = x.view(-1, 6*108*192)
    x = F.relu(self.fc1(x))
    # relu
    hidden_result = self.hidden_layer(x)
    relu_result = F.relu(hidden_result)
    predict_result = self.predict_layer(relu_result)
    return predict_result

```

```

[ ]: def testAccuracy():

    net.eval()
    accuracy = 0.0
    total = 0.0
    device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")
    #device = torch.device("cpu")

    with torch.no_grad():
        for data in valid_loader:
            images, labels = data
            images, labels = images.to(device), labels.to(device)
            # run the model on the test set to predict labels
            outputs = net(images)
            # the label with the highest energy will be our prediction
            _, predicted = torch.max(outputs.data, 1)
            total += labels.size(0)
            accuracy += (predicted == labels).sum().item()

```

```

    # compute the accuracy over all test images
    accuracy = (100 * accuracy / total)
    return accuracy)

def train(num_epochs):

    best_accuracy = 0.0

    # Define your execution device
    device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")
    #device = torch.device("cpu")

    # Convert model parameters and buffers to CPU or Cuda
    net.to(device)
    for epoch in range(num_epochs): # loop over the dataset multiple times
        running_loss = 0.0
        running_acc = 0.0

        for i, (images, labels) in enumerate(train_loader, 0):
            #images, labels = dataiter.next()
            #
            # get the inputs
            images = Variable(images.to(device))
            labels = Variable(labels.to(device))
            prediction = net(images)
            print(prediction, labels)
            print(prediction.data.cpu().numpy()[0])
            print(labels.data.cpu().numpy())
            #
            #
            loss = loss_func(prediction, labels)

            #
            # 0
            optimizer.zero_grad()
            #
            loss.backward()
            # loss
            optimizer.step()
            print('epoch:', i, 'loss:', loss.item())
            if i % 1000 == 999:
                # print every 1000 (twice per epoch)
                print('%d, %5d loss: %.3f' %
                      (epoch + 1, i + 1, running_loss / 1000))
                # zero the loss
                running_loss = 0.0

```

```

        # if i % 2 == 0:
        #     # plot and show learning process
        #     plt.cla()
        #     #
        #     plt.plot(labels.cpu().numpy(), labels.cpu().numpy(),
↪c='blue', lw='3')
        #     #
        #     plt.scatter(labels.data.cpu().numpy()[0], prediction.data.
↪cpu().numpy()[0])
        #     plt.text(0.5, 0, 'Loss=%.4f' % loss.data.cpu().numpy(),
↪fontdict={'size': 20, 'color': 'red'})
        #     plt.pause(0.1)
        # #
        # # Compute and print the average accuracy fo this epoch when tested
↪over all 10000 test images
        accuracy = testAccuracy()
        print('For epoch', epoch+1, 'the test accuracy over the whole test set
↪is %d %%' % (accuracy))

    print('Finished Training')
    #
    # if i % 2 == 0:
    #     #
    #     plt.cla()
    #     #
    #     plt.plot(images.numpy(), labels.numpy(), c='blue', lw='3')
    #     #
    #     plt.scatter(images.numpy(), labels.numpy(), c='orange')
    #     #
    #     plt.plot(images.numpy(), prediction.data.numpy(), c='red', lw='2')
    #     plt.text(-0.5, -65, 'Time=%d Loss=%.4f' % (i, loss.data.numpy()),
↪fontdict={'size': 15, 'color': 'red'})
    #     plt.pause(0.1)

```

```

[ ]: Vdataiter=iter(valid_loader)
      vimg, vlabels = Vdataiter.next()
      print(vimg.shape)
      print(vlabels.shape)
      print(vlabels)

```

```

torch.Size([1, 1, 432, 769])
torch.Size([1])
tensor([128.])

```

4.1

```
[ ]: def predict(model, device):
    Vdataiter=iter(valid_loader)
    vimg, vlabels = Vdataiter.next()
    model.to(device)
    with torch.no_grad():
        vimg=vimg.to(device)
        out = model(vimg)
        #_, pre = torch.max(out.data, 1)
    return out, vlabels
```

5

```
[ ]: #          1
INPUT_FEATURE_DIM = 3000
#
NEURON_NUM = 300
OUTPUT_FEATURE_DIM = 1
#
LEARNING_RATE = 0.001

#
net = Net(n_feature=INPUT_FEATURE_DIM, n_hidden=NEURON_NUM,
    ↪n_output=OUTPUT_FEATURE_DIM)
print(net)
#
#
optimizer = torch.optim.Adam(net.parameters(), lr=LEARNING_RATE)
#
loss_func = torch.nn.MSELoss() #                               #nn.MSELoss()
# Define the loss function with Classification Cross-Entropy loss and an
    ↪optimizer with Adam optimizer
```

```
Net(
  (conv1): Conv2d(1, 3, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn1): BatchNorm2d(3, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
  (conv2): Conv2d(3, 6, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(6, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
  (pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  (fc1): Linear(in_features=124416, out_features=3000, bias=True)
  (hidden_layer): Linear(in_features=3000, out_features=300, bias=True)
  (predict_layer): Linear(in_features=300, out_features=1, bias=True)
)
```

6

```
[ ]: # Vdataiter=iter(valid_loader)
# vimg, vlabels = Vdataiter.next()
# print(vimg.shape)
# print(vlabels.shape)
# print(vlabels)
if __name__ == "__main__":

    torch.cuda.empty_cache()

    device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")
    #device = torch.device("cpu")
    # Let's build our model
    train(30)
    print('Finished Training')

    # Test which classes performed well
    testAccuracy()

    # Let's load the model we just created and test the accuracy per label
    # model = net()
    # path = "myFirstModel.pth"
    # model.load_state_dict(torch.load(path))

    print(predict(net, device))
```

```
tensor([[0.0971]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[0.09707762]
[69.]
```

c:\Users\bridge1\anaconda3\envs\pytorch\lib\site-packages\torch\nn\modules\loss.py:530: UserWarning: Using a target size (torch.Size([1])) that is different to the input size (torch.Size([1, 1])). This will likely lead to incorrect results due to broadcasting. Please ensure they have the same size.

```
    return F.mse_loss(input, target, reduction=self.reduction)
```

```
epoch: 0 loss: 4747.61279296875
tensor([[183.0497]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[183.04974]
[169.]
epoch: 1 loss: 197.39529418945312
tensor([[97.8564]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[97.85644]
[140.]
```



```

epoch: 2 loss: 1776.079833984375
tensor([[269.3916]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[269.39157]
[163.]
epoch: 3 loss: 11319.166015625
tensor([[143.8645]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[143.86455]
[135.]
epoch: 4 loss: 78.58020782470703
tensor([[50.0015]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[50.001476]
[74.]
epoch: 5 loss: 575.9291381835938
tensor([[14.0526]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[14.052554]
[159.]
epoch: 6 loss: 21009.763671875
tensor([[68.3105]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[68.310455]
[35.]
epoch: 7 loss: 1109.58642578125
tensor([[88.8911]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[88.891136]
[114.]
epoch: 8 loss: 630.4550170898438
tensor([[119.2843]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[119.28429]
[186.]
epoch: 9 loss: 4450.986328125
tensor([[186.2352]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[186.23523]
[198.]
epoch: 10 loss: 138.40982055664062
tensor([[260.2426]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[260.2426]
[189.]
epoch: 11 loss: 5075.51025390625
tensor([[254.5493]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')

```

```

[254.54935]
[175.]
epoch: 12 loss: 6328.0986328125
tensor([[205.3853]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[205.38535]
[150.]
epoch: 13 loss: 3067.536376953125
tensor([[146.9189]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[146.91891]
[91.]
epoch: 14 loss: 3126.925048828125
tensor([[88.6134]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[88.61339]
[102.]
epoch: 15 loss: 179.20138549804688
tensor([[53.3151]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[53.31507]
[98.]
epoch: 16 loss: 1996.742919921875
tensor([[37.2268]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[37.226845]
[75.]
epoch: 17 loss: 1426.811279296875
tensor([[32.2102]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[32.21018]
[52.]
epoch: 18 loss: 391.6370544433594
tensor([[32.8000]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[32.800003]
[22.]
epoch: 19 loss: 116.64006805419922
tensor([[30.2280]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[30.227951]
[19.]
epoch: 20 loss: 126.06688690185547
tensor([[24.3729]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[24.372913]
[58.]
epoch: 21 loss: 1130.781005859375

```

```

tensor([[29.3054]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[29.30537]
[92.]
epoch: 22 loss: 3930.61669921875
tensor([[40.6766]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[40.67663]
[38.]
epoch: 23 loss: 7.164338111877441
tensor([[50.2678]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[50.267788]
[168.]
epoch: 24 loss: 13860.873046875
tensor([[73.4169]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[73.41687]
[57.]
epoch: 25 loss: 269.51361083984375
tensor([[92.9646]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[92.964554]
[93.]
epoch: 26 loss: 0.0012564307544380426
tensor([[112.5673]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[112.567276]
[153.]
epoch: 27 loss: 1634.80517578125
tensor([[145.9976]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[145.99763]
[152.]
epoch: 28 loss: 36.028385162353516
tensor([[184.7684]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[184.76842]
[173.]
epoch: 29 loss: 138.4956512451172
tensor([[196.5951]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[196.59512]
[13.]
epoch: 30 loss: 33707.16796875
tensor([[167.9616]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[167.9616]

```

```

[64.]
epoch: 31 loss: 10808.0126953125
tensor([[135.9892]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[135.98918]
[184.]
epoch: 32 loss: 2305.03857421875
tensor([[106.0503]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[106.05032]
[120.]
epoch: 33 loss: 194.59347534179688
tensor([[77.0486]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[77.048615]
[0.]
epoch: 34 loss: 5936.48876953125
tensor([[75.5835]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[75.58352]
[148.]
epoch: 35 loss: 5244.146484375
tensor([[71.3409]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[71.3409]
[155.]
epoch: 36 loss: 6998.845703125
tensor([[66.0278]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[66.02777]
[103.]
epoch: 37 loss: 1366.9456787109375
tensor([[90.9884]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[90.9884]
[172.]
epoch: 38 loss: 6562.87890625
tensor([[70.3202]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[70.320206]
[18.]
epoch: 39 loss: 2737.40380859375
tensor([[76.4670]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[76.46697]
[26.]
epoch: 40 loss: 2546.915283203125
tensor([[125.6797]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],

```

```

device='cuda:0')
[125.67968]
[145.]
epoch: 41 loss: 373.2747802734375
tensor([[82.6801]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[82.68012]
[36.]
epoch: 42 loss: 2179.033935546875
tensor([[139.3945]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[139.39453]
[177.]
epoch: 43 loss: 1414.1712646484375
tensor([[70.5654]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[70.565414]
[24.]
epoch: 44 loss: 2168.337890625
tensor([[111.9110]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[111.91099]
[87.]
epoch: 45 loss: 620.5573120117188
tensor([[51.4496]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[51.449608]
[37.]
epoch: 46 loss: 208.79116821289062
tensor([[84.1569]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[84.156944]
[101.]
epoch: 47 loss: 283.68853759765625
tensor([[55.3557]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[55.355682]
[80.]
epoch: 48 loss: 607.3424072265625
tensor([[28.8674]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[28.867437]
[43.]
epoch: 49 loss: 199.7293243408203
tensor([[96.7780]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[96.77803]
[107.]

```

```

epoch: 50 loss: 104.48866271972656
tensor([[37.3508]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[37.35076]
[32.]
epoch: 51 loss: 28.630647659301758
tensor([[91.7702]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[91.77021]
[78.]
epoch: 52 loss: 189.6186981201172
tensor([[35.0642]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[35.064167]
[16.]
epoch: 53 loss: 363.4424743652344
tensor([[28.9647]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[28.964699]
[6.]
epoch: 54 loss: 527.3773803710938
tensor([[20.3691]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[20.369114]
[25.]
epoch: 55 loss: 21.445106506347656
tensor([[12.7424]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[12.742408]
[14.]
epoch: 56 loss: 1.581538200378418
tensor([[105.0311]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[105.03109]
[117.]
epoch: 57 loss: 143.2548065185547
tensor([[141.4469]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[141.44685]
[146.]
epoch: 58 loss: 20.731142044067383
tensor([[187.2363]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[187.23634]
[194.]
epoch: 59 loss: 45.74705123901367
tensor([[17.2206]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')

```

```

[17.220587]
[21.]
epoch: 60 loss: 14.283964157104492
tensor([[227.2344]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[227.23444]
[170.]
epoch: 61 loss: 3275.78076171875
tensor([[185.2378]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[185.2378]
[171.]
epoch: 62 loss: 202.71475219726562
tensor([[97.8965]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[97.896484]
[124.]
epoch: 63 loss: 681.3935546875
tensor([[46.3930]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[46.392956]
[94.]
epoch: 64 loss: 2266.4306640625
tensor([[8.2812]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[8.281169]
[20.]
epoch: 65 loss: 137.3310089111328
tensor([[131.2978]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[131.29776]
[180.]
epoch: 66 loss: 2371.908203125
tensor([[164.5153]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[164.51526]
[197.]
epoch: 67 loss: 1055.2584228515625
tensor([[151.6069]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[151.60695]
[110.]
epoch: 68 loss: 1731.13818359375
tensor([[197.7937]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[197.79369]
[147.]
epoch: 69 loss: 2579.99853515625

```

```

tensor([[184.5161]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[184.51608]
[149.]
epoch: 70 loss: 1261.39208984375
tensor([[120.8539]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[120.85393]
[104.]
epoch: 71 loss: 284.05487060546875
tensor([[27.6119]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[27.611876]
[8.]
epoch: 72 loss: 384.62567138671875
tensor([[104.9595]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[104.959526]
[166.]
epoch: 73 loss: 3725.939453125
tensor([[19.1012]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[19.101192]
[23.]
epoch: 74 loss: 15.200699806213379
tensor([[30.5952]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[30.595182]
[51.]
epoch: 75 loss: 416.3565673828125
tensor([[22.8077]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[22.807709]
[9.]
epoch: 76 loss: 190.65281677246094
tensor([[23.0752]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[23.075233]
[1.]
epoch: 77 loss: 487.31591796875
tensor([[22.6119]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[22.611887]
[29.]
epoch: 78 loss: 40.807987213134766
tensor([[19.3836]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[19.383612]

```



```

[15.]
epoch: 79 loss: 19.21605110168457
tensor([[20.5829]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[20.582932]
[30.]
epoch: 80 loss: 88.6811752319336
tensor([[89.7193]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[89.719345]
[134.]
epoch: 81 loss: 1960.7763671875
tensor([[46.2360]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[46.236]
[65.]
epoch: 82 loss: 352.08770751953125
tensor([[76.4389]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[76.43894]
[89.]
epoch: 83 loss: 157.78018188476562
tensor([[68.3864]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[68.3864]
[61.]
epoch: 84 loss: 54.55887985229492
tensor([[62.7885]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[62.788532]
[47.]
epoch: 85 loss: 249.2777557373047
tensor([[163.9963]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[163.99626]
[164.]
epoch: 86 loss: 1.3975659385323524e-05
tensor([[49.8925]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[49.892487]
[27.]
epoch: 87 loss: 524.06591796875
tensor([[42.4410]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[42.44101]
[5.]
epoch: 88 loss: 1401.8292236328125
tensor([[157.6511]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],

```

```

device='cuda:0')
[157.65114]
[156.]
epoch: 89 loss: 2.726257801055908
tensor([[93.8720]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[93.871956]
[88.]
epoch: 90 loss: 34.47986602783203
tensor([[118.6464]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[118.64641]
[119.]
epoch: 91 loss: 0.12502723932266235
tensor([[64.7206]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[64.72057]
[71.]
epoch: 92 loss: 39.43119812011719
tensor([[15.3273]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[15.327284]
[17.]
epoch: 93 loss: 2.7979793548583984
tensor([[32.2251]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[32.225143]
[46.]
epoch: 94 loss: 189.74667358398438
tensor([[75.4102]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[75.410194]
[83.]
epoch: 95 loss: 57.60514831542969
tensor([[126.5956]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[126.59564]
[129.]
epoch: 96 loss: 5.780937194824219
tensor([[161.5286]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[161.5286]
[174.]
epoch: 97 loss: 155.53594970703125
tensor([[182.9799]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[182.97993]
[185.]

```

```

epoch: 98 loss: 4.080663681030273
tensor([[176.5501]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[176.55013]
[144.]
epoch: 99 loss: 1059.5106201171875
tensor([[163.3009]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[163.30089]
[136.]
epoch: 100 loss: 745.3385009765625
tensor([[102.9043]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[102.904335]
[90.]
epoch: 101 loss: 166.52186584472656
tensor([[92.3283]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[92.328316]
[95.]
epoch: 102 loss: 7.13789701461792
tensor([[140.3692]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[140.36922]
[205.]
epoch: 103 loss: 4177.13818359375
tensor([[145.9685]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[145.96846]
[190.]
epoch: 104 loss: 1938.7764892578125
tensor([[49.0669]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[49.066875]
[50.]
epoch: 105 loss: 0.8707213997840881
tensor([[181.9098]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[181.90977]
[167.]
epoch: 106 loss: 222.3013916015625
tensor([[215.1792]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[215.17921]
[203.]
epoch: 107 loss: 148.33326721191406
tensor([[134.3376]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')

```

```

[134.3376]
[100.]
epoch: 108 loss: 1179.07080078125
tensor([[205.3279]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[205.32791]
[182.]
epoch: 109 loss: 544.1914672851562
tensor([[21.5471]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[21.547077]
[10.]
epoch: 110 loss: 133.33499145507812
tensor([[109.8235]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[109.82351]
[113.]
epoch: 111 loss: 10.090093612670898
tensor([[95.1673]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[95.16727]
[111.]
epoch: 112 loss: 250.67544555664062
tensor([[112.5419]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[112.54188]
[133.]
epoch: 113 loss: 418.5347595214844
tensor([[150.8797]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[150.8797]
[176.]
epoch: 114 loss: 631.0294799804688
tensor([[81.2679]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[81.26793]
[96.]
epoch: 115 loss: 217.03392028808594
tensor([[19.3916]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[19.391611]
[2.]
epoch: 116 loss: 302.4681396484375
tensor([[199.7970]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[199.79704]
[193.]
epoch: 117 loss: 46.199790954589844

```

```

tensor([[177.6684]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[177.66841]
[154.]
epoch: 118 loss: 560.1936645507812
tensor([[144.7763]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[144.77626]
[125.]
epoch: 119 loss: 391.1004638671875
tensor([[29.2500]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[29.250002]
[44.]
epoch: 120 loss: 217.56243896484375
tensor([[157.5522]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[157.55222]
[160.]
epoch: 121 loss: 5.9916486740112305
tensor([[172.9886]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[172.98865]
[195.]
epoch: 122 loss: 484.4996337890625
tensor([[145.5867]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[145.5867]
[151.]
epoch: 123 loss: 29.30381202697754
tensor([[26.3491]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[26.349058]
[39.]
epoch: 124 loss: 160.04632568359375
tensor([[188.0996]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[188.09956]
[191.]
epoch: 125 loss: 8.412530899047852
tensor([[156.1925]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[156.1925]
[139.]
epoch: 126 loss: 295.58221435546875
tensor([[121.5124]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[121.51235]

```

```

[106.]
epoch: 127 loss: 240.633056640625
tensor([[55.7816]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[55.781593]
[56.]
epoch: 128 loss: 0.047701478004455566
tensor([[127.6111]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[127.611084]
[122.]
epoch: 129 loss: 31.484264373779297
tensor([[181.3850]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[181.385]
[200.]
epoch: 130 loss: 346.5184326171875
tensor([[135.4349]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[135.43488]
[137.]
epoch: 131 loss: 2.4496147632598877
tensor([[142.3766]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[142.37662]
[143.]
epoch: 132 loss: 0.38860583305358887
tensor([[54.6332]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[54.633247]
[60.]
epoch: 133 loss: 28.802034378051758
tensor([[24.9198]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[24.919836]
[11.]
epoch: 134 loss: 193.7618408203125
tensor([[45.3779]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[45.377934]
[49.]
epoch: 135 loss: 13.119365692138672
tensor([[80.6959]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[80.69591]
[84.]
epoch: 136 loss: 10.91702651977539
tensor([[114.5555]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],

```

```

device='cuda:0')
[114.55546]
[108.]
epoch: 137 loss: 42.974029541015625
tensor([[52.1309]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[52.130875]
[54.]
epoch: 138 loss: 3.4936296939849854
tensor([[83.9819]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[83.98187]
[86.]
epoch: 139 loss: 4.072838306427002
tensor([[68.0203]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[68.02032]
[70.]
epoch: 140 loss: 3.91914439201355
tensor([[192.2730]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[192.27296]
[183.]
epoch: 141 loss: 85.98786926269531
tensor([[24.4281]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[24.428059]
[7.]
epoch: 142 loss: 303.7372131347656
tensor([[56.2119]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[56.211933]
[62.]
epoch: 143 loss: 33.50171661376953
tensor([[36.6909]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[36.69093]
[45.]
epoch: 144 loss: 69.04065704345703
tensor([[77.1113]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[77.11127]
[85.]
epoch: 145 loss: 62.23210525512695
tensor([[24.3972]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[24.397186]
[28.]

```

```

epoch: 146 loss: 12.980266571044922
tensor([[136.0919]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[136.09192]
[132.]
epoch: 147 loss: 16.74380111694336
tensor([[49.3880]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[49.387966]
[53.]
epoch: 148 loss: 13.046788215637207
tensor([[123.1670]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[123.166985]
[118.]
epoch: 149 loss: 26.697729110717773
tensor([[111.1680]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[111.16796]
[109.]
epoch: 150 loss: 4.7000555992126465
tensor([[21.3804]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[21.38039]
[12.]
epoch: 151 loss: 87.9917221069336
tensor([[186.5453]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[186.5453]
[199.]
epoch: 152 loss: 155.11947631835938
tensor([[105.8049]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[105.804924]
[105.]
epoch: 153 loss: 0.6479026675224304
tensor([[78.7321]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[78.73214]
[82.]
epoch: 154 loss: 10.678911209106445
tensor([[20.3810]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[20.381021]
[4.]
epoch: 155 loss: 268.3378601074219
tensor([[122.6073]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')

```



```

[122.60732]
[116.]
epoch: 156 loss: 43.65671157836914
tensor([[146.1033]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[146.10332]
[142.]
epoch: 157 loss: 16.83721160888672
tensor([[128.4507]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[128.4507]
[126.]
epoch: 158 loss: 6.005924701690674
tensor([[187.0812]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[187.0812]
[196.]
epoch: 159 loss: 79.54486083984375
tensor([[43.7435]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[43.74355]
[55.]
epoch: 160 loss: 126.70767974853516
tensor([[73.1567]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[73.15668]
[81.]
epoch: 161 loss: 61.5177116394043
tensor([[143.5635]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[143.56354]
[138.]
epoch: 162 loss: 30.952951431274414
tensor([[21.9661]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[21.966106]
[33.]
epoch: 163 loss: 121.74681091308594
tensor([[199.4823]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[199.48227]
[192.]
epoch: 164 loss: 55.984352111816406
tensor([[188.8684]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[188.86836]
[179.]
epoch: 165 loss: 97.38457489013672

```

```

tensor([[44.1410]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[44.141014]
[48.]
epoch: 166 loss: 14.891772270202637
tensor([[168.6446]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[168.64458]
[165.]
epoch: 167 loss: 13.282941818237305
tensor([[15.7855]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[15.785498]
[3.]
epoch: 168 loss: 163.4689483642578
tensor([[120.4318]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[120.431816]
[123.]
epoch: 169 loss: 6.595568656921387
tensor([[131.4076]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[131.40762]
[141.]
epoch: 170 loss: 92.01368713378906
tensor([[69.3044]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[69.30438]
[76.]
epoch: 171 loss: 44.831295013427734
tensor([[153.4518]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[153.45177]
[162.]
epoch: 172 loss: 73.0722885131836
tensor([[78.7202]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[78.720215]
[79.]
epoch: 173 loss: 0.07827973365783691
tensor([[71.8197]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[71.819664]
[68.]
epoch: 174 loss: 14.58983325958252
tensor([[137.7991]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[137.79912]

```

```

[130.]
epoch: 175 loss: 60.82624053955078
tensor([[193.4194]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[193.41945]
[204.]
epoch: 176 loss: 111.94805908203125
tensor([[182.2412]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[182.24118]
[178.]
epoch: 177 loss: 17.987611770629883
tensor([[110.7458]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[110.745834]
[97.]
epoch: 178 loss: 188.94796752929688
tensor([[75.4304]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[75.43042]
[67.]
epoch: 179 loss: 71.07198333740234
tensor([[106.1325]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[106.13251]
[99.]
epoch: 180 loss: 50.87266159057617
tensor([[75.7090]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[75.70902]
[77.]
epoch: 181 loss: 1.6666228771209717
tensor([[23.1718]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[23.171759]
[34.]
epoch: 182 loss: 117.25080871582031
tensor([[178.7114]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[178.7114]
[201.]
epoch: 183 loss: 496.7818908691406
tensor([[113.0604]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[113.06041]
[115.]
epoch: 184 loss: 3.7620110511779785
tensor([[160.0181]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],

```

```

device='cuda:0')
[160.01807]
[157.]
epoch: 185 loss: 9.108724594116211
tensor([[117.8063]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[117.80625]
[112.]
epoch: 186 loss: 33.712554931640625
For epoch 1 the test accuracy over the whole test set is 0 %
tensor([[17.7672]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[17.767235]
[4.]
epoch: 0 loss: 189.53675842285156
tensor([[10.3436]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[10.343629]
[22.]
epoch: 1 loss: 135.87098693847656
tensor([[131.1168]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[131.11676]
[129.]
epoch: 2 loss: 4.480673789978027
tensor([[121.2147]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[121.214714]
[120.]
epoch: 3 loss: 1.4755302667617798
tensor([[255.6103]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[255.61026]
[204.]
epoch: 4 loss: 2663.618896484375
tensor([[55.4918]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[55.491776]
[96.]
epoch: 5 loss: 1640.916259765625
tensor([[2.5512]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[2.5512245]
[1.]
epoch: 6 loss: 2.406297445297241
tensor([[46.2726]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[46.27256]

```

```

[100.]
epoch: 7 loss: 2886.6376953125
tensor([[75.3928]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[75.392815]
[116.]
epoch: 8 loss: 1648.9434814453125
tensor([[19.3705]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[19.37047]
[9.]
epoch: 9 loss: 107.54664611816406
tensor([[55.7563]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[55.756268]
[52.]
epoch: 10 loss: 14.109545707702637
tensor([[48.8595]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[48.859497]
[38.]
epoch: 11 loss: 117.92868041992188
tensor([[85.7380]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[85.73796]
[78.]
epoch: 12 loss: 59.87603759765625
tensor([[35.0112]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[35.011177]
[26.]
epoch: 13 loss: 81.20130920410156
tensor([[122.9252]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[122.92518]
[146.]
epoch: 14 loss: 532.4473876953125
tensor([[27.0228]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[27.022846]
[0.]
epoch: 15 loss: 730.2341918945312
tensor([[19.6661]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[19.666063]
[3.]
epoch: 16 loss: 277.7576599121094
tensor([[131.1378]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],

```

```

device='cuda:0')
[131.13782]
[145.]
epoch: 17 loss: 192.1601104736328
tensor([[65.3302]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[65.33021]
[82.]
epoch: 18 loss: 277.8819580078125
tensor([[187.6877]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[187.68774]
[165.]
epoch: 19 loss: 514.7337646484375
tensor([[2.5434]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[2.5433671]
[16.]
epoch: 20 loss: 181.08096313476562
tensor([[28.9995]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[28.999527]
[43.]
epoch: 21 loss: 196.01324462890625
tensor([[18.3145]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[18.314466]
[15.]
epoch: 22 loss: 10.985688209533691
tensor([[95.7578]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[95.75778]
[80.]
epoch: 23 loss: 248.3076934814453
tensor([[151.4616]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[151.4616]
[132.]
epoch: 24 loss: 378.7536315917969
tensor([[71.8455]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[71.845535]
[71.]
epoch: 25 loss: 0.7149298787117004
tensor([[149.9814]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[149.98145]
[169.]

```

```

epoch: 26 loss: 361.7054138183594
tensor([[149.2706]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[149.27065]
[170.]
epoch: 27 loss: 429.7061462402344
tensor([[172.3366]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[172.3366]
[183.]
epoch: 28 loss: 113.70823669433594
tensor([[48.7463]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[48.746265]
[44.]
epoch: 29 loss: 22.527034759521484
tensor([[229.0120]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[229.01202]
[190.]
epoch: 30 loss: 1521.93798828125
tensor([[68.5587]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[68.558716]
[60.]
epoch: 31 loss: 73.25161743164062
tensor([[65.7814]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[65.78138]
[69.]
epoch: 32 loss: 10.359517097473145
tensor([[50.7340]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[50.733982]
[62.]
epoch: 33 loss: 126.92315673828125
tensor([[81.9031]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[81.90314]
[102.]
epoch: 34 loss: 403.8838806152344
tensor([[20.8173]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[20.817265]
[33.]
epoch: 35 loss: 148.41903686523438
tensor([[16.5899]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')

```

```

[16.58993]
[24.]
epoch: 36 loss: 54.90914535522461
tensor([[122.9302]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[122.93023]
[154.]
epoch: 37 loss: 965.3306884765625
tensor([[25.0639]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[25.063921]
[23.]
epoch: 38 loss: 4.259769916534424
tensor([[141.4595]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[141.45955]
[149.]
epoch: 39 loss: 56.858402252197266
tensor([[35.7839]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[35.783928]
[6.]
epoch: 40 loss: 887.0823364257812
tensor([[99.5254]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[99.52538]
[85.]
epoch: 41 loss: 210.98675537109375
tensor([[162.2592]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[162.25925]
[148.]
epoch: 42 loss: 203.3261260986328
tensor([[202.0579]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[202.05794]
[203.]
epoch: 43 loss: 0.8874815106391907
tensor([[67.8000]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[67.79998]
[57.]
epoch: 44 loss: 116.63957214355469
tensor([[141.5630]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[141.56297]
[144.]
epoch: 45 loss: 5.93910026550293

```



```

tensor([[107.2592]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[107.25921]
[109.]
epoch: 46 loss: 3.0303544998168945
tensor([[126.1856]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[126.18561]
[133.]
epoch: 47 loss: 46.43593978881836
tensor([[22.6834]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[22.683382]
[11.]
epoch: 48 loss: 136.5014190673828
tensor([[183.5138]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[183.51384]
[201.]
epoch: 49 loss: 305.76580810546875
tensor([[56.8859]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[56.885895]
[55.]
epoch: 50 loss: 3.5565991401672363
tensor([[69.0983]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[69.09829]
[67.]
epoch: 51 loss: 4.40281867980957
tensor([[165.7251]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[165.72513]
[156.]
epoch: 52 loss: 94.57811737060547
tensor([[140.2130]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[140.21304]
[135.]
epoch: 53 loss: 27.175819396972656
tensor([[22.3725]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[22.37251]
[29.]
epoch: 54 loss: 43.92361068725586
tensor([[176.0378]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[176.03783]

```

```

[177.]
epoch: 55 loss: 0.925777792930603
tensor([[195.6381]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[195.63808]
[198.]
epoch: 56 loss: 5.578681468963623
tensor([[35.6855]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[35.68551]
[45.]
epoch: 57 loss: 86.75975036621094
tensor([[111.7198]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[111.719795]
[110.]
epoch: 58 loss: 2.957695722579956
tensor([[108.9600]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[108.96003]
[105.]
epoch: 59 loss: 15.68183422088623
tensor([[24.7141]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[24.714127]
[30.]
epoch: 60 loss: 27.940458297729492
tensor([[60.6103]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[60.61033]
[61.]
epoch: 61 loss: 0.15184374153614044
tensor([[148.1544]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[148.15437]
[140.]
epoch: 62 loss: 66.49380493164062
tensor([[209.4327]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[209.43268]
[200.]
epoch: 63 loss: 88.97541809082031
tensor([[114.3556]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[114.355606]
[118.]
epoch: 64 loss: 13.281606674194336
tensor([[74.8375]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],

```

```

device='cuda:0')
[74.83749]
[87.]
epoch: 65 loss: 147.92674255371094
tensor([[180.6032]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[180.6032]
[194.]
epoch: 66 loss: 179.47438049316406
tensor([[132.2829]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[132.28293]
[139.]
epoch: 67 loss: 45.119049072265625
tensor([[118.7374]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[118.73742]
[117.]
epoch: 68 loss: 3.018625259399414
tensor([[159.1112]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[159.11118]
[151.]
epoch: 69 loss: 65.79116821289062
tensor([[197.0032]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[197.00316]
[185.]
epoch: 70 loss: 144.07582092285156
tensor([[109.5982]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[109.59818]
[107.]
epoch: 71 loss: 6.750553131103516
tensor([[19.6552]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[19.655224]
[10.]
epoch: 72 loss: 93.22335052490234
tensor([[187.1021]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[187.1021]
[193.]
epoch: 73 loss: 34.7852668762207
tensor([[163.8718]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[163.87184]
[171.]

```

```

epoch: 74 loss: 50.810646057128906
tensor([[107.7415]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[107.74152]
[114.]
epoch: 75 loss: 39.16852569580078
tensor([[26.7211]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[26.721085]
[32.]
epoch: 76 loss: 27.866947174072266
tensor([[89.5701]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[89.570145]
[93.]
epoch: 77 loss: 11.763907432556152
tensor([[168.1527]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[168.15274]
[164.]
epoch: 78 loss: 17.24525260925293
tensor([[157.9482]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[157.94815]
[153.]
epoch: 79 loss: 24.484193801879883
tensor([[109.9110]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[109.91101]
[108.]
epoch: 80 loss: 3.6519620418548584
tensor([[45.4516]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[45.451637]
[47.]
epoch: 81 loss: 2.3974270820617676
tensor([[89.7593]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[89.75928]
[91.]
epoch: 82 loss: 1.5393927097320557
tensor([[105.4146]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[105.41462]
[104.]
epoch: 83 loss: 2.001148223876953
tensor([[113.4112]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')

```

```

[113.41125]
[113.]
epoch: 84 loss: 0.1691243052482605
tensor([[178.3648]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[178.36484]
[178.]
epoch: 85 loss: 0.1331065148115158
tensor([[181.5511]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[181.55112]
[184.]
epoch: 86 loss: 5.997028350830078
tensor([[195.2472]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[195.24722]
[196.]
epoch: 87 loss: 0.5666733384132385
tensor([[158.9353]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[158.93529]
[157.]
epoch: 88 loss: 3.745337724685669
tensor([[32.9013]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[32.901325]
[36.]
epoch: 89 loss: 9.601785659790039
tensor([[21.0776]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[21.077585]
[20.]
epoch: 90 loss: 1.1611899137496948
tensor([[164.1090]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[164.10902]
[162.]
epoch: 91 loss: 4.447982311248779
tensor([[136.0046]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[136.00464]
[137.]
epoch: 92 loss: 0.9907441735267639
tensor([[96.5980]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[96.598]
[97.]
epoch: 93 loss: 0.16160434484481812

```

```

tensor([[140.8310]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[140.83104]
[142.]
epoch: 94 loss: 1.3664687871932983
tensor([[99.4820]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[99.481995]
[99.]
epoch: 95 loss: 0.23231881856918335
tensor([[21.9230]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[21.923002]
[21.]
epoch: 96 loss: 0.8519331216812134
tensor([[106.2774]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[106.27736]
[106.]
epoch: 97 loss: 0.0769280195236206
tensor([[197.2454]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[197.24544]
[197.]
epoch: 98 loss: 0.060239627957344055
tensor([[19.9521]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[19.952065]
[18.]
epoch: 99 loss: 3.810555934906006
tensor([[82.1840]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[82.18401]
[88.]
epoch: 100 loss: 33.82569885253906
tensor([[127.2251]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[127.225075]
[130.]
epoch: 101 loss: 7.700210094451904
tensor([[79.7804]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[79.78043]
[81.]
epoch: 102 loss: 1.487342119216919
tensor([[22.0676]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[22.067608]

```

```

[8.]
epoch: 103 loss: 197.89759826660156
tensor([[79.9061]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[79.906136]
[79.]
epoch: 104 loss: 0.8210816383361816
tensor([[26.4791]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[26.47912]
[27.]
epoch: 105 loss: 0.27131572365760803
tensor([[36.4601]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[36.460148]
[39.]
epoch: 106 loss: 6.4508490562438965
tensor([[107.2724]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[107.27244]
[103.]
epoch: 107 loss: 18.253726959228516
tensor([[18.6839]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[18.683928]
[7.]
epoch: 108 loss: 136.51416015625
tensor([[174.8493]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[174.8493]
[172.]
epoch: 109 loss: 8.118534088134766
tensor([[178.1019]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[178.1019]
[179.]
epoch: 110 loss: 0.8065868616104126
tensor([[76.2017]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[76.20173]
[83.]
epoch: 111 loss: 46.21649169921875
tensor([[172.1795]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[172.17952]
[176.]
epoch: 112 loss: 14.596070289611816
tensor([[83.7961]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],

```

```

device='cuda:0')
[83.79605]
[90.]
epoch: 113 loss: 38.488983154296875
tensor([[173.3337]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[173.3337]
[173.]
epoch: 114 loss: 0.11135198920965195
tensor([[171.0297]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[171.02966]
[168.]
epoch: 115 loss: 9.178858757019043
tensor([[126.2348]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[126.234795]
[125.]
epoch: 116 loss: 1.5247176885604858
tensor([[24.9735]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[24.973534]
[35.]
epoch: 117 loss: 100.530029296875
tensor([[168.0424]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[168.04245]
[163.]
epoch: 118 loss: 25.426301956176758
tensor([[94.1396]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[94.1396]
[94.]
epoch: 119 loss: 0.019488902762532234
tensor([[13.8478]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[13.847758]
[17.]
epoch: 120 loss: 9.936627388000488
tensor([[71.8746]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[71.87457]
[74.]
epoch: 121 loss: 4.5174407958984375
tensor([[74.8754]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[74.87538]
[76.]

```



```

epoch: 122 loss: 1.264766812324524
tensor([[171.0168]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[171.01675]
[167.]
epoch: 123 loss: 16.134313583374023
tensor([[127.6754]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[127.675354]
[126.]
epoch: 124 loss: 2.8068110942840576
tensor([[180.4096]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[180.40962]
[180.]
epoch: 125 loss: 0.16779033839702606
tensor([[157.6247]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[157.6247]
[159.]
epoch: 126 loss: 1.891464352607727
tensor([[122.4065]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[122.40651]
[124.]
epoch: 127 loss: 2.539212226867676
tensor([[184.8502]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[184.85022]
[189.]
epoch: 128 loss: 17.22067642211914
tensor([[121.5378]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[121.53779]
[122.]
epoch: 129 loss: 0.21363957226276398
tensor([[149.1735]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[149.17354]
[150.]
epoch: 130 loss: 0.6830390691757202
tensor([[104.2121]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[104.212074]
[101.]
epoch: 131 loss: 10.317420959472656
tensor([[113.6438]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')

```

```

[113.6438]
[111.]
epoch: 132 loss: 6.9896721839904785
tensor([[93.2457]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[93.24572]
[92.]
epoch: 133 loss: 1.5518181324005127
tensor([[65.3239]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[65.32387]
[65.]
epoch: 134 loss: 0.1048903539776802
tensor([[55.5307]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[55.53069]
[56.]
epoch: 135 loss: 0.22025258839130402
tensor([[57.4149]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[57.4149]
[58.]
epoch: 136 loss: 0.34233999252319336
tensor([[17.3954]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[17.395445]
[14.]
epoch: 137 loss: 11.529046058654785
tensor([[129.6535]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[129.65346]
[134.]
epoch: 138 loss: 18.892431259155273
tensor([[189.2534]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[189.25343]
[192.]
epoch: 139 loss: 7.543629169464111
tensor([[113.6260]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[113.626]
[112.]
epoch: 140 loss: 2.643874168395996
tensor([[18.4578]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[18.457802]
[12.]
epoch: 141 loss: 41.70320510864258

```

```

tensor([[90.4403]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[90.44028]
[89.]
epoch: 142 loss: 2.0743980407714844
tensor([[186.8747]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[186.87466]
[182.]
epoch: 143 loss: 23.762351989746094
tensor([[15.7904]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[15.790393]
[5.]
epoch: 144 loss: 116.43257904052734
tensor([[184.2369]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[184.23686]
[186.]
epoch: 145 loss: 3.1086549758911133
tensor([[112.2860]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[112.285995]
[115.]
epoch: 146 loss: 7.365820407867432
tensor([[36.1113]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[36.111305]
[46.]
epoch: 147 loss: 97.78628540039062
tensor([[58.1138]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[58.11381]
[64.]
epoch: 148 loss: 34.647216796875
tensor([[13.0803]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[13.0802965]
[19.]
epoch: 149 loss: 35.04288864135742
tensor([[31.5370]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[31.537046]
[37.]
epoch: 150 loss: 29.843862533569336
tensor([[53.1229]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[53.122875]

```

```

[49.]
epoch: 151 loss: 16.99810028076172
tensor([[105.0114]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[105.01135]
[95.]
epoch: 152 loss: 100.22718048095703
tensor([[128.7879]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[128.78795]
[119.]
epoch: 153 loss: 95.80393981933594
tensor([[76.5702]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[76.57022]
[75.]
epoch: 154 loss: 2.4655938148498535
tensor([[132.2707]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[132.27069]
[136.]
epoch: 155 loss: 13.907746315002441
tensor([[64.2027]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[64.20274]
[68.]
epoch: 156 loss: 14.419157028198242
tensor([[192.6217]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[192.62173]
[205.]
epoch: 157 loss: 153.22145080566406
tensor([[44.6369]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[44.636932]
[48.]
epoch: 158 loss: 11.310223579406738
tensor([[202.4556]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[202.45557]
[199.]
epoch: 159 loss: 11.940938949584961
tensor([[71.0251]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[71.02513]
[70.]
epoch: 160 loss: 1.050894021987915
tensor([[141.4058]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],

```

```

device='cuda:0')
[141.4058]
[138.]
epoch: 161 loss: 11.59952449798584
tensor([[144.0950]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[144.09497]
[141.]
epoch: 162 loss: 9.57884407043457
tensor([[16.2291]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[16.229132]
[2.]
epoch: 163 loss: 202.4681854248047
tensor([[190.0506]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[190.05058]
[191.]
epoch: 164 loss: 0.9013928771018982
tensor([[48.1268]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[48.12683]
[54.]
epoch: 165 loss: 34.49411392211914
tensor([[44.0607]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[44.06066]
[51.]
epoch: 166 loss: 48.154422760009766
tensor([[93.1971]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[93.1971]
[98.]
epoch: 167 loss: 23.067869186401367
tensor([[20.5009]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[20.500887]
[28.]
epoch: 168 loss: 56.236698150634766
tensor([[160.7672]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[160.76717]
[155.]
epoch: 169 loss: 33.26020431518555
tensor([[150.7888]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[150.78876]
[147.]

```

```

epoch: 170 loss: 14.354681968688965
tensor([[178.7281]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[178.72807]
[174.]
epoch: 171 loss: 22.35467529296875
tensor([[73.3815]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[73.381485]
[77.]
epoch: 172 loss: 13.093650817871094
tensor([[18.7429]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[18.742897]
[25.]
epoch: 173 loss: 39.151336669921875
tensor([[28.9257]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[28.925665]
[34.]
epoch: 174 loss: 25.748876571655273
tensor([[79.6663]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[79.66634]
[84.]
epoch: 175 loss: 18.780576705932617
tensor([[54.8670]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[54.866966]
[53.]
epoch: 176 loss: 3.485563039779663
tensor([[121.7933]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[121.7933]
[123.]
epoch: 177 loss: 1.4561325311660767
tensor([[21.8953]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[21.895338]
[13.]
epoch: 178 loss: 79.12703704833984
tensor([[172.8278]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[172.82782]
[175.]
epoch: 179 loss: 4.718366622924805
tensor([[85.4704]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')

```

```

[85.4704]
[86.]
epoch: 180 loss: 0.2804783284664154
tensor([[151.0227]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[151.02269]
[152.]
epoch: 181 loss: 0.95513516664505
tensor([[196.6824]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[196.68237]
[195.]
epoch: 182 loss: 2.830379009246826
tensor([[142.9060]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[142.90599]
[143.]
epoch: 183 loss: 0.008837766945362091
tensor([[55.6361]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[55.636063]
[50.]
epoch: 184 loss: 31.765201568603516
tensor([[165.4558]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[165.45584]
[166.]
epoch: 185 loss: 0.2961089611053467
tensor([[158.9098]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[158.90984]
[160.]
epoch: 186 loss: 1.18845796585083
For epoch 2 the test accuracy over the whole test set is 0 %
tensor([[201.7358]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[201.73582]
[204.]
epoch: 0 loss: 5.126490116119385
tensor([[163.6777]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[163.67767]
[164.]
epoch: 1 loss: 0.10389447212219238
tensor([[22.4309]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[22.430891]
[4.]

```

```

epoch: 2 loss: 339.69775390625
tensor([[104.8394]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[104.83938]
[103.]
epoch: 3 loss: 3.383312702178955
tensor([[136.0085]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[136.00853]
[139.]
epoch: 4 loss: 8.948894500732422
tensor([[139.6689]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[139.66891]
[143.]
epoch: 5 loss: 11.096128463745117
tensor([[108.3475]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[108.34746]
[110.]
epoch: 6 loss: 2.730895519256592
tensor([[42.4921]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[42.492123]
[45.]
epoch: 7 loss: 6.2894487380981445
tensor([[55.7740]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[55.773968]
[54.]
epoch: 8 loss: 3.1469614505767822
tensor([[152.3885]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[152.38853]
[151.]
epoch: 9 loss: 1.9280282258987427
tensor([[78.3187]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[78.31867]
[79.]
epoch: 10 loss: 0.46420758962631226
tensor([[149.6444]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[149.64441]
[148.]
epoch: 11 loss: 2.7040815353393555
tensor([[82.7046]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')

```



```

[82.70457]
[84.]
epoch: 12 loss: 1.6781467199325562
tensor([[93.1240]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[93.12397]
[93.]
epoch: 13 loss: 0.0153685687109828
tensor([[119.5761]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[119.57611]
[117.]
epoch: 14 loss: 6.63634729385376
tensor([[190.6112]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[190.61119]
[186.]
epoch: 15 loss: 21.263080596923828
tensor([[11.3989]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[11.398925]
[13.]
epoch: 16 loss: 2.563441753387451
tensor([[86.0084]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[86.008354]
[91.]
epoch: 17 loss: 24.916528701782227
tensor([[124.6922]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[124.69219]
[130.]
epoch: 18 loss: 28.17282485961914
tensor([[11.2172]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[11.217217]
[19.]
epoch: 19 loss: 60.57170486450195
tensor([[20.7783]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[20.778296]
[29.]
epoch: 20 loss: 67.59642791748047
tensor([[172.2170]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[172.217]
[170.]
epoch: 21 loss: 4.915067672729492

```

```

tensor([[181.9998]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[181.99979]
[178.]
epoch: 22 loss: 15.998291015625
tensor([[96.3759]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[96.37592]
[95.]
epoch: 23 loss: 1.8931645154953003
tensor([[197.3300]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[197.33005]
[195.]
epoch: 24 loss: 5.429121971130371
tensor([[17.9559]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[17.955906]
[17.]
epoch: 25 loss: 0.9137561321258545
tensor([[18.3029]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[18.302889]
[18.]
epoch: 26 loss: 0.0917416661977768
tensor([[154.0100]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[154.00996]
[159.]
epoch: 27 loss: 24.90045928955078
tensor([[188.8384]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[188.8384]
[194.]
epoch: 28 loss: 26.642175674438477
tensor([[161.1951]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[161.19507]
[162.]
epoch: 29 loss: 0.6479149460792542
tensor([[22.3540]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[22.354038]
[15.]
epoch: 30 loss: 54.081878662109375
tensor([[81.5115]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[81.51151]

```

```

[77.]
epoch: 31 loss: 20.35374641418457
tensor([[69.7597]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[69.759705]
[65.]
epoch: 32 loss: 22.654787063598633
tensor([[179.3232]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[179.3232]
[176.]
epoch: 33 loss: 11.043634414672852
tensor([[99.2724]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[99.27241]
[98.]
epoch: 34 loss: 1.619020938873291
tensor([[116.3168]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[116.316765]
[119.]
epoch: 35 loss: 7.199750900268555
tensor([[134.2188]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[134.21878]
[140.]
epoch: 36 loss: 33.42250061035156
tensor([[52.2908]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[52.290756]
[52.]
epoch: 37 loss: 0.08453918248414993
tensor([[52.0749]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[52.07486]
[51.]
epoch: 38 loss: 1.1553231477737427
tensor([[73.3898]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[73.3898]
[74.]
epoch: 39 loss: 0.3723427951335907
tensor([[169.3878]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[169.38782]
[168.]
epoch: 40 loss: 1.926037073135376
tensor([[193.5092]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],

```

```

device='cuda:0')
[193.5092]
[192.]
epoch: 41 loss: 2.2776877880096436
tensor([[16.6950]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[16.694986]
[11.]
epoch: 42 loss: 32.43286895751953
tensor([[77.1099]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[77.10987]
[81.]
epoch: 43 loss: 15.13310432434082
tensor([[56.6830]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[56.68301]
[58.]
epoch: 44 loss: 1.7344623804092407
tensor([[18.1529]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[18.152863]
[24.]
epoch: 45 loss: 34.18901824951172
tensor([[34.3526]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[34.352566]
[39.]
epoch: 46 loss: 21.598644256591797
tensor([[161.5672]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[161.56715]
[160.]
epoch: 47 loss: 2.4559714794158936
tensor([[106.5543]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[106.55427]
[104.]
epoch: 48 loss: 6.524284362792969
tensor([[63.8278]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[63.827812]
[62.]
epoch: 49 loss: 3.34089732170105
tensor([[172.8805]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[172.8805]
[172.]

```

```

epoch: 50 loss: 0.7752681970596313
tensor([[100.2299]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[100.22985]
[100.]
epoch: 51 loss: 0.05283137783408165
tensor([[180.4794]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[180.47937]
[182.]
epoch: 52 loss: 2.3123152256011963
tensor([[148.0097]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[148.0097]
[150.]
epoch: 53 loss: 3.961275815963745
tensor([[17.6420]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[17.642023]
[1.]
epoch: 54 loss: 276.9569396972656
tensor([[34.8039]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[34.803936]
[38.]
epoch: 55 loss: 10.214824676513672
tensor([[15.0902]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[15.090153]
[2.]
epoch: 56 loss: 171.3520965576172
tensor([[13.2571]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[13.257069]
[9.]
epoch: 57 loss: 18.12263298034668
tensor([[138.8620]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[138.86201]
[147.]
epoch: 58 loss: 66.226806640625
tensor([[101.3938]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[101.39378]
[107.]
epoch: 59 loss: 31.429662704467773
tensor([[56.3640]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')

```

```

[56.364033]
[61.]
epoch: 60 loss: 21.4921932220459
tensor([[10.4987]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[10.4987135]
[22.]
epoch: 61 loss: 132.2795867919922
tensor([[206.0439]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[206.04393]
[197.]
epoch: 62 loss: 81.79267120361328
tensor([[53.7507]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[53.750675]
[55.]
epoch: 63 loss: 1.5608124732971191
tensor([[137.2782]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[137.27818]
[135.]
epoch: 64 loss: 5.190117835998535
tensor([[8.9999]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[8.999905]
[7.]
epoch: 65 loss: 3.9996185302734375
tensor([[136.3427]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[136.3427]
[138.]
epoch: 66 loss: 2.746652841567993
tensor([[101.1648]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[101.16483]
[102.]
epoch: 67 loss: 0.6975038051605225
tensor([[127.4558]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[127.45577]
[129.]
epoch: 68 loss: 2.384638786315918
tensor([[8.6260]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[8.626001]
[5.]
epoch: 69 loss: 13.147886276245117

```

```

tensor([[170.6209]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[170.6209]
[171.]
epoch: 70 loss: 0.14372031390666962
tensor([[156.0467]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[156.04668]
[156.]
epoch: 71 loss: 0.0021787083242088556
tensor([[199.3786]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[199.37859]
[200.]
epoch: 72 loss: 0.3861555755138397
tensor([[196.4119]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[196.41194]
[196.]
epoch: 73 loss: 0.16969582438468933
tensor([[155.6432]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[155.64322]
[155.]
epoch: 74 loss: 0.4137306809425354
tensor([[46.1442]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[46.14419]
[50.]
epoch: 75 loss: 14.867257118225098
tensor([[136.0711]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[136.07108]
[136.]
epoch: 76 loss: 0.005051718093454838
tensor([[88.8471]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[88.84705]
[90.]
epoch: 77 loss: 1.3292856216430664
tensor([[9.4985]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[9.498485]
[16.]
epoch: 78 loss: 42.26970291137695
tensor([[170.7995]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[170.79955]

```

```

[167.]
epoch: 79 loss: 14.436544418334961
tensor([[107.9434]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[107.94341]
[105.]
epoch: 80 loss: 8.663679122924805
tensor([[87.3159]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[87.31586]
[88.]
epoch: 81 loss: 0.4680517315864563
tensor([[10.8739]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[10.873893]
[6.]
epoch: 82 loss: 23.754831314086914
tensor([[78.2304]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[78.23039]
[80.]
epoch: 83 loss: 3.1315109729766846
tensor([[96.2529]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[96.25287]
[97.]
epoch: 84 loss: 0.558205246925354
tensor([[10.9703]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[10.970258]
[14.]
epoch: 85 loss: 9.179338455200195
tensor([[120.7496]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[120.74957]
[122.]
epoch: 86 loss: 1.563568353652954
tensor([[117.7331]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[117.73312]
[118.]
epoch: 87 loss: 0.07122291624546051
tensor([[63.3869]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[63.386932]
[64.]
epoch: 88 loss: 0.37585192918777466
tensor([[84.1420]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],

```



```

device='cuda:0')
[84.14202]
[86.]
epoch: 89 loss: 3.452085256576538
tensor([[26.4110]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[26.410994]
[32.]
epoch: 90 loss: 31.23699188232422
tensor([[62.9267]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[62.926704]
[60.]
epoch: 91 loss: 8.565598487854004
tensor([[198.0258]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[198.02583]
[199.]
epoch: 92 loss: 0.9490010738372803
tensor([[71.1371]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[71.13708]
[67.]
epoch: 93 loss: 17.115407943725586
tensor([[72.8823]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[72.88234]
[69.]
epoch: 94 loss: 15.072559356689453
tensor([[133.0215]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[133.02145]
[132.]
epoch: 95 loss: 1.0433679819107056
tensor([[58.0873]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[58.087337]
[57.]
epoch: 96 loss: 1.1823028326034546
tensor([[180.7995]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[180.7995]
[193.]
epoch: 97 loss: 148.8522186279297
tensor([[70.9719]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[70.97194]
[71.]

```

```

epoch: 98 loss: 0.0007874148432165384
tensor([[126.4461]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[126.446144]
[120.]
epoch: 99 loss: 41.552772521972656
tensor([[21.9279]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[21.927929]
[25.]
epoch: 100 loss: 9.437621116638184
tensor([[35.4082]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[35.408207]
[36.]
epoch: 101 loss: 0.3502190411090851
tensor([[106.3891]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[106.38912]
[96.]
epoch: 102 loss: 107.93385314941406
tensor([[95.3075]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[95.30752]
[89.]
epoch: 103 loss: 39.78478240966797
tensor([[212.2420]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[212.24196]
[203.]
epoch: 104 loss: 85.4137954711914
tensor([[82.7801]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[82.78008]
[87.]
epoch: 105 loss: 17.807701110839844
tensor([[73.9922]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[73.99216]
[82.]
epoch: 106 loss: 64.12554931640625
tensor([[66.7116]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[66.711555]
[76.]
epoch: 107 loss: 86.27519989013672
tensor([[42.7343]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')

```

```

[42.73428]
[49.]
epoch: 108 loss: 39.25925064086914
tensor([[61.1474]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[61.147408]
[68.]
epoch: 109 loss: 46.95802307128906
tensor([[13.8504]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[13.850356]
[8.]
epoch: 110 loss: 34.22666549682617
tensor([[118.3038]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[118.30376]
[125.]
epoch: 111 loss: 44.83967590332031
tensor([[23.1415]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[23.141487]
[26.]
epoch: 112 loss: 8.171095848083496
tensor([[163.5497]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[163.54965]
[166.]
epoch: 113 loss: 6.004204750061035
tensor([[88.0591]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[88.05906]
[85.]
epoch: 114 loss: 9.357842445373535
tensor([[183.2739]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[183.27386]
[179.]
epoch: 115 loss: 18.265920639038086
tensor([[147.0544]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[147.05444]
[141.]
epoch: 116 loss: 36.65628433227539
tensor([[21.1512]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[21.151226]
[3.]
epoch: 117 loss: 329.4670104980469

```

```

tensor([[22.4106]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[22.410582]
[23.]
epoch: 118 loss: 0.3474140763282776
tensor([[43.1997]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[43.199703]
[44.]
epoch: 119 loss: 0.6404749155044556
tensor([[170.8148]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[170.81477]
[183.]
epoch: 120 loss: 148.479736328125
tensor([[162.3108]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[162.31084]
[169.]
epoch: 121 loss: 44.744903564453125
tensor([[149.2618]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[149.26181]
[149.]
epoch: 122 loss: 0.06854463368654251
tensor([[88.5768]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[88.57677]
[83.]
epoch: 123 loss: 31.100330352783203
tensor([[117.2199]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[117.21995]
[109.]
epoch: 124 loss: 67.56754302978516
tensor([[119.3428]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[119.34278]
[113.]
epoch: 125 loss: 40.230873107910156
tensor([[146.8826]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[146.88264]
[146.]
epoch: 126 loss: 0.7790615558624268
tensor([[122.0781]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[122.078125]

```

```

[126.]
epoch: 127 loss: 15.381103515625
tensor([[72.1705]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[72.17052]
[78.]
epoch: 128 loss: 33.982872009277344
tensor([[121.2316]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[121.23158]
[124.]
epoch: 129 loss: 7.664134502410889
tensor([[113.3902]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[113.39018]
[114.]
epoch: 130 loss: 0.37187740206718445
tensor([[27.4277]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[27.427721]
[37.]
epoch: 131 loss: 91.62852478027344
tensor([[22.1124]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[22.112366]
[30.]
epoch: 132 loss: 62.21477508544922
tensor([[28.9120]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[28.912046]
[33.]
epoch: 133 loss: 16.71136474609375
tensor([[16.0861]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[16.086132]
[10.]
epoch: 134 loss: 37.0410041809082
tensor([[203.4141]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[203.41406]
[180.]
epoch: 135 loss: 548.2183227539062
tensor([[119.1257]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[119.12567]
[112.]
epoch: 136 loss: 50.77519226074219
tensor([[105.8349]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],

```

```

device='cuda:0')
[105.834885]
[108.]
epoch: 137 loss: 4.687724590301514
tensor([[97.1975]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[97.1975]
[106.]
epoch: 138 loss: 77.48397064208984
tensor([[19.1736]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[19.17357]
[28.]
epoch: 139 loss: 77.90585327148438
tensor([[116.3029]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[116.30287]
[137.]
epoch: 140 loss: 428.3711242675781
tensor([[123.6792]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[123.67921]
[142.]
epoch: 141 loss: 335.6514587402344
tensor([[54.6410]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[54.640995]
[53.]
epoch: 142 loss: 2.6928646564483643
tensor([[74.6713]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[74.671265]
[70.]
epoch: 143 loss: 21.82071304321289
tensor([[193.0085]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[193.00854]
[191.]
epoch: 144 loss: 4.034252643585205
tensor([[203.9005]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[203.90047]
[198.]
epoch: 145 loss: 34.81550979614258
tensor([[180.3007]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[180.30072]
[173.]

```

```

epoch: 146 loss: 53.300514221191406
tensor([[172.9895]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[172.98947]
[174.]
epoch: 147 loss: 1.0211679935455322
tensor([[123.9478]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[123.94776]
[123.]
epoch: 148 loss: 0.8982519507408142
tensor([[110.6256]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[110.62564]
[115.]
epoch: 149 loss: 19.13501739501953
tensor([[29.6218]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[29.621803]
[27.]
epoch: 150 loss: 6.873852252960205
tensor([[21.3473]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[21.347263]
[12.]
epoch: 151 loss: 87.37133026123047
tensor([[155.9123]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[155.91234]
[184.]
epoch: 152 loss: 788.916748046875
tensor([[190.7062]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[190.7062]
[201.]
epoch: 153 loss: 105.96216583251953
tensor([[158.9136]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[158.91362]
[134.]
epoch: 154 loss: 620.6884765625
tensor([[64.3700]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[64.36996]
[47.]
epoch: 155 loss: 301.71539306640625
tensor([[122.7852]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')

```

```

[122.78525]
[101.]
epoch: 156 loss: 474.5970153808594
tensor([[203.2818]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[203.28181]
[189.]
epoch: 157 loss: 203.97023010253906
tensor([[145.9267]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[145.92668]
[152.]
epoch: 158 loss: 36.885196685791016
tensor([[143.6087]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[143.60875]
[163.]
epoch: 159 loss: 376.0205993652344
tensor([[133.6904]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[133.69037]
[154.]
epoch: 160 loss: 412.48114013671875
tensor([[131.6329]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[131.63287]
[145.]
epoch: 161 loss: 178.68006896972656
tensor([[9.3353]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[9.3352585]
[20.]
epoch: 162 loss: 113.73670959472656
tensor([[171.5250]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[171.52502]
[157.]
epoch: 163 loss: 210.97633361816406
tensor([[78.2835]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[78.28354]
[75.]
epoch: 164 loss: 10.78162670135498
tensor([[107.5412]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[107.5412]
[99.]
epoch: 165 loss: 72.95207214355469

```



```

tensor([[30.1853]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[30.185299]
[35.]
epoch: 166 loss: 23.181346893310547
tensor([[39.0282]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[39.028164]
[43.]
epoch: 167 loss: 15.775482177734375
tensor([[185.1194]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[185.11942]
[165.]
epoch: 168 loss: 404.7908630371094
tensor([[117.1818]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[117.181786]
[116.]
epoch: 169 loss: 1.3966171741485596
tensor([[47.3205]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[47.320538]
[56.]
epoch: 170 loss: 75.33306884765625
tensor([[142.5999]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[142.59991]
[153.]
epoch: 171 loss: 108.16177368164062
tensor([[168.1011]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[168.10109]
[177.]
epoch: 172 loss: 79.19060516357422
tensor([[208.2591]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[208.25906]
[205.]
epoch: 173 loss: 10.621496200561523
tensor([[136.9614]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[136.9614]
[133.]
epoch: 174 loss: 15.692652702331543
tensor([[145.6469]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[145.64691]

```

```

[144.]
epoch: 175 loss: 2.712317943572998
tensor([[42.0789]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[42.07891]
[46.]
epoch: 176 loss: 15.374939918518066
tensor([[30.5092]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[30.509165]
[34.]
epoch: 177 loss: 12.185930252075195
tensor([[87.0929]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[87.09292]
[92.]
epoch: 178 loss: 24.079450607299805
tensor([[22.5359]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[22.535866]
[21.]
epoch: 179 loss: 2.3588836193084717
tensor([[192.3401]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[192.34013]
[185.]
epoch: 180 loss: 53.8775634765625
tensor([[108.7203]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[108.72034]
[111.]
epoch: 181 loss: 5.196863651275635
tensor([[46.4737]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[46.47366]
[48.]
epoch: 182 loss: 2.3297152519226074
tensor([[186.1747]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[186.17467]
[190.]
epoch: 183 loss: 14.633170127868652
tensor([[20.7049]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[20.704868]
[0.]
epoch: 184 loss: 428.6915588378906
tensor([[170.9841]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],

```

```

device='cuda:0')
[170.98413]
[175.]
epoch: 185 loss: 16.12720489501953
tensor([[88.4797]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[88.47965]
[94.]
epoch: 186 loss: 30.4742374420166
For epoch 3 the test accuracy over the whole test set is 0 %
tensor([[190.7039]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[190.70389]
[192.]
epoch: 0 loss: 1.6799064874649048
tensor([[47.7477]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[47.747684]
[49.]
epoch: 1 loss: 1.5682941675186157
tensor([[35.6151]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[35.61507]
[38.]
epoch: 2 loss: 5.687889575958252
tensor([[101.1716]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[101.17159]
[98.]
epoch: 3 loss: 10.059000015258789
tensor([[198.1620]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[198.16203]
[184.]
epoch: 4 loss: 200.5631866455078
tensor([[170.0811]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[170.08107]
[167.]
epoch: 5 loss: 9.492992401123047
tensor([[95.8378]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[95.83777]
[103.]
epoch: 6 loss: 51.29755783081055
tensor([[78.8400]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[78.84]

```

```

[90.]
epoch: 7 loss: 124.54568481445312
tensor([[154.8239]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[154.82385]
[168.]
epoch: 8 loss: 173.61085510253906
tensor([[184.3505]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[184.3505]
[195.]
epoch: 9 loss: 113.41197204589844
tensor([[165.6351]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[165.63506]
[162.]
epoch: 10 loss: 13.213628768920898
tensor([[130.3016]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[130.30159]
[122.]
epoch: 11 loss: 68.91639709472656
tensor([[23.0826]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[23.0826]
[24.]
epoch: 12 loss: 0.8416234254837036
tensor([[102.5982]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[102.59823]
[93.]
epoch: 13 loss: 92.12599182128906
tensor([[215.4028]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[215.40285]
[198.]
epoch: 14 loss: 302.8591003417969
tensor([[61.8585]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[61.85848]
[61.]
epoch: 15 loss: 0.7369853854179382
tensor([[164.1107]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[164.11069]
[173.]
epoch: 16 loss: 79.01988220214844
tensor([[67.3722]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],

```

```

device='cuda:0')
[67.37219]
[76.]
epoch: 17 loss: 74.4390640258789
tensor([[145.2780]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[145.27795]
[156.]
epoch: 18 loss: 114.96226501464844
tensor([[173.1118]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[173.11179]
[183.]
epoch: 19 loss: 97.77677917480469
tensor([[147.3479]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[147.34795]
[143.]
epoch: 20 loss: 18.90463638305664
tensor([[86.4270]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[86.42705]
[82.]
epoch: 21 loss: 19.598751068115234
tensor([[97.7384]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[97.73838]
[91.]
epoch: 22 loss: 45.40576934814453
tensor([[106.3886]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[106.38863]
[99.]
epoch: 23 loss: 54.5919075012207
tensor([[56.8461]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[56.846104]
[53.]
epoch: 24 loss: 14.792513847351074
tensor([[57.1342]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[57.13422]
[55.]
epoch: 25 loss: 4.554895401000977
tensor([[156.2994]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[156.29938]
[152.]

```

```

epoch: 26 loss: 18.48464584350586
tensor([[80.7722]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[80.77225]
[86.]
epoch: 27 loss: 27.329397201538086
tensor([[170.1954]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[170.19543]
[175.]
epoch: 28 loss: 23.08384895324707
tensor([[144.6852]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[144.68517]
[147.]
epoch: 29 loss: 5.358458995819092
tensor([[48.7170]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[48.717007]
[52.]
epoch: 30 loss: 10.778044700622559
tensor([[13.7854]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[13.785387]
[21.]
epoch: 31 loss: 52.05064010620117
tensor([[111.3437]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[111.34368]
[112.]
epoch: 32 loss: 0.4307541847229004
tensor([[15.1324]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[15.132424]
[18.]
epoch: 33 loss: 8.222990036010742
tensor([[158.0796]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[158.0796]
[150.]
epoch: 34 loss: 65.28002166748047
tensor([[196.1899]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[196.18993]
[190.]
epoch: 35 loss: 38.315185546875
tensor([[185.2044]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')

```

```

[185.2044]
[185.]
epoch: 36 loss: 0.041782114654779434
tensor([[98.8807]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[98.880745]
[105.]
epoch: 37 loss: 37.445281982421875
tensor([[159.7140]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[159.714]
[164.]
epoch: 38 loss: 18.3697566986084
tensor([[58.3488]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[58.348835]
[62.]
epoch: 39 loss: 13.331006050109863
tensor([[25.4671]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[25.46707]
[29.]
epoch: 40 loss: 12.481596946716309
tensor([[43.3559]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[43.355915]
[46.]
epoch: 41 loss: 6.991185188293457
tensor([[141.9692]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[141.96921]
[141.]
epoch: 42 loss: 0.9393637180328369
tensor([[32.5622]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[32.56224]
[32.]
epoch: 43 loss: 0.31611448526382446
tensor([[194.9343]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[194.93431]
[196.]
epoch: 44 loss: 1.1356931924819946
tensor([[61.1055]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[61.10554]
[57.]
epoch: 45 loss: 16.85546875

```

```

tensor([[165.2210]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[165.22104]
[160.]
epoch: 46 loss: 27.259246826171875
tensor([[103.0796]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[103.07963]
[102.]
epoch: 47 loss: 1.165596604347229
tensor([[20.0353]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[20.035295]
[9.]
epoch: 48 loss: 121.77774810791016
tensor([[186.0901]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[186.09009]
[194.]
epoch: 49 loss: 62.56671142578125
tensor([[96.1144]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[96.1144]
[100.]
epoch: 50 loss: 15.09786605834961
tensor([[189.1985]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[189.19852]
[193.]
epoch: 51 loss: 14.451273918151855
tensor([[21.2999]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[21.299938]
[22.]
epoch: 52 loss: 0.49008652567863464
tensor([[111.4853]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[111.485275]
[108.]
epoch: 53 loss: 12.147143363952637
tensor([[20.3998]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[20.399792]
[19.]
epoch: 54 loss: 1.9594168663024902
tensor([[136.8029]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[136.80292]

```



```

[130.]
epoch: 55 loss: 46.279685974121094
tensor([[16.0409]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[16.04087]
[7.]
epoch: 56 loss: 81.73734283447266
tensor([[189.8418]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[189.84181]
[189.]
epoch: 57 loss: 0.7086476683616638
tensor([[134.9748]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[134.97476]
[140.]
epoch: 58 loss: 25.25301742553711
tensor([[143.6721]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[143.67209]
[149.]
epoch: 59 loss: 28.386640548706055
tensor([[121.1618]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[121.161835]
[124.]
epoch: 60 loss: 8.055182456970215
tensor([[150.5241]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[150.52414]
[146.]
epoch: 61 loss: 20.467838287353516
tensor([[209.6567]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[209.6567]
[203.]
epoch: 62 loss: 44.311553955078125
tensor([[4.9434]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[4.943409]
[3.]
epoch: 63 loss: 3.7768383026123047
tensor([[162.3074]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[162.30743]
[163.]
epoch: 64 loss: 0.47964754700660706
tensor([[192.2341]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],

```

```

device='cuda:0')
[192.23409]
[199.]
epoch: 65 loss: 45.77760314941406
tensor([[100.1991]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[100.19909]
[106.]
epoch: 66 loss: 33.65056610107422
tensor([[203.6080]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[203.60803]
[200.]
epoch: 67 loss: 13.01789665222168
tensor([[162.9238]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[162.92377]
[155.]
epoch: 68 loss: 62.78608322143555
tensor([[93.5371]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[93.537056]
[97.]
epoch: 69 loss: 11.991981506347656
tensor([[51.4736]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[51.47361]
[56.]
epoch: 70 loss: 20.48820686340332
tensor([[4.7672]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[4.767154]
[16.]
epoch: 71 loss: 126.17681121826172
tensor([[79.0129]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[79.01289]
[83.]
epoch: 72 loss: 15.897016525268555
tensor([[181.5860]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[181.58604]
[176.]
epoch: 73 loss: 31.20389175415039
tensor([[181.8168]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[181.81677]
[178.]

```

```

epoch: 74 loss: 14.56775188446045
tensor([[77.2301]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[77.23014]
[80.]
epoch: 75 loss: 7.672120571136475
tensor([[29.2534]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[29.253368]
[34.]
epoch: 76 loss: 22.5305118560791
tensor([[22.4642]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[22.464195]
[27.]
epoch: 77 loss: 20.573524475097656
tensor([[81.0850]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[81.08496]
[84.]
epoch: 78 loss: 8.497452735900879
tensor([[162.8736]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[162.87357]
[165.]
epoch: 79 loss: 4.521722793579102
tensor([[142.3599]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[142.35991]
[142.]
epoch: 80 loss: 0.1295345276594162
tensor([[156.1022]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[156.10225]
[154.]
epoch: 81 loss: 4.4194512367248535
tensor([[21.2764]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[21.276445]
[15.]
epoch: 82 loss: 39.393768310546875
tensor([[169.6389]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[169.63893]
[171.]
epoch: 83 loss: 1.8525080680847168
tensor([[38.2020]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')

```

```

[38.202045]
[36.]
epoch: 84 loss: 4.84900426864624
tensor([[21.7429]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[21.742857]
[14.]
epoch: 85 loss: 59.95183563232422
tensor([[132.1965]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[132.19655]
[133.]
epoch: 86 loss: 0.6455343961715698
tensor([[36.4329]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[36.432873]
[35.]
epoch: 87 loss: 2.05312442779541
tensor([[171.7571]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[171.75711]
[179.]
epoch: 88 loss: 52.45944595336914
tensor([[29.3862]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[29.386211]
[28.]
epoch: 89 loss: 1.9215819835662842
tensor([[135.6588]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[135.65878]
[134.]
epoch: 90 loss: 2.751560926437378
tensor([[184.6514]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[184.65141]
[186.]
epoch: 91 loss: 1.8186869621276855
tensor([[176.8026]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[176.8026]
[174.]
epoch: 92 loss: 7.854550361633301
tensor([[120.6648]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[120.664825]
[115.]
epoch: 93 loss: 32.090248107910156

```

```

tensor([[182.5124]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[182.51244]
[182.]
epoch: 94 loss: 0.26259055733680725
tensor([[70.3018]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[70.30181]
[68.]
epoch: 95 loss: 5.298335075378418
tensor([[104.4581]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[104.45805]
[104.]
epoch: 96 loss: 0.20981308817863464
tensor([[183.5501]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[183.5501]
[191.]
epoch: 97 loss: 55.50109100341797
tensor([[107.6909]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[107.6909]
[109.]
epoch: 98 loss: 1.7137356996536255
tensor([[75.2224]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[75.22236]
[75.]
epoch: 99 loss: 0.04944339394569397
tensor([[141.3870]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[141.387]
[139.]
epoch: 100 loss: 5.697737693786621
tensor([[60.4703]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[60.470272]
[58.]
epoch: 101 loss: 6.102243900299072
tensor([[43.9566]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[43.956646]
[45.]
epoch: 102 loss: 1.0885876417160034
tensor([[138.4702]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[138.47015]

```

```

[136.]
epoch: 103 loss: 6.101659774780273
tensor([[133.2788]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[133.27878]
[132.]
epoch: 104 loss: 1.6352733373641968
tensor([[100.8480]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[100.848045]
[101.]
epoch: 105 loss: 0.023090215399861336
tensor([[15.4403]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[15.440296]
[20.]
epoch: 106 loss: 20.7908992767334
tensor([[59.2632]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[59.263195]
[60.]
epoch: 107 loss: 0.5428815484046936
tensor([[146.7897]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[146.78967]
[148.]
epoch: 108 loss: 1.464891791343689
tensor([[28.9876]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[28.987566]
[33.]
epoch: 109 loss: 16.099626541137695
tensor([[119.4645]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[119.464455]
[120.]
epoch: 110 loss: 0.2868088185787201
tensor([[112.5429]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[112.5429]
[113.]
epoch: 111 loss: 0.208940327167511
tensor([[85.3692]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[85.36924]
[87.]
epoch: 112 loss: 2.65937876701355
tensor([[16.2437]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],

```

```

device='cuda:0')
[16.243673]
[4.]
epoch: 113 loss: 149.90753173828125
tensor([[199.7904]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[199.79037]
[197.]
epoch: 114 loss: 7.786191463470459
tensor([[14.6281]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[14.628069]
[2.]
epoch: 115 loss: 159.46812438964844
tensor([[159.1483]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[159.14825]
[159.]
epoch: 116 loss: 0.02197936549782753
tensor([[62.6824]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[62.68236]
[67.]
epoch: 117 loss: 18.64200210571289
tensor([[72.9661]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[72.966125]
[79.]
epoch: 118 loss: 36.40764236450195
tensor([[118.3534]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[118.353424]
[123.]
epoch: 119 loss: 21.590667724609375
tensor([[169.5349]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[169.53493]
[172.]
epoch: 120 loss: 6.076582908630371
tensor([[48.5268]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[48.526848]
[51.]
epoch: 121 loss: 6.116481781005859
tensor([[18.3606]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[18.360647]
[26.]

```

```

epoch: 122 loss: 58.359710693359375
tensor([[97.4193]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[97.41935]
[95.]
epoch: 123 loss: 5.85325288772583
tensor([[13.2630]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[13.263015]
[13.]
epoch: 124 loss: 0.06917677819728851
tensor([[13.9669]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[13.966913]
[12.]
epoch: 125 loss: 3.8687477111816406
tensor([[59.8869]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[59.88688]
[54.]
epoch: 126 loss: 34.65534210205078
tensor([[76.2174]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[76.21743]
[70.]
epoch: 127 loss: 38.656436920166016
tensor([[13.3607]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[13.36067]
[10.]
epoch: 128 loss: 11.294103622436523
tensor([[77.4225]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[77.42255]
[74.]
epoch: 129 loss: 11.713823318481445
tensor([[10.5201]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[10.520141]
[8.]
epoch: 130 loss: 6.351109027862549
tensor([[36.5349]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[36.534935]
[43.]
epoch: 131 loss: 41.79706573486328
tensor([[120.7228]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')

```



```

[120.72285]
[117.]
epoch: 132 loss: 13.859589576721191
tensor([[30.0289]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[30.028904]
[39.]
epoch: 133 loss: 80.4805679321289
tensor([[130.4812]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[130.48123]
[129.]
epoch: 134 loss: 2.194047212600708
tensor([[156.2337]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[156.23366]
[153.]
epoch: 135 loss: 10.45654296875
tensor([[107.9892]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[107.989235]
[111.]
epoch: 136 loss: 9.064706802368164
tensor([[8.4309]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[8.430873]
[11.]
epoch: 137 loss: 6.600413799285889
tensor([[123.1266]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[123.12658]
[126.]
epoch: 138 loss: 8.256546974182129
tensor([[164.6306]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[164.63063]
[166.]
epoch: 139 loss: 1.8751728534698486
tensor([[157.1917]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[157.19167]
[157.]
epoch: 140 loss: 0.03673572093248367
tensor([[29.8655]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[29.865532]
[37.]
epoch: 141 loss: 50.900634765625

```

```

tensor([[200.9911]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[200.99107]
[201.]
epoch: 142 loss: 7.968046702444553e-05
tensor([[46.4358]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[46.435753]
[48.]
epoch: 143 loss: 2.446869134902954
tensor([[141.9965]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[141.99646]
[138.]
epoch: 144 loss: 15.971692085266113
tensor([[122.4489]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[122.44888]
[119.]
epoch: 145 loss: 11.894794464111328
tensor([[92.2068]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[92.20682]
[92.]
epoch: 146 loss: 0.0427735298871994
tensor([[93.7801]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[93.78011]
[94.]
epoch: 147 loss: 0.048350196331739426
tensor([[63.4037]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[63.403652]
[64.]
epoch: 148 loss: 0.35563069581985474
tensor([[18.8490]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[18.849026]
[23.]
epoch: 149 loss: 17.230587005615234
tensor([[113.2267]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[113.22666]
[114.]
epoch: 150 loss: 0.5980521440505981
tensor([[177.5944]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[177.5944]

```

```

[180.]
epoch: 151 loss: 5.786881923675537
tensor([[84.2445]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[84.244484]
[85.]
epoch: 152 loss: 0.5708044767379761
tensor([[66.3196]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[66.31958]
[65.]
epoch: 153 loss: 1.7412915229797363
tensor([[17.2992]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[17.299166]
[1.]
epoch: 154 loss: 265.6628112792969
tensor([[18.0702]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[18.070158]
[17.]
epoch: 155 loss: 1.1452381610870361
tensor([[136.8281]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[136.82814]
[137.]
epoch: 156 loss: 0.029535770416259766
tensor([[27.6451]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[27.645134]
[30.]
epoch: 157 loss: 5.545393943786621
tensor([[85.0485]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[85.04849]
[88.]
epoch: 158 loss: 8.711397171020508
tensor([[149.9696]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[149.9696]
[151.]
epoch: 159 loss: 1.0617148876190186
tensor([[76.2888]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[76.28881]
[77.]
epoch: 160 loss: 0.5057901740074158
tensor([[13.8579]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],

```

```

device='cuda:0')
[13.85788]
[0.]
epoch: 161 loss: 192.04083251953125
tensor([[49.7903]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[49.790314]
[50.]
epoch: 162 loss: 0.043968334794044495
tensor([[115.3535]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[115.35354]
[116.]
epoch: 163 loss: 0.41791245341300964
tensor([[170.0937]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[170.09372]
[170.]
epoch: 164 loss: 0.008783341385424137
tensor([[8.6153]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[8.6153345]
[5.]
epoch: 165 loss: 13.070643424987793
tensor([[106.0046]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[106.00456]
[110.]
epoch: 166 loss: 15.963521957397461
tensor([[40.8906]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[40.890644]
[47.]
epoch: 167 loss: 37.3242301940918
tensor([[178.8207]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[178.82066]
[177.]
epoch: 168 loss: 3.3148155212402344
tensor([[94.1014]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[94.10143]
[96.]
epoch: 169 loss: 3.604557514190674
tensor([[75.6525]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[75.652504]
[78.]

```

```

epoch: 170 loss: 5.510737419128418
tensor([[207.8200]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[207.81998]
[204.]
epoch: 171 loss: 14.592223167419434
tensor([[6.2148]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[6.214754]
[6.]
epoch: 172 loss: 0.046119324862957
tensor([[12.5383]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[12.538316]
[25.]
epoch: 173 loss: 155.2935791015625
tensor([[69.8262]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[69.82624]
[71.]
epoch: 174 loss: 1.377711296081543
tensor([[109.1885]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[109.18849]
[107.]
epoch: 175 loss: 4.789496421813965
tensor([[146.4540]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[146.45398]
[144.]
epoch: 176 loss: 6.022015571594238
tensor([[119.4914]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[119.49145]
[118.]
epoch: 177 loss: 2.2244155406951904
tensor([[68.7382]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[68.738205]
[69.]
epoch: 178 loss: 0.06853664666414261
tensor([[130.4520]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[130.45203]
[135.]
epoch: 179 loss: 20.684064865112305
tensor([[192.0868]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')

```

```

[192.08684]
[205.]
epoch: 180 loss: 166.749755859375
tensor([[176.1192]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[176.11917]
[169.]
epoch: 181 loss: 50.68259811401367
tensor([[46.4538]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[46.453842]
[44.]
epoch: 182 loss: 6.021341323852539
tensor([[158.6632]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[158.66324]
[145.]
epoch: 183 loss: 186.68408203125
tensor([[86.6133]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[86.61329]
[81.]
epoch: 184 loss: 31.50901222229004
tensor([[90.0715]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[90.071495]
[89.]
epoch: 185 loss: 1.1481016874313354
tensor([[123.0060]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[123.00599]
[125.]
epoch: 186 loss: 3.9760794639587402
For epoch 4 the test accuracy over the whole test set is 0 %
tensor([[28.8479]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[28.847853]
[36.]
epoch: 0 loss: 51.1532096862793
tensor([[190.3050]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[190.305]
[204.]
epoch: 1 loss: 187.55322265625
tensor([[169.7801]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[169.7801]
[171.]

```

```

epoch: 2 loss: 1.4881423711776733
tensor([[69.0808]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[69.08084]
[68.]
epoch: 3 loss: 1.168217420578003
tensor([[127.5476]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[127.547554]
[120.]
epoch: 4 loss: 56.965572357177734
tensor([[213.4194]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[213.41939]
[197.]
epoch: 5 loss: 269.5962829589844
tensor([[69.7788]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[69.77881]
[71.]
epoch: 6 loss: 1.4913084506988525
tensor([[45.3229]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[45.32293]
[49.]
epoch: 7 loss: 13.520848274230957
tensor([[8.6986]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[8.698634]
[4.]
epoch: 8 loss: 22.077163696289062
tensor([[171.8486]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[171.84865]
[192.]
epoch: 9 loss: 406.0769958496094
tensor([[131.8909]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[131.89085]
[144.]
epoch: 10 loss: 146.6314239501953
tensor([[12.0029]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[12.002901]
[14.]
epoch: 11 loss: 3.9884040355682373
tensor([[14.7782]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')

```

```

[14.778171]
[12.]
epoch: 12 loss: 7.718231678009033
tensor([[77.0013]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[77.00126]
[69.]
epoch: 13 loss: 64.0201416015625
tensor([[16.9483]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[16.948269]
[11.]
epoch: 14 loss: 35.38190460205078
tensor([[169.3416]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[169.34164]
[146.]
epoch: 15 loss: 544.8323364257812
tensor([[131.8297]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[131.82965]
[122.]
epoch: 16 loss: 96.62203979492188
tensor([[172.4209]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[172.42093]
[168.]
epoch: 17 loss: 19.544612884521484
tensor([[154.1546]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[154.15457]
[163.]
epoch: 18 loss: 78.24160766601562
tensor([[69.5910]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[69.590996]
[85.]
epoch: 19 loss: 237.43740844726562
tensor([[148.6331]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[148.63313]
[162.]
epoch: 20 loss: 178.6731414794922
tensor([[27.0940]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[27.09396]
[30.]
epoch: 21 loss: 8.445069313049316

```



```

tensor([[155.3165]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[155.3165]
[149.]
epoch: 22 loss: 39.89814376831055
tensor([[70.1282]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[70.12823]
[54.]
epoch: 23 loss: 260.1197204589844
tensor([[15.6782]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[15.678202]
[24.]
epoch: 24 loss: 69.25232696533203
tensor([[68.2476]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[68.24762]
[86.]
epoch: 25 loss: 315.1470031738281
tensor([[39.3284]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[39.32835]
[51.]
epoch: 26 loss: 136.2274169921875
tensor([[120.9005]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[120.90053]
[126.]
epoch: 27 loss: 26.004615783691406
tensor([[180.4456]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[180.4456]
[170.]
epoch: 28 loss: 109.11061096191406
tensor([[76.9310]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[76.931]
[77.]
epoch: 29 loss: 0.004761033691465855
tensor([[8.4569]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[8.456895]
[0.]
epoch: 30 loss: 71.51907348632812
tensor([[10.6661]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[10.666088]

```

```

[21.]
epoch: 31 loss: 106.78973388671875
tensor([[10.1046]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[10.104628]
[6.]
epoch: 32 loss: 16.84796714782715
tensor([[73.3094]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[73.30943]
[70.]
epoch: 33 loss: 10.952346801757812
tensor([[113.8088]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[113.808815]
[108.]
epoch: 34 loss: 33.742332458496094
tensor([[149.1812]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[149.18118]
[147.]
epoch: 35 loss: 4.757558822631836
tensor([[136.2471]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[136.24706]
[138.]
epoch: 36 loss: 3.0728158950805664
tensor([[42.0821]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[42.082085]
[45.]
epoch: 37 loss: 8.514229774475098
tensor([[133.3521]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[133.35211]
[134.]
epoch: 38 loss: 0.4197590947151184
tensor([[63.7639]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[63.76394]
[58.]
epoch: 39 loss: 33.222991943359375
tensor([[158.7994]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[158.79941]
[166.]
epoch: 40 loss: 51.84852600097656
tensor([[171.9999]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],

```

```

device='cuda:0')
[171.99992]
[177.]
epoch: 41 loss: 25.000762939453125
tensor([[194.8860]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[194.88596]
[189.]
epoch: 42 loss: 34.644474029541016
tensor([[152.5552]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[152.55522]
[143.]
epoch: 43 loss: 91.30226135253906
tensor([[127.0725]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[127.07254]
[125.]
epoch: 44 loss: 4.2954230308532715
tensor([[94.0281]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[94.02811]
[100.]
epoch: 45 loss: 35.663509368896484
tensor([[17.3717]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[17.37175]
[26.]
epoch: 46 loss: 74.44670104980469
tensor([[128.8886]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[128.88863]
[141.]
epoch: 47 loss: 146.6853790283203
tensor([[135.6422]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[135.64224]
[137.]
epoch: 48 loss: 1.843505620956421
tensor([[64.1139]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[64.1139]
[52.]
epoch: 49 loss: 146.74655151367188
tensor([[167.5228]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[167.5228]
[160.]

```

```

epoch: 50 loss: 56.59246826171875
tensor([[188.0271]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[188.02707]
[201.]
epoch: 51 loss: 168.29693603515625
tensor([[127.5969]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[127.59693]
[117.]
epoch: 52 loss: 112.29496002197266
tensor([[88.3928]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[88.39284]
[81.]
epoch: 53 loss: 54.65404510498047
tensor([[176.6877]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[176.68773]
[174.]
epoch: 54 loss: 7.223886489868164
tensor([[26.5077]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[26.5077]
[32.]
epoch: 55 loss: 30.165359497070312
tensor([[71.8286]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[71.8286]
[76.]
epoch: 56 loss: 17.40059471130371
tensor([[140.6393]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[140.63933]
[145.]
epoch: 57 loss: 19.015460968017578
tensor([[190.8865]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[190.88646]
[195.]
epoch: 58 loss: 16.92121696472168
tensor([[149.8842]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[149.88417]
[140.]
epoch: 59 loss: 97.69683074951172
tensor([[14.7680]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')

```

```

[14.768023]
[18.]
epoch: 60 loss: 10.4456787109375
tensor([[13.7192]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[13.719171]
[7.]
epoch: 61 loss: 45.147254943847656
tensor([[11.9474]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[11.947388]
[2.]
epoch: 62 loss: 98.95052337646484
tensor([[147.6895]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[147.68947]
[148.]
epoch: 63 loss: 0.09642988443374634
tensor([[99.7253]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[99.72529]
[102.]
epoch: 64 loss: 5.174313068389893
tensor([[184.3008]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[184.30083]
[190.]
epoch: 65 loss: 32.48057174682617
tensor([[123.9691]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[123.96914]
[124.]
epoch: 66 loss: 0.0009523952030576766
tensor([[83.9409]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[83.940926]
[87.]
epoch: 67 loss: 9.357935905456543
tensor([[91.5895]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[91.58954]
[90.]
epoch: 68 loss: 2.526632785797119
tensor([[199.0758]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[199.07576]
[182.]
epoch: 69 loss: 291.5815734863281

```

```

tensor([[24.5177]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[24.517656]
[37.]
epoch: 70 loss: 155.80889892578125
tensor([[147.7524]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[147.75238]
[155.]
epoch: 71 loss: 52.52798843383789
tensor([[153.4986]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[153.4986]
[164.]
epoch: 72 loss: 110.27947998046875
tensor([[37.9405]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[37.940487]
[47.]
epoch: 73 loss: 82.07477569580078
tensor([[26.7076]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[26.707586]
[34.]
epoch: 74 loss: 53.179298400878906
tensor([[68.9775]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[68.97753]
[65.]
epoch: 75 loss: 15.820755958557129
tensor([[159.4606]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[159.46062]
[142.]
epoch: 76 loss: 304.8731384277344
tensor([[39.5286]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[39.52862]
[39.]
epoch: 77 loss: 0.27944087982177734
tensor([[101.7614]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[101.76144]
[98.]
epoch: 78 loss: 14.148404121398926
tensor([[146.1402]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[146.14018]

```

```

[152.]
epoch: 79 loss: 34.337459564208984
tensor([[45.0471]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[45.047085]
[48.]
epoch: 80 loss: 8.719708442687988
tensor([[143.0432]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[143.04323]
[154.]
epoch: 81 loss: 120.05084991455078
tensor([[90.4239]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[90.42385]
[93.]
epoch: 82 loss: 6.6365437507629395
tensor([[187.8433]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[187.84332]
[199.]
epoch: 83 loss: 124.47145080566406
tensor([[81.8666]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[81.86657]
[75.]
epoch: 84 loss: 47.149776458740234
tensor([[43.2797]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[43.279682]
[38.]
epoch: 85 loss: 27.875043869018555
tensor([[133.5356]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[133.53558]
[114.]
epoch: 86 loss: 381.6390075683594
tensor([[107.6935]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[107.693474]
[95.]
epoch: 87 loss: 161.12428283691406
tensor([[177.6570]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[177.65698]
[167.]
epoch: 88 loss: 113.57127380371094
tensor([[9.4755]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],

```

```

device='cuda:0')
[9.475515]
[1.]
epoch: 89 loss: 71.83435821533203
tensor([[99.6581]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[99.6581]
[116.]
epoch: 90 loss: 267.0578308105469
tensor([[77.1424]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[77.14244]
[96.]
epoch: 91 loss: 355.6075439453125
tensor([[75.0764]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[75.076355]
[94.]
epoch: 92 loss: 358.1043395996094
tensor([[92.3423]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[92.34235]
[109.]
epoch: 93 loss: 277.4774169921875
tensor([[101.0485]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[101.04846]
[110.]
epoch: 94 loss: 80.13003540039062
tensor([[132.3471]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[132.34712]
[129.]
epoch: 95 loss: 11.203227043151855
tensor([[99.9179]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[99.91792]
[92.]
epoch: 96 loss: 62.693504333496094
tensor([[103.6230]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[103.623024]
[91.]
epoch: 97 loss: 159.34072875976562
tensor([[135.5797]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[135.57971]
[118.]

```



```

epoch: 98 loss: 309.0462646484375
tensor([[67.3147]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[67.314674]
[62.]
epoch: 99 loss: 28.245763778686523
tensor([[62.1415]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[62.14154]
[61.]
epoch: 100 loss: 1.3031147718429565
tensor([[132.5059]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[132.50592]
[133.]
epoch: 101 loss: 0.24411463737487793
tensor([[72.1736]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[72.173615]
[78.]
epoch: 102 loss: 33.94676971435547
tensor([[52.6680]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[52.66804]
[60.]
epoch: 103 loss: 53.75761795043945
tensor([[13.8527]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[13.852664]
[25.]
epoch: 104 loss: 124.26309967041016
tensor([[47.9789]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[47.97887]
[55.]
epoch: 105 loss: 49.296260833740234
tensor([[172.1889]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[172.18892]
[185.]
epoch: 106 loss: 164.12379455566406
tensor([[12.3686]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[12.368645]
[13.]
epoch: 107 loss: 0.3986094892024994
tensor([[161.8036]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')

```

```

[161.80363]
[153.]
epoch: 108 loss: 77.50398254394531
tensor([[15.0667]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[15.066659]
[8.]
epoch: 109 loss: 49.93766784667969
tensor([[68.7718]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[68.77181]
[67.]
epoch: 110 loss: 3.13931941986084
tensor([[85.4590]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[85.45903]
[83.]
epoch: 111 loss: 6.0468292236328125
tensor([[202.0922]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[202.09224]
[194.]
epoch: 112 loss: 65.48433685302734
tensor([[97.2611]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[97.261086]
[97.]
epoch: 113 loss: 0.06816564500331879
tensor([[133.8165]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[133.81651]
[139.]
epoch: 114 loss: 26.8685359954834
tensor([[105.3092]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[105.30918]
[111.]
epoch: 115 loss: 32.38541793823242
tensor([[12.9437]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[12.943661]
[10.]
epoch: 116 loss: 8.665138244628906
tensor([[51.4560]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[51.45596]
[53.]
epoch: 117 loss: 2.384061574935913

```

```

tensor([[193.2926]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[193.2926]
[203.]
epoch: 118 loss: 94.23356628417969
tensor([[151.3290]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[151.329]
[151.]
epoch: 119 loss: 0.10823754966259003
tensor([[163.0137]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[163.01367]
[156.]
epoch: 120 loss: 49.191593170166016
tensor([[82.4855]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[82.48551]
[80.]
epoch: 121 loss: 6.177768707275391
tensor([[33.7413]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[33.741253]
[35.]
epoch: 122 loss: 1.584444284439087
tensor([[86.3659]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[86.365906]
[84.]
epoch: 123 loss: 5.597509860992432
tensor([[201.4647]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[201.46466]
[193.]
epoch: 124 loss: 71.65048217773438
tensor([[151.0232]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[151.02316]
[150.]
epoch: 125 loss: 1.046862244606018
tensor([[168.1294]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[168.12944]
[172.]
epoch: 126 loss: 14.981232643127441
tensor([[167.8547]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[167.85474]

```

```

[175.]
epoch: 127 loss: 51.05479431152344
tensor([[10.8878]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[10.887843]
[3.]
epoch: 128 loss: 62.21807098388672
tensor([[38.7905]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[38.790485]
[44.]
epoch: 129 loss: 27.139041900634766
tensor([[198.9109]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[198.91092]
[205.]
epoch: 130 loss: 37.076904296875
tensor([[130.7754]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[130.7754]
[132.]
epoch: 131 loss: 1.4996308088302612
tensor([[40.1544]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[40.1544]
[43.]
epoch: 132 loss: 8.097439765930176
tensor([[81.7131]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[81.713135]
[79.]
epoch: 133 loss: 7.361100196838379
tensor([[13.2394]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[13.239361]
[15.]
epoch: 134 loss: 3.0998504161834717
tensor([[196.3357]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[196.33574]
[176.]
epoch: 135 loss: 413.54229736328125
tensor([[187.7052]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[187.70515]
[178.]
epoch: 136 loss: 94.19002532958984
tensor([[100.6222]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],

```

```

device='cuda:0')
[100.622154]
[105.]
epoch: 137 loss: 19.1655330657959
tensor([[183.0233]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[183.02327]
[200.]
epoch: 138 loss: 288.2093811035156
tensor([[103.2935]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[103.29353]
[115.]
epoch: 139 loss: 137.04136657714844
tensor([[10.9292]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[10.929181]
[20.]
epoch: 140 loss: 82.27975463867188
tensor([[197.5650]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[197.56505]
[198.]
epoch: 141 loss: 0.1891830563545227
tensor([[177.7623]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[177.76233]
[169.]
epoch: 142 loss: 76.77841186523438
tensor([[202.6194]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[202.61937]
[196.]
epoch: 143 loss: 43.8160514831543
tensor([[104.9984]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[104.998405]
[103.]
epoch: 144 loss: 3.993624448776245
tensor([[56.5979]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[56.597904]
[57.]
epoch: 145 loss: 0.16168102622032166
tensor([[100.0230]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[100.023]
[107.]

```

```

epoch: 146 loss: 48.67849349975586
tensor([[172.4343]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[172.43431]
[191.]
epoch: 147 loss: 344.684814453125
tensor([[173.7326]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[173.73264]
[183.]
epoch: 148 loss: 85.88404846191406
tensor([[20.7129]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[20.712889]
[16.]
epoch: 149 loss: 22.211320877075195
tensor([[85.4895]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[85.489525]
[74.]
epoch: 150 loss: 132.00918579101562
tensor([[198.5858]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[198.58583]
[173.]
epoch: 151 loss: 654.6347045898438
tensor([[173.8392]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[173.8392]
[159.]
epoch: 152 loss: 220.20193481445312
tensor([[101.5746]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[101.5746]
[101.]
epoch: 153 loss: 0.33016541600227356
tensor([[100.4043]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[100.40432]
[113.]
epoch: 154 loss: 158.65115356445312
tensor([[72.7789]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[72.77892]
[88.]
epoch: 155 loss: 231.68118286132812
tensor([[20.0772]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')

```

```

[20.077171]
[29.]
epoch: 156 loss: 79.61687469482422
tensor([[142.9183]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[142.91826]
[179.]
epoch: 157 loss: 1301.89208984375
tensor([[107.6816]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[107.681625]
[123.]
epoch: 158 loss: 234.65260314941406
tensor([[130.2892]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[130.28923]
[135.]
epoch: 159 loss: 22.19135093688965
tensor([[20.9463]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[20.946339]
[17.]
epoch: 160 loss: 15.573588371276855
tensor([[128.0092]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[128.00925]
[104.]
epoch: 161 loss: 576.4439086914062
tensor([[36.7128]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[36.71284]
[28.]
epoch: 162 loss: 75.9135971069336
tensor([[25.3675]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[25.367521]
[19.]
epoch: 163 loss: 40.545326232910156
tensor([[111.2467]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[111.246666]
[89.]
epoch: 164 loss: 494.9141540527344
tensor([[133.5473]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[133.54727]
[112.]
epoch: 165 loss: 464.284912109375

```

```

tensor([[107.7331]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[107.73314]
[99.]
epoch: 166 loss: 76.26771545410156
tensor([[50.7430]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[50.74301]
[56.]
epoch: 167 loss: 27.635929107666016
tensor([[10.3241]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[10.324148]
[23.]
epoch: 168 loss: 160.67721557617188
tensor([[168.6396]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[168.6396]
[186.]
epoch: 169 loss: 301.3833923339844
tensor([[143.8384]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[143.83842]
[165.]
epoch: 170 loss: 447.812255859375
tensor([[119.9042]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[119.90421]
[136.]
epoch: 171 loss: 259.0743713378906
tensor([[128.0644]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[128.06438]
[130.]
epoch: 172 loss: 3.7466371059417725
tensor([[54.0706]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[54.070564]
[50.]
epoch: 173 loss: 16.569494247436523
tensor([[145.6362]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[145.63625]
[119.]
epoch: 174 loss: 709.4895629882812
tensor([[20.4996]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[20.499594]

```



```

[5.]
epoch: 175 loss: 240.23741149902344
tensor([[211.9516]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[211.95155]
[180.]
epoch: 176 loss: 1020.9017333984375
tensor([[13.9480]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[13.9479685]
[33.]
epoch: 177 loss: 362.9799499511719
tensor([[38.2822]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[38.282158]
[64.]
epoch: 178 loss: 661.4074096679688
tensor([[1.9429]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[1.9429268]
[9.]
epoch: 179 loss: 49.80228042602539
tensor([[148.3666]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[148.36662]
[184.]
epoch: 180 loss: 1269.737548828125
tensor([[6.0255]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[6.025473]
[22.]
epoch: 181 loss: 255.18553161621094
tensor([[28.9045]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[28.904457]
[46.]
epoch: 182 loss: 292.2575988769531
tensor([[155.3041]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[155.30412]
[157.]
epoch: 183 loss: 2.8759989738464355
tensor([[111.6569]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[111.65687]
[106.]
epoch: 184 loss: 32.00015640258789
tensor([[32.2794]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],

```

```

device='cuda:0')
[32.279392]
[27.]
epoch: 185 loss: 27.87198257446289
tensor([[93.7697]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[93.76967]
[82.]
epoch: 186 loss: 138.5251007080078
For epoch 5 the test accuracy over the whole test set is 0 %
tensor([[138.4351]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[138.43509]
[123.]
epoch: 0 loss: 238.24197387695312
tensor([[104.1109]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[104.11091]
[108.]
epoch: 1 loss: 15.125032424926758
tensor([[14.0826]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[14.08261]
[13.]
epoch: 2 loss: 1.1720446348190308
tensor([[8.4715]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[8.471547]
[0.]
epoch: 3 loss: 71.7671127319336
tensor([[147.3786]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[147.37862]
[185.]
epoch: 4 loss: 1415.3685302734375
tensor([[42.8026]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[42.80264]
[57.]
epoch: 5 loss: 201.5650634765625
tensor([[230.1705]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[230.1705]
[205.]
epoch: 6 loss: 633.5541381835938
tensor([[12.8519]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[12.851942]

```

```

[5.]
epoch: 7 loss: 61.652992248535156
tensor([[104.2210]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[104.220985]
[109.]
epoch: 8 loss: 22.838979721069336
tensor([[174.6724]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[174.67242]
[166.]
epoch: 9 loss: 75.21094512939453
tensor([[56.8069]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[56.806942]
[67.]
epoch: 10 loss: 103.89842987060547
tensor([[183.7822]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[183.78218]
[183.]
epoch: 11 loss: 0.6118068099021912
tensor([[143.7706]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[143.77058]
[150.]
epoch: 12 loss: 38.80562210083008
tensor([[111.1917]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[111.1917]
[117.]
epoch: 13 loss: 33.736305236816406
tensor([[188.9422]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[188.94225]
[180.]
epoch: 14 loss: 79.96375274658203
tensor([[83.2752]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[83.2752]
[87.]
epoch: 15 loss: 13.874135971069336
tensor([[177.7440]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[177.74396]
[178.]
epoch: 16 loss: 0.06555774807929993
tensor([[17.4313]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],

```

```

device='cuda:0')
[17.43129]
[22.]
epoch: 17 loss: 20.87311363220215
tensor([[181.1708]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[181.17082]
[177.]
epoch: 18 loss: 17.3957576751709
tensor([[126.6706]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[126.67056]
[124.]
epoch: 19 loss: 7.131905555725098
tensor([[99.5483]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[99.54835]
[101.]
epoch: 20 loss: 2.107295036315918
tensor([[34.3503]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[34.350304]
[37.]
epoch: 21 loss: 7.020890712738037
tensor([[138.8193]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[138.81926]
[152.]
epoch: 22 loss: 173.73191833496094
tensor([[110.9549]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[110.95494]
[110.]
epoch: 23 loss: 0.9119119048118591
tensor([[147.1270]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[147.12701]
[149.]
epoch: 24 loss: 3.5080759525299072
tensor([[173.5744]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[173.57445]
[174.]
epoch: 25 loss: 0.18109482526779175
tensor([[108.5880]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[108.58796]
[92.]

```

```

epoch: 26 loss: 275.160400390625
tensor([[181.2419]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[181.24187]
[179.]
epoch: 27 loss: 5.025968074798584
tensor([[110.3197]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[110.319664]
[102.]
epoch: 28 loss: 69.21681213378906
tensor([[81.8824]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[81.88237]
[82.]
epoch: 29 loss: 0.013836817815899849
tensor([[16.7816]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[16.78164]
[10.]
epoch: 30 loss: 45.99062728881836
tensor([[21.1149]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[21.114868]
[27.]
epoch: 31 loss: 34.6347770690918
tensor([[90.5776]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[90.577614]
[105.]
epoch: 32 loss: 208.00521850585938
tensor([[17.6465]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[17.646484]
[24.]
epoch: 33 loss: 40.36716079711914
tensor([[128.8723]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[128.87227]
[144.]
epoch: 34 loss: 228.84825134277344
tensor([[79.5888]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[79.58882]
[80.]
epoch: 35 loss: 0.16906782984733582
tensor([[139.3704]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')

```

```

[139.37044]
[134.]
epoch: 36 loss: 28.84160041809082
tensor([[171.4714]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[171.4714]
[167.]
epoch: 37 loss: 19.99346351623535
tensor([[26.4267]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[26.426655]
[19.]
epoch: 38 loss: 55.15520095825195
tensor([[146.6633]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[146.66328]
[137.]
epoch: 39 loss: 93.37906646728516
tensor([[21.6223]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[21.62232]
[3.]
epoch: 40 loss: 346.7908020019531
tensor([[152.8336]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[152.83363]
[160.]
epoch: 41 loss: 51.3568115234375
tensor([[128.8461]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[128.84608]
[133.]
epoch: 42 loss: 17.25501251220703
tensor([[123.1808]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[123.180824]
[126.]
epoch: 43 loss: 7.947751522064209
tensor([[32.3294]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[32.32938]
[36.]
epoch: 44 loss: 13.473450660705566
tensor([[160.5312]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[160.5312]
[164.]
epoch: 45 loss: 12.032544136047363

```

```

tensor([[166.7837]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[166.78366]
[163.]
epoch: 46 loss: 14.316089630126953
tensor([[173.3276]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[173.32758]
[172.]
epoch: 47 loss: 1.762457251548767
tensor([[152.0628]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[152.06282]
[148.]
epoch: 48 loss: 16.50650978088379
tensor([[167.7629]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[167.76286]
[171.]
epoch: 49 loss: 10.47905445098877
tensor([[82.9665]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[82.96649]
[85.]
epoch: 50 loss: 4.135156154632568
tensor([[96.4887]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[96.48867]
[96.]
epoch: 51 loss: 0.23879870772361755
tensor([[46.8401]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[46.84015]
[48.]
epoch: 52 loss: 1.3452545404434204
tensor([[191.6262]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[191.62619]
[198.]
epoch: 53 loss: 40.625450134277344
tensor([[101.6376]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[101.637634]
[98.]
epoch: 54 loss: 13.232382774353027
tensor([[13.8990]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[13.89899]

```

```

[6.]
epoch: 55 loss: 62.394039154052734
tensor([[98.0394]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[98.0394]
[94.]
epoch: 56 loss: 16.31673812866211
tensor([[9.9632]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[9.963238]
[9.]
epoch: 57 loss: 0.927827000617981
tensor([[177.9828]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[177.9828]
[170.]
epoch: 58 loss: 63.72515106201172
tensor([[131.6634]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[131.66344]
[130.]
epoch: 59 loss: 2.7670223712921143
tensor([[138.8772]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[138.87718]
[143.]
epoch: 60 loss: 16.99762725830078
tensor([[192.2591]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[192.2591]
[204.]
epoch: 61 loss: 137.84886169433594
tensor([[10.4557]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[10.455654]
[28.]
epoch: 62 loss: 307.8040771484375
tensor([[3.8262]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[3.8262224]
[15.]
epoch: 63 loss: 124.85330200195312
tensor([[146.8883]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[146.88828]
[136.]
epoch: 64 loss: 118.55453491210938
tensor([[63.4181]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],

```



```

device='cuda:0')
[63.41811]
[64.]
epoch: 65 loss: 0.3385961055755615
tensor([[10.6907]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[10.690747]
[14.]
epoch: 66 loss: 10.951153755187988
tensor([[156.3444]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[156.34444]
[142.]
epoch: 67 loss: 205.7628631591797
tensor([[209.0575]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[209.05746]
[191.]
epoch: 68 loss: 326.072021484375
tensor([[29.1087]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[29.108734]
[38.]
epoch: 69 loss: 79.05461120605469
tensor([[8.5506]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[8.550562]
[2.]
epoch: 70 loss: 42.90986251831055
tensor([[143.0565]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[143.05652]
[169.]
epoch: 71 loss: 673.064208984375
tensor([[47.6997]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[47.69974]
[61.]
epoch: 72 loss: 176.89688110351562
tensor([[174.8339]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[174.83392]
[201.]
epoch: 73 loss: 684.66357421875
tensor([[97.2697]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[97.2697]
[107.]

```

```

epoch: 74 loss: 94.67875671386719
tensor([[110.7833]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[110.783325]
[111.]
epoch: 75 loss: 0.04694797098636627
tensor([[18.2832]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[18.283167]
[1.]
epoch: 76 loss: 298.7078552246094
tensor([[233.9940]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[233.994]
[190.]
epoch: 77 loss: 1935.4722900390625
tensor([[204.6111]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[204.61111]
[182.]
epoch: 78 loss: 511.26251220703125
tensor([[129.5148]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[129.51477]
[138.]
epoch: 79 loss: 71.9991226196289
tensor([[60.9819]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[60.981865]
[76.]
epoch: 80 loss: 225.5443878173828
tensor([[78.9779]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[78.977875]
[103.]
epoch: 81 loss: 577.0625
tensor([[138.9024]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[138.90242]
[184.]
epoch: 82 loss: 2033.791748046875
tensor([[77.4569]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[77.45695]
[97.]
epoch: 83 loss: 381.930908203125
tensor([[21.0185]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')

```

```

[21.018528]
[26.]
epoch: 84 loss: 24.8150634765625
tensor([[109.4083]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[109.40829]
[115.]
epoch: 85 loss: 31.267253875732422
tensor([[202.1895]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[202.18947]
[203.]
epoch: 86 loss: 0.6569615006446838
tensor([[60.1736]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[60.17365]
[47.]
epoch: 87 loss: 173.54502868652344
tensor([[140.0039]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[140.00389]
[118.]
epoch: 88 loss: 484.17120361328125
tensor([[184.4988]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[184.49884]
[159.]
epoch: 89 loss: 650.1908569335938
tensor([[105.0142]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[105.01421]
[95.]
epoch: 90 loss: 100.28446960449219
tensor([[47.5029]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[47.502884]
[46.]
epoch: 91 loss: 2.258660078048706
tensor([[54.5248]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[54.524807]
[55.]
epoch: 92 loss: 0.22580841183662415
tensor([[56.3015]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[56.301495]
[62.]
epoch: 93 loss: 32.472965240478516

```

```

tensor([[14.3095]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[14.309515]
[21.]
epoch: 94 loss: 44.76258850097656
tensor([[87.1660]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[87.16599]
[112.]
epoch: 95 loss: 616.7279052734375
tensor([[123.5540]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[123.55398]
[157.]
epoch: 96 loss: 1118.6363525390625
tensor([[58.3594]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[58.359436]
[60.]
epoch: 97 loss: 2.6914501190185547
tensor([[43.9085]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[43.908478]
[43.]
epoch: 98 loss: 0.8253318667411804
tensor([[189.4310]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[189.43098]
[196.]
epoch: 99 loss: 43.1519660949707
tensor([[96.6535]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[96.65348]
[84.]
epoch: 100 loss: 160.11056518554688
tensor([[104.7448]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[104.74477]
[89.]
epoch: 101 loss: 247.8976593017578
tensor([[107.3244]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[107.324425]
[91.]
epoch: 102 loss: 266.4868469238281
tensor([[28.0770]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[28.077042]

```

```

[12.]
epoch: 103 loss: 258.47125244140625
tensor([[211.6207]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[211.62074]
[199.]
epoch: 104 loss: 159.28314208984375
tensor([[60.9110]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[60.911045]
[58.]
epoch: 105 loss: 8.474183082580566
tensor([[51.0266]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[51.026646]
[54.]
epoch: 106 loss: 8.840835571289062
tensor([[71.9350]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[71.93499]
[88.]
epoch: 107 loss: 258.08453369140625
tensor([[162.0650]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[162.06503]
[192.]
epoch: 108 loss: 896.1022338867188
tensor([[128.8587]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[128.8587]
[147.]
epoch: 109 loss: 329.10662841796875
tensor([[67.2879]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[67.287865]
[71.]
epoch: 110 loss: 13.779948234558105
tensor([[123.1844]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[123.184425]
[119.]
epoch: 111 loss: 17.509414672851562
tensor([[215.8076]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[215.80763]
[186.]
epoch: 112 loss: 888.4949340820312
tensor([[90.3484]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],

```

```

device='cuda:0')
[90.34839]
[86.]
epoch: 113 loss: 18.908483505249023
tensor([[83.6277]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[83.62768]
[83.]
epoch: 114 loss: 0.39397957921028137
tensor([[75.5171]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[75.51712]
[78.]
epoch: 115 loss: 6.16469144821167
tensor([[12.9910]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[12.991005]
[7.]
epoch: 116 loss: 35.89213943481445
tensor([[112.5841]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[112.584076]
[122.]
epoch: 117 loss: 88.65962982177734
tensor([[61.1097]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[61.109653]
[68.]
epoch: 118 loss: 47.47687530517578
tensor([[46.2733]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[46.27328]
[52.]
epoch: 119 loss: 32.795310974121094
tensor([[70.1410]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[70.141]
[77.]
epoch: 120 loss: 47.0458984375
tensor([[145.4595]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[145.45953]
[151.]
epoch: 121 loss: 30.696767807006836
tensor([[191.8140]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[191.81404]
[193.]

```

```

epoch: 122 loss: 1.4064984321594238
tensor([[139.7238]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[139.72383]
[141.]
epoch: 123 loss: 1.628606915473938
tensor([[52.3981]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[52.3981]
[50.]
epoch: 124 loss: 5.750892162322998
tensor([[76.7639]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[76.763916]
[74.]
epoch: 125 loss: 7.6392316818237305
tensor([[182.3415]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[182.34149]
[173.]
epoch: 126 loss: 87.26346588134766
tensor([[14.9655]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[14.965503]
[4.]
epoch: 127 loss: 120.24224853515625
tensor([[128.4845]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[128.48451]
[125.]
epoch: 128 loss: 12.141826629638672
tensor([[168.4827]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[168.48273]
[165.]
epoch: 129 loss: 12.129387855529785
tensor([[87.0792]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[87.0792]
[90.]
epoch: 130 loss: 8.531067848205566
tensor([[54.2146]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[54.214592]
[53.]
epoch: 131 loss: 1.4752336740493774
tensor([[148.5407]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')

```

```

[148.54073]
[153.]
epoch: 132 loss: 19.885128021240234
tensor([[178.6875]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[178.68752]
[189.]
epoch: 133 loss: 106.34734344482422
tensor([[28.2213]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[28.221252]
[34.]
epoch: 134 loss: 33.393924713134766
tensor([[18.4902]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[18.490236]
[25.]
epoch: 135 loss: 42.37702178955078
tensor([[14.2914]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[14.291394]
[18.]
epoch: 136 loss: 13.753756523132324
tensor([[153.7921]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[153.79205]
[145.]
epoch: 137 loss: 77.30020141601562
tensor([[205.1454]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[205.14543]
[197.]
epoch: 138 loss: 66.34805297851562
tensor([[69.6300]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[69.630035]
[65.]
epoch: 139 loss: 21.437227249145508
tensor([[128.5728]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[128.5728]
[132.]
epoch: 140 loss: 11.745701789855957
tensor([[133.1995]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[133.1995]
[139.]
epoch: 141 loss: 33.645877838134766

```



```

tensor([[77.0435]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[77.04347]
[79.]
epoch: 142 loss: 3.828000783920288
tensor([[43.2499]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[43.24985]
[45.]
epoch: 143 loss: 3.063020706176758
tensor([[69.8845]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[69.884476]
[69.]
epoch: 144 loss: 0.7822972536087036
tensor([[104.4135]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[104.41354]
[106.]
epoch: 145 loss: 2.516843557357788
tensor([[134.8920]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[134.89203]
[140.]
epoch: 146 loss: 26.09136962890625
tensor([[154.5888]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[154.58876]
[156.]
epoch: 147 loss: 1.991597294807434
tensor([[189.0500]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[189.05002]
[194.]
epoch: 148 loss: 24.5023193359375
tensor([[23.1223]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[23.122322]
[23.]
epoch: 149 loss: 0.014962691813707352
tensor([[33.0362]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[33.036243]
[29.]
epoch: 150 loss: 16.291261672973633
tensor([[21.4922]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[21.492214]

```

```

[16.]
epoch: 151 loss: 30.164417266845703
tensor([[34.4438]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[34.443768]
[30.]
epoch: 152 loss: 19.7470703125
tensor([[18.2587]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[18.258722]
[8.]
epoch: 153 loss: 105.24138641357422
tensor([[170.6310]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[170.63103]
[155.]
epoch: 154 loss: 244.32901000976562
tensor([[118.7313]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[118.73127]
[116.]
epoch: 155 loss: 7.459835052490234
tensor([[26.4580]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[26.458021]
[32.]
epoch: 156 loss: 30.713529586791992
tensor([[8.8239]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[8.823943]
[11.]
epoch: 157 loss: 4.735223293304443
tensor([[33.8629]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[33.862885]
[44.]
epoch: 158 loss: 102.7611083984375
tensor([[78.1806]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[78.18062]
[93.]
epoch: 159 loss: 219.61407470703125
tensor([[29.9031]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[29.903086]
[39.]
epoch: 160 loss: 82.75385284423828
tensor([[174.1745]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],

```

```

device='cuda:0')
[174.17448]
[200.]
epoch: 161 loss: 666.957275390625
tensor([[70.5265]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[70.52652]
[70.]
epoch: 162 loss: 0.2772230803966522
tensor([[152.1328]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[152.13277]
[146.]
epoch: 163 loss: 37.6108283996582
tensor([[90.3123]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[90.31227]
[81.]
epoch: 164 loss: 86.71839141845703
tensor([[128.6037]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[128.60373]
[114.]
epoch: 165 loss: 213.26890563964844
tensor([[58.0492]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[58.049164]
[49.]
epoch: 166 loss: 81.88736724853516
tensor([[188.0990]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[188.099]
[168.]
epoch: 167 loss: 403.9697570800781
tensor([[103.5937]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[103.59367]
[100.]
epoch: 168 loss: 12.914490699768066
tensor([[114.8986]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[114.89855]
[120.]
epoch: 169 loss: 26.02477264404297
tensor([[146.4420]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[146.44199]
[154.]

```

```

epoch: 170 loss: 57.123573303222656
tensor([[187.0060]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[187.00603]
[195.]
epoch: 171 loss: 63.903602600097656
tensor([[161.6204]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[161.62044]
[175.]
epoch: 172 loss: 179.0126953125
tensor([[51.0992]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[51.099155]
[56.]
epoch: 173 loss: 24.018278121948242
tensor([[93.8326]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[93.83258]
[99.]
epoch: 174 loss: 26.70222282409668
tensor([[13.7498]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[13.749828]
[17.]
epoch: 175 loss: 10.563615798950195
tensor([[75.9889]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[75.988945]
[75.]
epoch: 176 loss: 0.978012204170227
tensor([[120.4728]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[120.47282]
[113.]
epoch: 177 loss: 55.842987060546875
tensor([[17.4207]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[17.420746]
[20.]
epoch: 178 loss: 6.652552127838135
tensor([[113.3877]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[113.387665]
[104.]
epoch: 179 loss: 88.12825012207031
tensor([[137.7922]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')

```

```

[137.79216]
[129.]
epoch: 180 loss: 77.30207824707031
tensor([[191.0155]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[191.01553]
[176.]
epoch: 181 loss: 225.46624755859375
tensor([[130.8027]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[130.8027]
[135.]
epoch: 182 loss: 17.617294311523438
tensor([[24.6159]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[24.61587]
[33.]
epoch: 183 loss: 70.29364013671875
tensor([[25.0423]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[25.042316]
[35.]
epoch: 184 loss: 99.15546417236328
tensor([[147.8976]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[147.89764]
[162.]
epoch: 185 loss: 198.87644958496094
tensor([[42.3498]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[42.349785]
[51.]
epoch: 186 loss: 74.82622528076172
For epoch 6 the test accuracy over the whole test set is 0 %
tensor([[124.5350]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[124.53505]
[138.]
epoch: 0 loss: 181.30490112304688
tensor([[160.7084]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[160.70837]
[163.]
epoch: 1 loss: 5.25154972076416
tensor([[77.9361]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[77.93608]
[79.]

```

```

epoch: 2 loss: 1.131923794746399
tensor([[111.1238]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[111.1238]
[103.]
epoch: 3 loss: 65.99616241455078
tensor([[38.8952]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[38.895206]
[35.]
epoch: 4 loss: 15.172633171081543
tensor([[66.0077]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[66.007744]
[57.]
epoch: 5 loss: 81.13945007324219
tensor([[176.8778]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[176.87778]
[156.]
epoch: 6 loss: 435.881591796875
tensor([[23.3906]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[23.390556]
[23.]
epoch: 7 loss: 0.15253424644470215
tensor([[140.1802]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[140.18019]
[143.]
epoch: 8 loss: 7.951322555541992
tensor([[81.0407]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[81.0407]
[90.]
epoch: 9 loss: 80.26900482177734
tensor([[80.4860]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[80.486015]
[91.]
epoch: 10 loss: 110.54387664794922
tensor([[108.6584]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[108.65839]
[122.]
epoch: 11 loss: 177.9986572265625
tensor([[17.5360]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')

```

```

[17.535994]
[16.]
epoch: 12 loss: 2.359276294708252
tensor([[147.1868]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[147.18675]
[153.]
epoch: 13 loss: 33.79384994506836
tensor([[177.4049]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[177.40492]
[178.]
epoch: 14 loss: 0.3541172444820404
tensor([[70.6362]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[70.63619]
[70.]
epoch: 15 loss: 0.40474066138267517
tensor([[144.2816]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[144.28156]
[142.]
epoch: 16 loss: 5.205493927001953
tensor([[66.4052]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[66.40521]
[61.]
epoch: 17 loss: 29.21632194519043
tensor([[211.7361]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[211.73615]
[193.]
epoch: 18 loss: 351.0431213378906
tensor([[116.9437]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[116.94372]
[120.]
epoch: 19 loss: 9.340860366821289
tensor([[144.4268]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[144.4268]
[150.]
epoch: 20 loss: 31.060518264770508
tensor([[103.4109]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[103.410866]
[112.]
epoch: 21 loss: 73.77322387695312

```

```

tensor([[66.4010]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[66.40096]
[68.]
epoch: 22 loss: 2.556919813156128
tensor([[20.0132]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[20.013172]
[19.]
epoch: 23 loss: 1.0265177488327026
tensor([[122.9504]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[122.9504]
[133.]
epoch: 24 loss: 100.99443054199219
tensor([[197.2273]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[197.22731]
[200.]
epoch: 25 loss: 7.687808990478516
tensor([[101.4924]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[101.49245]
[102.]
epoch: 26 loss: 0.25761014223098755
tensor([[23.3943]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[23.394331]
[20.]
epoch: 27 loss: 11.521482467651367
tensor([[160.3332]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[160.3332]
[155.]
epoch: 28 loss: 28.44308853149414
tensor([[155.9648]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[155.9648]
[152.]
epoch: 29 loss: 15.719622611999512
tensor([[113.6851]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[113.68513]
[113.]
epoch: 30 loss: 0.4693993628025055
tensor([[69.5763]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[69.57634]

```



```

[64.]
epoch: 31 loss: 31.095563888549805
tensor([[205.2271]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[205.22708]
[205.]
epoch: 32 loss: 0.05156591534614563
tensor([[76.4218]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[76.42184]
[76.]
epoch: 33 loss: 0.1779463291168213
tensor([[20.7708]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[20.770763]
[17.]
epoch: 34 loss: 14.218656539916992
tensor([[136.3279]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[136.3279]
[146.]
epoch: 35 loss: 93.54959106445312
tensor([[175.3978]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[175.39781]
[183.]
epoch: 36 loss: 57.79326248168945
tensor([[100.6933]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[100.69328]
[101.]
epoch: 37 loss: 0.0940752699971199
tensor([[21.4332]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[21.433208]
[18.]
epoch: 38 loss: 11.786920547485352
tensor([[84.6194]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[84.61939]
[81.]
epoch: 39 loss: 13.100001335144043
tensor([[71.6923]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[71.69227]
[65.]
epoch: 40 loss: 44.78645706176758
tensor([[40.6390]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],

```

```

device='cuda:0')
[40.639008]
[39.]
epoch: 41 loss: 2.6863458156585693
tensor([[43.9883]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[43.988274]
[43.]
epoch: 42 loss: 0.9766847491264343
tensor([[143.2233]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[143.22331]
[137.]
epoch: 43 loss: 38.729618072509766
tensor([[145.4708]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[145.47081]
[141.]
epoch: 44 loss: 19.988142013549805
tensor([[167.8869]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[167.88689]
[164.]
epoch: 45 loss: 15.107887268066406
tensor([[164.7364]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[164.73636]
[165.]
epoch: 46 loss: 0.06950676441192627
tensor([[111.7667]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[111.76666]
[118.]
epoch: 47 loss: 38.854496002197266
tensor([[72.7944]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[72.794365]
[78.]
epoch: 48 loss: 27.098636627197266
tensor([[82.0944]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[82.09437]
[89.]
epoch: 49 loss: 47.68775177001953
tensor([[13.6563]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[13.656262]
[1.]

```

```

epoch: 50 loss: 160.1809844970703
tensor([[191.3949]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[191.39488]
[192.]
epoch: 51 loss: 0.3661675453186035
tensor([[28.5348]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[28.53485]
[33.]
epoch: 52 loss: 19.937572479248047
tensor([[70.5237]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[70.52368]
[75.]
epoch: 53 loss: 20.037425994873047
tensor([[174.2903]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[174.29031]
[177.]
epoch: 54 loss: 7.342399597167969
tensor([[13.8407]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[13.840738]
[6.]
epoch: 55 loss: 61.477176666259766
tensor([[55.9913]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[55.991287]
[56.]
epoch: 56 loss: 7.591233588755131e-05
tensor([[130.2067]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[130.20671]
[130.]
epoch: 57 loss: 0.04272936284542084
tensor([[138.1484]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[138.14844]
[136.]
epoch: 58 loss: 4.61578369140625
tensor([[107.9493]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[107.94928]
[106.]
epoch: 59 loss: 3.799691677093506
tensor([[109.6330]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')

```

```

[109.63302]
[108.]
epoch: 60 loss: 2.6667494773864746
tensor([[54.5360]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[54.53602]
[54.]
epoch: 61 loss: 0.28731569647789
tensor([[70.3445]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[70.34454]
[71.]
epoch: 62 loss: 0.4296232759952545
tensor([[27.6235]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[27.623486]
[32.]
epoch: 63 loss: 19.153879165649414
tensor([[172.7314]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[172.73137]
[169.]
epoch: 64 loss: 13.923114776611328
tensor([[123.7640]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[123.76399]
[124.]
epoch: 65 loss: 0.055699631571769714
tensor([[11.9672]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[11.96717]
[8.]
epoch: 66 loss: 15.738435745239258
tensor([[188.0463]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[188.04633]
[184.]
epoch: 67 loss: 16.372751235961914
tensor([[135.6926]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[135.69257]
[140.]
epoch: 68 loss: 18.553987503051758
tensor([[189.3710]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[189.37103]
[190.]
epoch: 69 loss: 0.3955998420715332

```

```

tensor([[94.7335]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[94.733505]
[99.]
epoch: 70 loss: 18.202978134155273
tensor([[139.4051]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[139.40514]
[145.]
epoch: 71 loss: 31.302501678466797
tensor([[156.2466]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[156.24664]
[159.]
epoch: 72 loss: 7.580974578857422
tensor([[201.9408]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[201.94075]
[196.]
epoch: 73 loss: 35.292510986328125
tensor([[68.2388]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[68.238785]
[69.]
epoch: 74 loss: 0.5794485807418823
tensor([[107.9161]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[107.91607]
[109.]
epoch: 75 loss: 1.1749063730239868
tensor([[183.7643]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[183.76434]
[182.]
epoch: 76 loss: 3.1129071712493896
tensor([[111.9926]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[111.99259]
[114.]
epoch: 77 loss: 4.029687404632568
tensor([[73.1319]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[73.131935]
[74.]
epoch: 78 loss: 0.7535366415977478
tensor([[82.9420]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[82.94199]

```

```

[84.]
epoch: 79 loss: 1.1193772554397583
tensor([[33.2404]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[33.240437]
[37.]
epoch: 80 loss: 14.134317398071289
tensor([[133.0902]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[133.09023]
[132.]
epoch: 81 loss: 1.1885910034179688
tensor([[45.1430]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[45.14296]
[46.]
epoch: 82 loss: 0.7345182299613953
tensor([[145.7254]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[145.72537]
[144.]
epoch: 83 loss: 2.976909637451172
tensor([[87.5071]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[87.50711]
[86.]
epoch: 84 loss: 2.2713823318481445
tensor([[14.4706]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[14.4706]
[12.]
epoch: 85 loss: 6.103865146636963
tensor([[87.7199]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[87.71989]
[87.]
epoch: 86 loss: 0.5182369947433472
tensor([[42.6435]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[42.643536]
[44.]
epoch: 87 loss: 1.8399956226348877
tensor([[186.6800]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[186.68]
[186.]
epoch: 88 loss: 0.4623900353908539
tensor([[146.0784]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],

```

```

device='cuda:0')
[146.0784]
[148.]
epoch: 89 loss: 3.6925477981567383
tensor([[167.0757]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[167.07574]
[170.]
epoch: 90 loss: 8.55126953125
tensor([[106.2368]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[106.236824]
[104.]
epoch: 91 loss: 5.003381729125977
tensor([[14.7577]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[14.757669]
[14.]
epoch: 92 loss: 0.5740630030632019
tensor([[14.0129]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[14.01286]
[9.]
epoch: 93 loss: 25.128768920898438
tensor([[12.6962]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[12.696192]
[3.]
epoch: 94 loss: 94.0161361694336
tensor([[116.9393]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[116.93935]
[117.]
epoch: 95 loss: 0.003678869688883424
tensor([[18.5718]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[18.571848]
[24.]
epoch: 96 loss: 29.464834213256836
tensor([[12.5193]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[12.519292]
[11.]
epoch: 97 loss: 2.3082478046417236
tensor([[97.0491]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[97.04908]
[96.]

```

```

epoch: 98 loss: 1.1005686521530151
tensor([[203.0479]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[203.04791]
[199.]
epoch: 99 loss: 16.385597229003906
tensor([[11.1331]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[11.133084]
[10.]
epoch: 100 loss: 1.2838799953460693
tensor([[186.0494]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[186.04941]
[189.]
epoch: 101 loss: 8.70599365234375
tensor([[187.8475]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[187.84749]
[191.]
epoch: 102 loss: 9.938329696655273
tensor([[20.2621]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[20.262089]
[27.]
epoch: 103 loss: 45.39944839477539
tensor([[13.7229]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[13.722936]
[22.]
epoch: 104 loss: 68.50979614257812
tensor([[57.1766]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[57.176624]
[55.]
epoch: 105 loss: 4.7376933097839355
tensor([[186.4097]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[186.40967]
[185.]
epoch: 106 loss: 1.987163782119751
tensor([[162.3623]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[162.36232]
[162.]
epoch: 107 loss: 0.1312757432460785
tensor([[12.1785]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')

```



```

[12.178511]
[13.]
epoch: 108 loss: 0.6748447418212891
tensor([[11.4498]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[11.449822]
[7.]
epoch: 109 loss: 19.800920486450195
tensor([[128.0057]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[128.00568]
[129.]
epoch: 110 loss: 0.9886797070503235
tensor([[115.7275]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[115.72754]
[116.]
epoch: 111 loss: 0.0742349624633789
tensor([[110.9516]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[110.95162]
[110.]
epoch: 112 loss: 0.905584454536438
tensor([[10.2002]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[10.200181]
[4.]
epoch: 113 loss: 38.44224548339844
tensor([[114.5571]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[114.55709]
[115.]
epoch: 114 loss: 0.19616860151290894
tensor([[137.8362]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[137.83621]
[139.]
epoch: 115 loss: 1.3544021844863892
tensor([[121.4785]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[121.47848]
[123.]
epoch: 116 loss: 2.315030813217163
tensor([[148.8482]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[148.8482]
[149.]
epoch: 117 loss: 0.02304155007004738

```

```

tensor([[58.0700]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[58.069984]
[58.]
epoch: 118 loss: 0.004897821228951216
tensor([[208.8499]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[208.84985]
[203.]
epoch: 119 loss: 34.220787048339844
tensor([[126.2258]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[126.22578]
[126.]
epoch: 120 loss: 0.05097510665655136
tensor([[60.6199]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[60.619946]
[62.]
epoch: 121 loss: 1.904550313949585
tensor([[195.9361]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[195.93607]
[198.]
epoch: 122 loss: 4.259824752807617
tensor([[2.6100]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[2.6099515]
[2.]
epoch: 123 loss: 0.37204083800315857
tensor([[85.1685]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[85.16855]
[88.]
epoch: 124 loss: 8.017117500305176
tensor([[160.7440]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[160.74405]
[167.]
epoch: 125 loss: 39.13692092895508
tensor([[92.3399]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[92.339874]
[92.]
epoch: 126 loss: 0.11551451683044434
tensor([[205.9734]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[205.97343]

```

```

[204.]
epoch: 127 loss: 3.8944435119628906
tensor([[32.1058]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[32.105793]
[38.]
epoch: 128 loss: 34.741676330566406
tensor([[16.5014]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[16.501394]
[25.]
epoch: 129 loss: 72.22630310058594
tensor([[34.8566]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[34.85658]
[36.]
epoch: 130 loss: 1.307412028312683
tensor([[142.2924]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[142.29245]
[135.]
epoch: 131 loss: 53.17982482910156
tensor([[177.2333]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[177.23329]
[173.]
epoch: 132 loss: 17.920757293701172
tensor([[168.9892]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[168.98918]
[166.]
epoch: 133 loss: 8.935206413269043
tensor([[59.9347]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[59.934723]
[52.]
epoch: 134 loss: 62.9598274230957
tensor([[50.5531]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[50.553135]
[47.]
epoch: 135 loss: 12.624767303466797
tensor([[92.3398]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[92.339836]
[94.]
epoch: 136 loss: 2.7561440467834473
tensor([[161.2559]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],

```

```

device='cuda:0')
[161.25586]
[176.]
epoch: 137 loss: 217.38967895507812
tensor([[160.3948]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[160.39482]
[172.]
epoch: 138 loss: 134.68017578125
tensor([[56.3355]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[56.335518]
[51.]
epoch: 139 loss: 28.467750549316406
tensor([[14.4390]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[14.438996]
[5.]
epoch: 140 loss: 89.09465026855469
tensor([[99.9001]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[99.90012]
[93.]
epoch: 141 loss: 47.6117057800293
tensor([[30.6729]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[30.672892]
[30.]
epoch: 142 loss: 0.45278313755989075
tensor([[184.4762]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[184.47617]
[174.]
epoch: 143 loss: 109.75004577636719
tensor([[182.8645]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[182.86449]
[175.]
epoch: 144 loss: 61.85015106201172
tensor([[206.2151]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[206.2151]
[201.]
epoch: 145 loss: 27.19729995727539
tensor([[152.7352]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[152.73523]
[157.]

```

```

epoch: 146 loss: 18.18826675415039
tensor([[10.8968]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[10.896843]
[15.]
epoch: 147 loss: 16.83589744567871
tensor([[43.1815]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[43.181484]
[48.]
epoch: 148 loss: 23.218093872070312
tensor([[20.7132]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[20.713184]
[28.]
epoch: 149 loss: 53.09768295288086
tensor([[156.1527]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[156.15266]
[179.]
epoch: 150 loss: 522.000732421875
tensor([[100.9619]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[100.96193]
[111.]
epoch: 151 loss: 100.76286315917969
tensor([[60.7803]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[60.78025]
[60.]
epoch: 152 loss: 0.608790934085846
tensor([[101.9125]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[101.912544]
[100.]
epoch: 153 loss: 3.657825469970703
tensor([[204.9530]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[204.95296]
[194.]
epoch: 154 loss: 119.96726989746094
tensor([[54.9042]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[54.90417]
[50.]
epoch: 155 loss: 24.050893783569336
tensor([[177.4538]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')

```

```

[177.45381]
[171.]
epoch: 156 loss: 41.65168380737305
tensor([[11.5691]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[11.569073]
[0.]
epoch: 157 loss: 133.84344482421875
tensor([[146.9479]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[146.94786]
[147.]
epoch: 158 loss: 0.002718504751101136
tensor([[116.1998]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[116.1998]
[119.]
epoch: 159 loss: 7.841127872467041
tensor([[121.8008]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[121.80078]
[125.]
epoch: 160 loss: 10.235000610351562
tensor([[30.4927]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[30.492691]
[34.]
epoch: 161 loss: 12.301216125488281
tensor([[83.4813]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[83.481346]
[85.]
epoch: 162 loss: 2.306309461593628
tensor([[108.4978]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[108.49778]
[107.]
epoch: 163 loss: 2.24334454536438
tensor([[50.2267]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[50.226734]
[49.]
epoch: 164 loss: 1.504876732826233
tensor([[55.6477]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[55.647697]
[53.]
epoch: 165 loss: 7.01030158996582

```

```

tensor([[163.3577]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[163.35774]
[168.]
epoch: 166 loss: 21.550556182861328
tensor([[159.0739]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[159.07387]
[160.]
epoch: 167 loss: 0.8577208518981934
tensor([[79.2512]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[79.25122]
[77.]
epoch: 168 loss: 5.067994594573975
tensor([[22.0471]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[22.047092]
[26.]
epoch: 169 loss: 15.62547779083252
tensor([[14.7813]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[14.781343]
[21.]
epoch: 170 loss: 38.671688079833984
tensor([[88.1336]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[88.13364]
[83.]
epoch: 171 loss: 26.354223251342773
tensor([[85.1275]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[85.12754]
[80.]
epoch: 172 loss: 26.291671752929688
tensor([[139.4618]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[139.46176]
[134.]
epoch: 173 loss: 29.830839157104492
tensor([[98.2792]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[98.27922]
[97.]
epoch: 174 loss: 1.636405348777771
tensor([[92.4932]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[92.49324]

```

```

[95.]
epoch: 175 loss: 6.283843994140625
tensor([[64.4280]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[64.428]
[67.]
epoch: 176 loss: 6.615176677703857
tensor([[185.8518]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[185.85182]
[197.]
epoch: 177 loss: 124.28187561035156
tensor([[189.0084]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[189.0084]
[195.]
epoch: 178 loss: 35.8991813659668
tensor([[106.0172]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[106.0172]
[105.]
epoch: 179 loss: 1.0346890687942505
tensor([[158.6129]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[158.61293]
[151.]
epoch: 180 loss: 57.95670700073242
tensor([[190.6442]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[190.64423]
[180.]
epoch: 181 loss: 113.29954528808594
tensor([[81.0491]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[81.04913]
[82.]
epoch: 182 loss: 0.9041475057601929
tensor([[24.3278]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[24.327751]
[29.]
epoch: 183 loss: 21.82990837097168
tensor([[39.4017]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[39.40174]
[45.]
epoch: 184 loss: 31.340503692626953
tensor([[94.2272]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],

```



```

device='cuda:0')
[94.22719]
[98.]
epoch: 185 loss: 14.234109878540039
tensor([[151.6443]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[151.64433]
[154.]
epoch: 186 loss: 5.549167633056641
For epoch 7 the test accuracy over the whole test set is 0 %
tensor([[183.3725]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[183.37251]
[184.]
epoch: 0 loss: 0.393740177154541
tensor([[11.4471]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[11.447121]
[0.]
epoch: 1 loss: 131.0365753173828
tensor([[108.2195]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[108.21947]
[112.]
epoch: 2 loss: 14.292428970336914
tensor([[39.5352]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[39.535217]
[44.]
epoch: 3 loss: 19.934284210205078
tensor([[33.3624]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[33.362408]
[37.]
epoch: 4 loss: 13.232077598571777
tensor([[13.6697]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[13.66971]
[12.]
epoch: 5 loss: 2.7879319190979004
tensor([[12.4109]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[12.410869]
[5.]
epoch: 6 loss: 54.92097473144531
tensor([[13.7743]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[13.774316]

```

```

[13.]
epoch: 7 loss: 0.5995650291442871
tensor([[145.9027]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[145.90273]
[143.]
epoch: 8 loss: 8.425813674926758
tensor([[81.6770]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[81.67704]
[81.]
epoch: 9 loss: 0.4583832919597626
tensor([[168.9980]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[168.99797]
[166.]
epoch: 10 loss: 8.98782730102539
tensor([[156.8817]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[156.88165]
[155.]
epoch: 11 loss: 3.5406174659729004
tensor([[155.4583]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[155.4583]
[156.]
epoch: 12 loss: 0.2934413552284241
tensor([[112.1450]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[112.14502]
[115.]
epoch: 13 loss: 8.15091323852539
tensor([[201.4969]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[201.49689]
[204.]
epoch: 14 loss: 6.265573501586914
tensor([[76.3482]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[76.348236]
[79.]
epoch: 15 loss: 7.031851768493652
tensor([[102.6591]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[102.659134]
[104.]
epoch: 16 loss: 1.797921895980835
tensor([[90.3674]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],

```

```

device='cuda:0')
[90.36738]
[92.]
epoch: 17 loss: 2.6654539108276367
tensor([[11.4893]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[11.489344]
[15.]
epoch: 18 loss: 12.324707984924316
tensor([[118.4575]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[118.45746]
[116.]
epoch: 19 loss: 6.039102077484131
tensor([[11.8666]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[11.866634]
[11.]
epoch: 20 loss: 0.751055121421814
tensor([[107.6430]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[107.64299]
[105.]
epoch: 21 loss: 6.985396862030029
tensor([[117.1725]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[117.17246]
[118.]
epoch: 22 loss: 0.6848183870315552
tensor([[84.6504]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[84.650406]
[86.]
epoch: 23 loss: 1.8214043378829956
tensor([[8.4910]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[8.490935]
[3.]
epoch: 24 loss: 30.151010513305664
tensor([[58.9320]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[58.932014]
[58.]
epoch: 25 loss: 0.8686509728431702
tensor([[143.5199]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[143.51991]
[148.]

```

```

epoch: 26 loss: 20.071182250976562
tensor([[73.1143]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[73.11434]
[74.]
epoch: 27 loss: 0.7843905687332153
tensor([[11.0484]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[11.048366]
[21.]
epoch: 28 loss: 99.03502655029297
tensor([[33.2999]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[33.29993]
[36.]
epoch: 29 loss: 7.290374755859375
tensor([[116.0716]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[116.07156]
[110.]
epoch: 30 loss: 36.863887786865234
tensor([[167.4247]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[167.4247]
[159.]
epoch: 31 loss: 70.97553253173828
tensor([[171.4253]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[171.42528]
[167.]
epoch: 32 loss: 19.58308219909668
tensor([[79.0071]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[79.00709]
[78.]
epoch: 33 loss: 1.0142256021499634
tensor([[25.6095]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[25.609474]
[29.]
epoch: 34 loss: 11.495665550231934
tensor([[23.9981]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[23.99807]
[28.]
epoch: 35 loss: 16.015445709228516
tensor([[171.5434]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')

```

```

[171.54343]
[183.]
epoch: 36 loss: 131.25308227539062
tensor([[86.3032]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[86.30322]
[91.]
epoch: 37 loss: 22.059717178344727
tensor([[103.5852]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[103.585205]
[107.]
epoch: 38 loss: 11.6608247756958
tensor([[12.9530]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[12.952998]
[7.]
epoch: 39 loss: 35.43818664550781
tensor([[94.8315]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[94.83145]
[94.]
epoch: 40 loss: 0.6913114786148071
tensor([[51.3404]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[51.34041]
[48.]
epoch: 41 loss: 11.15832805633545
tensor([[68.2149]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[68.21494]
[62.]
epoch: 42 loss: 38.62551498413086
tensor([[86.0193]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[86.01934]
[83.]
epoch: 43 loss: 9.116416931152344
tensor([[98.9187]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[98.91867]
[96.]
epoch: 44 loss: 8.518638610839844
tensor([[54.4370]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[54.437042]
[51.]
epoch: 45 loss: 11.81325912475586

```

```

tensor([[77.1028]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[77.10284]
[76.]
epoch: 46 loss: 1.216248631477356
tensor([[122.7180]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[122.71797]
[125.]
epoch: 47 loss: 5.207652568817139
tensor([[198.4778]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[198.47783]
[197.]
epoch: 48 loss: 2.18397855758667
tensor([[18.3530]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[18.353035]
[24.]
epoch: 49 loss: 31.888214111328125
tensor([[96.0474]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[96.04738]
[99.]
epoch: 50 loss: 8.717973709106445
tensor([[135.9851]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[135.9851]
[142.]
epoch: 51 loss: 36.178932189941406
tensor([[184.6577]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[184.65773]
[185.]
epoch: 52 loss: 0.11714868247509003
tensor([[132.9813]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[132.98134]
[135.]
epoch: 53 loss: 4.074994087219238
tensor([[150.8703]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[150.87029]
[150.]
epoch: 54 loss: 0.7573960423469543
tensor([[21.9983]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[21.998262]

```

```

[25.]
epoch: 55 loss: 9.010428428649902
tensor([[168.1391]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[168.13907]
[162.]
epoch: 56 loss: 37.68816375732422
tensor([[213.2357]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[213.23573]
[203.]
epoch: 57 loss: 104.77023315429688
tensor([[139.4207]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[139.42073]
[141.]
epoch: 58 loss: 2.494091749191284
tensor([[118.8905]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[118.890495]
[123.]
epoch: 59 loss: 16.888029098510742
tensor([[48.9819]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[48.981903]
[50.]
epoch: 60 loss: 1.036521315574646
tensor([[54.0368]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[54.036842]
[55.]
epoch: 61 loss: 0.9276726841926575
tensor([[11.9147]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[11.914655]
[8.]
epoch: 62 loss: 15.324522018432617
tensor([[175.0134]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[175.01341]
[190.]
epoch: 63 loss: 224.59780883789062
tensor([[109.3806]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[109.380585]
[117.]
epoch: 64 loss: 58.05548858642578
tensor([[144.5659]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],

```

```

device='cuda:0')
[144.56593]
[149.]
epoch: 65 loss: 19.660947799682617
tensor([[146.1958]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[146.19576]
[144.]
epoch: 66 loss: 4.821340084075928
tensor([[140.9118]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[140.91185]
[134.]
epoch: 67 loss: 47.7736701965332
tensor([[97.0093]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[97.009346]
[90.]
epoch: 68 loss: 49.13093185424805
tensor([[148.4226]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[148.4226]
[138.]
epoch: 69 loss: 108.63074493408203
tensor([[109.4708]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[109.47082]
[102.]
epoch: 70 loss: 55.813114166259766
tensor([[15.4994]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[15.499357]
[18.]
epoch: 71 loss: 6.253214359283447
tensor([[30.9453]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[30.945278]
[33.]
epoch: 72 loss: 4.221881866455078
tensor([[77.0009]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[77.0009]
[80.]
epoch: 73 loss: 8.994599342346191
tensor([[42.7645]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[42.764492]
[47.]

```



```

epoch: 74 loss: 17.93952751159668
tensor([[129.5107]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[129.51067]
[137.]
epoch: 75 loss: 56.090126037597656
tensor([[168.2190]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[168.21896]
[176.]
epoch: 76 loss: 60.54452896118164
tensor([[11.3424]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[11.34237]
[4.]
epoch: 77 loss: 53.910396575927734
tensor([[116.9398]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[116.93984]
[119.]
epoch: 78 loss: 4.244250297546387
tensor([[69.8884]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[69.88839]
[69.]
epoch: 79 loss: 0.7892360687255859
tensor([[132.9695]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[132.96945]
[130.]
epoch: 80 loss: 8.817645072937012
tensor([[112.1069]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[112.1069]
[108.]
epoch: 81 loss: 16.866653442382812
tensor([[69.0528]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[69.052795]
[67.]
epoch: 82 loss: 4.2139692306518555
tensor([[30.2388]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[30.238789]
[32.]
epoch: 83 loss: 3.101865530014038
tensor([[75.9138]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')

```

```

[75.913795]
[75.]
epoch: 84 loss: 0.8350221514701843
tensor([[214.3793]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[214.37929]
[199.]
epoch: 85 loss: 236.52249145507812
tensor([[137.5839]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[137.58394]
[139.]
epoch: 86 loss: 2.005229949951172
tensor([[182.5151]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[182.51506]
[186.]
epoch: 87 loss: 12.144804000854492
tensor([[38.5450]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[38.544987]
[46.]
epoch: 88 loss: 55.57722473144531
tensor([[181.8468]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[181.84677]
[193.]
epoch: 89 loss: 124.3945083618164
tensor([[75.2001]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[75.200066]
[84.]
epoch: 90 loss: 77.4388427734375
tensor([[163.5538]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[163.55382]
[173.]
epoch: 91 loss: 89.23036193847656
tensor([[110.4630]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[110.463]
[113.]
epoch: 92 loss: 6.436381816864014
tensor([[13.0931]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[13.093132]
[10.]
epoch: 93 loss: 9.567465782165527

```

```

tensor([[191.0966]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[191.0966]
[178.]
epoch: 94 loss: 171.52102661132812
tensor([[66.0581]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[66.05805]
[60.]
epoch: 95 loss: 36.69999313354492
tensor([[162.8490]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[162.84901]
[153.]
epoch: 96 loss: 97.00308227539062
tensor([[95.2439]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[95.24385]
[93.]
epoch: 97 loss: 5.034865856170654
tensor([[179.4593]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[179.45927]
[179.]
epoch: 98 loss: 0.2109328806400299
tensor([[84.6538]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[84.653755]
[89.]
epoch: 99 loss: 18.88984489440918
tensor([[132.4983]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[132.49832]
[140.]
epoch: 100 loss: 56.27518081665039
tensor([[161.6776]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[161.67764]
[171.]
epoch: 101 loss: 86.90634155273438
tensor([[68.5598]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[68.55978]
[71.]
epoch: 102 loss: 5.9546918869018555
tensor([[149.4730]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[149.47299]

```

```

[152.]
epoch: 103 loss: 6.385769844055176
tensor([[85.1143]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[85.114296]
[85.]
epoch: 104 loss: 0.013063566759228706
tensor([[13.5291]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[13.529131]
[14.]
epoch: 105 loss: 0.22171767055988312
tensor([[153.8058]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[153.80579]
[147.]
epoch: 106 loss: 46.3187255859375
tensor([[171.7102]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[171.7102]
[164.]
epoch: 107 loss: 59.447261810302734
tensor([[25.2329]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[25.232891]
[27.]
epoch: 108 loss: 3.122673988342285
tensor([[70.9670]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[70.96705]
[68.]
epoch: 109 loss: 8.803378105163574
tensor([[10.4347]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[10.434711]
[1.]
epoch: 110 loss: 89.01377868652344
tensor([[12.9088]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[12.908812]
[17.]
epoch: 111 loss: 16.737823486328125
tensor([[74.9466]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[74.94664]
[77.]
epoch: 112 loss: 4.216287136077881
tensor([[122.8218]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],

```

```

device='cuda:0')
[122.821815]
[129.]
epoch: 113 loss: 38.16996383666992
tensor([[82.8820]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[82.881996]
[87.]
epoch: 114 loss: 16.957956314086914
tensor([[31.9244]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[31.924423]
[35.]
epoch: 115 loss: 9.459172248840332
tensor([[157.6994]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[157.69937]
[163.]
epoch: 116 loss: 28.096664428710938
tensor([[58.4570]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[58.45702]
[56.]
epoch: 117 loss: 6.0369462966918945
tensor([[23.4202]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[23.420225]
[26.]
epoch: 118 loss: 6.655238151550293
tensor([[172.3149]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[172.31493]
[170.]
epoch: 119 loss: 5.358882904052734
tensor([[15.2529]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[15.252947]
[20.]
epoch: 120 loss: 22.534513473510742
tensor([[75.7113]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[75.711296]
[70.]
epoch: 121 loss: 32.61890411376953
tensor([[202.3066]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[202.30661]
[195.]

```

```

epoch: 122 loss: 53.38655090332031
tensor([[200.7362]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[200.73619]
[198.]
epoch: 123 loss: 7.4867401123046875
tensor([[155.7973]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[155.79732]
[157.]
epoch: 124 loss: 1.4464452266693115
tensor([[128.9558]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[128.95576]
[132.]
epoch: 125 loss: 9.26736831665039
tensor([[96.1350]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[96.13498]
[97.]
epoch: 126 loss: 0.7482609152793884
tensor([[184.9214]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[184.9214]
[196.]
epoch: 127 loss: 122.73533630371094
tensor([[187.5561]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[187.55612]
[194.]
epoch: 128 loss: 41.52356719970703
tensor([[122.1517]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[122.1517]
[120.]
epoch: 129 loss: 4.629825115203857
tensor([[169.6604]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[169.66042]
[165.]
epoch: 130 loss: 21.71947479248047
tensor([[60.6331]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[60.633087]
[53.]
epoch: 131 loss: 58.26401901245117
tensor([[208.6930]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')

```

```

[208.69304]
[200.]
epoch: 132 loss: 75.56892395019531
tensor([[207.2241]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[207.22408]
[205.]
epoch: 133 loss: 4.9465107917785645
tensor([[111.8046]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[111.80459]
[109.]
epoch: 134 loss: 7.865715503692627
tensor([[81.9104]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[81.91036]
[82.]
epoch: 135 loss: 0.008034927770495415
tensor([[148.0862]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[148.08617]
[154.]
epoch: 136 loss: 34.973426818847656
tensor([[172.9699]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[172.96994]
[182.]
epoch: 137 loss: 81.54197692871094
tensor([[183.1586]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[183.15857]
[189.]
epoch: 138 loss: 34.122310638427734
tensor([[170.8176]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[170.81758]
[172.]
epoch: 139 loss: 1.3981143236160278
tensor([[11.9674]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[11.967415]
[2.]
epoch: 140 loss: 99.34935760498047
tensor([[101.1394]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[101.13943]
[95.]
epoch: 141 loss: 37.69256591796875

```

```

tensor([[188.0375]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[188.03749]
[177.]
epoch: 142 loss: 121.82620239257812
tensor([[137.4588]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[137.45883]
[133.]
epoch: 143 loss: 19.881181716918945
tensor([[174.6545]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[174.65451]
[174.]
epoch: 144 loss: 0.428384006023407
tensor([[39.5036]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[39.503605]
[45.]
epoch: 145 loss: 30.210359573364258
tensor([[94.2218]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[94.22184]
[100.]
epoch: 146 loss: 33.38713455200195
tensor([[115.1354]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[115.13537]
[122.]
epoch: 147 loss: 47.1231689453125
tensor([[129.2595]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[129.25948]
[136.]
epoch: 148 loss: 45.434669494628906
tensor([[32.5788]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[32.57885]
[39.]
epoch: 149 loss: 41.231170654296875
tensor([[99.9594]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[99.95944]
[101.]
epoch: 150 loss: 1.0827606916427612
tensor([[9.6192]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[9.619214]

```



```

[9.]
epoch: 151 loss: 0.3834260404109955
tensor([[11.7659]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[11.765852]
[16.]
epoch: 152 loss: 17.928009033203125
tensor([[36.5289]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[36.528866]
[38.]
epoch: 153 loss: 2.164235830307007
tensor([[58.3488]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[58.348824]
[54.]
epoch: 154 loss: 18.91226577758789
tensor([[204.4657]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[204.46574]
[191.]
epoch: 155 loss: 181.32626342773438
tensor([[151.3382]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[151.33821]
[146.]
epoch: 156 loss: 28.496498107910156
tensor([[66.5888]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[66.58882]
[65.]
epoch: 157 loss: 2.524353504180908
tensor([[9.8159]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[9.815877]
[6.]
epoch: 158 loss: 14.560916900634766
tensor([[15.6903]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[15.690339]
[23.]
epoch: 159 loss: 53.4311408996582
tensor([[104.0250]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[104.02505]
[114.]
epoch: 160 loss: 99.49967956542969
tensor([[25.6388]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],

```

```

device='cuda:0')
[25.638762]
[30.]
epoch: 161 loss: 19.020401000976562
tensor([[112.6819]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[112.68194]
[124.]
epoch: 162 loss: 128.09852600097656
tensor([[60.2977]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[60.297714]
[61.]
epoch: 163 loss: 0.49320530891418457
tensor([[49.6477]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[49.647736]
[49.]
epoch: 164 loss: 0.4195614159107208
tensor([[176.1459]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[176.14594]
[192.]
epoch: 165 loss: 251.35137939453125
tensor([[37.6315]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[37.631516]
[34.]
epoch: 166 loss: 13.187905311584473
tensor([[74.5380]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[74.537964]
[64.]
epoch: 167 loss: 111.0486831665039
tensor([[49.3258]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[49.325832]
[43.]
epoch: 168 loss: 40.01615524291992
tensor([[183.8649]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[183.86488]
[169.]
epoch: 169 loss: 220.96475219726562
tensor([[121.0586]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[121.058624]
[111.]

```

```
epoch: 170 loss: 101.17591857910156
tensor([[153.1193]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[153.11932]
[151.]
epoch: 171 loss: 4.491533279418945
tensor([[154.6142]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[154.61418]
[160.]
epoch: 172 loss: 29.00704002380371
tensor([[136.7389]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[136.73889]
[145.]
epoch: 173 loss: 68.24591064453125
tensor([[97.6052]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[97.60519]
[103.]
epoch: 174 loss: 29.104013442993164
tensor([[81.7785]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[81.77851]
[88.]
epoch: 175 loss: 38.70692443847656
tensor([[165.2835]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[165.28346]
[175.]
epoch: 176 loss: 94.41110229492188
tensor([[104.8343]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[104.8343]
[106.]
epoch: 177 loss: 1.358863115310669
tensor([[52.3857]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[52.385674]
[52.]
epoch: 178 loss: 0.14874406158924103
tensor([[188.2228]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[188.22278]
[180.]
epoch: 179 loss: 67.61408233642578
tensor([[133.6028]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
```

```

[133.60283]
[126.]
epoch: 180 loss: 57.803009033203125
tensor([[210.7874]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[210.78743]
[201.]
epoch: 181 loss: 95.79378509521484
tensor([[56.3713]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[56.3713]
[57.]
epoch: 182 loss: 0.39526399970054626
tensor([[16.1776]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[16.177635]
[22.]
epoch: 183 loss: 33.899932861328125
tensor([[155.9321]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[155.93213]
[168.]
epoch: 184 loss: 145.63351440429688
tensor([[92.6091]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[92.609116]
[98.]
epoch: 185 loss: 29.061634063720703
tensor([[14.4702]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[14.470215]
[19.]
epoch: 186 loss: 20.518953323364258
For epoch 8 the test accuracy over the whole test set is 0 %
tensor([[46.4940]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[46.49398]
[49.]
epoch: 0 loss: 6.280134201049805
tensor([[91.1084]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[91.10839]
[93.]
epoch: 1 loss: 3.5781853199005127
tensor([[198.3384]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[198.33836]
[205.]

```

```

epoch: 2 loss: 44.37739944458008
tensor([[196.1276]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[196.1276]
[193.]
epoch: 3 loss: 9.781844139099121
tensor([[145.1159]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[145.11586]
[140.]
epoch: 4 loss: 26.17202377319336
tensor([[201.9534]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[201.95343]
[195.]
epoch: 5 loss: 48.35019302368164
tensor([[120.2929]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[120.29292]
[116.]
epoch: 6 loss: 18.429187774658203
tensor([[61.2888]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[61.28883]
[58.]
epoch: 7 loss: 10.816401481628418
tensor([[48.8688]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[48.868763]
[48.]
epoch: 8 loss: 0.7547491192817688
tensor([[130.8800]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[130.88004]
[136.]
epoch: 9 loss: 26.21403694152832
tensor([[119.7592]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[119.75919]
[124.]
epoch: 10 loss: 17.984439849853516
tensor([[132.7055]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[132.7055]
[139.]
epoch: 11 loss: 39.620662689208984
tensor([[27.5481]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')

```

```

[27.548111]
[27.]
epoch: 12 loss: 0.30042561888694763
tensor([[134.1614]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[134.16142]
[137.]
epoch: 13 loss: 8.057520866394043
tensor([[111.5481]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[111.54813]
[111.]
epoch: 14 loss: 0.3004423677921295
tensor([[60.2420]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[60.242012]
[56.]
epoch: 15 loss: 17.994665145874023
tensor([[196.0043]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[196.00427]
[196.]
epoch: 16 loss: 1.825392246246338e-05
tensor([[130.3285]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[130.32852]
[126.]
epoch: 17 loss: 18.736101150512695
tensor([[83.6309]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[83.63085]
[80.]
epoch: 18 loss: 13.183084487915039
tensor([[97.1807]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[97.18069]
[94.]
epoch: 19 loss: 10.116769790649414
tensor([[181.7117]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[181.71172]
[183.]
epoch: 20 loss: 1.659676432609558
tensor([[25.3610]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[25.36101]
[25.]
epoch: 21 loss: 0.13032792508602142

```

```

tensor([[19.7678]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[19.76779]
[21.]
epoch: 22 loss: 1.5183418989181519
tensor([[195.9743]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[195.97427]
[201.]
epoch: 23 loss: 25.257925033569336
tensor([[15.5782]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[15.57825]
[5.]
epoch: 24 loss: 111.89936828613281
tensor([[144.9855]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[144.98553]
[146.]
epoch: 25 loss: 1.0291398763656616
tensor([[92.4883]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[92.488304]
[92.]
epoch: 26 loss: 0.23844093084335327
tensor([[160.6468]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[160.6468]
[160.]
epoch: 27 loss: 0.41835644841194153
tensor([[99.3698]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[99.36981]
[98.]
epoch: 28 loss: 1.876384973526001
tensor([[69.6893]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[69.68934]
[71.]
epoch: 29 loss: 1.7178330421447754
tensor([[26.9601]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[26.960117]
[29.]
epoch: 30 loss: 4.161121368408203
tensor([[100.4503]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[100.45034]

```

```

[100.]
epoch: 31 loss: 0.20280635356903076
tensor([[89.7082]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[89.708244]
[90.]
epoch: 32 loss: 0.08512137830257416
tensor([[59.5723]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[59.572296]
[61.]
epoch: 33 loss: 2.0383384227752686
tensor([[72.1803]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[72.18026]
[74.]
epoch: 34 loss: 3.3114547729492188
tensor([[52.8247]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[52.82467]
[54.]
epoch: 35 loss: 1.3814032077789307
tensor([[18.9170]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[18.91695]
[23.]
epoch: 36 loss: 16.671295166015625
tensor([[117.7098]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[117.709816]
[114.]
epoch: 37 loss: 13.762734413146973
tensor([[24.0532]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[24.05322]
[26.]
epoch: 38 loss: 3.789949417114258
tensor([[15.1843]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[15.184298]
[15.]
epoch: 39 loss: 0.033965591341257095
tensor([[156.7781]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[156.7781]
[151.]
epoch: 40 loss: 33.38651657104492
tensor([[179.4425]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],

```



```

device='cuda:0')
[179.44246]
[176.]
epoch: 41 loss: 11.85052490234375
tensor([[55.3982]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[55.398228]
[57.]
epoch: 42 loss: 2.5656745433807373
tensor([[105.3260]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[105.32598]
[107.]
epoch: 43 loss: 2.8023390769958496
tensor([[11.3159]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[11.315899]
[2.]
epoch: 44 loss: 86.78597259521484
tensor([[14.9856]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[14.985617]
[20.]
epoch: 45 loss: 25.144039154052734
tensor([[29.7825]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[29.78249]
[34.]
epoch: 46 loss: 17.787391662597656
tensor([[12.2571]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[12.2570915]
[9.]
epoch: 47 loss: 10.60864543914795
tensor([[115.3867]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[115.3867]
[123.]
epoch: 48 loss: 57.962284088134766
tensor([[91.0551]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[91.05507]
[97.]
epoch: 49 loss: 35.34220504760742
tensor([[127.9533]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[127.95326]
[134.]

```

```
epoch: 50 loss: 36.56303787231445
tensor([[49.0765]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[49.07648]
[50.]
epoch: 51 loss: 0.8528875708580017
tensor([[104.1895]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[104.18947]
[101.]
epoch: 52 loss: 10.172708511352539
tensor([[168.9568]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[168.95679]
[166.]
epoch: 53 loss: 8.742589950561523
tensor([[48.3422]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[48.34221]
[47.]
epoch: 54 loss: 1.8015246391296387
tensor([[44.6951]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[44.69513]
[44.]
epoch: 55 loss: 0.4832048714160919
tensor([[104.6381]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[104.63808]
[99.]
epoch: 56 loss: 31.78791046142578
tensor([[36.4442]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[36.444202]
[36.]
epoch: 57 loss: 0.19731579720973969
tensor([[170.1548]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[170.1548]
[169.]
epoch: 58 loss: 1.333564043045044
tensor([[144.1770]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[144.17697]
[142.]
epoch: 59 loss: 4.739204406738281
tensor([[178.2755]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
```

```

[178.27551]
[180.]
epoch: 60 loss: 2.9738564491271973
tensor([[167.9629]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[167.96288]
[171.]
epoch: 61 loss: 9.224125862121582
tensor([[97.2482]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[97.248245]
[96.]
epoch: 62 loss: 1.5581161975860596
tensor([[32.6880]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[32.68801]
[33.]
epoch: 63 loss: 0.0973370298743248
tensor([[123.0816]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[123.08159]
[122.]
epoch: 64 loss: 1.1698342561721802
tensor([[185.4003]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[185.40034]
[186.]
epoch: 65 loss: 0.35958629846572876
tensor([[14.7006]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[14.700648]
[14.]
epoch: 66 loss: 0.49090805649757385
tensor([[171.9825]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[171.9825]
[175.]
epoch: 67 loss: 9.105317115783691
tensor([[85.3524]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[85.35242]
[85.]
epoch: 68 loss: 0.12419773638248444
tensor([[131.7627]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[131.76271]
[130.]
epoch: 69 loss: 3.1071486473083496

```

```

tensor([[12.7461]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[12.746076]
[8.]
epoch: 70 loss: 22.52523422241211
tensor([[171.9080]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[171.90799]
[172.]
epoch: 71 loss: 0.008465931750833988
tensor([[68.6168]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[68.6168]
[68.]
epoch: 72 loss: 0.38044026494026184
tensor([[63.9239]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[63.92391]
[64.]
epoch: 73 loss: 0.005789957009255886
tensor([[77.8840]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[77.884]
[78.]
epoch: 74 loss: 0.013455376960337162
tensor([[110.8746]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[110.874565]
[110.]
epoch: 75 loss: 0.7648641467094421
tensor([[25.1396]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[25.139605]
[28.]
epoch: 76 loss: 8.181861877441406
tensor([[120.7588]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[120.75882]
[120.]
epoch: 77 loss: 0.5758071541786194
tensor([[39.0814]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[39.081367]
[43.]
epoch: 78 loss: 15.355680465698242
tensor([[69.2157]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[69.215675]

```

```

[70.]
epoch: 79 loss: 0.6151651740074158
tensor([[151.0838]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[151.08376]
[150.]
epoch: 80 loss: 1.1745259761810303
tensor([[27.6896]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[27.689552]
[30.]
epoch: 81 loss: 5.338168621063232
tensor([[66.8600]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[66.86001]
[67.]
epoch: 82 loss: 0.019597692415118217
tensor([[202.4998]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[202.49983]
[200.]
epoch: 83 loss: 6.2491607666015625
tensor([[172.9325]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[172.93246]
[174.]
epoch: 84 loss: 1.1396318674087524
tensor([[113.1491]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[113.14912]
[113.]
epoch: 85 loss: 0.02223573625087738
tensor([[102.9554]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[102.955414]
[102.]
epoch: 86 loss: 0.9128155708312988
tensor([[161.9547]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[161.95471]
[164.]
epoch: 87 loss: 4.183203220367432
tensor([[12.9998]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[12.999753]
[11.]
epoch: 88 loss: 3.999011993408203
tensor([[116.7480]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],

```

```

device='cuda:0')
[116.74796]
[117.]
epoch: 89 loss: 0.06352267414331436
tensor([[13.0389]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[13.038866]
[12.]
epoch: 90 loss: 1.0792427062988281
tensor([[156.6943]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[156.69434]
[157.]
epoch: 91 loss: 0.0934305191040039
tensor([[190.5355]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[190.53546]
[190.]
epoch: 92 loss: 0.2867189347743988
tensor([[118.7340]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[118.73397]
[119.]
epoch: 93 loss: 0.07077161967754364
tensor([[163.3297]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[163.32965]
[165.]
epoch: 94 loss: 2.7900662422180176
tensor([[133.7160]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[133.71605]
[133.]
epoch: 95 loss: 0.5127264261245728
tensor([[88.9307]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[88.9307]
[89.]
epoch: 96 loss: 0.004802183713763952
tensor([[86.2727]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[86.27268]
[86.]
epoch: 97 loss: 0.0743555799126625
tensor([[12.3839]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[12.383905]
[10.]

```

```

epoch: 98 loss: 5.683004856109619
tensor([[167.8999]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[167.89989]
[168.]
epoch: 99 loss: 0.010022595524787903
tensor([[159.7805]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[159.78053]
[159.]
epoch: 100 loss: 0.6092315316200256
tensor([[110.0878]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[110.087845]
[109.]
epoch: 101 loss: 1.1834064722061157
tensor([[130.1472]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[130.14725]
[129.]
epoch: 102 loss: 1.316176414489746
tensor([[143.7611]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[143.76111]
[144.]
epoch: 103 loss: 0.05706919729709625
tensor([[202.4022]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[202.40224]
[204.]
epoch: 104 loss: 2.552846908569336
tensor([[155.2752]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[155.27522]
[156.]
epoch: 105 loss: 0.525301992893219
tensor([[195.3767]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[195.37674]
[194.]
epoch: 106 loss: 1.8954116106033325
tensor([[182.0937]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[182.09372]
[182.]
epoch: 107 loss: 0.008783341385424137
tensor([[125.3636]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')

```

```

[125.363594]
[125.]
epoch: 108 loss: 0.13220064342021942
tensor([[82.2215]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[82.221466]
[84.]
epoch: 109 loss: 3.1631829738616943
tensor([[177.2504]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[177.25038]
[179.]
epoch: 110 loss: 3.0611650943756104
tensor([[73.6649]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[73.66491]
[75.]
epoch: 111 loss: 1.782467007637024
tensor([[132.9407]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[132.94067]
[132.]
epoch: 112 loss: 0.884867250919342
tensor([[53.0517]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[53.051735]
[53.]
epoch: 113 loss: 0.0026765023358166218
tensor([[202.2494]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[202.24942]
[199.]
epoch: 114 loss: 10.558731079101562
tensor([[51.5605]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[51.560528]
[52.]
epoch: 115 loss: 0.1931358128786087
tensor([[81.0568]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[81.05677]
[82.]
epoch: 116 loss: 0.8896822333335876
tensor([[76.1648]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[76.16484]
[77.]
epoch: 117 loss: 0.6974910497665405

```



```

tensor([[196.0995]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[196.0995]
[197.]
epoch: 118 loss: 0.8108956217765808
tensor([[15.2821]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[15.282135]
[22.]
epoch: 119 loss: 45.12971115112305
tensor([[8.6616]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[8.661627]
[1.]
epoch: 120 loss: 58.70052719116211
tensor([[152.4326]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[152.43265]
[153.]
epoch: 121 loss: 0.32188862562179565
tensor([[154.7082]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[154.7082]
[155.]
epoch: 122 loss: 0.08514363318681717
tensor([[154.4242]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[154.42421]
[154.]
epoch: 123 loss: 0.17995378375053406
tensor([[43.7998]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[43.79982]
[45.]
epoch: 124 loss: 1.4404321908950806
tensor([[147.1508]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[147.15082]
[147.]
epoch: 125 loss: 0.022746030241250992
tensor([[139.4918]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[139.49176]
[138.]
epoch: 126 loss: 2.225348711013794
tensor([[107.2502]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[107.250206]

```

```

[106.]
epoch: 127 loss: 1.5630149841308594
tensor([[56.3671]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[56.367096]
[55.]
epoch: 128 loss: 1.8689513206481934
tensor([[188.3985]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[188.39847]
[189.]
epoch: 129 loss: 0.36184072494506836
tensor([[104.0833]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[104.08328]
[104.]
epoch: 130 loss: 0.006935969926416874
tensor([[149.6494]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[149.64941]
[152.]
epoch: 131 loss: 5.525254249572754
tensor([[86.3270]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[86.32697]
[87.]
epoch: 132 loss: 0.4529654085636139
tensor([[111.3611]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[111.36111]
[112.]
epoch: 133 loss: 0.4081844389438629
tensor([[65.7248]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[65.72478]
[65.]
epoch: 134 loss: 0.525301992893219
tensor([[45.1145]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[45.114544]
[46.]
epoch: 135 loss: 0.7840324640274048
tensor([[52.1297]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[52.12969]
[51.]
epoch: 136 loss: 1.2761955261230469
tensor([[119.1353]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],

```

```

device='cuda:0')
[119.1353]
[118.]
epoch: 137 loss: 1.288905382156372
tensor([[116.7248]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[116.72478]
[115.]
epoch: 138 loss: 2.974856376647949
tensor([[18.9960]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[18.99605]
[24.]
epoch: 139 loss: 25.03951644897461
tensor([[109.0785]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[109.07847]
[108.]
epoch: 140 loss: 1.1630939245224
tensor([[12.9975]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[12.997523]
[16.]
epoch: 141 loss: 9.01486587524414
tensor([[198.0055]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[198.00546]
[198.]
epoch: 142 loss: 2.9840506613254547e-05
tensor([[76.9630]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[76.96299]
[76.]
epoch: 143 loss: 0.9273493885993958
tensor([[167.9131]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[167.91315]
[170.]
epoch: 144 loss: 4.354955673217773
tensor([[145.2022]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[145.2022]
[145.]
epoch: 145 loss: 0.040882501751184464
tensor([[166.0117]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[166.01169]
[167.]

```

```

epoch: 146 loss: 0.9767601490020752
tensor([[148.6100]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[148.61002]
[148.]
epoch: 147 loss: 0.37211936712265015
tensor([[16.1593]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[16.159323]
[18.]
epoch: 148 loss: 3.3880927562713623
tensor([[85.3774]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[85.37738]
[83.]
epoch: 149 loss: 5.651937484741211
tensor([[188.3751]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[188.37509]
[184.]
epoch: 150 loss: 19.14142608642578
tensor([[16.0702]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[16.070152]
[17.]
epoch: 151 loss: 0.8646167516708374
tensor([[105.1425]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[105.14251]
[105.]
epoch: 152 loss: 0.020308947190642357
tensor([[69.6024]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[69.60239]
[69.]
epoch: 153 loss: 0.3628694713115692
tensor([[165.5616]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[165.56158]
[173.]
epoch: 154 loss: 55.33002471923828
tensor([[138.6821]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[138.68207]
[141.]
epoch: 155 loss: 5.372809410095215
tensor([[79.8623]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')

```

```

[79.86225]
[79.]
epoch: 156 loss: 0.7434772849082947
tensor([[13.5627]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[13.562658]
[3.]
epoch: 157 loss: 111.56974792480469
tensor([[13.2507]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[13.250651]
[4.]
epoch: 158 loss: 85.57454681396484
tensor([[62.5806]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[62.580597]
[60.]
epoch: 159 loss: 6.659480571746826
tensor([[105.2483]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[105.24829]
[103.]
epoch: 160 loss: 5.054812431335449
tensor([[9.9756]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[9.975571]
[6.]
epoch: 161 loss: 15.80516242980957
tensor([[193.9871]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[193.9871]
[191.]
epoch: 162 loss: 8.92280387878418
tensor([[92.3283]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[92.32831]
[95.]
epoch: 163 loss: 7.137937545776367
tensor([[83.4636]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[83.463554]
[88.]
epoch: 164 loss: 20.57933807373047
tensor([[160.1984]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[160.19838]
[163.]
epoch: 165 loss: 7.849077224731445

```

```

tensor([[25.2501]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[25.25008]
[32.]
epoch: 166 loss: 45.56142044067383
tensor([[142.4863]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[142.48627]
[143.]
epoch: 167 loss: 0.26392149925231934
tensor([[30.7815]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[30.781544]
[37.]
epoch: 168 loss: 38.66919708251953
tensor([[137.7347]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[137.73474]
[135.]
epoch: 169 loss: 7.478809356689453
tensor([[80.4407]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[80.44073]
[81.]
epoch: 170 loss: 0.3127860128879547
tensor([[192.6754]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[192.67543]
[185.]
epoch: 171 loss: 58.9122314453125
tensor([[150.4362]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[150.4362]
[149.]
epoch: 172 loss: 2.0626790523529053
tensor([[161.2727]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[161.27267]
[162.]
epoch: 173 loss: 0.52900230884552
tensor([[88.2304]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[88.23037]
[91.]
epoch: 174 loss: 7.6708526611328125
tensor([[32.9944]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[32.994446]

```

```

[38.]
epoch: 175 loss: 25.055572509765625
tensor([[4.7098]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[4.7098117]
[0.]
epoch: 176 loss: 22.18232536315918
tensor([[168.6936]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[168.69356]
[178.]
epoch: 177 loss: 86.60987091064453
tensor([[10.0462]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[10.046153]
[19.]
epoch: 178 loss: 80.17137145996094
tensor([[191.3363]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[191.33629]
[192.]
epoch: 179 loss: 0.440513014793396
tensor([[202.9140]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[202.91397]
[203.]
epoch: 180 loss: 0.007400997914373875
tensor([[35.7566]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[35.756596]
[35.]
epoch: 181 loss: 0.5724369287490845
tensor([[67.8045]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[67.80453]
[62.]
epoch: 182 loss: 33.692535400390625
tensor([[13.0761]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[13.076147]
[13.]
epoch: 183 loss: 0.005798377562314272
tensor([[41.3776]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[41.377632]
[39.]
epoch: 184 loss: 5.653134822845459
tensor([[11.3739]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],

```

```

device='cuda:0')
[11.37386]
[7.]
epoch: 185 loss: 19.13065528869629
tensor([[188.2639]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[188.26392]
[177.]
epoch: 186 loss: 126.87580108642578
For epoch 9 the test accuracy over the whole test set is 0 %
tensor([[8.5904]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[8.590382]
[2.]
epoch: 0 loss: 43.433128356933594
tensor([[140.4370]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[140.43695]
[138.]
epoch: 1 loss: 5.9387288093566895
tensor([[126.7352]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[126.7352]
[129.]
epoch: 2 loss: 5.129323482513428
tensor([[43.3858]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[43.38579]
[47.]
epoch: 3 loss: 13.062500953674316
tensor([[177.0709]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[177.07089]
[189.]
epoch: 4 loss: 142.30360412597656
tensor([[9.6688]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[9.668849]
[20.]
epoch: 5 loss: 106.73268127441406
tensor([[109.2873]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[109.28732]
[113.]
epoch: 6 loss: 13.783970832824707
tensor([[167.1730]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[167.17305]

```



```

[170.]
epoch: 7 loss: 7.991646766662598
tensor([[182.3087]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[182.30872]
[180.]
epoch: 8 loss: 5.330168724060059
tensor([[59.6053]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[59.60531]
[56.]
epoch: 9 loss: 12.998250007629395
tensor([[12.2969]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[12.296934]
[15.]
epoch: 10 loss: 7.306565284729004
tensor([[11.1639]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[11.163924]
[9.]
epoch: 11 loss: 4.682568073272705
tensor([[210.9336]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[210.93356]
[197.]
epoch: 12 loss: 194.14418029785156
tensor([[144.0879]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[144.08786]
[137.]
epoch: 13 loss: 50.237762451171875
tensor([[161.1195]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[161.11949]
[162.]
epoch: 14 loss: 0.775295078754425
tensor([[133.9626]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[133.96259]
[139.]
epoch: 15 loss: 25.375545501708984
tensor([[137.4579]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[137.45787]
[148.]
epoch: 16 loss: 111.1364974975586
tensor([[40.9137]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],

```

```

device='cuda:0')
[40.91375]
[44.]
epoch: 17 loss: 9.524940490722656
tensor([[111.1887]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[111.18867]
[117.]
epoch: 18 loss: 33.77158737182617
tensor([[154.4737]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[154.47368]
[164.]
epoch: 19 loss: 90.75080108642578
tensor([[101.5283]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[101.528336]
[100.]
epoch: 20 loss: 2.3358097076416016
tensor([[17.0413]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[17.041267]
[21.]
epoch: 21 loss: 15.671564102172852
tensor([[13.2757]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[13.2756815]
[8.]
epoch: 22 loss: 27.832815170288086
tensor([[162.0539]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[162.05392]
[151.]
epoch: 23 loss: 122.18924713134766
tensor([[195.4540]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[195.45404]
[192.]
epoch: 24 loss: 11.93039608001709
tensor([[159.5518]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[159.55177]
[152.]
epoch: 25 loss: 57.02927780151367
tensor([[87.7155]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[87.71553]
[82.]

```

```

epoch: 26 loss: 32.66728591918945
tensor([[71.4872]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[71.48718]
[68.]
epoch: 27 loss: 12.160442352294922
tensor([[89.6596]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[89.65956]
[91.]
epoch: 28 loss: 1.796776294708252
tensor([[168.1493]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[168.14929]
[200.]
epoch: 29 loss: 1014.4675903320312
tensor([[9.8476]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[9.847597]
[4.]
epoch: 30 loss: 34.194393157958984
tensor([[128.6953]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[128.69534]
[125.]
epoch: 31 loss: 13.655560493469238
tensor([[66.1335]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[66.133514]
[62.]
epoch: 32 loss: 17.085941314697266
tensor([[68.3081]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[68.30812]
[65.]
epoch: 33 loss: 10.943662643432617
tensor([[6.9230]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[6.9230285]
[1.]
epoch: 34 loss: 35.08226776123047
tensor([[5.6530]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[5.6529665]
[5.]
epoch: 35 loss: 0.4263652563095093
tensor([[50.6861]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')

```

```

[50.686108]
[55.]
epoch: 36 loss: 18.60966682434082
tensor([[31.9591]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[31.959085]
[39.]
epoch: 37 loss: 49.57447814941406
tensor([[183.8186]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[183.81862]
[171.]
epoch: 38 loss: 164.31698608398438
tensor([[17.4750]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[17.475021]
[27.]
epoch: 39 loss: 90.7252197265625
tensor([[105.8934]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[105.89336]
[111.]
epoch: 40 loss: 26.077810287475586
tensor([[102.2585]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[102.25849]
[107.]
epoch: 41 loss: 22.481903076171875
tensor([[23.4045]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[23.40448]
[28.]
epoch: 42 loss: 21.118804931640625
tensor([[106.2413]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[106.241295]
[104.]
epoch: 43 loss: 5.023402690887451
tensor([[182.2521]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[182.25214]
[178.]
epoch: 44 loss: 18.08066177368164
tensor([[132.7525]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[132.75246]
[132.]
epoch: 45 loss: 0.5661910176277161

```

```

tensor([[158.2151]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[158.21512]
[163.]
epoch: 46 loss: 22.895092010498047
tensor([[34.8042]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[34.80422]
[32.]
epoch: 47 loss: 7.863640308380127
tensor([[22.7614]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[22.761362]
[23.]
epoch: 48 loss: 0.05694805830717087
tensor([[95.1346]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[95.13461]
[92.]
epoch: 49 loss: 9.825798988342285
tensor([[63.9291]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[63.92911]
[60.]
epoch: 50 loss: 15.43791675567627
tensor([[149.9261]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[149.92612]
[156.]
epoch: 51 loss: 36.89205551147461
tensor([[140.0076]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[140.00755]
[142.]
epoch: 52 loss: 3.9698445796966553
tensor([[19.2201]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[19.220144]
[19.]
epoch: 53 loss: 0.048463501036167145
tensor([[176.4416]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[176.44164]
[174.]
epoch: 54 loss: 5.961582183837891
tensor([[83.2228]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[83.22276]

```

```

[78.]
epoch: 55 loss: 27.277254104614258
tensor([[107.0880]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[107.08796]
[101.]
epoch: 56 loss: 37.06324768066406
tensor([[169.0490]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[169.04897]
[169.]
epoch: 57 loss: 0.002397615695372224
tensor([[42.6344]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[42.634438]
[45.]
epoch: 58 loss: 5.595885753631592
tensor([[160.6552]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[160.65515]
[165.]
epoch: 59 loss: 18.877710342407227
tensor([[15.0230]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[15.022997]
[17.]
epoch: 60 loss: 3.908541202545166
tensor([[171.9164]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[171.91643]
[176.]
epoch: 61 loss: 16.67556381225586
tensor([[31.6049]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[31.604895]
[33.]
epoch: 62 loss: 1.9463189840316772
tensor([[100.1423]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[100.142296]
[98.]
epoch: 63 loss: 4.589431285858154
tensor([[19.7681]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[19.768124]
[22.]
epoch: 64 loss: 4.981272220611572
tensor([[70.5030]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],

```

```

device='cuda:0')
[70.50299]
[71.]
epoch: 65 loss: 0.2470182180404663
tensor([[188.5886]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[188.58861]
[177.]
epoch: 66 loss: 134.29583740234375
tensor([[23.6631]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[23.66312]
[25.]
epoch: 67 loss: 1.7872474193572998
tensor([[69.3855]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[69.38549]
[74.]
epoch: 68 loss: 21.293699264526367
tensor([[201.4957]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[201.49573]
[199.]
epoch: 69 loss: 6.228655815124512
tensor([[150.0670]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[150.06702]
[155.]
epoch: 70 loss: 24.334325790405273
tensor([[90.5818]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[90.58179]
[96.]
epoch: 71 loss: 29.357030868530273
tensor([[131.0943]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[131.09431]
[136.]
epoch: 72 loss: 24.065750122070312
tensor([[131.8719]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[131.87192]
[133.]
epoch: 73 loss: 1.2725696563720703
tensor([[144.8911]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[144.89108]
[140.]

```

```

epoch: 74 loss: 23.922691345214844
tensor([[145.7609]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[145.76088]
[141.]
epoch: 75 loss: 22.665973663330078
tensor([[19.6454]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[19.645386]
[10.]
epoch: 76 loss: 93.03346252441406
tensor([[78.0520]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[78.05196]
[79.]
epoch: 77 loss: 0.8987725973129272
tensor([[189.8636]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[189.86362]
[198.]
epoch: 78 loss: 66.20072937011719
tensor([[16.5963]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[16.596313]
[11.]
epoch: 79 loss: 31.318723678588867
tensor([[99.0081]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[99.00812]
[99.]
epoch: 80 loss: 6.58966600894928e-05
tensor([[85.2973]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[85.297295]
[86.]
epoch: 81 loss: 0.49379485845565796
tensor([[168.7650]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[168.76503]
[166.]
epoch: 82 loss: 7.645390510559082
tensor([[67.9475]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[67.947464]
[69.]
epoch: 83 loss: 1.1078320741653442
tensor([[173.6193]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')

```



```

[173.61928]
[172.]
epoch: 84 loss: 2.622061014175415
tensor([[85.9362]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[85.93618]
[87.]
epoch: 85 loss: 1.131712794303894
tensor([[188.7497]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[188.7497]
[185.]
epoch: 86 loss: 14.060211181640625
tensor([[10.7300]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[10.730023]
[12.]
epoch: 87 loss: 1.6128406524658203
tensor([[49.5660]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[49.565983]
[53.]
epoch: 88 loss: 11.792473793029785
tensor([[91.5700]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[91.570045]
[95.]
epoch: 89 loss: 11.764588356018066
tensor([[146.3167]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[146.3167]
[150.]
epoch: 90 loss: 13.566726684570312
tensor([[172.4174]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[172.41737]
[175.]
epoch: 91 loss: 6.669958591461182
tensor([[24.9821]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[24.982128]
[26.]
epoch: 92 loss: 1.0360630750656128
tensor([[36.6028]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[36.60281]
[34.]
epoch: 93 loss: 6.7746195793151855

```

```

tensor([[112.7262]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[112.72616]
[102.]
epoch: 94 loss: 115.05046844482422
tensor([[52.8205]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[52.82047]
[50.]
epoch: 95 loss: 7.955044746398926
tensor([[57.9579]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[57.95791]
[57.]
epoch: 96 loss: 0.9175889492034912
tensor([[187.6948]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[187.69481]
[183.]
epoch: 97 loss: 22.041231155395508
tensor([[6.2942]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[6.294151]
[3.]
epoch: 98 loss: 10.85142993927002
tensor([[9.1449]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[9.144933]
[16.]
epoch: 99 loss: 46.991947174072266
tensor([[43.3034]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[43.303444]
[51.]
epoch: 100 loss: 59.236976623535156
tensor([[95.6261]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[95.62614]
[106.]
epoch: 101 loss: 107.61703491210938
tensor([[28.7681]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[28.768053]
[35.]
epoch: 102 loss: 38.837162017822266
tensor([[80.9372]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[80.93718]

```

```

[88.]
epoch: 103 loss: 49.8834342956543
tensor([[27.0486]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[27.048573]
[29.]
epoch: 104 loss: 3.8080692291259766
tensor([[193.1313]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[193.13126]
[194.]
epoch: 105 loss: 0.7547159790992737
tensor([[113.5586]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[113.55863]
[108.]
epoch: 106 loss: 30.898387908935547
tensor([[71.8801]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[71.880135]
[67.]
epoch: 107 loss: 23.81571388244629
tensor([[127.9237]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[127.92372]
[123.]
epoch: 108 loss: 24.243032455444336
tensor([[64.8256]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[64.825554]
[61.]
epoch: 109 loss: 14.634862899780273
tensor([[85.5193]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[85.5193]
[84.]
epoch: 110 loss: 2.3082797527313232
tensor([[11.5328]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[11.532758]
[0.]
epoch: 111 loss: 133.00450134277344
tensor([[166.7564]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[166.75644]
[179.]
epoch: 112 loss: 149.90478515625
tensor([[47.8377]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],

```

```

device='cuda:0')
[47.83774]
[48.]
epoch: 113 loss: 0.0263277068734169
tensor([[146.0808]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[146.08081]
[147.]
epoch: 114 loss: 0.8449092507362366
tensor([[37.1306]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[37.130642]
[37.]
epoch: 115 loss: 0.01706731505692005
tensor([[85.5355]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[85.53547]
[83.]
epoch: 116 loss: 6.428603172302246
tensor([[36.2788]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[36.27876]
[36.]
epoch: 117 loss: 0.07770658284425735
tensor([[92.7835]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[92.78351]
[89.]
epoch: 118 loss: 14.314934730529785
tensor([[137.3691]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[137.36906]
[130.]
epoch: 119 loss: 54.30310821533203
tensor([[27.6976]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[27.69761]
[30.]
epoch: 120 loss: 5.300995826721191
tensor([[216.9988]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[216.99881]
[204.]
epoch: 121 loss: 168.96905517578125
tensor([[202.4978]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[202.49783]
[201.]

```

```

epoch: 122 loss: 2.243504524230957
tensor([[86.6629]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[86.66293]
[97.]
epoch: 123 loss: 106.85494995117188
tensor([[57.4537]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[57.4537]
[70.]
epoch: 124 loss: 157.4096221923828
tensor([[138.7835]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[138.78351]
[159.]
epoch: 125 loss: 408.70654296875
tensor([[83.5043]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[83.50431]
[94.]
epoch: 126 loss: 110.15949249267578
tensor([[5.6204]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[5.620444]
[6.]
epoch: 127 loss: 0.14406289160251617
tensor([[196.3348]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[196.33481]
[196.]
epoch: 128 loss: 0.1120966300368309
tensor([[93.4298]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[93.42976]
[90.]
epoch: 129 loss: 11.763279914855957
tensor([[164.9847]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[164.98471]
[154.]
epoch: 130 loss: 120.66387176513672
tensor([[116.7563]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[116.756256]
[109.]
epoch: 131 loss: 60.159507751464844
tensor([[123.4194]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')

```

```

[123.419365]
[119.]
epoch: 132 loss: 19.530786514282227
tensor([[66.2648]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[66.26484]
[64.]
epoch: 133 loss: 5.1294965744018555
tensor([[162.7235]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[162.72353]
[168.]
epoch: 134 loss: 27.841176986694336
tensor([[122.2959]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[122.29591]
[126.]
epoch: 135 loss: 13.720254898071289
tensor([[129.0421]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[129.04214]
[134.]
epoch: 136 loss: 24.5803279876709
tensor([[176.1468]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[176.14676]
[182.]
epoch: 137 loss: 34.26042938232422
tensor([[12.3604]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[12.360406]
[13.]
epoch: 138 loss: 0.4090805947780609
tensor([[10.2612]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[10.26119]
[7.]
epoch: 139 loss: 10.63536262512207
tensor([[131.4621]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[131.46205]
[124.]
epoch: 140 loss: 55.682212829589844
tensor([[81.6546]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[81.654594]
[76.]
epoch: 141 loss: 31.974437713623047

```

```

tensor([[155.6199]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[155.61989]
[149.]
epoch: 142 loss: 43.82292175292969
tensor([[105.8344]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[105.83445]
[103.]
epoch: 143 loss: 8.03410530090332
tensor([[142.1299]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[142.1299]
[143.]
epoch: 144 loss: 0.7570773959159851
tensor([[43.8302]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[43.830215]
[46.]
epoch: 145 loss: 4.707964897155762
tensor([[80.2875]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[80.287544]
[85.]
epoch: 146 loss: 22.207239151000977
tensor([[8.2846]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[8.284619]
[14.]
epoch: 147 loss: 32.66557693481445
tensor([[71.6659]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[71.665924]
[77.]
epoch: 148 loss: 28.45236587524414
tensor([[181.4285]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[181.42851]
[191.]
epoch: 149 loss: 91.61337280273438
tensor([[77.3633]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[77.36334]
[80.]
epoch: 150 loss: 6.951963901519775
tensor([[204.4945]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[204.49454]

```

```

[203.]
epoch: 151 loss: 2.2336418628692627
tensor([[162.7343]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[162.73427]
[157.]
epoch: 152 loss: 32.881832122802734
tensor([[114.5899]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[114.58985]
[112.]
epoch: 153 loss: 6.707330226898193
tensor([[120.4323]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[120.432335]
[118.]
epoch: 154 loss: 5.916253089904785
tensor([[55.8512]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[55.8512]
[52.]
epoch: 155 loss: 14.831742286682129
tensor([[76.9017]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[76.90172]
[75.]
epoch: 156 loss: 3.6165318489074707
tensor([[168.1578]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[168.15776]
[167.]
epoch: 157 loss: 1.3404096364974976
tensor([[191.1846]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[191.18463]
[190.]
epoch: 158 loss: 1.4033514261245728
tensor([[142.0871]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[142.0871]
[146.]
epoch: 159 loss: 15.310808181762695
tensor([[167.4487]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[167.44872]
[173.]
epoch: 160 loss: 30.816762924194336
tensor([[54.7350]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],

```



```

device='cuda:0')
[54.735035]
[54.]
epoch: 161 loss: 0.5402763485908508
tensor([[111.6883]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[111.68833]
[114.]
epoch: 162 loss: 5.343810558319092
tensor([[15.4361]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[15.436131]
[18.]
epoch: 163 loss: 6.573426723480225
tensor([[191.4850]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[191.48495]
[184.]
epoch: 164 loss: 56.0245475769043
tensor([[109.7212]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[109.72115]
[110.]
epoch: 165 loss: 0.07775550335645676
tensor([[93.2803]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[93.280266]
[93.]
epoch: 166 loss: 0.07854892313480377
tensor([[119.7369]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[119.736885]
[122.]
epoch: 167 loss: 5.121689319610596
tensor([[22.6579]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[22.657938]
[24.]
epoch: 168 loss: 1.8011304140090942
tensor([[132.2634]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[132.26344]
[135.]
epoch: 169 loss: 7.488744258880615
tensor([[58.7922]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[58.792156]
[58.]

```

```

epoch: 170 loss: 0.6275115013122559
tensor([[117.5061]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[117.50611]
[120.]
epoch: 171 loss: 6.219481468200684
tensor([[115.0209]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[115.0209]
[116.]
epoch: 172 loss: 0.958642840385437
tensor([[105.8715]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[105.8715]
[105.]
epoch: 173 loss: 0.7595089673995972
tensor([[163.7042]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[163.70422]
[160.]
epoch: 174 loss: 13.721272468566895
tensor([[198.7918]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[198.79184]
[195.]
epoch: 175 loss: 14.378047943115234
tensor([[193.7926]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[193.79263]
[193.]
epoch: 176 loss: 0.6282671689987183
tensor([[182.5109]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[182.51088]
[186.]
epoch: 177 loss: 12.173961639404297
tensor([[39.0845]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[39.08448]
[38.]
epoch: 178 loss: 1.1760975122451782
tensor([[196.7291]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[196.72914]
[205.]
epoch: 179 loss: 68.4071044921875
tensor([[43.7165]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')

```

```

[43.716496]
[43.]
epoch: 180 loss: 0.5133658051490784
tensor([[82.7972]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[82.79716]
[81.]
epoch: 181 loss: 3.229774236679077
tensor([[147.4176]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[147.4176]
[145.]
epoch: 182 loss: 5.844801902770996
tensor([[155.3828]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[155.3828]
[153.]
epoch: 183 loss: 5.677722930908203
tensor([[114.0545]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[114.0545]
[115.]
epoch: 184 loss: 0.8939763903617859
tensor([[142.3972]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[142.39722]
[144.]
epoch: 185 loss: 2.5689139366149902
tensor([[49.2193]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[49.21931]
[49.]
epoch: 186 loss: 0.04809720814228058
For epoch 10 the test accuracy over the whole test set is 0 %
tensor([[141.9129]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[141.91286]
[142.]
epoch: 0 loss: 0.007593892514705658
tensor([[79.9545]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[79.95446]
[80.]
epoch: 1 loss: 0.0020738784223794937
tensor([[14.3737]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[14.373739]
[11.]

```

```

epoch: 2 loss: 11.382116317749023
tensor([[98.3573]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[98.35732]
[98.]
epoch: 3 loss: 0.1276795119047165
tensor([[9.5196]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[9.51964]
[5.]
epoch: 4 loss: 20.42714500427246
tensor([[191.6734]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[191.67342]
[190.]
epoch: 5 loss: 2.800321578979492
tensor([[61.7392]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[61.739155]
[68.]
epoch: 6 loss: 39.19818115234375
tensor([[45.8557]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[45.855747]
[51.]
epoch: 7 loss: 26.463336944580078
tensor([[141.5174]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[141.51744]
[145.]
epoch: 8 loss: 12.128218650817871
tensor([[33.3373]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[33.33733]
[35.]
epoch: 9 loss: 2.764472007751465
tensor([[178.0624]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[178.06236]
[171.]
epoch: 10 loss: 49.87696838378906
tensor([[115.2286]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[115.22857]
[113.]
epoch: 11 loss: 4.966519832611084
tensor([[88.0840]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')

```

```

[88.08401]
[87.]
epoch: 12 loss: 1.1750717163085938
tensor([[85.9008]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[85.90076]
[86.]
epoch: 13 loss: 0.009849205613136292
tensor([[156.5323]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[156.5323]
[159.]
epoch: 14 loss: 6.089529037475586
tensor([[48.9402]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[48.94023]
[49.]
epoch: 15 loss: 0.003572294721379876
tensor([[60.5012]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[60.501225]
[60.]
epoch: 16 loss: 0.251226007938385
tensor([[8.7544]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[8.754446]
[10.]
epoch: 17 loss: 1.5514047145843506
tensor([[83.8394]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[83.839355]
[85.]
epoch: 18 loss: 1.3470957279205322
tensor([[129.4218]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[129.42178]
[133.]
epoch: 19 loss: 12.803633689880371
tensor([[48.3773]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[48.377346]
[47.]
epoch: 20 loss: 1.8970820903778076
tensor([[12.1196]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[12.119643]
[15.]
epoch: 21 loss: 8.296455383300781

```

```

tensor([[21.9628]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[21.962772]
[24.]
epoch: 22 loss: 4.150296211242676
tensor([[204.8703]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[204.87029]
[200.]
epoch: 23 loss: 23.719676971435547
tensor([[91.4131]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[91.413124]
[88.]
epoch: 24 loss: 11.649415969848633
tensor([[92.0974]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[92.09738]
[90.]
epoch: 25 loss: 4.399009704589844
tensor([[50.7181]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[50.718136]
[48.]
epoch: 26 loss: 7.3882622718811035
tensor([[19.0209]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[19.020872]
[22.]
epoch: 27 loss: 8.875203132629395
tensor([[38.2782]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[38.278225]
[38.]
epoch: 28 loss: 0.07740911841392517
tensor([[65.3765]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[65.37652]
[67.]
epoch: 29 loss: 2.635693073272705
tensor([[142.2726]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[142.27258]
[157.]
epoch: 30 loss: 216.8968048095703
tensor([[55.3398]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[55.33982]

```

```

[53.]
epoch: 31 loss: 5.474761486053467
tensor([[126.2779]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[126.27787]
[129.]
epoch: 32 loss: 7.4099907875061035
tensor([[48.1434]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[48.14337]
[44.]
epoch: 33 loss: 17.16752815246582
tensor([[38.2139]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[38.213905]
[32.]
epoch: 34 loss: 38.61262130737305
tensor([[148.9896]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[148.98962]
[140.]
epoch: 35 loss: 80.81333923339844
tensor([[157.0715]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[157.07152]
[151.]
epoch: 36 loss: 36.86333084106445
tensor([[99.3444]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[99.34436]
[96.]
epoch: 37 loss: 11.184745788574219
tensor([[83.4012]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[83.401245]
[84.]
epoch: 38 loss: 0.3585073947906494
tensor([[60.7102]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[60.710167]
[62.]
epoch: 39 loss: 1.6636693477630615
tensor([[10.2618]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[10.261786]
[4.]
epoch: 40 loss: 39.20996856689453
tensor([[92.3810]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],

```

```

device='cuda:0')
[92.38096]
[102.]
epoch: 41 loss: 92.52595520019531
tensor([[162.2550]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[162.25496]
[179.]
epoch: 42 loss: 280.3963928222656
tensor([[162.4437]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[162.4437]
[170.]
epoch: 43 loss: 57.09774398803711
tensor([[110.5993]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[110.59935]
[109.]
epoch: 44 loss: 2.557920455932617
tensor([[195.5543]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[195.55428]
[180.]
epoch: 45 loss: 241.93548583984375
tensor([[87.2503]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[87.25033]
[82.]
epoch: 46 loss: 27.56594467163086
tensor([[80.8629]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[80.86291]
[77.]
epoch: 47 loss: 14.922053337097168
tensor([[162.0106]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[162.0106]
[153.]
epoch: 48 loss: 81.19100189208984
tensor([[200.4853]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[200.48526]
[195.]
epoch: 49 loss: 30.088077545166016
tensor([[76.0522]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[76.05216]
[83.]

```



```

epoch: 50 loss: 48.2724494934082
tensor([[69.1798]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[69.179794]
[79.]
epoch: 51 loss: 96.43643951416016
tensor([[5.6034]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[5.6034465]
[6.]
epoch: 52 loss: 0.15725469589233398
tensor([[171.1937]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[171.19371]
[191.]
epoch: 53 loss: 392.28912353515625
tensor([[71.9559]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[71.9559]
[81.]
epoch: 54 loss: 81.79570770263672
tensor([[10.1814]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[10.181448]
[12.]
epoch: 55 loss: 3.30713152885437
tensor([[159.5280]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[159.52797]
[156.]
epoch: 56 loss: 12.44656753540039
tensor([[98.8928]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[98.89279]
[95.]
epoch: 57 loss: 15.153827667236328
tensor([[16.9105]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[16.91048]
[20.]
epoch: 58 loss: 9.545130729675293
tensor([[160.0315]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[160.03151]
[147.]
epoch: 59 loss: 169.8202362060547
tensor([[213.2322]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')

```

```

[213.23224]
[196.]
epoch: 60 loss: 296.9500427246094
tensor([[64.2157]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[64.21571]
[64.]
epoch: 61 loss: 0.04653231427073479
tensor([[42.6641]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[42.66408]
[46.]
epoch: 62 loss: 11.128352165222168
tensor([[12.3005]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[12.30047]
[14.]
epoch: 63 loss: 2.8884010314941406
tensor([[102.8662]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[102.86618]
[124.]
epoch: 64 loss: 446.6383361816406
tensor([[87.9472]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[87.94721]
[106.]
epoch: 65 loss: 325.90313720703125
tensor([[172.2144]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[172.2144]
[205.]
epoch: 66 loss: 1074.8955078125
tensor([[112.4722]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[112.47219]
[118.]
epoch: 67 loss: 30.556673049926758
tensor([[108.9073]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[108.907265]
[100.]
epoch: 68 loss: 79.33936309814453
tensor([[110.0770]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[110.07704]
[91.]
epoch: 69 loss: 363.9335021972656

```

```

tensor([[132.1931]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[132.19315]
[107.]
epoch: 70 loss: 634.694580078125
tensor([[49.2261]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[49.22614]
[39.]
epoch: 71 loss: 104.57392120361328
tensor([[199.8941]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[199.89409]
[173.]
epoch: 72 loss: 723.2919921875
tensor([[149.2406]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[149.24057]
[144.]
epoch: 73 loss: 27.46357536315918
tensor([[132.6661]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[132.66606]
[143.]
epoch: 74 loss: 106.790283203125
tensor([[12.2467]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[12.246683]
[18.]
epoch: 75 loss: 33.10065460205078
tensor([[41.5812]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[41.581184]
[57.]
epoch: 76 loss: 237.7398681640625
tensor([[113.2837]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[113.28374]
[149.]
epoch: 77 loss: 1275.6514892578125
tensor([[76.9369]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[76.93692]
[101.]
epoch: 78 loss: 579.0317993164062
tensor([[91.5723]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[91.57227]

```

```

[112.]
epoch: 79 loss: 417.2920227050781
tensor([[104.6301]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[104.63014]
[115.]
epoch: 80 loss: 107.53395080566406
tensor([[95.0579]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[95.057915]
[94.]
epoch: 81 loss: 1.1191835403442383
tensor([[211.3440]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[211.34398]
[189.]
epoch: 82 loss: 499.2533874511719
tensor([[188.3197]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[188.3197]
[172.]
epoch: 83 loss: 266.3326721191406
tensor([[37.2225]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[37.222458]
[30.]
epoch: 84 loss: 52.16389846801758
tensor([[140.1850]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[140.18497]
[137.]
epoch: 85 loss: 10.144015312194824
tensor([[99.4721]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[99.47211]
[99.]
epoch: 86 loss: 0.22288495302200317
tensor([[185.8744]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[185.87442]
[203.]
epoch: 87 loss: 293.2854919433594
tensor([[153.1707]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[153.17065]
[167.]
epoch: 88 loss: 191.2508087158203
tensor([[47.6744]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],

```

```

device='cuda:0')
[47.674408]
[43.]
epoch: 89 loss: 21.85009002685547
tensor([[23.4267]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[23.426714]
[13.]
epoch: 90 loss: 108.71636199951172
tensor([[150.8270]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[150.82697]
[146.]
epoch: 91 loss: 23.29959487915039
tensor([[29.0254]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[29.02536]
[23.]
epoch: 92 loss: 36.30496597290039
tensor([[170.3634]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[170.36343]
[162.]
epoch: 93 loss: 69.94702911376953
tensor([[18.3864]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[18.386408]
[7.]
epoch: 94 loss: 129.65028381347656
tensor([[174.3934]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[174.39339]
[175.]
epoch: 95 loss: 0.3679795265197754
tensor([[15.4898]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[15.489824]
[9.]
epoch: 96 loss: 42.117820739746094
tensor([[34.3027]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[34.302746]
[33.]
epoch: 97 loss: 1.6971466541290283
tensor([[91.4877]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[91.487724]
[92.]

```

```

epoch: 98 loss: 0.26242637634277344
tensor([[114.3280]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[114.32796]
[117.]
epoch: 99 loss: 7.13981294631958
tensor([[10.6177]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[10.617742]
[16.]
epoch: 100 loss: 28.968706130981445
tensor([[67.7201]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[67.72009]
[70.]
epoch: 101 loss: 5.197977066040039
tensor([[164.3343]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[164.33432]
[177.]
epoch: 102 loss: 160.41944885253906
tensor([[193.2866]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[193.28662]
[197.]
epoch: 103 loss: 13.789182662963867
tensor([[157.7720]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[157.77202]
[155.]
epoch: 104 loss: 7.684086322784424
tensor([[114.2613]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[114.261345]
[108.]
epoch: 105 loss: 39.20444107055664
tensor([[149.6463]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[149.64633]
[141.]
epoch: 106 loss: 74.7590560913086
tensor([[67.0644]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[67.06443]
[65.]
epoch: 107 loss: 4.261872291564941
tensor([[121.9124]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')

```

```

[121.912384]
[116.]
epoch: 108 loss: 34.95628356933594
tensor([[112.3724]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[112.372406]
[110.]
epoch: 109 loss: 5.628310203552246
tensor([[86.1839]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[86.18394]
[89.]
epoch: 110 loss: 7.930210590362549
tensor([[121.8500]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[121.85003]
[123.]
epoch: 111 loss: 1.322433352470398
tensor([[169.8629]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[169.86292]
[176.]
epoch: 112 loss: 37.66381072998047
tensor([[202.2498]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[202.24985]
[201.]
epoch: 113 loss: 1.5621185302734375
tensor([[50.1510]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[50.15095]
[56.]
epoch: 114 loss: 34.211368560791016
tensor([[136.5123]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[136.5123]
[139.]
epoch: 115 loss: 6.188658237457275
tensor([[201.9963]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[201.9963]
[198.]
epoch: 116 loss: 15.97047233581543
tensor([[94.2544]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[94.25442]
[97.]
epoch: 117 loss: 7.538223743438721

```

```

tensor([[66.5590]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[66.55896]
[71.]
epoch: 118 loss: 19.722837448120117
tensor([[209.4588]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[209.45883]
[204.]
epoch: 119 loss: 29.798845291137695
tensor([[150.3462]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[150.34622]
[152.]
epoch: 120 loss: 2.7349820137023926
tensor([[136.0272]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[136.02718]
[138.]
epoch: 121 loss: 3.8920350074768066
tensor([[152.3563]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[152.35628]
[154.]
epoch: 122 loss: 2.7018237113952637
tensor([[174.7359]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[174.73593]
[178.]
epoch: 123 loss: 10.654144287109375
tensor([[92.1489]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[92.14889]
[93.]
epoch: 124 loss: 0.7243922352790833
tensor([[124.4753]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[124.4753]
[122.]
epoch: 125 loss: 6.1271281242370605
tensor([[7.8572]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[7.8571763]
[21.]
epoch: 126 loss: 172.73382568359375
tensor([[76.5589]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[76.558945]

```



```

[75.]
epoch: 127 loss: 2.4303085803985596
tensor([[17.7094]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[17.70938]
[25.]
epoch: 128 loss: 53.15315246582031
tensor([[203.4547]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[203.45474]
[194.]
epoch: 129 loss: 89.39215087890625
tensor([[52.0348]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[52.03483]
[50.]
epoch: 130 loss: 4.140525817871094
tensor([[192.5170]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[192.517]
[192.]
epoch: 131 loss: 0.2672872245311737
tensor([[163.6471]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[163.64706]
[169.]
epoch: 132 loss: 28.653921127319336
tensor([[7.4973]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[7.4972997]
[0.]
epoch: 133 loss: 56.209503173828125
tensor([[59.7169]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[59.71688]
[58.]
epoch: 134 loss: 2.9476797580718994
tensor([[103.0864]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[103.086365]
[104.]
epoch: 135 loss: 0.8347293734550476
tensor([[127.6097]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[127.60971]
[132.]
epoch: 136 loss: 19.274641036987305
tensor([[177.4319]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],

```

```

device='cuda:0')
[177.4319]
[186.]
epoch: 137 loss: 73.41233825683594
tensor([[8.5874]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[8.587418]
[1.]
epoch: 138 loss: 57.568904876708984
tensor([[138.7264]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[138.72636]
[136.]
epoch: 139 loss: 7.433061599731445
tensor([[203.1114]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[203.11139]
[199.]
epoch: 140 loss: 16.903520584106445
tensor([[36.0118]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[36.011776]
[36.]
epoch: 141 loss: 0.00013867348025087267
tensor([[74.2587]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[74.25874]
[69.]
epoch: 142 loss: 27.654380798339844
tensor([[137.5589]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[137.55891]
[134.]
epoch: 143 loss: 12.66586971282959
tensor([[35.5668]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[35.566795]
[37.]
epoch: 144 loss: 2.0540754795074463
tensor([[161.7088]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[161.7088]
[163.]
epoch: 145 loss: 1.667194128036499
tensor([[129.2314]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[129.23145]
[130.]

```

```

epoch: 146 loss: 0.5906763076782227
tensor([[73.7619]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[73.761925]
[74.]
epoch: 147 loss: 0.056679826229810715
tensor([[22.0507]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[22.050701]
[26.]
epoch: 148 loss: 15.59696102142334
tensor([[6.2864]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[6.2863946]
[3.]
epoch: 149 loss: 10.800389289855957
tensor([[11.2998]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[11.299758]
[19.]
epoch: 150 loss: 59.29372787475586
tensor([[75.5432]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[75.54324]
[76.]
epoch: 151 loss: 0.20862658321857452
tensor([[117.7719]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[117.771904]
[119.]
epoch: 152 loss: 1.508219838142395
tensor([[110.5583]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[110.55835]
[111.]
epoch: 153 loss: 0.1950550675392151
tensor([[179.0343]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[179.0343]
[183.]
epoch: 154 loss: 15.726762771606445
tensor([[166.3810]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[166.38097]
[168.]
epoch: 155 loss: 2.6212704181671143
tensor([[107.4278]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')

```

```

[107.42777]
[103.]
epoch: 156 loss: 19.60516929626465
tensor([[33.9950]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[33.994995]
[34.]
epoch: 157 loss: 2.504885196685791e-05
tensor([[119.2769]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[119.27688]
[114.]
epoch: 158 loss: 27.84544563293457
tensor([[151.1823]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[151.18225]
[148.]
epoch: 159 loss: 10.126721382141113
tensor([[183.7860]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[183.786]
[182.]
epoch: 160 loss: 3.189779758453369
tensor([[57.2213]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[57.22127]
[55.]
epoch: 161 loss: 4.934047222137451
tensor([[162.7931]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[162.79309]
[166.]
epoch: 162 loss: 10.284266471862793
tensor([[124.1991]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[124.19907]
[125.]
epoch: 163 loss: 0.6414827704429626
tensor([[124.5657]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[124.565735]
[126.]
epoch: 164 loss: 2.0571165084838867
tensor([[53.8357]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[53.83573]
[54.]
epoch: 165 loss: 0.026984138414263725

```

```

tensor([[157.6578]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[157.65776]
[160.]
epoch: 166 loss: 5.486085414886475
tensor([[192.5863]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[192.58627]
[193.]
epoch: 167 loss: 0.17116987705230713
tensor([[25.8356]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[25.835613]
[27.]
epoch: 168 loss: 1.355796456336975
tensor([[15.5707]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[15.570657]
[17.]
epoch: 169 loss: 2.0430221557617188
tensor([[14.5306]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[14.530585]
[8.]
epoch: 170 loss: 42.64854431152344
tensor([[11.4264]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[11.426425]
[2.]
epoch: 171 loss: 88.85749053955078
tensor([[83.8278]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[83.82776]
[78.]
epoch: 172 loss: 33.962772369384766
tensor([[109.2591]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[109.259056]
[105.]
epoch: 173 loss: 18.139558792114258
tensor([[119.6943]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[119.69428]
[120.]
epoch: 174 loss: 0.09346316754817963
tensor([[22.9686]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[22.968578]

```

```

[29.]
epoch: 175 loss: 36.378047943115234
tensor([[158.7949]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[158.79492]
[165.]
epoch: 176 loss: 38.502994537353516
tensor([[46.4340]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[46.434036]
[52.]
epoch: 177 loss: 30.979951858520508
tensor([[181.6285]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[181.62851]
[184.]
epoch: 178 loss: 5.623967170715332
tensor([[133.7028]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[133.70279]
[135.]
epoch: 179 loss: 1.6827555894851685
tensor([[39.3211]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[39.321056]
[45.]
epoch: 180 loss: 32.25040054321289
tensor([[170.5782]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[170.57816]
[164.]
epoch: 181 loss: 43.27212905883789
tensor([[192.8268]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[192.82678]
[185.]
epoch: 182 loss: 61.25851821899414
tensor([[151.3045]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[151.3045]
[150.]
epoch: 183 loss: 1.7017316818237305
tensor([[170.5188]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[170.51884]
[174.]
epoch: 184 loss: 12.11844253540039
tensor([[57.9851]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],

```

```

device='cuda:0')
[57.985115]
[61.]
epoch: 185 loss: 9.089530944824219
tensor([[21.9062]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[21.906229]
[28.]
epoch: 186 loss: 37.1340446472168
For epoch 11 the test accuracy over the whole test set is 0 %
tensor([[94.4290]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[94.42897]
[98.]
epoch: 0 loss: 12.752252578735352
tensor([[101.1464]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[101.146385]
[104.]
epoch: 1 loss: 8.143117904663086
tensor([[167.5199]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[167.51991]
[174.]
epoch: 2 loss: 41.99153137207031
tensor([[99.9732]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[99.97318]
[99.]
epoch: 3 loss: 0.9470845460891724
tensor([[159.1293]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[159.12933]
[156.]
epoch: 4 loss: 9.792728424072266
tensor([[73.1436]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[73.1436]
[69.]
epoch: 5 loss: 17.169424057006836
tensor([[66.7727]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[66.7727]
[62.]
epoch: 6 loss: 22.778640747070312
tensor([[29.8020]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[29.801966]

```

```

[30.]
epoch: 7 loss: 0.03921758010983467
tensor([[88.1985]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[88.198524]
[85.]
epoch: 8 loss: 10.230558395385742
tensor([[157.5719]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[157.57195]
[154.]
epoch: 9 loss: 12.758792877197266
tensor([[117.8932]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[117.89323]
[116.]
epoch: 10 loss: 3.5843069553375244
tensor([[139.2193]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[139.21925]
[141.]
epoch: 11 loss: 3.171057939529419
tensor([[111.0967]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[111.09667]
[113.]
epoch: 12 loss: 3.622657299041748
tensor([[79.8792]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[79.879166]
[82.]
epoch: 13 loss: 4.49793815612793
tensor([[180.2884]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[180.28839]
[186.]
epoch: 14 loss: 32.622474670410156
tensor([[86.0344]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[86.03443]
[88.]
epoch: 15 loss: 3.863459587097168
tensor([[10.3316]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[10.331553]
[16.]
epoch: 16 loss: 32.13128662109375
tensor([[84.4137]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],

```



```

device='cuda:0')
[84.41369]
[83.]
epoch: 17 loss: 1.9985156059265137
tensor([[111.1485]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[111.14849]
[108.]
epoch: 18 loss: 9.912995338439941
tensor([[51.0642]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[51.064163]
[49.]
epoch: 19 loss: 4.260769844055176
tensor([[26.1434]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[26.143381]
[27.]
epoch: 20 loss: 0.7337958812713623
tensor([[88.8722]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[88.87224]
[86.]
epoch: 21 loss: 8.249752044677734
tensor([[113.8217]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[113.82171]
[110.]
epoch: 22 loss: 14.605457305908203
tensor([[10.5369]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[10.536896]
[10.]
epoch: 23 loss: 0.28825706243515015
tensor([[149.2894]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[149.28937]
[148.]
epoch: 24 loss: 1.6624690294265747
tensor([[52.6758]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[52.675777]
[52.]
epoch: 25 loss: 0.45667514204978943
tensor([[6.1965]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[6.1964684]
[2.]

```

```

epoch: 26 loss: 17.6103458404541
tensor([[72.5960]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[72.59597]
[75.]
epoch: 27 loss: 5.779359340667725
tensor([[58.5348]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[58.53476]
[60.]
epoch: 28 loss: 2.1469297409057617
tensor([[11.9394]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[11.93941]
[19.]
epoch: 29 loss: 49.8519287109375
tensor([[44.9471]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[44.947124]
[48.]
epoch: 30 loss: 9.320049285888672
tensor([[103.7200]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[103.71999]
[107.]
epoch: 31 loss: 10.758441925048828
tensor([[108.5383]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[108.53831]
[111.]
epoch: 32 loss: 6.05993127822876
tensor([[8.8850]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[8.885024]
[6.]
epoch: 33 loss: 8.3233642578125
tensor([[21.1799]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[21.179895]
[24.]
epoch: 34 loss: 7.9529900550842285
tensor([[175.9913]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[175.99133]
[175.]
epoch: 35 loss: 0.982741117477417
tensor([[139.3956]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')

```

```

[139.39555]
[137.]
epoch: 36 loss: 5.738677024841309
tensor([[99.2357]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[99.23572]
[97.]
epoch: 37 loss: 4.998434066772461
tensor([[103.3461]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[103.34613]
[101.]
epoch: 38 loss: 5.504327774047852
tensor([[34.8849]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[34.88493]
[35.]
epoch: 39 loss: 0.013241183944046497
tensor([[68.4766]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[68.4766]
[67.]
epoch: 40 loss: 2.180349588394165
tensor([[65.0047]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[65.00473]
[64.]
epoch: 41 loss: 1.0094828605651855
tensor([[193.7462]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[193.74619]
[197.]
epoch: 42 loss: 10.587309837341309
tensor([[85.4765]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[85.47648]
[87.]
epoch: 43 loss: 2.321117639541626
tensor([[137.1371]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[137.1371]
[139.]
epoch: 44 loss: 3.470395565032959
tensor([[28.0313]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[28.03128]
[28.]
epoch: 45 loss: 0.0009784707799553871

```

```

tensor([[170.9241]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[170.92413]
[172.]
epoch: 46 loss: 1.1574891805648804
tensor([[14.4055]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[14.405524]
[14.]
epoch: 47 loss: 0.16444991528987885
tensor([[199.6280]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[199.62802]
[195.]
epoch: 48 loss: 21.418581008911133
tensor([[60.3111]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[60.311134]
[57.]
epoch: 49 loss: 10.963610649108887
tensor([[127.5388]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[127.53877]
[126.]
epoch: 50 loss: 2.367820978164673
tensor([[28.7549]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[28.754927]
[29.]
epoch: 51 loss: 0.060060933232307434
tensor([[87.8161]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[87.81612]
[90.]
epoch: 52 loss: 4.769347667694092
tensor([[54.5870]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[54.586994]
[54.]
epoch: 53 loss: 0.3445621430873871
tensor([[178.1650]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[178.16496]
[183.]
epoch: 54 loss: 23.37758445739746
tensor([[201.3947]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[201.39468]

```

```

[205.]
epoch: 55 loss: 12.99830436706543
tensor([[25.3685]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[25.36849]
[26.]
epoch: 56 loss: 0.39880460500717163
tensor([[131.2081]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[131.20808]
[130.]
epoch: 57 loss: 1.4594671726226807
tensor([[173.2239]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[173.22386]
[169.]
epoch: 58 loss: 17.841007232666016
tensor([[153.4112]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[153.41116]
[151.]
epoch: 59 loss: 5.813708782196045
tensor([[138.4399]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[138.43994]
[136.]
epoch: 60 loss: 5.953313827514648
tensor([[193.0864]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[193.08644]
[190.]
epoch: 61 loss: 9.526118278503418
tensor([[118.0650]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[118.06499]
[120.]
epoch: 62 loss: 3.744274616241455
tensor([[96.1809]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[96.18089]
[100.]
epoch: 63 loss: 14.585578918457031
tensor([[195.7547]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[195.75475]
[200.]
epoch: 64 loss: 18.022186279296875
tensor([[10.0813]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],

```

```

device='cuda:0')
[10.081304]
[8.]
epoch: 65 loss: 4.331824779510498
tensor([[66.9283]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[66.92829]
[68.]
epoch: 66 loss: 1.1485594511032104
tensor([[115.2517]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[115.251686]
[117.]
epoch: 67 loss: 3.0566015243530273
tensor([[13.5700]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[13.569967]
[15.]
epoch: 68 loss: 2.0449936389923096
tensor([[134.4156]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[134.41556]
[134.]
epoch: 69 loss: 0.1726883351802826
tensor([[147.9449]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[147.94495]
[147.]
epoch: 70 loss: 0.8929234743118286
tensor([[113.6097]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[113.60971]
[112.]
epoch: 71 loss: 2.5911684036254883
tensor([[146.9488]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[146.94878]
[145.]
epoch: 72 loss: 3.7977287769317627
tensor([[165.9995]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[165.99947]
[163.]
epoch: 73 loss: 8.996795654296875
tensor([[73.8670]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[73.867035]
[74.]

```

```

epoch: 74 loss: 0.017679713666439056
tensor([[105.3559]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[105.3559]
[106.]
epoch: 75 loss: 0.4148601293563843
tensor([[9.3775]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[9.377499]
[5.]
epoch: 76 loss: 19.162494659423828
tensor([[75.2435]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[75.243515]
[77.]
epoch: 77 loss: 3.0852394104003906
tensor([[170.5854]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[170.58545]
[176.]
epoch: 78 loss: 29.317359924316406
tensor([[93.9782]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[93.978195]
[96.]
epoch: 79 loss: 4.0876946449279785
tensor([[42.0728]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[42.072803]
[44.]
epoch: 80 loss: 3.7140862941741943
tensor([[91.2005]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[91.20054]
[91.]
epoch: 81 loss: 0.04021574556827545
tensor([[116.9638]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[116.96376]
[115.]
epoch: 82 loss: 3.8563547134399414
tensor([[175.0219]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[175.02185]
[173.]
epoch: 83 loss: 4.087879657745361
tensor([[93.7991]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')

```

```

[93.79915]
[92.]
epoch: 84 loss: 3.236935615539551
tensor([[42.0078]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[42.007774]
[43.]
epoch: 85 loss: 0.9845117330551147
tensor([[119.5828]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[119.582825]
[118.]
epoch: 86 loss: 2.5053341388702393
tensor([[181.5381]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[181.53806]
[179.]
epoch: 87 loss: 6.441725254058838
tensor([[164.8990]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[164.89899]
[165.]
epoch: 88 loss: 0.010203663259744644
tensor([[146.5560]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[146.556]
[150.]
epoch: 89 loss: 11.861137390136719
tensor([[193.7982]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[193.79816]
[194.]
epoch: 90 loss: 0.04074070230126381
tensor([[200.0128]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[200.01282]
[199.]
epoch: 91 loss: 1.025799036026001
tensor([[77.0373]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[77.03726]
[78.]
epoch: 92 loss: 0.926864504814148
tensor([[140.0033]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[140.00327]
[144.]
epoch: 93 loss: 15.97388744354248

```



```

tensor([[6.8587]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[6.858668]
[3.]
epoch: 94 loss: 14.889317512512207
tensor([[153.7764]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[153.77635]
[155.]
epoch: 95 loss: 1.4973145723342896
tensor([[193.6036]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[193.60359]
[191.]
epoch: 96 loss: 6.778690814971924
tensor([[31.7595]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[31.759483]
[32.]
epoch: 97 loss: 0.057848263531923294
tensor([[51.4969]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[51.49691]
[50.]
epoch: 98 loss: 2.2407398223876953
tensor([[194.9022]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[194.90218]
[193.]
epoch: 99 loss: 3.6182732582092285
tensor([[139.8498]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[139.8498]
[142.]
epoch: 100 loss: 4.623392581939697
tensor([[178.0155]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[178.01553]
[180.]
epoch: 101 loss: 3.938107490539551
tensor([[5.2603]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[5.2602696]
[4.]
epoch: 102 loss: 1.5882796049118042
tensor([[37.1878]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[37.187798]

```

```

[39.]
epoch: 103 loss: 3.2840776443481445
tensor([[46.7143]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[46.714264]
[47.]
epoch: 104 loss: 0.08164510875940323
tensor([[36.8539]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[36.85388]
[38.]
epoch: 105 loss: 1.3135868310928345
tensor([[104.7022]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[104.70219]
[103.]
epoch: 106 loss: 2.8974392414093018
tensor([[77.5327]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[77.53269]
[76.]
epoch: 107 loss: 2.349144697189331
tensor([[60.2684]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[60.268364]
[58.]
epoch: 108 loss: 5.145474910736084
tensor([[133.4106]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[133.41061]
[133.]
epoch: 109 loss: 0.16860386729240417
tensor([[79.5479]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[79.54785]
[79.]
epoch: 110 loss: 0.3001413345336914
tensor([[83.4871]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[83.48705]
[84.]
epoch: 111 loss: 0.26311472058296204
tensor([[15.3996]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[15.399596]
[23.]
epoch: 112 loss: 57.766136169433594
tensor([[125.0311]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],

```

```

device='cuda:0')
[125.03108]
[125.]
epoch: 113 loss: 0.0009661002550274134
tensor([[138.9081]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[138.90813]
[140.]
epoch: 114 loss: 1.1921870708465576
tensor([[162.2218]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[162.2218]
[162.]
epoch: 115 loss: 0.049196019768714905
tensor([[206.7757]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[206.77568]
[204.]
epoch: 116 loss: 7.704402446746826
tensor([[13.1653]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[13.165302]
[20.]
epoch: 117 loss: 46.71309280395508
tensor([[57.9167]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[57.916725]
[56.]
epoch: 118 loss: 3.673835277557373
tensor([[112.0505]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[112.05054]
[114.]
epoch: 119 loss: 3.800405502319336
tensor([[165.9308]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[165.9308]
[170.]
epoch: 120 loss: 16.558378219604492
tensor([[95.0460]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[95.04602]
[95.]
epoch: 121 loss: 0.0021178871393203735
tensor([[119.5094]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[119.50936]
[119.]

```

```

epoch: 122 loss: 0.2594488859176636
tensor([[178.2422]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[178.24222]
[178.]
epoch: 123 loss: 0.0586695671081543
tensor([[48.4882]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[48.48819]
[46.]
epoch: 124 loss: 6.1910881996154785
tensor([[95.8259]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[95.82585]
[93.]
epoch: 125 loss: 7.98543643951416
tensor([[107.6227]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[107.622665]
[105.]
epoch: 126 loss: 6.878373622894287
tensor([[38.4947]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[38.49466]
[37.]
epoch: 127 loss: 2.234006881713867
tensor([[137.9541]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[137.95413]
[138.]
epoch: 128 loss: 0.0021038660779595375
tensor([[156.6892]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[156.68916]
[157.]
epoch: 129 loss: 0.09661950916051865
tensor([[202.3867]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[202.3867]
[203.]
epoch: 130 loss: 0.376132607460022
tensor([[52.8275]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[52.827496]
[51.]
epoch: 131 loss: 3.339740037918091
tensor([[71.0079]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')

```

```

[71.00793]
[71.]
epoch: 132 loss: 6.283639231696725e-05
tensor([[182.0889]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[182.0889]
[185.]
epoch: 133 loss: 8.474516868591309
tensor([[33.1030]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[33.10299]
[33.]
epoch: 134 loss: 0.010606775060296059
tensor([[183.9421]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[183.94214]
[184.]
epoch: 135 loss: 0.003347933292388916
tensor([[107.8918]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[107.89184]
[109.]
epoch: 136 loss: 1.228022813796997
tensor([[54.9932]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[54.993217]
[53.]
epoch: 137 loss: 3.9729158878326416
tensor([[123.1897]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[123.189735]
[122.]
epoch: 138 loss: 1.4154703617095947
tensor([[93.7838]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[93.78377]
[94.]
epoch: 139 loss: 0.04675640910863876
tensor([[24.5788]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[24.578833]
[25.]
epoch: 140 loss: 0.1773819625377655
tensor([[102.3427]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[102.34269]
[102.]
epoch: 141 loss: 0.11743610352277756

```

```

tensor([[15.3673]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[15.367346]
[12.]
epoch: 142 loss: 11.339017868041992
tensor([[132.2757]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[132.27571]
[132.]
epoch: 143 loss: 0.07601658999919891
tensor([[135.0603]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[135.06027]
[135.]
epoch: 144 loss: 0.003632740117609501
tensor([[60.2909]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[60.290855]
[61.]
epoch: 145 loss: 0.5028860569000244
tensor([[42.4691]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[42.469063]
[45.]
epoch: 146 loss: 6.405642986297607
tensor([[148.9043]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[148.90425]
[149.]
epoch: 147 loss: 0.009167851880192757
tensor([[153.9406]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[153.94064]
[153.]
epoch: 148 loss: 0.8848098516464233
tensor([[19.3290]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[19.328999]
[22.]
epoch: 149 loss: 7.134248733520508
tensor([[9.5016]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[9.501624]
[0.]
epoch: 150 loss: 90.2808609008789
tensor([[123.2825]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[123.282455]

```

```

[123.]
epoch: 151 loss: 0.07978107780218124
tensor([[53.9545]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[53.954506]
[55.]
epoch: 152 loss: 1.0930578708648682
tensor([[176.7324]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[176.73238]
[177.]
epoch: 153 loss: 0.07162255048751831
tensor([[183.2669]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[183.2669]
[182.]
epoch: 154 loss: 1.6050527095794678
tensor([[163.4422]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[163.4422]
[164.]
epoch: 155 loss: 0.31114116311073303
tensor([[169.3237]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[169.32373]
[171.]
epoch: 156 loss: 2.8098795413970947
tensor([[12.5124]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[12.51235]
[17.]
epoch: 157 loss: 20.139001846313477
tensor([[194.0308]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[194.03082]
[192.]
epoch: 158 loss: 4.124240875244141
tensor([[85.3419]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[85.34189]
[89.]
epoch: 159 loss: 13.381780624389648
tensor([[9.7693]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[9.769304]
[9.]
epoch: 160 loss: 0.5918290615081787
tensor([[32.8366]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],

```

```

device='cuda:0')
[32.836594]
[36.]
epoch: 161 loss: 10.007140159606934
tensor([[142.5561]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[142.55606]
[143.]
epoch: 162 loss: 0.1970820277929306
tensor([[16.1062]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[16.10616]
[21.]
epoch: 163 loss: 23.949678421020508
tensor([[162.8761]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[162.87605]
[160.]
epoch: 164 loss: 8.271679878234863
tensor([[13.3256]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[13.325641]
[13.]
epoch: 165 loss: 0.10604184865951538
tensor([[34.0292]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[34.029243]
[34.]
epoch: 166 loss: 0.0008551804930903018
tensor([[191.5492]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[191.54922]
[189.]
epoch: 167 loss: 6.498547554016113
tensor([[150.4603]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[150.46033]
[152.]
epoch: 168 loss: 2.3705925941467285
tensor([[202.0970]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[202.09697]
[201.]
epoch: 169 loss: 1.203342318534851
tensor([[156.5871]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[156.58713]
[159.]

```



```

epoch: 170 loss: 5.821952819824219
tensor([[11.5155]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[11.515498]
[7.]
epoch: 171 loss: 20.38972282409668
tensor([[65.1556]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[65.15565]
[65.]
epoch: 172 loss: 0.024226075038313866
tensor([[141.6421]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[141.6421]
[146.]
epoch: 173 loss: 18.991247177124023
tensor([[17.1515]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[17.151543]
[18.]
epoch: 174 loss: 0.7198798656463623
tensor([[165.6930]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[165.69304]
[167.]
epoch: 175 loss: 1.7081471681594849
tensor([[82.6854]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[82.68545]
[81.]
epoch: 176 loss: 2.8407340049743652
tensor([[169.7631]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[169.7631]
[168.]
epoch: 177 loss: 3.1085474491119385
tensor([[201.7140]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[201.71404]
[198.]
epoch: 178 loss: 13.794055938720703
tensor([[81.5067]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[81.50672]
[80.]
epoch: 179 loss: 2.270209550857544
tensor([[193.9787]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')

```

```

[193.97865]
[196.]
epoch: 180 loss: 4.085844039916992
tensor([[13.6172]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[13.61717]
[11.]
epoch: 181 loss: 6.849580764770508
tensor([[123.0048]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[123.00475]
[124.]
epoch: 182 loss: 0.9905163645744324
tensor([[70.2339]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[70.233894]
[70.]
epoch: 183 loss: 0.05470656603574753
tensor([[125.9704]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[125.97042]
[129.]
epoch: 184 loss: 9.178349494934082
tensor([[7.1582]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[7.158168]
[1.]
epoch: 185 loss: 37.923030853271484
tensor([[162.9796]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[162.97964]
[166.]
epoch: 186 loss: 9.12254524230957
For epoch 12 the test accuracy over the whole test set is 0 %
tensor([[138.1416]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[138.1416]
[139.]
epoch: 0 loss: 0.7368478775024414
tensor([[164.7679]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[164.7679]
[164.]
epoch: 1 loss: 0.5896682143211365
tensor([[11.8365]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[11.836502]
[15.]

```

```

epoch: 2 loss: 10.007719039916992
tensor([[151.4933]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[151.49329]
[149.]
epoch: 3 loss: 6.216475963592529
tensor([[182.5836]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[182.58362]
[179.]
epoch: 4 loss: 12.84231948852539
tensor([[53.9881]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[53.988083]
[52.]
epoch: 5 loss: 3.9524736404418945
tensor([[91.7676]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[91.76756]
[92.]
epoch: 6 loss: 0.054027020931243896
tensor([[145.6529]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[145.65291]
[148.]
epoch: 7 loss: 5.508839130401611
tensor([[141.9371]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[141.93709]
[144.]
epoch: 8 loss: 4.255605697631836
tensor([[76.3739]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[76.37388]
[77.]
epoch: 9 loss: 0.39202815294265747
tensor([[167.4360]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[167.436]
[169.]
epoch: 10 loss: 2.4460813999176025
tensor([[84.7596]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[84.75955]
[85.]
epoch: 11 loss: 0.05781523883342743
tensor([[105.0260]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')

```

```

[105.026024]
[104.]
epoch: 12 loss: 1.0527249574661255
tensor([[62.6394]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[62.639374]
[61.]
epoch: 13 loss: 2.687546491622925
tensor([[56.8993]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[56.899303]
[55.]
epoch: 14 loss: 3.607353448867798
tensor([[131.7688]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[131.76877]
[130.]
epoch: 15 loss: 3.1285412311553955
tensor([[181.1002]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[181.10023]
[180.]
epoch: 16 loss: 1.2105170488357544
tensor([[16.7282]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[16.72822]
[23.]
epoch: 17 loss: 39.33522415161133
tensor([[9.3075]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[9.307455]
[14.]
epoch: 18 loss: 22.019977569580078
tensor([[88.8858]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[88.88583]
[90.]
epoch: 19 loss: 1.2413665056228638
tensor([[194.2794]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[194.27939]
[195.]
epoch: 20 loss: 0.5192810297012329
tensor([[70.3381]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[70.33807]
[70.]
epoch: 21 loss: 0.11429384350776672

```

```

tensor([[37.0479]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[37.04789]
[38.]
epoch: 22 loss: 0.9065139889717102
tensor([[176.6745]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[176.67455]
[177.]
epoch: 23 loss: 0.10592076927423477
tensor([[190.6156]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[190.6156]
[189.]
epoch: 24 loss: 2.6101653575897217
tensor([[59.8977]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[59.89774]
[58.]
epoch: 25 loss: 3.601414918899536
tensor([[168.3087]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[168.30875]
[168.]
epoch: 26 loss: 0.09532430022954941
tensor([[115.3653]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[115.36535]
[115.]
epoch: 27 loss: 0.1334797590970993
tensor([[68.8190]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[68.81902]
[68.]
epoch: 28 loss: 0.6707988977432251
tensor([[201.9770]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[201.97699]
[205.]
epoch: 29 loss: 9.138590812683105
tensor([[26.5812]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[26.581203]
[26.]
epoch: 30 loss: 0.33779746294021606
tensor([[159.2867]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[159.2867]

```

```

[159.]
epoch: 31 loss: 0.08219539374113083
tensor([[146.8350]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[146.83499]
[147.]
epoch: 32 loss: 0.027227818965911865
tensor([[84.3897]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[84.38974]
[84.]
epoch: 33 loss: 0.15189726650714874
tensor([[193.5452]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[193.5452]
[192.]
epoch: 34 loss: 2.387632369995117
tensor([[43.5375]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[43.537514]
[44.]
epoch: 35 loss: 0.21389354765415192
tensor([[23.8652]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[23.865244]
[24.]
epoch: 36 loss: 0.018159203231334686
tensor([[167.0232]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[167.0232]
[167.]
epoch: 37 loss: 0.0005379319190979004
tensor([[165.9001]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[165.90015]
[166.]
epoch: 38 loss: 0.00997072458267212
tensor([[157.5982]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[157.59822]
[157.]
epoch: 39 loss: 0.35786816477775574
tensor([[175.8198]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[175.81984]
[178.]
epoch: 40 loss: 4.7530999183654785
tensor([[128.9259]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],

```

```

device='cuda:0')
[128.9259]
[129.]
epoch: 41 loss: 0.005490317940711975
tensor([[141.2349]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[141.23491]
[141.]
epoch: 42 loss: 0.05518226698040962
tensor([[111.3088]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[111.3088]
[111.]
epoch: 43 loss: 0.09535728394985199
tensor([[61.6521]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[61.652084]
[60.]
epoch: 44 loss: 2.7293827533721924
tensor([[205.3358]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[205.33578]
[203.]
epoch: 45 loss: 5.455891132354736
tensor([[204.9803]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[204.98026]
[204.]
epoch: 46 loss: 0.9609001278877258
tensor([[193.5096]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[193.50961]
[196.]
epoch: 47 loss: 6.202027320861816
tensor([[168.3733]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[168.37326]
[172.]
epoch: 48 loss: 13.153239250183105
tensor([[10.4733]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[10.473326]
[2.]
epoch: 49 loss: 71.79724884033203
tensor([[133.9937]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[133.99374]
[134.]

```

```

epoch: 50 loss: 3.9138831198215485e-05
tensor([[10.9330]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[10.933023]
[5.]
epoch: 51 loss: 35.200767517089844
tensor([[7.9511]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[7.9511065]
[0.]
epoch: 52 loss: 63.220096588134766
tensor([[44.6079]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[44.60787]
[47.]
epoch: 53 loss: 5.722294807434082
tensor([[54.5150]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[54.515026]
[56.]
epoch: 54 loss: 2.2051475048065186
tensor([[101.5794]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[101.57943]
[102.]
epoch: 55 loss: 0.17687943577766418
tensor([[30.7602]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[30.760246]
[34.]
epoch: 56 loss: 10.496004104614258
tensor([[41.1712]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[41.17125]
[45.]
epoch: 57 loss: 14.659331321716309
tensor([[52.7091]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[52.70906]
[54.]
epoch: 58 loss: 1.6665244102478027
tensor([[52.2791]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[52.279137]
[53.]
epoch: 59 loss: 0.5196439623832703
tensor([[183.5142]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')

```



```

[183.51424]
[176.]
epoch: 60 loss: 56.463748931884766
tensor([[99.0667]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[99.0667]
[98.]
epoch: 61 loss: 1.1378569602966309
tensor([[100.6894]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[100.68938]
[101.]
epoch: 62 loss: 0.09648675471544266
tensor([[190.6168]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[190.61679]
[191.]
epoch: 63 loss: 0.14684931933879852
tensor([[146.5518]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[146.55183]
[150.]
epoch: 64 loss: 11.889847755432129
tensor([[180.0480]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[180.04803]
[183.]
epoch: 65 loss: 8.714098930358887
tensor([[70.5329]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[70.53294]
[74.]
epoch: 66 loss: 12.020479202270508
tensor([[73.0090]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[73.00901]
[75.]
epoch: 67 loss: 3.9640400409698486
tensor([[77.6444]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[77.64441]
[78.]
epoch: 68 loss: 0.12644483149051666
tensor([[32.4976]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[32.497623]
[35.]
epoch: 69 loss: 6.26188850402832

```

```

tensor([[120.5900]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[120.590034]
[116.]
epoch: 70 loss: 21.068416595458984
tensor([[2.5370]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[2.5370467]
[10.]
epoch: 71 loss: 55.695674896240234
tensor([[112.4527]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[112.45274]
[108.]
epoch: 72 loss: 19.82692527770996
tensor([[151.4015]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[151.40146]
[146.]
epoch: 73 loss: 29.175756454467773
tensor([[1.1519]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[1.1519011]
[4.]
epoch: 74 loss: 8.111666679382324
tensor([[36.1992]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[36.199154]
[37.]
epoch: 75 loss: 0.641354501247406
tensor([[8.6181]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[8.618055]
[16.]
epoch: 76 loss: 54.493106842041016
tensor([[62.6868]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[62.686768]
[62.]
epoch: 77 loss: 0.471649706363678
tensor([[113.9024]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[113.90242]
[118.]
epoch: 78 loss: 16.7901611328125
tensor([[69.1153]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[69.115326]

```

```

[69.]
epoch: 79 loss: 0.013300069607794285
tensor([[177.8106]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[177.81062]
[185.]
epoch: 80 loss: 51.687137603759766
tensor([[12.4890]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[12.489014]
[11.]
epoch: 81 loss: 2.2171616554260254
tensor([[12.1572]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[12.157195]
[9.]
epoch: 82 loss: 9.967881202697754
tensor([[91.5051]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[91.5051]
[88.]
epoch: 83 loss: 12.285700798034668
tensor([[103.9196]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[103.91964]
[99.]
epoch: 84 loss: 24.20285415649414
tensor([[9.0292]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[9.0292015]
[1.]
epoch: 85 loss: 64.46807861328125
tensor([[148.3343]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[148.33432]
[145.]
epoch: 86 loss: 11.117690086364746
tensor([[198.9045]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[198.90446]
[197.]
epoch: 87 loss: 3.626985788345337
tensor([[163.6383]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[163.63834]
[165.]
epoch: 88 loss: 1.854128360748291
tensor([[82.0580]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],

```

```

device='cuda:0')
[82.05801]
[83.]
epoch: 89 loss: 0.8873521685600281
tensor([[194.1551]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[194.15506]
[200.]
epoch: 90 loss: 34.163326263427734
tensor([[148.2115]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[148.21155]
[152.]
epoch: 91 loss: 14.35236930847168
tensor([[154.2597]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[154.25969]
[154.]
epoch: 92 loss: 0.06743855029344559
tensor([[13.8634]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[13.86336]
[19.]
epoch: 93 loss: 26.385066986083984
tensor([[60.9395]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[60.93951]
[57.]
epoch: 94 loss: 15.519742012023926
tensor([[210.9345]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[210.93452]
[199.]
epoch: 95 loss: 142.43287658691406
tensor([[151.9902]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[151.99016]
[151.]
epoch: 96 loss: 0.980413019657135
tensor([[77.5016]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[77.50158]
[79.]
epoch: 97 loss: 2.2452645301818848
tensor([[86.2099]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[86.209946]
[91.]

```

```

epoch: 98 loss: 22.94462013244629
tensor([[24.9017]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[24.901678]
[27.]
epoch: 99 loss: 4.402955055236816
tensor([[9.9509]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[9.950938]
[13.]
epoch: 100 loss: 9.296777725219727
tensor([[159.5809]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[159.58093]
[175.]
epoch: 101 loss: 237.7476348876953
tensor([[13.0325]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[13.032451]
[17.]
epoch: 102 loss: 15.741447448730469
tensor([[46.3276]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[46.327633]
[46.]
epoch: 103 loss: 0.10734332352876663
tensor([[154.6423]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[154.64227]
[153.]
epoch: 104 loss: 2.6970603466033936
tensor([[19.3844]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[19.384422]
[21.]
epoch: 105 loss: 2.610091209411621
tensor([[37.3143]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[37.314293]
[33.]
epoch: 106 loss: 18.613122940063477
tensor([[34.5156]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[34.515553]
[30.]
epoch: 107 loss: 20.390214920043945
tensor([[101.3510]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')

```

```

[101.35099]
[93.]
epoch: 108 loss: 69.7390365600586
tensor([[11.8444]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[11.844436]
[3.]
epoch: 109 loss: 78.22404479980469
tensor([[85.5875]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[85.587456]
[80.]
epoch: 110 loss: 31.219661712646484
tensor([[127.6913]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[127.69129]
[122.]
epoch: 111 loss: 32.390804290771484
tensor([[41.0432]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[41.04319]
[43.]
epoch: 112 loss: 3.8291053771972656
tensor([[62.0591]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[62.059128]
[65.]
epoch: 113 loss: 8.64872932434082
tensor([[89.5780]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[89.578026]
[95.]
epoch: 114 loss: 29.397804260253906
tensor([[106.6228]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[106.62276]
[113.]
epoch: 115 loss: 40.669227600097656
tensor([[70.7792]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[70.779175]
[76.]
epoch: 116 loss: 27.257015228271484
tensor([[185.5723]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[185.57231]
[186.]
epoch: 117 loss: 0.1829175353050232

```

```

tensor([[78.1020]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[78.101974]
[81.]
epoch: 118 loss: 8.398551940917969
tensor([[157.9671]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[157.96715]
[155.]
epoch: 119 loss: 8.803966522216797
tensor([[99.0740]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[99.07396]
[100.]
epoch: 120 loss: 0.8575512766838074
tensor([[177.1456]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[177.14557]
[173.]
epoch: 121 loss: 17.185741424560547
tensor([[81.7595]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[81.75946]
[82.]
epoch: 122 loss: 0.05785927549004555
tensor([[14.1281]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[14.128142]
[20.]
epoch: 123 loss: 34.47871398925781
tensor([[86.3447]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[86.34468]
[87.]
epoch: 124 loss: 0.42944326996803284
tensor([[119.2757]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[119.2757]
[119.]
epoch: 125 loss: 0.07601238042116165
tensor([[27.5349]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[27.534891]
[28.]
epoch: 126 loss: 0.2163262665271759
tensor([[111.8200]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[111.82002]

```

```

[112.]
epoch: 127 loss: 0.03239187225699425
tensor([[97.3451]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[97.345116]
[97.]
epoch: 128 loss: 0.11910481750965118
tensor([[86.2619]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[86.26192]
[86.]
epoch: 129 loss: 0.06860057264566422
tensor([[29.6529]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[29.652912]
[29.]
epoch: 130 loss: 0.4262942671775818
tensor([[125.0786]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[125.07858]
[124.]
epoch: 131 loss: 1.1633408069610596
tensor([[33.0440]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[33.044]
[32.]
epoch: 132 loss: 1.0899332761764526
tensor([[161.1356]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[161.13562]
[162.]
epoch: 133 loss: 0.74715256690979
tensor([[50.7978]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[50.797756]
[49.]
epoch: 134 loss: 3.2319273948669434
tensor([[11.1235]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[11.1235285]
[8.]
epoch: 135 loss: 9.756430625915527
tensor([[94.5745]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[94.57455]
[96.]
epoch: 136 loss: 2.031916856765747
tensor([[132.1117]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],

```



```

device='cuda:0')
[132.11168]
[135.]
epoch: 137 loss: 8.342397689819336
tensor([[105.3075]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[105.307465]
[107.]
epoch: 138 loss: 2.864675998687744
tensor([[196.4175]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[196.41754]
[198.]
epoch: 139 loss: 2.5041749477386475
tensor([[171.3959]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[171.3959]
[170.]
epoch: 140 loss: 1.948549509048462
tensor([[173.5677]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[173.56766]
[171.]
epoch: 141 loss: 6.592864990234375
tensor([[105.1925]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[105.19254]
[103.]
epoch: 142 loss: 4.807244777679443
tensor([[194.9885]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[194.98854]
[193.]
epoch: 143 loss: 3.954293966293335
tensor([[156.0156]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[156.0156]
[156.]
epoch: 144 loss: 0.00024318788200616837
tensor([[48.7744]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[48.77439]
[48.]
epoch: 145 loss: 0.5996816754341125
tensor([[178.4991]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[178.49908]
[182.]

```

```

epoch: 146 loss: 12.256409645080566
tensor([[134.4686]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[134.4686]
[137.]
epoch: 147 loss: 6.407999038696289
tensor([[23.8792]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[23.87925]
[25.]
epoch: 148 loss: 1.256081461906433
tensor([[116.5524]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[116.55243]
[117.]
epoch: 149 loss: 0.20031961798667908
tensor([[186.3120]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[186.31201]
[184.]
epoch: 150 loss: 5.34539794921875
tensor([[12.8849]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[12.884877]
[12.]
epoch: 151 loss: 0.7830076813697815
tensor([[110.0524]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[110.05241]
[110.]
epoch: 152 loss: 0.002746421378105879
tensor([[9.1335]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[9.133488]
[6.]
epoch: 153 loss: 9.818745613098145
tensor([[88.7848]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[88.784836]
[89.]
epoch: 154 loss: 0.04629562795162201
tensor([[140.1805]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[140.1805]
[140.]
epoch: 155 loss: 0.032578885555267334
tensor([[126.6137]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')

```

```

[126.61366]
[126.]
epoch: 156 loss: 0.37658193707466125
tensor([[135.6959]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[135.69589]
[136.]
epoch: 157 loss: 0.09248147159814835
tensor([[7.7293]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[7.7293224]
[7.]
epoch: 158 loss: 0.5319111943244934
tensor([[122.8794]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[122.87941]
[123.]
epoch: 159 loss: 0.014541998505592346
tensor([[17.5583]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[17.55827]
[22.]
epoch: 160 loss: 19.72896957397461
tensor([[104.9402]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[104.94021]
[106.]
epoch: 161 loss: 1.1231582164764404
tensor([[125.0557]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[125.05567]
[125.]
epoch: 162 loss: 0.0030993372201919556
tensor([[108.0687]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[108.06873]
[109.]
epoch: 163 loss: 0.867257833480835
tensor([[14.4833]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[14.48329]
[18.]
epoch: 164 loss: 12.3672513961792
tensor([[133.7507]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[133.75069]
[133.]
epoch: 165 loss: 0.563530445098877

```

```

tensor([[69.1182]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[69.11819]
[67.]
epoch: 166 loss: 4.486715793609619
tensor([[66.0805]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[66.0805]
[64.]
epoch: 167 loss: 4.328470706939697
tensor([[143.7694]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[143.76944]
[143.]
epoch: 168 loss: 0.5920374393463135
tensor([[35.9576]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[35.957626]
[36.]
epoch: 169 loss: 0.0017955268267542124
tensor([[171.7524]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[171.75244]
[174.]
epoch: 170 loss: 5.051519393920898
tensor([[52.9363]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[52.93628]
[51.]
epoch: 171 loss: 3.7491774559020996
tensor([[72.1373]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[72.13732]
[71.]
epoch: 172 loss: 1.2935001850128174
tensor([[132.0647]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[132.06468]
[132.]
epoch: 173 loss: 0.004183761775493622
tensor([[104.7197]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[104.71966]
[105.]
epoch: 174 loss: 0.07859169691801071
tensor([[160.4427]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[160.44275]

```

```

[160.]
epoch: 175 loss: 0.19602669775485992
tensor([[113.2203]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[113.220276]
[114.]
epoch: 176 loss: 0.6079697012901306
tensor([[137.7329]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[137.73286]
[138.]
epoch: 177 loss: 0.07136143743991852
tensor([[141.5278]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[141.52785]
[142.]
epoch: 178 loss: 0.22292818129062653
tensor([[201.4588]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[201.45879]
[201.]
epoch: 179 loss: 0.2104846090078354
tensor([[119.6741]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[119.67415]
[120.]
epoch: 180 loss: 0.10617916285991669
tensor([[163.9048]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[163.90475]
[163.]
epoch: 181 loss: 0.8185809850692749
tensor([[193.8188]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[193.81876]
[194.]
epoch: 182 loss: 0.032849349081516266
tensor([[93.2331]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[93.23315]
[94.]
epoch: 183 loss: 0.5880640149116516
tensor([[37.3858]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[37.385757]
[39.]
epoch: 184 loss: 2.605778932571411
tensor([[50.5098]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],

```

```

device='cuda:0')
[50.50979]
[50.]
epoch: 185 loss: 0.2598843276500702
tensor([[190.2631]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[190.26306]
[190.]
epoch: 186 loss: 0.06920136511325836
For epoch 13 the test accuracy over the whole test set is 0 %
tensor([[15.2425]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[15.242506]
[16.]
epoch: 0 loss: 0.5737971067428589
tensor([[145.6869]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[145.68692]
[146.]
epoch: 1 loss: 0.09801898151636124
tensor([[142.3198]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[142.31982]
[142.]
epoch: 2 loss: 0.10228753089904785
tensor([[101.4449]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[101.44488]
[101.]
epoch: 3 loss: 0.19791610538959503
tensor([[202.5856]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[202.58556]
[203.]
epoch: 4 loss: 0.1717638075351715
tensor([[7.6344]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[7.6344423]
[2.]
epoch: 5 loss: 31.74694061279297
tensor([[32.6702]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[32.670162]
[33.]
epoch: 6 loss: 0.1087929755449295
tensor([[144.9565]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[144.95654]

```

```

[145.]
epoch: 7 loss: 0.0018885135650634766
tensor([[186.0605]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[186.06053]
[185.]
epoch: 8 loss: 1.1247273683547974
tensor([[15.9859]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[15.985945]
[20.]
epoch: 9 loss: 16.112640380859375
tensor([[77.3303]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[77.33034]
[78.]
epoch: 10 loss: 0.4484478235244751
tensor([[12.2462]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[12.246236]
[11.]
epoch: 11 loss: 1.5531038045883179
tensor([[78.3684]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[78.36844]
[79.]
epoch: 12 loss: 0.39886966347694397
tensor([[47.5382]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[47.53823]
[48.]
epoch: 13 loss: 0.213230699300766
tensor([[168.6763]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[168.67633]
[170.]
epoch: 14 loss: 1.7521008253097534
tensor([[165.9025]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[165.90254]
[166.]
epoch: 15 loss: 0.00949803926050663
tensor([[199.2312]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[199.23119]
[200.]
epoch: 16 loss: 0.5910751223564148
tensor([[184.1420]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],

```

```

device='cuda:0')
[184.14201]
[182.]
epoch: 17 loss: 4.588222026824951
tensor([[12.2164]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[12.216358]
[10.]
epoch: 18 loss: 4.912243366241455
tensor([[93.2740]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[93.273964]
[93.]
epoch: 19 loss: 0.07505623251199722
tensor([[9.3544]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[9.354438]
[8.]
epoch: 20 loss: 1.83450186252594
tensor([[83.7814]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[83.78137]
[84.]
epoch: 21 loss: 0.047798171639442444
tensor([[56.0870]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[56.086987]
[55.]
epoch: 22 loss: 1.18153977394104
tensor([[35.7541]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[35.75405]
[37.]
epoch: 23 loss: 1.5523884296417236
tensor([[25.8600]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[25.860039]
[27.]
epoch: 24 loss: 1.2995116710662842
tensor([[84.8149]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[84.81486]
[86.]
epoch: 25 loss: 1.4045628309249878
tensor([[110.7639]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[110.76391]
[112.]

```



```

epoch: 26 loss: 1.527922511100769
tensor([[133.6118]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[133.61182]
[134.]
epoch: 27 loss: 0.15068650245666504
tensor([[199.1099]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[199.10988]
[199.]
epoch: 28 loss: 0.012073293328285217
tensor([[118.4214]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[118.4214]
[118.]
epoch: 29 loss: 0.1775796264410019
tensor([[111.6543]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[111.6543]
[111.]
epoch: 30 loss: 0.4281044006347656
tensor([[176.8312]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[176.83122]
[176.]
epoch: 31 loss: 0.6909309029579163
tensor([[157.0140]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[157.01402]
[156.]
epoch: 32 loss: 1.0282423496246338
tensor([[55.4010]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[55.401]
[54.]
epoch: 33 loss: 1.9628037214279175
tensor([[10.7223]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[10.722293]
[12.]
epoch: 34 loss: 1.632535457611084
tensor([[29.1137]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[29.113705]
[30.]
epoch: 35 loss: 0.7855194211006165
tensor([[28.1761]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')

```

```

[28.176123]
[29.]
epoch: 36 loss: 0.6787738800048828
tensor([[70.2706]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[70.27058]
[71.]
epoch: 37 loss: 0.5320586562156677
tensor([[7.1853]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[7.18534]
[7.]
epoch: 38 loss: 0.034350886940956116
tensor([[56.3297]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[56.329723]
[56.]
epoch: 39 loss: 0.10871749371290207
tensor([[31.4494]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[31.44939]
[32.]
epoch: 40 loss: 0.30317091941833496
tensor([[158.7172]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[158.71722]
[162.]
epoch: 41 loss: 10.776617050170898
tensor([[141.5250]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[141.52496]
[144.]
epoch: 42 loss: 6.1258063316345215
tensor([[38.6712]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[38.67121]
[39.]
epoch: 43 loss: 0.10810204595327377
tensor([[139.2968]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[139.29684]
[137.]
epoch: 44 loss: 5.275494575500488
tensor([[174.2899]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[174.28992]
[173.]
epoch: 45 loss: 1.6638858318328857

```

```

tensor([[60.1398]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[60.139767]
[57.]
epoch: 46 loss: 9.858135223388672
tensor([[199.1881]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[199.18811]
[197.]
epoch: 47 loss: 4.787827014923096
tensor([[81.3217]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[81.321655]
[80.]
epoch: 48 loss: 1.7467726469039917
tensor([[94.6892]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[94.68922]
[95.]
epoch: 49 loss: 0.09658631682395935
tensor([[127.8551]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[127.85513]
[130.]
epoch: 50 loss: 4.600454330444336
tensor([[121.8329]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[121.83295]
[123.]
epoch: 51 loss: 1.3620132207870483
tensor([[14.5394]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[14.539427]
[18.]
epoch: 52 loss: 11.975566864013672
tensor([[4.9257]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[4.925718]
[1.]
epoch: 53 loss: 15.411260604858398
tensor([[14.9748]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[14.974787]
[19.]
epoch: 54 loss: 16.202341079711914
tensor([[189.0097]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[189.00967]

```

```

[193.]
epoch: 55 loss: 15.922700881958008
tensor([[182.9825]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[182.98251]
[183.]
epoch: 56 loss: 0.00030578020960092545
tensor([[66.5434]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[66.54344]
[65.]
epoch: 57 loss: 2.3822124004364014
tensor([[174.4821]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[174.48213]
[171.]
epoch: 58 loss: 12.125243186950684
tensor([[9.4374]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[9.437384]
[6.]
epoch: 59 loss: 11.815606117248535
tensor([[181.9807]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[181.98065]
[179.]
epoch: 60 loss: 8.884285926818848
tensor([[158.8318]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[158.83183]
[157.]
epoch: 61 loss: 3.355611801147461
tensor([[34.4083]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[34.40833]
[35.]
epoch: 62 loss: 0.35007455945014954
tensor([[87.3337]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[87.333725]
[90.]
epoch: 63 loss: 7.109022617340088
tensor([[105.6397]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[105.63966]
[110.]
epoch: 64 loss: 19.01253318786621
tensor([[147.4330]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],

```

```

device='cuda:0')
[147.43301]
[153.]
epoch: 65 loss: 30.991334915161133
tensor([[7.2233]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[7.223317]
[5.]
epoch: 66 loss: 4.94313907623291
tensor([[20.9330]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[20.933006]
[23.]
epoch: 67 loss: 4.272462844848633
tensor([[120.7219]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[120.72191]
[119.]
epoch: 68 loss: 2.9649691581726074
tensor([[77.5402]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[77.54023]
[75.]
epoch: 69 loss: 6.452767372131348
tensor([[70.9218]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[70.9218]
[68.]
epoch: 70 loss: 8.536908149719238
tensor([[142.8035]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[142.80354]
[139.]
epoch: 71 loss: 14.466939926147461
tensor([[12.4703]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[12.470296]
[14.]
epoch: 72 loss: 2.3399946689605713
tensor([[184.6848]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[184.68478]
[184.]
epoch: 73 loss: 0.4689290523529053
tensor([[103.2082]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[103.20822]
[104.]

```

```

epoch: 74 loss: 0.62691330909729
tensor([[40.8526]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[40.85265]
[43.]
epoch: 75 loss: 4.611113548278809
tensor([[60.7841]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[60.784077]
[62.]
epoch: 76 loss: 1.4784694910049438
tensor([[79.2026]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[79.202614]
[81.]
epoch: 77 loss: 3.2305970191955566
tensor([[125.8447]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[125.84467]
[129.]
epoch: 78 loss: 9.956087112426758
tensor([[48.8842]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[48.884247]
[49.]
epoch: 79 loss: 0.013398797251284122
tensor([[63.7478]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[63.74779]
[64.]
epoch: 80 loss: 0.06360923498868942
tensor([[74.2466]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[74.246635]
[74.]
epoch: 81 loss: 0.0608290396630764
tensor([[149.4191]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[149.41914]
[150.]
epoch: 82 loss: 0.3373940587043762
tensor([[137.7186]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[137.71861]
[136.]
epoch: 83 loss: 2.953629493713379
tensor([[72.2823]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')

```

```

[72.28234]
[70.]
epoch: 84 loss: 5.209080219268799
tensor([[99.9792]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[99.97923]
[98.]
epoch: 85 loss: 3.9173624515533447
tensor([[116.1897]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[116.18965]
[115.]
epoch: 86 loss: 1.4152706861495972
tensor([[91.3288]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[91.32883]
[91.]
epoch: 87 loss: 0.10812713205814362
tensor([[121.1728]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[121.17279]
[122.]
epoch: 88 loss: 0.6842755079269409
tensor([[85.7811]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[85.78108]
[87.]
epoch: 89 loss: 1.4857606887817383
tensor([[156.8838]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[156.88377]
[160.]
epoch: 90 loss: 9.71086597442627
tensor([[199.2832]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[199.28325]
[205.]
epoch: 91 loss: 32.681243896484375
tensor([[93.6856]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[93.68561]
[92.]
epoch: 92 loss: 2.8412740230560303
tensor([[7.1904]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[7.190353]
[3.]
epoch: 93 loss: 17.559057235717773

```

```

tensor([[132.3963]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[132.39632]
[124.]
epoch: 94 loss: 70.49813079833984
tensor([[55.1740]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[55.173958]
[52.]
epoch: 95 loss: 10.074007987976074
tensor([[102.5469]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[102.54687]
[99.]
epoch: 96 loss: 12.580267906188965
tensor([[11.3060]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[11.306029]
[15.]
epoch: 97 loss: 13.645419120788574
tensor([[176.7990]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[176.79903]
[175.]
epoch: 98 loss: 3.2364962100982666
tensor([[58.1144]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[58.114365]
[61.]
epoch: 99 loss: 8.326891899108887
tensor([[131.7353]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[131.73526]
[135.]
epoch: 100 loss: 10.658527374267578
tensor([[160.7771]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[160.77705]
[163.]
epoch: 101 loss: 4.9414896965026855
tensor([[170.7159]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[170.71585]
[174.]
epoch: 102 loss: 10.785635948181152
tensor([[147.0607]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[147.06073]

```



```

[148.]
epoch: 103 loss: 0.8822281956672668
tensor([[193.8521]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[193.85208]
[189.]
epoch: 104 loss: 23.542692184448242
tensor([[24.1116]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[24.11157]
[26.]
epoch: 105 loss: 3.566166400909424
tensor([[107.6708]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[107.67075]
[109.]
epoch: 106 loss: 1.766896367073059
tensor([[209.8442]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[209.84421]
[204.]
epoch: 107 loss: 34.15476608276367
tensor([[148.9248]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[148.92477]
[151.]
epoch: 108 loss: 4.306562423706055
tensor([[92.6040]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[92.604004]
[97.]
epoch: 109 loss: 19.32478141784668
tensor([[144.8479]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[144.84795]
[149.]
epoch: 110 loss: 17.239551544189453
tensor([[189.8787]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[189.87872]
[192.]
epoch: 111 loss: 4.499815464019775
tensor([[43.8206]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[43.82059]
[46.]
epoch: 112 loss: 4.749823570251465
tensor([[194.3548]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],

```

```

device='cuda:0')
[194.35475]
[191.]
epoch: 113 loss: 11.254358291625977
tensor([[114.9372]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[114.93724]
[114.]
epoch: 114 loss: 0.8784199357032776
tensor([[4.4059]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[4.4059052]
[0.]
epoch: 115 loss: 19.41200065612793
tensor([[19.3581]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[19.358131]
[22.]
epoch: 116 loss: 6.9794697761535645
tensor([[157.1719]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[157.17186]
[154.]
epoch: 117 loss: 10.060694694519043
tensor([[116.0931]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[116.09314]
[116.]
epoch: 118 loss: 0.008674994111061096
tensor([[93.0130]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[93.01301]
[94.]
epoch: 119 loss: 0.9741529822349548
tensor([[162.6942]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[162.69424]
[164.]
epoch: 120 loss: 1.7049977779388428
tensor([[101.5064]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[101.5064]
[103.]
epoch: 121 loss: 2.230837821960449
tensor([[22.7506]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[22.750574]
[24.]

```

```

epoch: 122 loss: 1.5610650777816772
tensor([[47.2265]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[47.226547]
[47.]
epoch: 123 loss: 0.051323652267456055
tensor([[174.9230]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[174.92302]
[178.]
epoch: 124 loss: 9.467809677124023
tensor([[201.0123]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[201.01234]
[201.]
epoch: 125 loss: 0.00015238323248922825
tensor([[154.1657]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[154.16568]
[152.]
epoch: 126 loss: 4.690169334411621
tensor([[17.9717]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[17.971651]
[21.]
epoch: 127 loss: 9.170897483825684
tensor([[80.3892]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[80.389175]
[77.]
epoch: 128 loss: 11.486510276794434
tensor([[63.6998]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[63.699844]
[60.]
epoch: 129 loss: 13.688848495483398
tensor([[87.2119]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[87.21194]
[85.]
epoch: 130 loss: 4.892664909362793
tensor([[176.0470]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[176.047]
[177.]
epoch: 131 loss: 0.9082145690917969
tensor([[10.4608]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')

```

```

[10.460787]
[9.]
epoch: 132 loss: 2.1338980197906494
tensor([[142.3184]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[142.31837]
[143.]
epoch: 133 loss: 0.464613139629364
tensor([[130.8626]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[130.86263]
[132.]
epoch: 134 loss: 1.2936216592788696
tensor([[118.4279]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[118.42794]
[120.]
epoch: 135 loss: 2.4713714122772217
tensor([[43.3699]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[43.369907]
[44.]
epoch: 136 loss: 0.397016704082489
tensor([[87.2631]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[87.26311]
[88.]
epoch: 137 loss: 0.5430108308792114
tensor([[69.1449]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[69.1449]
[69.]
epoch: 138 loss: 0.02099527418613434
tensor([[88.9328]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[88.93276]
[89.]
epoch: 139 loss: 0.004520928952842951
tensor([[169.5245]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[169.5245]
[167.]
epoch: 140 loss: 6.373128414154053
tensor([[170.2756]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[170.27562]
[169.]
epoch: 141 loss: 1.6272051334381104

```

```

tensor([[168.1729]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[168.17288]
[168.]
epoch: 142 loss: 0.02988821268081665
tensor([[74.8753]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[74.87527]
[76.]
epoch: 143 loss: 1.2650243043899536
tensor([[28.0367]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[28.036736]
[28.]
epoch: 144 loss: 0.001349499449133873
tensor([[190.5465]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[190.54651]
[196.]
epoch: 145 loss: 29.74056625366211
tensor([[66.0587]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[66.05875]
[67.]
epoch: 146 loss: 0.8859584331512451
tensor([[113.8996]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[113.89963]
[113.]
epoch: 147 loss: 0.8093299865722656
tensor([[169.6402]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[169.64018]
[165.]
epoch: 148 loss: 21.531293869018555
tensor([[142.8662]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[142.8662]
[140.]
epoch: 149 loss: 8.21507740020752
tensor([[126.1780]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[126.178]
[125.]
epoch: 150 loss: 1.3876873254776
tensor([[6.0158]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[6.0158243]

```

```

[4.]
epoch: 151 loss: 4.063547611236572
tensor([[101.8706]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[101.8706]
[105.]
epoch: 152 loss: 9.793157577514648
tensor([[165.9121]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[165.9121]
[172.]
epoch: 153 loss: 37.06241226196289
tensor([[13.5152]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[13.515151]
[17.]
epoch: 154 loss: 12.144172668457031
tensor([[140.8725]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[140.87253]
[141.]
epoch: 155 loss: 0.016249090433120728
tensor([[35.1967]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[35.196735]
[34.]
epoch: 156 loss: 1.4321755170822144
tensor([[59.6358]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[59.635773]
[58.]
epoch: 157 loss: 2.6757524013519287
tensor([[142.9834]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[142.9834]
[138.]
epoch: 158 loss: 24.834259033203125
tensor([[46.0025]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[46.002525]
[45.]
epoch: 159 loss: 1.0050569772720337
tensor([[52.3182]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[52.31819]
[51.]
epoch: 160 loss: 1.7376289367675781
tensor([[35.8436]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],

```

```

device='cuda:0')
[35.843586]
[36.]
epoch: 161 loss: 0.024465348571538925
tensor([[37.2472]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[37.247215]
[38.]
epoch: 162 loss: 0.5666848421096802
tensor([[190.0958]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[190.09584]
[190.]
epoch: 163 loss: 0.009185392409563065
tensor([[11.9384]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[11.938444]
[13.]
epoch: 164 loss: 1.1269007921218872
tensor([[48.8195]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[48.81952]
[50.]
epoch: 165 loss: 1.3935352563858032
tensor([[130.5739]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[130.57393]
[133.]
epoch: 166 loss: 5.885821342468262
tensor([[124.6762]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[124.676186]
[126.]
epoch: 167 loss: 1.7524845600128174
tensor([[100.9617]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[100.96174]
[102.]
epoch: 168 loss: 1.077986717224121
tensor([[116.7243]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[116.7243]
[117.]
epoch: 169 loss: 0.07601238042116165
tensor([[162.5975]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[162.59755]
[159.]

```

```

epoch: 170 loss: 12.942361831665039
tensor([[81.8046]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[81.80458]
[82.]
epoch: 171 loss: 0.03818870708346367
tensor([[196.3586]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[196.3586]
[195.]
epoch: 172 loss: 1.8457852602005005
tensor([[106.6983]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[106.69834]
[108.]
epoch: 173 loss: 1.6943151950836182
tensor([[104.9095]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[104.909546]
[106.]
epoch: 174 loss: 1.1890901327133179
tensor([[197.2068]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[197.20682]
[198.]
epoch: 175 loss: 0.6291382908821106
tensor([[95.1949]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[95.19488]
[96.]
epoch: 176 loss: 0.648222029209137
tensor([[106.1570]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[106.157]
[107.]
epoch: 177 loss: 0.7106528878211975
tensor([[83.0516]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[83.051636]
[83.]
epoch: 178 loss: 0.0026662498712539673
tensor([[99.7733]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[99.773315]
[100.]
epoch: 179 loss: 0.051385894417762756
tensor([[148.0405]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')

```



```

[148.0405]
[147.]
epoch: 180 loss: 1.0826336145401
tensor([[26.0848]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[26.08479]
[25.]
epoch: 181 loss: 1.1767678260803223
tensor([[195.6158]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[195.61575]
[194.]
epoch: 182 loss: 2.6106584072113037
tensor([[155.8309]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[155.8309]
[155.]
epoch: 183 loss: 0.6903982758522034
tensor([[178.3655]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[178.36548]
[180.]
epoch: 184 loss: 2.6716604232788086
tensor([[54.2191]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[54.219112]
[53.]
epoch: 185 loss: 1.4862350225448608
tensor([[184.1928]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[184.19284]
[186.]
epoch: 186 loss: 3.2658252716064453
For epoch 14 the test accuracy over the whole test set is 0 %
tensor([[104.7091]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[104.70912]
[105.]
epoch: 0 loss: 0.08461018651723862
tensor([[117.6663]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[117.6663]
[118.]
epoch: 1 loss: 0.11135708540678024
tensor([[177.8470]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[177.84705]
[178.]

```

```

epoch: 2 loss: 0.023394957184791565
tensor([[125.9625]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[125.9625]
[124.]
epoch: 3 loss: 3.851412296295166
tensor([[141.4865]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[141.48651]
[141.]
epoch: 4 loss: 0.23669317364692688
tensor([[30.2341]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[30.234093]
[29.]
epoch: 5 loss: 1.5229848623275757
tensor([[61.4496]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[61.44955]
[61.]
epoch: 6 loss: 0.20209576189517975
tensor([[34.7590]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[34.758957]
[34.]
epoch: 7 loss: 0.5760155916213989
tensor([[69.4495]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[69.44946]
[70.]
epoch: 8 loss: 0.3030911087989807
tensor([[82.9544]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[82.95441]
[84.]
epoch: 9 loss: 1.0932652950286865
tensor([[133.7893]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[133.78932]
[135.]
epoch: 10 loss: 1.465741515159607
tensor([[162.2452]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[162.24524]
[162.]
epoch: 11 loss: 0.06014229357242584
tensor([[135.5568]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')

```

```

[135.5568]
[136.]
epoch: 12 loss: 0.19643226265907288
tensor([[63.0965]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[63.096504]
[64.]
epoch: 13 loss: 0.816304624080658
tensor([[117.2327]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[117.232704]
[117.]
epoch: 14 loss: 0.054151225835084915
tensor([[171.1856]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[171.18562]
[170.]
epoch: 15 loss: 1.4057023525238037
tensor([[205.7583]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[205.75826]
[205.]
epoch: 16 loss: 0.574950635433197
tensor([[56.4971]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[56.497074]
[56.]
epoch: 17 loss: 0.2470826953649521
tensor([[61.5604]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[61.560383]
[62.]
epoch: 18 loss: 0.1932632476091385
tensor([[113.0043]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[113.004326]
[114.]
epoch: 19 loss: 0.991366982460022
tensor([[121.2684]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[121.26838]
[122.]
epoch: 20 loss: 0.5352689623832703
tensor([[48.9919]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[48.991898]
[49.]
epoch: 21 loss: 6.564916111528873e-05

```

```

tensor([[184.3542]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[184.35425]
[185.]
epoch: 22 loss: 0.4169955849647522
tensor([[53.8388]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[53.83881]
[53.]
epoch: 23 loss: 0.703602135181427
tensor([[104.4691]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[104.46909]
[104.]
epoch: 24 loss: 0.22004854679107666
tensor([[16.5948]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[16.594837]
[17.]
epoch: 25 loss: 0.16415689885616302
tensor([[71.2079]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[71.20795]
[71.]
epoch: 26 loss: 0.04324186220765114
tensor([[64.7570]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[64.757]
[65.]
epoch: 27 loss: 0.059047162532806396
tensor([[124.3424]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[124.34243]
[123.]
epoch: 28 loss: 1.8021186590194702
tensor([[204.2428]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[204.24284]
[204.]
epoch: 29 loss: 0.058973029255867004
tensor([[155.6587]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[155.65868]
[156.]
epoch: 30 loss: 0.11650197207927704
tensor([[76.8102]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[76.81024]

```

```

[77.]
epoch: 31 loss: 0.036008212715387344
tensor([[18.1955]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[18.19552]
[20.]
epoch: 32 loss: 3.2561466693878174
tensor([[114.5280]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[114.527985]
[115.]
epoch: 33 loss: 0.22279852628707886
tensor([[102.9268]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[102.92676]
[103.]
epoch: 34 loss: 0.005364418029785156
tensor([[137.1274]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[137.12743]
[138.]
epoch: 35 loss: 0.7613851428031921
tensor([[159.1980]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[159.19803]
[159.]
epoch: 36 loss: 0.03921531140804291
tensor([[27.7359]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[27.735882]
[27.]
epoch: 37 loss: 0.5415220260620117
tensor([[146.7620]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[146.76196]
[147.]
epoch: 38 loss: 0.05666166543960571
tensor([[94.8663]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[94.86632]
[94.]
epoch: 39 loss: 0.7505064606666565
tensor([[150.0019]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[150.00189]
[150.]
epoch: 40 loss: 3.5800039768218994e-06
tensor([[193.4816]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],

```

```

device='cuda:0')
[193.48164]
[193.]
epoch: 41 loss: 0.23198063671588898
tensor([[158.2232]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[158.22322]
[157.]
epoch: 42 loss: 1.4962692260742188
tensor([[73.7462]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[73.74623]
[74.]
epoch: 43 loss: 0.06439866870641708
tensor([[136.0231]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[136.02312]
[137.]
epoch: 44 loss: 0.9543002843856812
tensor([[174.4276]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[174.42758]
[177.]
epoch: 45 loss: 6.617335319519043
tensor([[18.7387]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[18.738739]
[18.]
epoch: 46 loss: 0.5457353591918945
tensor([[55.5538]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[55.55376]
[54.]
epoch: 47 loss: 2.4141716957092285
tensor([[82.3495]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[82.34948]
[81.]
epoch: 48 loss: 1.82109534740448
tensor([[150.2805]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[150.28049]
[149.]
epoch: 49 loss: 1.639647126197815
tensor([[38.0947]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[38.09472]
[38.]

```

```

epoch: 50 loss: 0.008971676230430603
tensor([[86.3280]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[86.32804]
[86.]
epoch: 51 loss: 0.10761094838380814
tensor([[44.5550]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[44.554996]
[45.]
epoch: 52 loss: 0.19802811741828918
tensor([[192.3107]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[192.31065]
[192.]
epoch: 53 loss: 0.09650571644306183
tensor([[153.8738]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[153.87381]
[154.]
epoch: 54 loss: 0.015923961997032166
tensor([[125.8439]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[125.84394]
[126.]
epoch: 55 loss: 0.024354495108127594
tensor([[51.3664]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[51.36637]
[52.]
epoch: 56 loss: 0.40148550271987915
tensor([[198.0382]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[198.03818]
[200.]
epoch: 57 loss: 3.848747491836548
tensor([[105.9282]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[105.928154]
[106.]
epoch: 58 loss: 0.005161848850548267
tensor([[167.3971]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[167.3971]
[166.]
epoch: 59 loss: 1.9518736600875854
tensor([[195.8977]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')

```

```

[195.89769]
[195.]
epoch: 60 loss: 0.8058469891548157
tensor([[59.5506]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[59.550617]
[60.]
epoch: 61 loss: 0.20194488763809204
tensor([[6.1322]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[6.132195]
[1.]
epoch: 62 loss: 26.339426040649414
tensor([[15.6943]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[15.694285]
[16.]
epoch: 63 loss: 0.09346142411231995
tensor([[6.7478]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[6.747847]
[4.]
epoch: 64 loss: 7.550663471221924
tensor([[41.2542]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[41.254246]
[44.]
epoch: 65 loss: 7.539166450500488
tensor([[178.2987]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[178.29869]
[180.]
epoch: 66 loss: 2.8944530487060547
tensor([[11.7962]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[11.796197]
[15.]
epoch: 67 loss: 10.26435375213623
tensor([[48.8224]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[48.82242]
[50.]
epoch: 68 loss: 1.386698842048645
tensor([[91.7096]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[91.70964]
[92.]
epoch: 69 loss: 0.08430863916873932

```



```

tensor([[44.9451]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[44.945087]
[46.]
epoch: 70 loss: 1.1128405332565308
tensor([[110.9110]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[110.91103]
[109.]
epoch: 71 loss: 3.6520495414733887
tensor([[24.2436]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[24.243618]
[25.]
epoch: 72 loss: 0.5721136927604675
tensor([[176.3047]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[176.3047]
[172.]
epoch: 73 loss: 18.530466079711914
tensor([[14.9822]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[14.982176]
[19.]
epoch: 74 loss: 16.142911911010742
tensor([[21.0052]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[21.005205]
[23.]
epoch: 75 loss: 3.9792065620422363
tensor([[118.2265]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[118.226456]
[119.]
epoch: 76 loss: 0.5983707904815674
tensor([[187.9871]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[187.98706]
[190.]
epoch: 77 loss: 4.051925182342529
tensor([[87.6261]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[87.62606]
[88.]
epoch: 78 loss: 0.13983075320720673
tensor([[85.3975]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[85.397545]

```

```

[85.]
epoch: 79 loss: 0.15804190933704376
tensor([[149.7300]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[149.73001]
[151.]
epoch: 80 loss: 1.6128721237182617
tensor([[164.2215]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[164.22145]
[163.]
epoch: 81 loss: 1.491942048072815
tensor([[103.1166]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[103.116585]
[101.]
epoch: 82 loss: 4.479931354522705
tensor([[114.0319]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[114.03187]
[113.]
epoch: 83 loss: 1.0647515058517456
tensor([[69.7647]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[69.76466]
[68.]
epoch: 84 loss: 3.1140379905700684
tensor([[128.7815]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[128.78152]
[130.]
epoch: 85 loss: 1.4846822023391724
tensor([[82.0887]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[82.088684]
[82.]
epoch: 86 loss: 0.007864866405725479
tensor([[86.5070]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[86.50702]
[87.]
epoch: 87 loss: 0.24303022027015686
tensor([[167.0216]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[167.0216]
[171.]
epoch: 88 loss: 15.827615737915039
tensor([[162.8571]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],

```

```

device='cuda:0')
[162.85709]
[164.]
epoch: 89 loss: 1.3062520027160645
tensor([[156.1748]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[156.17479]
[155.]
epoch: 90 loss: 1.3801301717758179
tensor([[23.6846]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[23.684572]
[22.]
epoch: 91 loss: 2.8377835750579834
tensor([[110.1630]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[110.16301]
[107.]
epoch: 92 loss: 10.004630088806152
tensor([[86.0644]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[86.06439]
[83.]
epoch: 93 loss: 9.390499114990234
tensor([[8.4782]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[8.478189]
[2.]
epoch: 94 loss: 41.96693801879883
tensor([[15.8073]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[15.807286]
[13.]
epoch: 95 loss: 7.880856037139893
tensor([[203.8230]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[203.82295]
[203.]
epoch: 96 loss: 0.6772504448890686
tensor([[73.8421]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[73.8421]
[76.]
epoch: 97 loss: 4.656523704528809
tensor([[31.5350]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[31.535006]
[33.]

```

```

epoch: 98 loss: 2.1462087631225586
tensor([[65.3005]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[65.30046]
[69.]
epoch: 99 loss: 13.686590194702148
tensor([[134.8812]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[134.88116]
[139.]
epoch: 100 loss: 16.964805603027344
tensor([[166.1941]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[166.19406]
[168.]
epoch: 101 loss: 3.2614147663116455
tensor([[30.9692]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[30.96922]
[32.]
epoch: 102 loss: 1.0625090599060059
tensor([[142.6524]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[142.6524]
[140.]
epoch: 103 loss: 7.035251140594482
tensor([[176.9025]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[176.90254]
[173.]
epoch: 104 loss: 15.22983455657959
tensor([[8.6583]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[8.658308]
[14.]
epoch: 105 loss: 28.533672332763672
tensor([[178.9662]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[178.96623]
[176.]
epoch: 106 loss: 8.798534393310547
tensor([[40.9960]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[40.99597]
[43.]
epoch: 107 loss: 4.016129493713379
tensor([[86.5378]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')

```

```

[86.537834]
[89.]
epoch: 108 loss: 6.062260627746582
tensor([[171.6374]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[171.6374]
[175.]
epoch: 109 loss: 11.307042121887207
tensor([[196.1561]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[196.15611]
[198.]
epoch: 110 loss: 3.3999204635620117
tensor([[182.9674]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[182.96739]
[182.]
epoch: 111 loss: 0.9358472228050232
tensor([[10.1108]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[10.1107645]
[10.]
epoch: 112 loss: 0.012268775142729282
tensor([[6.8741]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[6.8741]
[7.]
epoch: 113 loss: 0.01585075818002224
tensor([[40.9223]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[40.922333]
[39.]
epoch: 114 loss: 3.6953632831573486
tensor([[113.0025]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[113.00254]
[110.]
epoch: 115 loss: 9.015250205993652
tensor([[39.3747]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[39.37466]
[37.]
epoch: 116 loss: 5.639012336730957
tensor([[136.7696]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[136.76959]
[134.]
epoch: 117 loss: 7.6706414222717285

```

```

tensor([[31.4540]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[31.45402]
[30.]
epoch: 118 loss: 2.11417293548584
tensor([[79.6765]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[79.67648]
[80.]
epoch: 119 loss: 0.10466314852237701
tensor([[196.5144]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[196.5144]
[199.]
epoch: 120 loss: 6.178185939788818
tensor([[4.3161]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[4.316103]
[3.]
epoch: 121 loss: 1.732127070426941
tensor([[92.7667]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[92.76667]
[95.]
epoch: 122 loss: 4.987761974334717
tensor([[140.7059]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[140.70593]
[143.]
epoch: 123 loss: 5.262745380401611
tensor([[97.3805]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[97.38047]
[98.]
epoch: 124 loss: 0.38381707668304443
tensor([[204.9232]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[204.9232]
[201.]
epoch: 125 loss: 15.391517639160156
tensor([[55.6694]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[55.669434]
[55.]
epoch: 126 loss: 0.44814133644104004
tensor([[114.5777]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[114.57773]

```

```

[116.]
epoch: 127 loss: 2.0228569507598877
tensor([[144.0550]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[144.05502]
[146.]
epoch: 128 loss: 3.7829346656799316
tensor([[89.0233]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[89.02327]
[90.]
epoch: 129 loss: 0.9540021419525146
tensor([[191.9426]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[191.94261]
[191.]
epoch: 130 loss: 0.8885167837142944
tensor([[10.2308]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[10.230827]
[12.]
epoch: 131 loss: 3.129971981048584
tensor([[97.5031]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[97.5031]
[97.]
epoch: 132 loss: 0.25310713052749634
tensor([[165.9491]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[165.94907]
[165.]
epoch: 133 loss: 0.9007265567779541
tensor([[125.6122]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[125.612236]
[125.]
epoch: 134 loss: 0.37483295798301697
tensor([[111.0516]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[111.05155]
[112.]
epoch: 135 loss: 0.8995539546012878
tensor([[194.9862]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[194.98618]
[196.]
epoch: 136 loss: 1.0278400182724
tensor([[178.8743]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],

```

```

device='cuda:0')
[178.87428]
[179.]
epoch: 137 loss: 0.015804804861545563
tensor([[144.1044]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[144.10439]
[144.]
epoch: 138 loss: 0.010896306484937668
tensor([[99.9423]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[99.94228]
[100.]
epoch: 139 loss: 0.003331179264932871
tensor([[120.5210]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[120.52101]
[120.]
epoch: 140 loss: 0.2714528441429138
tensor([[6.3771]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[6.3770695]
[5.]
epoch: 141 loss: 1.8963203430175781
tensor([[152.4090]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[152.40901]
[152.]
epoch: 142 loss: 0.16729068756103516
tensor([[75.9685]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[75.968475]
[75.]
epoch: 143 loss: 0.9379444718360901
tensor([[129.4231]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[129.42307]
[129.]
epoch: 144 loss: 0.1789841502904892
tensor([[166.5577]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[166.55768]
[167.]
epoch: 145 loss: 0.19564855098724365
tensor([[23.2461]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[23.246128]
[24.]

```



```

epoch: 146 loss: 0.5683228969573975
tensor([[171.9464]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[171.94637]
[174.]
epoch: 147 loss: 4.2174153327941895
tensor([[99.6072]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[99.60717]
[99.]
epoch: 148 loss: 0.3686555325984955
tensor([[47.7246]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[47.724556]
[47.]
epoch: 149 loss: 0.5249813795089722
tensor([[134.2710]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[134.27095]
[133.]
epoch: 150 loss: 1.6153147220611572
tensor([[35.2552]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[35.25518]
[35.]
epoch: 151 loss: 0.0651170164346695
tensor([[186.9109]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[186.91092]
[184.]
epoch: 152 loss: 8.473450660705566
tensor([[47.9128]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[47.912823]
[48.]
epoch: 153 loss: 0.007599877659231424
tensor([[91.1981]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[91.19807]
[93.]
epoch: 154 loss: 3.2469635009765625
tensor([[128.9745]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[128.97453]
[132.]
epoch: 155 loss: 9.153450012207031
tensor([[193.2098]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')

```

```

[193.20982]
[197.]
epoch: 156 loss: 14.365437507629395
tensor([[110.2749]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[110.27487]
[111.]
epoch: 157 loss: 0.5258108973503113
tensor([[52.4793]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[52.47925]
[51.]
epoch: 158 loss: 2.1881861686706543
tensor([[152.1841]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[152.1841]
[148.]
epoch: 159 loss: 17.506669998168945
tensor([[28.8637]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[28.863659]
[28.]
epoch: 160 loss: 0.7459067106246948
tensor([[26.6138]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[26.6138]
[26.]
epoch: 161 loss: 0.37675049901008606
tensor([[145.8889]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[145.88889]
[142.]
epoch: 162 loss: 15.123430252075195
tensor([[1.8195]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[1.8194649]
[0.]
epoch: 163 loss: 3.310452699661255
tensor([[7.1138]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[7.113846]
[9.]
epoch: 164 loss: 3.557577610015869
tensor([[105.4872]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[105.48715]
[108.]
epoch: 165 loss: 6.314404487609863

```

```

tensor([[184.4446]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[184.44456]
[186.]
epoch: 166 loss: 2.4193785190582275
tensor([[4.9950]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[4.9949527]
[8.]
epoch: 167 loss: 9.030309677124023
tensor([[9.4137]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[9.413701]
[11.]
epoch: 168 loss: 2.5163443088531494
tensor([[56.2213]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[56.221306]
[57.]
epoch: 169 loss: 0.6063646078109741
tensor([[66.2831]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[66.28312]
[67.]
epoch: 170 loss: 0.5139181017875671
tensor([[153.6229]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[153.62292]
[153.]
epoch: 171 loss: 0.38803529739379883
tensor([[162.0110]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[162.01096]
[160.]
epoch: 172 loss: 4.043943405151367
tensor([[190.4933]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[190.4933]
[189.]
epoch: 173 loss: 2.2299489974975586
tensor([[8.2279]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[8.227915]
[6.]
epoch: 174 loss: 4.96360445022583
tensor([[35.3496]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[35.34964]

```

```

[36.]
epoch: 175 loss: 0.4229682683944702
tensor([[77.4609]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[77.46086]
[79.]
epoch: 176 loss: 2.368948221206665
tensor([[56.9830]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[56.982964]
[58.]
epoch: 177 loss: 1.0343631505966187
tensor([[93.5771]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[93.57715]
[96.]
epoch: 178 loss: 5.870209693908691
tensor([[18.9687]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[18.968748]
[21.]
epoch: 179 loss: 4.125984191894531
tensor([[89.6772]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[89.67721]
[91.]
epoch: 180 loss: 1.7497788667678833
tensor([[179.4080]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[179.40804]
[183.]
epoch: 181 loss: 12.902210235595703
tensor([[193.1037]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[193.10371]
[194.]
epoch: 182 loss: 0.8033286333084106
tensor([[147.0877]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[147.08766]
[145.]
epoch: 183 loss: 4.358331680297852
tensor([[82.9788]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[82.97882]
[78.]
epoch: 184 loss: 24.78865623474121
tensor([[173.6713]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],

```

```

device='cuda:0')
[173.67126]
[169.]
epoch: 185 loss: 21.82071304321289
tensor([[105.2168]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[105.21682]
[102.]
epoch: 186 loss: 10.347929000854492
For epoch 15 the test accuracy over the whole test set is 0 %
tensor([[99.8918]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[99.891785]
[100.]
epoch: 0 loss: 0.011710558086633682
tensor([[180.3593]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[180.35933]
[185.]
epoch: 1 loss: 21.535825729370117
tensor([[61.7783]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[61.77825]
[61.]
epoch: 2 loss: 0.6056756377220154
tensor([[38.4201]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[38.42005]
[38.]
epoch: 3 loss: 0.17644332349300385
tensor([[136.8789]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[136.8789]
[142.]
epoch: 4 loss: 26.225601196289062
tensor([[85.6488]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[85.64882]
[86.]
epoch: 5 loss: 0.12332811951637268
tensor([[145.6520]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[145.65196]
[147.]
epoch: 6 loss: 1.8172056674957275
tensor([[188.9082]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[188.90817]

```

```

[184.]
epoch: 7 loss: 24.090158462524414
tensor([[75.4703]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[75.47028]
[74.]
epoch: 8 loss: 2.161733627319336
tensor([[89.0010]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[89.00104]
[88.]
epoch: 9 loss: 1.002076268196106
tensor([[192.4689]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[192.4689]
[192.]
epoch: 10 loss: 0.21986964344978333
tensor([[168.6377]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[168.63774]
[172.]
epoch: 11 loss: 11.304784774780273
tensor([[168.8801]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[168.88013]
[171.]
epoch: 12 loss: 4.493861675262451
tensor([[19.0322]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[19.032244]
[19.]
epoch: 13 loss: 0.001039658091031015
tensor([[156.3824]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[156.38245]
[155.]
epoch: 14 loss: 1.9111577272415161
tensor([[183.4764]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[183.47636]
[180.]
epoch: 15 loss: 12.085107803344727
tensor([[69.7783]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[69.77828]
[69.]
epoch: 16 loss: 0.6057231426239014
tensor([[179.2260]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],

```

```

device='cuda:0')
[179.22597]
[179.]
epoch: 17 loss: 0.05106126889586449
tensor([[19.4916]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[19.491617]
[21.]
epoch: 18 loss: 2.2752187252044678
tensor([[102.1735]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[102.17354]
[105.]
epoch: 19 loss: 7.98888635635376
tensor([[18.6205]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[18.620522]
[20.]
epoch: 20 loss: 1.9029607772827148
tensor([[35.3570]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[35.356953]
[36.]
epoch: 21 loss: 0.41350987553596497
tensor([[119.2581]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[119.258064]
[122.]
epoch: 22 loss: 7.518211364746094
tensor([[174.3626]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[174.36256]
[177.]
epoch: 23 loss: 6.95606803894043
tensor([[111.3514]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[111.3514]
[111.]
epoch: 24 loss: 0.1234835609793663
tensor([[96.6127]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[96.61266]
[94.]
epoch: 25 loss: 6.826009273529053
tensor([[199.1917]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[199.19165]
[193.]

```

```

epoch: 26 loss: 38.3365364074707
tensor([[136.1358]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[136.13582]
[135.]
epoch: 27 loss: 1.290083646774292
tensor([[28.3303]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[28.330284]
[27.]
epoch: 28 loss: 1.7696558237075806
tensor([[53.0656]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[53.06557]
[53.]
epoch: 29 loss: 0.004299534019082785
tensor([[77.4367]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[77.43666]
[81.]
epoch: 30 loss: 12.697386741638184
tensor([[124.8019]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[124.801865]
[133.]
epoch: 31 loss: 67.20942687988281
tensor([[104.7688]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[104.76878]
[110.]
epoch: 32 loss: 27.365625381469727
tensor([[84.5820]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[84.58199]
[85.]
epoch: 33 loss: 0.17472976446151733
tensor([[202.2968]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[202.29678]
[203.]
epoch: 34 loss: 0.49451351165771484
tensor([[62.9199]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[62.919933]
[58.]
epoch: 35 loss: 24.20574378967285
tensor([[103.6710]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')

```



```

[103.67103]
[97.]
epoch: 36 loss: 44.50261688232422
tensor([[55.5580]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[55.557983]
[50.]
epoch: 37 loss: 30.89118003845215
tensor([[49.3857]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[49.38567]
[46.]
epoch: 38 loss: 11.462759017944336
tensor([[101.9483]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[101.94833]
[99.]
epoch: 39 loss: 8.692671775817871
tensor([[93.2022]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[93.20222]
[93.]
epoch: 40 loss: 0.04089175537228584
tensor([[127.0875]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[127.08746]
[129.]
epoch: 41 loss: 3.6577963829040527
tensor([[83.6018]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[83.601845]
[87.]
epoch: 42 loss: 11.54745864868164
tensor([[97.3579]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[97.35795]
[101.]
epoch: 43 loss: 13.264540672302246
tensor([[139.7398]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[139.73976]
[143.]
epoch: 44 loss: 10.629156112670898
tensor([[11.5802]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[11.580217]
[9.]
epoch: 45 loss: 6.657521724700928

```

```

tensor([[76.3298]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[76.32979]
[78.]
epoch: 46 loss: 2.789607524871826
tensor([[181.6013]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[181.60126]
[178.]
epoch: 47 loss: 12.969054222106934
tensor([[119.9876]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[119.987625]
[118.]
epoch: 48 loss: 3.950653553009033
tensor([[159.7873]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[159.78732]
[157.]
epoch: 49 loss: 7.769169330596924
tensor([[199.4867]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[199.48671]
[200.]
epoch: 50 loss: 0.26346704363822937
tensor([[130.7721]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[130.7721]
[132.]
epoch: 51 loss: 1.5077513456344604
tensor([[37.5389]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[37.538944]
[43.]
epoch: 52 loss: 29.823129653930664
tensor([[110.7925]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[110.792465]
[114.]
epoch: 53 loss: 10.28827953338623
tensor([[52.4466]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[52.446564]
[56.]
epoch: 54 loss: 12.626909255981445
tensor([[203.2985]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[203.29846]

```

```

[204.]
epoch: 55 loss: 0.49215567111968994
tensor([[145.6594]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[145.65944]
[144.]
epoch: 56 loss: 2.7537381649017334
tensor([[70.0391]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[70.03907]
[70.]
epoch: 57 loss: 0.001526474952697754
tensor([[78.3380]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[78.33803]
[77.]
epoch: 58 loss: 1.7903188467025757
tensor([[171.7949]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[171.79488]
[168.]
epoch: 59 loss: 14.401084899902344
tensor([[121.4611]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[121.4611]
[119.]
epoch: 60 loss: 6.057002067565918
tensor([[174.0112]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[174.01123]
[176.]
epoch: 61 loss: 3.9552042484283447
tensor([[133.4835]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[133.48346]
[134.]
epoch: 62 loss: 0.26681411266326904
tensor([[13.2774]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[13.277412]
[13.]
epoch: 63 loss: 0.07695765048265457
tensor([[147.3197]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[147.31966]
[152.]
epoch: 64 loss: 21.905616760253906
tensor([[89.4648]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],

```

```

device='cuda:0')
[89.4648]
[90.]
epoch: 65 loss: 0.28644120693206787
tensor([[8.6059]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[8.605893]
[7.]
epoch: 66 loss: 2.578892707824707
tensor([[167.2086]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[167.20856]
[167.]
epoch: 67 loss: 0.04349607601761818
tensor([[69.3759]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[69.37593]
[68.]
epoch: 68 loss: 1.8931854963302612
tensor([[195.8849]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[195.88493]
[194.]
epoch: 69 loss: 3.552974224090576
tensor([[164.8889]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[164.8889]
[163.]
epoch: 70 loss: 3.567945957183838
tensor([[24.5393]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[24.53929]
[25.]
epoch: 71 loss: 0.21225418150424957
tensor([[37.9276]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[37.927612]
[39.]
epoch: 72 loss: 1.1500153541564941
tensor([[107.2125]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[107.212524]
[108.]
epoch: 73 loss: 0.6201177835464478
tensor([[161.9359]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[161.93591]
[166.]

```

```

epoch: 74 loss: 16.516801834106445
tensor([[27.8923]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[27.892298]
[28.]
epoch: 75 loss: 0.011599775403738022
tensor([[146.0356]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[146.0356]
[146.]
epoch: 76 loss: 0.0012672713492065668
tensor([[33.7878]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[33.787758]
[33.]
epoch: 77 loss: 0.6205624938011169
tensor([[142.7713]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[142.77133]
[139.]
epoch: 78 loss: 14.222943305969238
tensor([[94.0888]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[94.0888]
[92.]
epoch: 79 loss: 4.363079071044922
tensor([[202.7690]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[202.76895]
[205.]
epoch: 80 loss: 4.9775776863098145
tensor([[96.0149]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[96.014915]
[95.]
epoch: 81 loss: 1.0300533771514893
tensor([[120.9979]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[120.9979]
[120.]
epoch: 82 loss: 0.9958082437515259
tensor([[153.3602]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[153.36018]
[154.]
epoch: 83 loss: 0.40936487913131714
tensor([[70.1309]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')

```

```

[70.13087]
[71.]
epoch: 84 loss: 0.7553921937942505
tensor([[159.3278]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[159.32776]
[160.]
epoch: 85 loss: 0.4519082307815552
tensor([[12.3905]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[12.390494]
[15.]
epoch: 86 loss: 6.8095197677612305
tensor([[147.5835]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[147.58354]
[148.]
epoch: 87 loss: 0.17343737185001373
tensor([[7.3440]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[7.3440223]
[8.]
epoch: 88 loss: 0.4303067624568939
tensor([[16.3074]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[16.307413]
[18.]
epoch: 89 loss: 2.8648505210876465
tensor([[83.2866]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[83.2866]
[82.]
epoch: 90 loss: 1.6553349494934082
tensor([[26.6376]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[26.637648]
[26.]
epoch: 91 loss: 0.40659448504447937
tensor([[13.4676]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[13.467625]
[11.]
epoch: 92 loss: 6.089171409606934
tensor([[67.6129]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[67.61292]
[67.]
epoch: 93 loss: 0.37567418813705444

```

```

tensor([[104.4774]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[104.4774]
[102.]
epoch: 94 loss: 6.137519359588623
tensor([[174.7878]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[174.78781]
[174.]
epoch: 95 loss: 0.620646595954895
tensor([[50.2982]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[50.298164]
[51.]
epoch: 96 loss: 0.49257326126098633
tensor([[54.0158]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[54.015812]
[55.]
epoch: 97 loss: 0.9686262011528015
tensor([[110.6153]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[110.615295]
[112.]
epoch: 98 loss: 1.9174067974090576
tensor([[107.3004]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[107.30042]
[109.]
epoch: 99 loss: 2.8885631561279297
tensor([[105.1536]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[105.153595]
[106.]
epoch: 100 loss: 0.7164014577865601
tensor([[198.6632]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[198.66321]
[199.]
epoch: 101 loss: 0.11342884600162506
tensor([[14.3367]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[14.336666]
[14.]
epoch: 102 loss: 0.11334406584501266
tensor([[184.4762]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[184.47615]

```

```

[183.]
epoch: 103 loss: 2.179020404815674
tensor([[44.1847]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[44.18474]
[45.]
epoch: 104 loss: 0.6646518707275391
tensor([[186.7214]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[186.72136]
[186.]
epoch: 105 loss: 0.5203591585159302
tensor([[7.1323]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[7.1322775]
[4.]
epoch: 106 loss: 9.811161994934082
tensor([[7.8592]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[7.859225]
[6.]
epoch: 107 loss: 3.456716775894165
tensor([[45.5106]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[45.51063]
[47.]
epoch: 108 loss: 2.2182183265686035
tensor([[153.9007]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[153.9007]
[156.]
epoch: 109 loss: 4.407078266143799
tensor([[126.5213]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[126.52133]
[126.]
epoch: 110 loss: 0.2717868387699127
tensor([[3.7553]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[3.7553368]
[3.]
epoch: 111 loss: 0.5705336332321167
tensor([[88.2065]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[88.20645]
[89.]
epoch: 112 loss: 0.6297193765640259
tensor([[194.0647]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],

```



```

device='cuda:0')
[194.0647]
[195.]
epoch: 113 loss: 0.8747912049293518
tensor([[126.7958]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[126.795845]
[124.]
epoch: 114 loss: 7.816749572753906
tensor([[183.2699]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[183.2699]
[182.]
epoch: 115 loss: 1.6126395463943481
tensor([[165.1064]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[165.10645]
[165.]
epoch: 116 loss: 0.011330604553222656
tensor([[46.3256]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[46.325645]
[48.]
epoch: 117 loss: 2.8034632205963135
tensor([[33.1534]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[33.153397]
[35.]
epoch: 118 loss: 3.4099440574645996
tensor([[58.4120]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[58.41204]
[60.]
epoch: 119 loss: 2.5216147899627686
tensor([[193.1377]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[193.13771]
[196.]
epoch: 120 loss: 8.192700386047363
tensor([[168.7300]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[168.73004]
[169.]
epoch: 121 loss: 0.07287758588790894
tensor([[126.4832]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[126.48319]
[123.]

```

```

epoch: 122 loss: 12.13262939453125
tensor([[177.3924]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[177.39244]
[175.]
epoch: 123 loss: 5.723773002624512
tensor([[21.6336]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[21.63362]
[24.]
epoch: 124 loss: 5.599757671356201
tensor([[29.8066]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[29.806576]
[30.]
epoch: 125 loss: 0.03741293027997017
tensor([[144.5209]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[144.5209]
[145.]
epoch: 126 loss: 0.22953246533870697
tensor([[32.3086]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[32.308624]
[32.]
epoch: 127 loss: 0.09524893760681152
tensor([[147.8842]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[147.88423]
[150.]
epoch: 128 loss: 4.476476192474365
tensor([[115.3930]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[115.39296]
[115.]
epoch: 129 loss: 0.1544172465801239
tensor([[117.3672]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[117.36723]
[117.]
epoch: 130 loss: 0.13486027717590332
tensor([[113.7623]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[113.76226]
[113.]
epoch: 131 loss: 0.5810409784317017
tensor([[62.8128]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')

```

```

[62.812786]
[62.]
epoch: 132 loss: 0.6606212258338928
tensor([[0.9075]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[0.9075364]
[2.]
epoch: 133 loss: 1.1934767961502075
tensor([[1.0941]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[1.0941247]
[1.]
epoch: 134 loss: 0.008859454654157162
tensor([[168.2118]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[168.2118]
[170.]
epoch: 135 loss: 3.1976332664489746
tensor([[35.4934]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[35.49343]
[34.]
epoch: 136 loss: 2.2303364276885986
tensor([[22.1050]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[22.104996]
[23.]
epoch: 137 loss: 0.8010326623916626
tensor([[66.1165]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[66.1165]
[65.]
epoch: 138 loss: 1.2465741634368896
tensor([[99.9076]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[99.90756]
[98.]
epoch: 139 loss: 3.638793706893921
tensor([[77.1521]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[77.15207]
[76.]
epoch: 140 loss: 1.3272632360458374
tensor([[80.4297]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[80.42967]
[80.]
epoch: 141 loss: 0.18461823463439941

```

```

tensor([[49.4374]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[49.437424]
[49.]
epoch: 142 loss: 0.19133949279785156
tensor([[124.8182]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[124.81815]
[125.]
epoch: 143 loss: 0.03306819126009941
tensor([[10.5991]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[10.59906]
[12.]
epoch: 144 loss: 1.9626327753067017
tensor([[186.1975]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[186.19754]
[189.]
epoch: 145 loss: 7.853780269622803
tensor([[139.6276]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[139.62758]
[140.]
epoch: 146 loss: 0.13869759440422058
tensor([[190.9297]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[190.92972]
[190.]
epoch: 147 loss: 0.8643755912780762
tensor([[5.7866]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[5.7866073]
[5.]
epoch: 148 loss: 0.6187509894371033
tensor([[205.9353]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[205.9353]
[201.]
epoch: 149 loss: 24.357213973999023
tensor([[196.5908]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[196.59082]
[197.]
epoch: 150 loss: 0.16742801666259766
tensor([[160.5925]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[160.59247]

```

```

[164.]
epoch: 151 loss: 11.611272811889648
tensor([[100.9434]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[100.943375]
[104.]
epoch: 152 loss: 9.342958450317383
tensor([[42.1457]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[42.145744]
[44.]
epoch: 153 loss: 3.4382641315460205
tensor([[189.1934]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[189.19339]
[191.]
epoch: 154 loss: 3.2638401985168457
tensor([[139.1855]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[139.18549]
[138.]
epoch: 155 loss: 1.405376672744751
tensor([[173.8699]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[173.86989]
[173.]
epoch: 156 loss: 0.756705641746521
tensor([[84.4905]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[84.49047]
[83.]
epoch: 157 loss: 2.221503496170044
tensor([[58.3622]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[58.36218]
[57.]
epoch: 158 loss: 1.855531096458435
tensor([[156.3851]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[156.38513]
[153.]
epoch: 159 loss: 11.459117889404297
tensor([[11.6443]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[11.644313]
[10.]
epoch: 160 loss: 2.7037646770477295
tensor([[140.2621]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],

```

```

device='cuda:0')
[140.26207]
[141.]
epoch: 161 loss: 0.5445411205291748
tensor([[105.1288]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[105.12878]
[107.]
epoch: 162 loss: 3.5014772415161133
tensor([[159.8882]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[159.88821]
[162.]
epoch: 163 loss: 4.459639549255371
tensor([[20.0563]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[20.056305]
[22.]
epoch: 164 loss: 3.7779505252838135
tensor([[73.4497]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[73.44968]
[75.]
epoch: 165 loss: 2.403502941131592
tensor([[137.6738]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[137.67377]
[137.]
epoch: 166 loss: 0.4539620876312256
tensor([[83.9180]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[83.918015]
[84.]
epoch: 167 loss: 0.00672161765396595
tensor([[52.9621]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[52.962063]
[52.]
epoch: 168 loss: 0.9255648851394653
tensor([[132.1179]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[132.11787]
[130.]
epoch: 169 loss: 4.485390663146973
tensor([[151.5645]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[151.56454]
[151.]

```

```

epoch: 170 loss: 0.31871068477630615
tensor([[199.1460]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[199.14597]
[198.]
epoch: 171 loss: 1.3132370710372925
tensor([[62.6393]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[62.639256]
[64.]
epoch: 172 loss: 1.8516255617141724
tensor([[3.0399]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[3.0399184]
[0.]
epoch: 173 loss: 9.241104125976562
tensor([[14.3762]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[14.376199]
[17.]
epoch: 174 loss: 6.88433313369751
tensor([[76.6700]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[76.670044]
[79.]
epoch: 175 loss: 5.428695201873779
tensor([[112.9800]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[112.979965]
[116.]
epoch: 176 loss: 9.120610237121582
tensor([[101.4780]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[101.478035]
[103.]
epoch: 177 loss: 2.316377639770508
tensor([[95.8510]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[95.850975]
[96.]
epoch: 178 loss: 0.0222084391862154
tensor([[159.4293]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[159.42926]
[159.]
epoch: 179 loss: 0.18426436185836792
tensor([[16.5068]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')

```

```

[16.506807]
[16.]
epoch: 180 loss: 0.2568536698818207
tensor([[150.9898]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[150.98984]
[149.]
epoch: 181 loss: 3.959453821182251
tensor([[93.8756]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[93.875626]
[91.]
epoch: 182 loss: 8.269222259521484
tensor([[39.1640]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[39.163967]
[37.]
epoch: 183 loss: 4.682753562927246
tensor([[136.4693]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[136.46931]
[136.]
epoch: 184 loss: 0.22025616466999054
tensor([[30.5384]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[30.538448]
[29.]
epoch: 185 loss: 2.366823196411133
tensor([[54.9203]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[54.920334]
[54.]
epoch: 186 loss: 0.8470144271850586
For epoch 16 the test accuracy over the whole test set is 0 %
tensor([[182.1463]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[182.14632]
[190.]
epoch: 0 loss: 61.68034362792969
tensor([[153.1819]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[153.18188]
[155.]
epoch: 1 loss: 3.3055429458618164
tensor([[83.3294]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[83.3294]
[81.]

```



```

epoch: 2 loss: 5.426100254058838
tensor([[72.1802]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[72.18021]
[69.]
epoch: 3 loss: 10.113712310791016
tensor([[39.7983]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[39.798325]
[38.]
epoch: 4 loss: 3.233971357345581
tensor([[88.5801]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[88.58007]
[84.]
epoch: 5 loss: 20.9770450592041
tensor([[10.6062]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[10.60619]
[11.]
epoch: 6 loss: 0.15508653223514557
tensor([[0.9796]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[0.97962373]
[0.]
epoch: 7 loss: 0.959662675857544
tensor([[148.4972]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[148.49718]
[140.]
epoch: 8 loss: 72.20201873779297
tensor([[106.1198]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[106.119835]
[105.]
epoch: 9 loss: 1.2540302276611328
tensor([[200.2824]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[200.28241]
[198.]
epoch: 10 loss: 5.2093939781188965
tensor([[164.4960]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[164.49597]
[171.]
epoch: 11 loss: 42.30238342285156
tensor([[20.0779]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')

```

```

[20.077868]
[26.]
epoch: 12 loss: 35.071651458740234
tensor([[4.7698]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[4.769781]
[12.]
epoch: 13 loss: 52.276065826416016
tensor([[177.1324]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[177.13242]
[185.]
epoch: 14 loss: 61.89888000488281
tensor([[192.6232]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[192.62323]
[195.]
epoch: 15 loss: 5.649035930633545
tensor([[28.3068]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[28.306839]
[32.]
epoch: 16 loss: 13.63943862915039
tensor([[220.0992]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[220.09918]
[204.]
epoch: 17 loss: 259.18365478515625
tensor([[12.2290]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[12.228991]
[14.]
epoch: 18 loss: 3.136474370956421
tensor([[50.5008]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[50.500755]
[57.]
epoch: 19 loss: 42.24018096923828
tensor([[73.1274]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[73.1274]
[82.]
epoch: 20 loss: 78.72297668457031
tensor([[34.2630]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[34.262978]
[39.]
epoch: 21 loss: 22.439380645751953

```

```

tensor([[100.3584]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[100.3584]
[109.]
epoch: 22 loss: 74.67727661132812
tensor([[46.8912]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[46.891155]
[48.]
epoch: 23 loss: 1.2295366525650024
tensor([[141.4885]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[141.48853]
[147.]
epoch: 24 loss: 30.376352310180664
tensor([[109.1367]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[109.13673]
[107.]
epoch: 25 loss: 4.56559944152832
tensor([[22.0254]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[22.025368]
[9.]
epoch: 26 loss: 169.6602020263672
tensor([[15.8804]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[15.880446]
[1.]
epoch: 27 loss: 221.4276885986328
tensor([[24.3765]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[24.376543]
[16.]
epoch: 28 loss: 70.16647338867188
tensor([[144.3752]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[144.37521]
[138.]
epoch: 29 loss: 40.643348693847656
tensor([[119.1299]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[119.12986]
[115.]
epoch: 30 loss: 17.055742263793945
tensor([[3.9707]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[3.970694]

```

```

[4.]
epoch: 31 loss: 0.0008588377968408167
tensor([[81.8444]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[81.84444]
[85.]
epoch: 32 loss: 9.957579612731934
tensor([[144.7906]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[144.79059]
[156.]
epoch: 33 loss: 125.65090942382812
tensor([[54.4826]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[54.482616]
[56.]
epoch: 34 loss: 2.302452802658081
tensor([[84.4764]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[84.47638]
[87.]
epoch: 35 loss: 6.368660926818848
tensor([[171.1071]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[171.10709]
[174.]
epoch: 36 loss: 8.368949890136719
tensor([[172.1076]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[172.10759]
[168.]
epoch: 37 loss: 16.87229347229004
tensor([[42.8828]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[42.882847]
[43.]
epoch: 38 loss: 0.013724864460527897
tensor([[69.9988]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[69.99882]
[67.]
epoch: 39 loss: 8.992905616760254
tensor([[143.4918]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[143.49176]
[136.]
epoch: 40 loss: 56.12647247314453
tensor([[ -2.6135]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],

```

```

device='cuda:0')
[-2.6135335]
[13.]
epoch: 41 loss: 243.78240966796875
tensor([[84.5049]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[84.50491]
[83.]
epoch: 42 loss: 2.2647640705108643
tensor([[163.5034]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[163.50343]
[162.]
epoch: 43 loss: 2.2603113651275635
tensor([[155.4085]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[155.40852]
[157.]
epoch: 44 loss: 2.532797336578369
tensor([[5.9311]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[5.931055]
[18.]
epoch: 45 loss: 145.65943908691406
tensor([[93.5636]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[93.563576]
[96.]
epoch: 46 loss: 5.936162948608398
tensor([[26.9899]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[26.989902]
[29.]
epoch: 47 loss: 4.040492057800293
tensor([[142.9747]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[142.97466]
[148.]
epoch: 48 loss: 25.254091262817383
tensor([[5.9812]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[5.9811716]
[5.]
epoch: 49 loss: 0.9626977443695068
tensor([[121.8018]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[121.80184]
[120.]

```

```

epoch: 50 loss: 3.246633529663086
tensor([[120.6577]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[120.657715]
[117.]
epoch: 51 loss: 13.378877639770508
tensor([[173.1196]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[173.11958]
[173.]
epoch: 52 loss: 0.01430012471973896
tensor([[57.8565]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[57.856503]
[53.]
epoch: 53 loss: 23.585617065429688
tensor([[178.4378]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[178.43779]
[180.]
epoch: 54 loss: 2.440500259399414
tensor([[13.1714]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[13.171352]
[8.]
epoch: 55 loss: 26.74288558959961
tensor([[185.9245]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[185.92448]
[189.]
epoch: 56 loss: 9.458797454833984
tensor([[54.7121]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[54.712112]
[52.]
epoch: 57 loss: 7.35555362701416
tensor([[12.8326]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[12.832567]
[7.]
epoch: 58 loss: 34.01884078979492
tensor([[9.3790]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[9.379029]
[2.]
epoch: 59 loss: 54.4500732421875
tensor([[25.8766]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')

```

```

[25.876644]
[27.]
epoch: 60 loss: 1.2619284391403198
tensor([[151.9047]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[151.9047]
[151.]
epoch: 61 loss: 0.8184705376625061
tensor([[132.2543]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[132.25432]
[133.]
epoch: 62 loss: 0.5560413002967834
tensor([[168.1271]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[168.12706]
[167.]
epoch: 63 loss: 1.2702641487121582
tensor([[12.4087]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[12.40868]
[17.]
epoch: 64 loss: 21.080219268798828
tensor([[65.0656]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[65.06555]
[70.]
epoch: 65 loss: 24.348779678344727
tensor([[56.6193]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[56.619278]
[61.]
epoch: 66 loss: 19.190725326538086
tensor([[134.1640]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[134.16403]
[134.]
epoch: 67 loss: 0.026906490325927734
tensor([[17.0929]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[17.092903]
[22.]
epoch: 68 loss: 24.079599380493164
tensor([[133.2479]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[133.24794]
[130.]
epoch: 69 loss: 10.549114227294922

```

```

tensor([[101.5601]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[101.56014]
[99.]
epoch: 70 loss: 6.554329872131348
tensor([[195.6256]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[195.62563]
[191.]
epoch: 71 loss: 21.396411895751953
tensor([[159.7934]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[159.79338]
[159.]
epoch: 72 loss: 0.6294530034065247
tensor([[67.9500]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[67.94996]
[71.]
epoch: 73 loss: 9.302751541137695
tensor([[20.9633]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[20.96331]
[24.]
epoch: 74 loss: 9.221485137939453
tensor([[116.8909]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[116.890945]
[125.]
epoch: 75 loss: 65.75676727294922
tensor([[137.6749]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[137.67491]
[146.]
epoch: 76 loss: 69.30709838867188
tensor([[195.2964]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[195.29642]
[205.]
epoch: 77 loss: 94.15951538085938
tensor([[188.9228]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[188.92284]
[179.]
epoch: 78 loss: 98.46267700195312
tensor([[137.7890]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[137.78905]

```



```

[126.]
epoch: 79 loss: 138.98162841796875
tensor([[152.9908]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[152.99083]
[143.]
epoch: 80 loss: 99.8166732788086
tensor([[114.4226]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[114.422585]
[112.]
epoch: 81 loss: 5.868916034698486
tensor([[135.7720]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[135.772]
[142.]
epoch: 82 loss: 38.78794479370117
tensor([[170.0656]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[170.06555]
[182.]
epoch: 83 loss: 142.43106079101562
tensor([[50.0805]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[50.080532]
[49.]
epoch: 84 loss: 1.1675496101379395
tensor([[100.7023]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[100.70234]
[102.]
epoch: 85 loss: 1.683923602104187
tensor([[47.4798]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[47.479797]
[45.]
epoch: 86 loss: 6.149394989013672
tensor([[78.0403]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[78.04033]
[75.]
epoch: 87 loss: 9.243599891662598
tensor([[21.0790]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[21.079048]
[21.]
epoch: 88 loss: 0.0062486110255122185
tensor([[182.6644]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],

```

```

device='cuda:0')
[182.66437]
[176.]
epoch: 89 loss: 44.413795471191406
tensor([[53.1971]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[53.19706]
[51.]
epoch: 90 loss: 4.827071189880371
tensor([[195.8374]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[195.83737]
[197.]
epoch: 91 loss: 1.3517042398452759
tensor([[157.5780]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[157.57796]
[160.]
epoch: 92 loss: 5.866291522979736
tensor([[84.1766]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[84.176605]
[88.]
epoch: 93 loss: 14.61834716796875
tensor([[17.5291]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[17.529076]
[19.]
epoch: 94 loss: 2.163618564605713
tensor([[161.2807]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[161.28073]
[163.]
epoch: 95 loss: 2.955885171890259
tensor([[78.0769]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[78.07691]
[80.]
epoch: 96 loss: 3.698267698287964
tensor([[173.4752]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[173.47519]
[172.]
epoch: 97 loss: 2.1761832237243652
tensor([[197.0871]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[197.08714]
[193.]

```

```

epoch: 98 loss: 16.704736709594727
tensor([[109.2679]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[109.267944]
[111.]
epoch: 99 loss: 3.000016927719116
tensor([[110.3329]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[110.33292]
[113.]
epoch: 100 loss: 7.113335609436035
tensor([[66.8947]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[66.89472]
[68.]
epoch: 101 loss: 1.2216395139694214
tensor([[55.5243]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[55.524303]
[55.]
epoch: 102 loss: 0.2748940885066986
tensor([[199.7037]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[199.7037]
[199.]
epoch: 103 loss: 0.4952004849910736
tensor([[63.5259]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[63.525917]
[64.]
epoch: 104 loss: 0.22475464642047882
tensor([[75.9385]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[75.93845]
[76.]
epoch: 105 loss: 0.0037879501469433308
tensor([[98.4846]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[98.48455]
[100.]
epoch: 106 loss: 2.2965872287750244
tensor([[107.4408]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[107.440834]
[108.]
epoch: 107 loss: 0.31266656517982483
tensor([[201.7617]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')

```

```

[201.76169]
[201.]
epoch: 108 loss: 0.5801689624786377
tensor([[145.0226]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[145.02257]
[145.]
epoch: 109 loss: 0.0005093032959848642
tensor([[38.9412]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[38.94122]
[37.]
epoch: 110 loss: 3.7683324813842773
tensor([[174.6411]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[174.64108]
[175.]
epoch: 111 loss: 0.12882158160209656
tensor([[98.0268]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[98.0268]
[97.]
epoch: 112 loss: 1.0543224811553955
tensor([[120.9241]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[120.92411]
[122.]
epoch: 113 loss: 1.1575384140014648
tensor([[116.9725]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[116.97254]
[118.]
epoch: 114 loss: 1.0556703805923462
tensor([[107.3705]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[107.370514]
[106.]
epoch: 115 loss: 1.8783084154129028
tensor([[87.6137]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[87.61368]
[86.]
epoch: 116 loss: 2.603956699371338
tensor([[103.8143]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[103.81435]
[103.]
epoch: 117 loss: 0.6631599068641663

```

```

tensor([[12.8124]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[12.812388]
[6.]
epoch: 118 loss: 46.408634185791016
tensor([[31.7643]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[31.764265]
[30.]
epoch: 119 loss: 3.112631320953369
tensor([[149.4292]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[149.42921]
[153.]
epoch: 120 loss: 12.750509262084961
tensor([[128.2254]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[128.22539]
[132.]
epoch: 121 loss: 14.247698783874512
tensor([[15.7880]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[15.787997]
[15.]
epoch: 122 loss: 0.6209396719932556
tensor([[95.8569]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[95.85692]
[95.]
epoch: 123 loss: 0.7343090176582336
tensor([[36.3963]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[36.396347]
[36.]
epoch: 124 loss: 0.15709097683429718
tensor([[156.7131]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[156.71309]
[152.]
epoch: 125 loss: 22.213207244873047
tensor([[60.8761]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[60.876057]
[60.]
epoch: 126 loss: 0.7674753069877625
tensor([[57.5706]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[57.570557]

```

```

[58.]
epoch: 127 loss: 0.1844215989112854
tensor([[45.3088]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[45.308823]
[47.]
epoch: 128 loss: 2.8600809574127197
tensor([[183.4989]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[183.49887]
[183.]
epoch: 129 loss: 0.24887213110923767
tensor([[165.0787]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[165.07867]
[166.]
epoch: 130 loss: 0.8488410115242004
tensor([[32.3273]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[32.32728]
[34.]
epoch: 131 loss: 2.7979984283447266
tensor([[95.8332]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[95.83319]
[98.]
epoch: 132 loss: 4.695061683654785
tensor([[76.1792]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[76.17919]
[78.]
epoch: 133 loss: 3.3153433799743652
tensor([[133.2507]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[133.25075]
[135.]
epoch: 134 loss: 3.0598835945129395
tensor([[156.3522]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[156.3522]
[154.]
epoch: 135 loss: 5.53286075592041
tensor([[195.7224]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[195.72237]
[194.]
epoch: 136 loss: 2.966545820236206
tensor([[116.1767]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],

```

```

device='cuda:0')
[116.176674]
[116.]
epoch: 137 loss: 0.031213663518428802
tensor([[15.2809]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[15.280883]
[20.]
epoch: 138 loss: 22.27006721496582
tensor([[143.3803]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[143.38028]
[144.]
epoch: 139 loss: 0.38405343890190125
tensor([[137.7839]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[137.7839]
[139.]
epoch: 140 loss: 1.4788869619369507
tensor([[185.0547]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[185.05467]
[184.]
epoch: 141 loss: 1.1123335361480713
tensor([[149.8189]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[149.81886]
[150.]
epoch: 142 loss: 0.03281064331531525
tensor([[164.5397]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[164.53966]
[164.]
epoch: 143 loss: 0.29123032093048096
tensor([[136.0200]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[136.02003]
[137.]
epoch: 144 loss: 0.9603317975997925
tensor([[165.1352]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[165.1352]
[165.]
epoch: 145 loss: 0.018277112394571304
tensor([[30.4098]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[30.40984]
[28.]

```

```

epoch: 146 loss: 5.8073272705078125
tensor([[90.5289]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[90.528915]
[89.]
epoch: 147 loss: 2.3375823497772217
tensor([[66.7400]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[66.74001]
[65.]
epoch: 148 loss: 3.0276455879211426
tensor([[165.9042]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[165.90417]
[169.]
epoch: 149 loss: 9.584134101867676
tensor([[147.7008]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[147.70082]
[149.]
epoch: 150 loss: 1.6878663301467896
tensor([[91.6200]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[91.619995]
[90.]
epoch: 151 loss: 2.6243841648101807
tensor([[52.4443]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[52.444317]
[50.]
epoch: 152 loss: 5.974684715270996
tensor([[93.7789]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[93.778885]
[92.]
epoch: 153 loss: 3.16443133354187
tensor([[140.9755]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[140.97552]
[141.]
epoch: 154 loss: 0.0005990304052829742
tensor([[55.3104]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[55.31041]
[54.]
epoch: 155 loss: 1.7171732187271118
tensor([[93.9072]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')

```



```

[93.90725]
[94.]
epoch: 156 loss: 0.008602664805948734
tensor([[42.4066]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[42.406628]
[44.]
epoch: 157 loss: 2.5388355255126953
tensor([[184.8387]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[184.83867]
[186.]
epoch: 158 loss: 1.348690152168274
tensor([[32.3490]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[32.34902]
[33.]
epoch: 159 loss: 0.42377743124961853
tensor([[176.3988]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[176.39877]
[177.]
epoch: 160 loss: 0.3614736795425415
tensor([[123.5910]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[123.590965]
[123.]
epoch: 161 loss: 0.34923994541168213
tensor([[93.0995]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[93.09953]
[93.]
epoch: 162 loss: 0.009906833991408348
tensor([[101.8734]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[101.87337]
[101.]
epoch: 163 loss: 0.7627704739570618
tensor([[125.4388]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[125.43877]
[124.]
epoch: 164 loss: 2.0700490474700928
tensor([[72.6481]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[72.64812]
[74.]
epoch: 165 loss: 1.8275874853134155

```

```

tensor([[200.8060]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[200.80605]
[200.]
epoch: 166 loss: 0.6497094035148621
tensor([[103.6004]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[103.60043]
[104.]
epoch: 167 loss: 0.15965351462364197
tensor([[11.6912]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[11.691151]
[10.]
epoch: 168 loss: 2.8599905967712402
tensor([[19.8832]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[19.883163]
[23.]
epoch: 169 loss: 9.714670181274414
tensor([[6.8381]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[6.8381267]
[3.]
epoch: 170 loss: 14.731216430664062
tensor([[194.0827]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[194.08266]
[196.]
epoch: 171 loss: 3.6762046813964844
tensor([[112.5428]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[112.54279]
[114.]
epoch: 172 loss: 2.1234514713287354
tensor([[60.1527]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[60.152733]
[62.]
epoch: 173 loss: 3.41239595413208
tensor([[111.2459]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[111.24595]
[110.]
epoch: 174 loss: 1.5523884296417236
tensor([[23.9764]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[23.976383]

```

```

[25.]
epoch: 175 loss: 1.0477913618087769
tensor([[184.1932]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[184.19318]
[178.]
epoch: 176 loss: 38.355430603027344
tensor([[119.0944]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[119.09438]
[119.]
epoch: 177 loss: 0.008908196352422237
tensor([[75.5673]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[75.56725]
[77.]
epoch: 178 loss: 2.0527637004852295
tensor([[88.3342]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[88.33417]
[91.]
epoch: 179 loss: 7.106663227081299
tensor([[76.5181]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[76.51812]
[79.]
epoch: 180 loss: 6.159729480743408
tensor([[33.4575]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[33.457493]
[35.]
epoch: 181 loss: 2.379328489303589
tensor([[195.6450]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[195.64496]
[203.]
epoch: 182 loss: 54.09663391113281
tensor([[46.2944]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[46.294437]
[46.]
epoch: 183 loss: 0.08669339120388031
tensor([[176.4415]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[176.44147]
[170.]
epoch: 184 loss: 41.49250030517578
tensor([[133.5284]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],

```

```

device='cuda:0')
[133.52837]
[129.]
epoch: 185 loss: 20.506099700927734
tensor([[195.8188]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[195.81883]
[192.]
epoch: 186 loss: 14.583480834960938
For epoch 17 the test accuracy over the whole test set is 0 %
tensor([[123.9744]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[123.97442]
[125.]
epoch: 0 loss: 1.0518171787261963
tensor([[9.4642]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[9.464163]
[9.]
epoch: 1 loss: 0.21544712781906128
tensor([[11.2570]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[11.25697]
[12.]
epoch: 2 loss: 0.552092969417572
tensor([[5.1774]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[5.177416]
[4.]
epoch: 3 loss: 1.3863080739974976
tensor([[5.2762]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[5.27617]
[5.]
epoch: 4 loss: 0.07626974582672119
tensor([[42.3183]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[42.318268]
[46.]
epoch: 5 loss: 13.55515193939209
tensor([[61.9516]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[61.95156]
[68.]
epoch: 6 loss: 36.583614349365234
tensor([[33.0031]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[33.00313]

```

```

[35.]
epoch: 7 loss: 3.9874823093414307
tensor([[53.6315]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[53.631542]
[57.]
epoch: 8 loss: 11.346508026123047
tensor([[84.7953]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[84.79527]
[93.]
epoch: 9 loss: 67.31755065917969
tensor([[166.3717]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[166.37166]
[183.]
epoch: 10 loss: 276.5017395019531
tensor([[164.9187]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[164.91875]
[165.]
epoch: 11 loss: 0.006602058187127113
tensor([[196.7459]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[196.7459]
[182.]
epoch: 12 loss: 217.44143676757812
tensor([[20.6582]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[20.658215]
[16.]
epoch: 13 loss: 21.698963165283203
tensor([[123.9269]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[123.92693]
[110.]
epoch: 14 loss: 193.95947265625
tensor([[172.8746]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[172.87457]
[166.]
epoch: 15 loss: 47.25975036621094
tensor([[121.8048]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[121.80484]
[123.]
epoch: 16 loss: 1.4284071922302246
tensor([[25.0494]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],

```

```

device='cuda:0')
[25.04941]
[24.]
epoch: 17 loss: 1.101261019706726
tensor([[102.5294]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[102.529366]
[116.]
epoch: 18 loss: 181.45799255371094
tensor([[61.0125]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[61.012486]
[64.]
epoch: 19 loss: 8.92524242401123
tensor([[141.9992]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[141.99924]
[167.]
epoch: 20 loss: 625.0381469726562
tensor([[40.2923]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[40.292313]
[36.]
epoch: 21 loss: 18.423948287963867
tensor([[115.7170]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[115.71703]
[112.]
epoch: 22 loss: 13.816337585449219
tensor([[135.7681]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[135.76805]
[126.]
epoch: 23 loss: 95.41482543945312
tensor([[25.3335]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[25.333508]
[23.]
epoch: 24 loss: 5.445257663726807
tensor([[4.8862]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[4.8862185]
[2.]
epoch: 25 loss: 8.330257415771484
tensor([[196.3524]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[196.35239]
[180.]

```

```

epoch: 26 loss: 267.4005432128906
tensor([[179.2112]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[179.21124]
[175.]
epoch: 27 loss: 17.73456573486328
tensor([[97.3226]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[97.3226]
[98.]
epoch: 28 loss: 0.45886898040771484
tensor([[95.9170]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[95.917015]
[105.]
epoch: 29 loss: 82.50061798095703
tensor([[52.9486]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[52.948586]
[56.]
epoch: 30 loss: 9.31113052368164
tensor([[92.1434]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[92.14345]
[109.]
epoch: 31 loss: 284.1433410644531
tensor([[35.5414]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[35.541393]
[39.]
epoch: 32 loss: 11.961960792541504
tensor([[51.2768]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[51.276756]
[53.]
epoch: 33 loss: 2.969568967819214
tensor([[116.8916]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[116.89163]
[135.]
epoch: 34 loss: 327.9129943847656
tensor([[97.4734]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[97.47337]
[99.]
epoch: 35 loss: 2.3305888175964355
tensor([[92.1199]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')

```

```

[92.119934]
[85.]
epoch: 36 loss: 50.69346237182617
tensor([[125.7745]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[125.7745]
[119.]
epoch: 37 loss: 45.893821716308594
tensor([[173.2724]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[173.2724]
[168.]
epoch: 38 loss: 27.798200607299805
tensor([[81.8220]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[81.821976]
[69.]
epoch: 39 loss: 164.40306091308594
tensor([[102.9966]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[102.996605]
[95.]
epoch: 40 loss: 63.9456901550293
tensor([[147.2079]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[147.20787]
[153.]
epoch: 41 loss: 33.548763275146484
tensor([[63.8872]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[63.887234]
[62.]
epoch: 42 loss: 3.5616512298583984
tensor([[178.3066]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[178.30656]
[200.]
epoch: 43 loss: 470.60516357421875
tensor([[78.7926]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[78.79258]
[76.]
epoch: 44 loss: 7.798501014709473
tensor([[122.3769]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[122.376945]
[115.]
epoch: 45 loss: 54.4193229675293

```



```

tensor([[142.2972]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[142.29723]
[132.]
epoch: 46 loss: 106.0328598022461
tensor([[91.6388]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[91.638824]
[87.]
epoch: 47 loss: 21.518692016601562
tensor([[107.4283]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[107.42825]
[104.]
epoch: 48 loss: 11.752920150756836
tensor([[146.3534]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[146.3534]
[145.]
epoch: 49 loss: 1.8316740989685059
tensor([[23.2960]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[23.295952]
[34.]
epoch: 50 loss: 114.57664489746094
tensor([[113.3795]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[113.37955]
[122.]
epoch: 51 loss: 74.31221008300781
tensor([[126.5784]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[126.578354]
[133.]
epoch: 52 loss: 41.2375373840332
tensor([[209.7837]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[209.78366]
[205.]
epoch: 53 loss: 22.883411407470703
tensor([[173.0487]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[173.04866]
[173.]
epoch: 54 loss: 0.0023678226862102747
tensor([[62.3572]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[62.357193]

```

```

[71.]
epoch: 55 loss: 74.69811248779297
tensor([[197.1713]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[197.17128]
[189.]
epoch: 56 loss: 66.76981353759766
tensor([[10.3590]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[10.359034]
[19.]
epoch: 57 loss: 74.66629791259766
tensor([[36.4894]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[36.48942]
[43.]
epoch: 58 loss: 42.38762664794922
tensor([[168.5007]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[168.50067]
[170.]
epoch: 59 loss: 2.247986316680908
tensor([[101.3429]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[101.34294]
[103.]
epoch: 60 loss: 2.7458436489105225
tensor([[6.9401]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[6.9401255]
[3.]
epoch: 61 loss: 15.524588584899902
tensor([[8.9747]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[8.97471]
[6.]
epoch: 62 loss: 8.848902702331543
tensor([[145.1879]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[145.18788]
[144.]
epoch: 63 loss: 1.4110623598098755
tensor([[80.4557]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[80.45568]
[79.]
epoch: 64 loss: 2.119006633758545
tensor([[108.2584]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],

```

```

device='cuda:0')
[108.25838]
[107.]
epoch: 65 loss: 1.5835129022598267
tensor([[89.2851]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[89.28506]
[89.]
epoch: 66 loss: 0.08125752955675125
tensor([[122.4792]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[122.47916]
[124.]
epoch: 67 loss: 2.312964916229248
tensor([[39.8770]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[39.877007]
[38.]
epoch: 68 loss: 3.52315354347229
tensor([[126.2234]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[126.2234]
[130.]
epoch: 69 loss: 14.262735366821289
tensor([[90.3122]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[90.31224]
[90.]
epoch: 70 loss: 0.09749419242143631
tensor([[193.5197]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[193.51965]
[198.]
epoch: 71 loss: 20.07350730895996
tensor([[30.3740]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[30.374048]
[26.]
epoch: 72 loss: 19.13229751586914
tensor([[173.7948]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[173.79485]
[172.]
epoch: 73 loss: 3.221470594406128
tensor([[134.2933]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[134.29327]
[129.]

```

```

epoch: 74 loss: 28.018749237060547
tensor([[188.3755]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[188.37555]
[185.]
epoch: 75 loss: 11.394332885742188
tensor([[69.6350]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[69.635025]
[67.]
epoch: 76 loss: 6.943356990814209
tensor([[171.1482]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[171.1482]
[178.]
epoch: 77 loss: 46.9472541809082
tensor([[52.1430]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[52.142975]
[50.]
epoch: 78 loss: 4.592341423034668
tensor([[136.5732]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[136.57315]
[138.]
epoch: 79 loss: 2.0358991622924805
tensor([[156.4838]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[156.48378]
[156.]
epoch: 80 loss: 0.23404300212860107
tensor([[140.8831]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[140.88315]
[141.]
epoch: 81 loss: 0.013654344715178013
tensor([[52.5706]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[52.570557]
[51.]
epoch: 82 loss: 2.4666481018066406
tensor([[11.7242]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[11.724248]
[11.]
epoch: 83 loss: 0.524535059928894
tensor([[109.0820]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')

```

```

[109.08198]
[108.]
epoch: 84 loss: 1.1706761121749878
tensor([[10.3832]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[10.383243]
[10.]
epoch: 85 loss: 0.14687488973140717
tensor([[2.8101]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[2.810061]
[1.]
epoch: 86 loss: 3.276320695877075
tensor([[149.3457]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[149.34575]
[149.]
epoch: 87 loss: 0.11954230070114136
tensor([[29.4271]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[29.427135]
[30.]
epoch: 88 loss: 0.32817378640174866
tensor([[114.8227]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[114.82274]
[117.]
epoch: 89 loss: 4.740467071533203
tensor([[35.4755]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[35.475502]
[37.]
epoch: 90 loss: 2.32409405708313
tensor([[45.5026]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[45.502598]
[47.]
epoch: 91 loss: 2.242213249206543
tensor([[91.3291]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[91.32909]
[92.]
epoch: 92 loss: 0.45012518763542175
tensor([[198.6938]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[198.69382]
[196.]
epoch: 93 loss: 7.256650924682617

```

```

tensor([[26.1150]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[26.115042]
[27.]
epoch: 94 loss: 0.7831511497497559
tensor([[43.9262]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[43.926193]
[45.]
epoch: 95 loss: 1.1530609130859375
tensor([[90.4306]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[90.43065]
[91.]
epoch: 96 loss: 0.32416078448295593
tensor([[42.8828]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[42.882767]
[44.]
epoch: 97 loss: 1.2482101917266846
tensor([[72.8062]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[72.806175]
[74.]
epoch: 98 loss: 1.425217628479004
tensor([[52.2119]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[52.21186]
[52.]
epoch: 99 loss: 0.04488493874669075
tensor([[57.9073]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[57.90727]
[58.]
epoch: 100 loss: 0.00859912671148777
tensor([[48.8607]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[48.86071]
[48.]
epoch: 101 loss: 0.7408219575881958
tensor([[114.1060]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[114.10605]
[113.]
epoch: 102 loss: 1.2233434915542603
tensor([[97.4776]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[97.47763]

```

```

[96.]
epoch: 103 loss: 2.183392286300659
tensor([[13.5291]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[13.5291195]
[18.]
epoch: 104 loss: 19.988773345947266
tensor([[24.1696]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[24.169596]
[25.]
epoch: 105 loss: 0.6895712614059448
tensor([[28.9197]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[28.919685]
[28.]
epoch: 106 loss: 0.8458211421966553
tensor([[135.9748]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[135.9748]
[137.]
epoch: 107 loss: 1.0510504245758057
tensor([[17.1239]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[17.12387]
[21.]
epoch: 108 loss: 15.02437686920166
tensor([[172.5948]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[172.59483]
[176.]
epoch: 109 loss: 11.595159530639648
tensor([[201.5421]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[201.54205]
[199.]
epoch: 110 loss: 6.462034702301025
tensor([[157.7794]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[157.77943]
[154.]
epoch: 111 loss: 14.284122467041016
tensor([[84.3228]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[84.32276]
[82.]
epoch: 112 loss: 5.39522123336792
tensor([[62.2335]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],

```

```

device='cuda:0')
[62.233494]
[61.]
epoch: 113 loss: 1.5215070247650146
tensor([[34.5055]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[34.50553]
[32.]
epoch: 114 loss: 6.277687072753906
tensor([[30.6718]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[30.671755]
[29.]
epoch: 115 loss: 2.794764280319214
tensor([[16.8139]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[16.813936]
[17.]
epoch: 116 loss: 0.03461972624063492
tensor([[14.4427]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[14.442672]
[13.]
epoch: 117 loss: 2.0813019275665283
tensor([[155.3250]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[155.32503]
[171.]
epoch: 118 loss: 245.7047576904297
tensor([[156.8727]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[156.87274]
[162.]
epoch: 119 loss: 26.28877830505371
tensor([[207.0453]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[207.04533]
[204.]
epoch: 120 loss: 9.27405834197998
tensor([[24.1597]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[24.159737]
[22.]
epoch: 121 loss: 4.664462089538574
tensor([[125.5925]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[125.59245]
[114.]

```



```

epoch: 122 loss: 134.3849639892578
tensor([[121.5077]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[121.50772]
[111.]
epoch: 123 loss: 110.41220092773438
tensor([[84.3719]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[84.37186]
[81.]
epoch: 124 loss: 11.369417190551758
tensor([[99.0071]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[99.00714]
[100.]
epoch: 125 loss: 0.9857687950134277
tensor([[92.4540]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[92.454]
[97.]
epoch: 126 loss: 20.666093826293945
tensor([[70.6004]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[70.600426]
[78.]
epoch: 127 loss: 54.753700256347656
tensor([[49.2740]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[49.274033]
[55.]
epoch: 128 loss: 32.78670120239258
tensor([[198.4663]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[198.46634]
[203.]
epoch: 129 loss: 20.554080963134766
tensor([[185.7973]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[185.7973]
[186.]
epoch: 130 loss: 0.04108637943863869
tensor([[165.3590]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[165.35896]
[160.]
epoch: 131 loss: 28.71848487854004
tensor([[83.4114]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')

```

```

[83.411415]
[86.]
epoch: 132 loss: 6.700771808624268
tensor([[106.3527]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[106.352684]
[106.]
epoch: 133 loss: 0.1243860200047493
tensor([[147.0549]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[147.05492]
[143.]
epoch: 134 loss: 16.442346572875977
tensor([[53.3563]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[53.356308]
[54.]
epoch: 135 loss: 0.41433942317962646
tensor([[15.7679]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[15.767879]
[15.]
epoch: 136 loss: 0.5896374583244324
tensor([[194.6079]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[194.60791]
[193.]
epoch: 137 loss: 2.5853750705718994
tensor([[57.2051]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[57.205124]
[60.]
epoch: 138 loss: 7.8113322257995605
tensor([[32.4547]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[32.454746]
[33.]
epoch: 139 loss: 0.29730165004730225
tensor([[143.3576]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[143.3576]
[147.]
epoch: 140 loss: 13.267041206359863
tensor([[147.4313]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[147.43127]
[150.]
epoch: 141 loss: 6.598351001739502

```

```

tensor([[173.0859]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[173.08589]
[174.]
epoch: 142 loss: 0.8355939388275146
tensor([[161.1716]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[161.17165]
[157.]
epoch: 143 loss: 17.402631759643555
tensor([[167.9305]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[167.9305]
[164.]
epoch: 144 loss: 15.448800086975098
tensor([[132.4911]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[132.49115]
[134.]
epoch: 145 loss: 2.2766287326812744
tensor([[144.0808]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[144.0808]
[146.]
epoch: 146 loss: 3.683346748352051
tensor([[86.0135]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[86.01349]
[88.]
epoch: 147 loss: 3.9462268352508545
tensor([[10.2371]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[10.237143]
[8.]
epoch: 148 loss: 5.004806995391846
tensor([[158.5932]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[158.59323]
[159.]
epoch: 149 loss: 0.16546085476875305
tensor([[100.1225]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[100.122475]
[101.]
epoch: 150 loss: 0.7700507044792175
tensor([[150.0982]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[150.0982]

```

```

[151.]
epoch: 151 loss: 0.8132331967353821
tensor([[75.8137]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[75.81366]
[75.]
epoch: 152 loss: 0.6620420813560486
tensor([[142.6431]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[142.64307]
[142.]
epoch: 153 loss: 0.41353440284729004
tensor([[196.2176]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[196.21759]
[195.]
epoch: 154 loss: 1.482526183128357
tensor([[81.7141]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[81.71408]
[80.]
epoch: 155 loss: 2.938072919845581
tensor([[155.6580]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[155.65804]
[155.]
epoch: 156 loss: 0.43301042914390564
tensor([[182.5476]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[182.54764]
[184.]
epoch: 157 loss: 2.1093554496765137
tensor([[50.6122]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[50.612225]
[49.]
epoch: 158 loss: 2.5992681980133057
tensor([[186.2733]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[186.27333]
[190.]
epoch: 159 loss: 13.88806438446045
tensor([[65.4010]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[65.40102]
[65.]
epoch: 160 loss: 0.16081401705741882
tensor([[193.2381]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],

```

```

device='cuda:0')
[193.23811]
[192.]
epoch: 161 loss: 1.532924771308899
tensor([[18.4758]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[18.475752]
[20.]
epoch: 162 loss: 2.3233323097229004
tensor([[119.9899]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[119.989876]
[120.]
epoch: 163 loss: 0.00010249955812469125
tensor([[8.4118]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[8.411795]
[7.]
epoch: 164 loss: 1.9931641817092896
tensor([[14.7000]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[14.699998]
[14.]
epoch: 165 loss: 0.4899970591068268
tensor([[3.9802]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[3.9802172]
[0.]
epoch: 166 loss: 15.84212875366211
tensor([[152.6608]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[152.6608]
[148.]
epoch: 167 loss: 21.72303009033203
tensor([[140.9850]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[140.98495]
[140.]
epoch: 168 loss: 0.9701360464096069
tensor([[177.2580]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[177.258]
[179.]
epoch: 169 loss: 3.034579277038574
tensor([[75.6751]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[75.67509]
[77.]

```

```

epoch: 170 loss: 1.7553945779800415
tensor([[149.0236]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[149.02359]
[152.]
epoch: 171 loss: 8.859016418457031
tensor([[191.5361]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[191.53606]
[194.]
epoch: 172 loss: 6.071017265319824
tensor([[82.5540]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[82.554]
[83.]
epoch: 173 loss: 0.1989152431488037
tensor([[167.1935]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[167.1935]
[163.]
epoch: 174 loss: 17.58541488647461
tensor([[201.5133]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[201.51329]
[197.]
epoch: 175 loss: 20.36979103088379
tensor([[189.6516]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[189.65157]
[191.]
epoch: 176 loss: 1.8182754516601562
tensor([[172.1016]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[172.10164]
[177.]
epoch: 177 loss: 23.993942260742188
tensor([[112.6885]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[112.68854]
[118.]
epoch: 178 loss: 28.211633682250977
tensor([[135.1595]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[135.15947]
[136.]
epoch: 179 loss: 0.7064913511276245
tensor([[173.5600]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')

```

```

[173.56004]
[169.]
epoch: 180 loss: 20.793994903564453
tensor([[95.6199]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[95.61992]
[94.]
epoch: 181 loss: 2.6241369247436523
tensor([[85.1323]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[85.13231]
[84.]
epoch: 182 loss: 1.2821235656738281
tensor([[102.7868]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[102.786804]
[102.]
epoch: 183 loss: 0.6190608739852905
tensor([[69.9565]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[69.956505]
[70.]
epoch: 184 loss: 0.0018918304704129696
tensor([[138.1048]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[138.1048]
[139.]
epoch: 185 loss: 0.8013877868652344
tensor([[203.0497]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[203.04971]
[201.]
epoch: 186 loss: 4.201323986053467
For epoch 18 the test accuracy over the whole test set is 0 %
tensor([[131.1024]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[131.10236]
[134.]
epoch: 0 loss: 8.396341323852539
tensor([[149.7168]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[149.71684]
[152.]
epoch: 1 loss: 5.212807655334473
tensor([[148.7735]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[148.77351]
[150.]

```

```

epoch: 2 loss: 1.5042684078216553
tensor([[205.2693]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[205.26935]
[205.]
epoch: 3 loss: 0.07254841923713684
tensor([[132.5721]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[132.57205]
[132.]
epoch: 4 loss: 0.32724350690841675
tensor([[189.6171]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[189.61714]
[186.]
epoch: 5 loss: 13.083714485168457
tensor([[147.9684]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[147.96835]
[147.]
epoch: 6 loss: 0.9377080798149109
tensor([[144.7409]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[144.74092]
[146.]
epoch: 7 loss: 1.5852798223495483
tensor([[0.4337]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[0.43369862]
[7.]
epoch: 8 loss: 43.11631393432617
tensor([[139.8464]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[139.84644]
[143.]
epoch: 9 loss: 9.944969177246094
tensor([[143.0806]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[143.08064]
[145.]
epoch: 10 loss: 3.6839325428009033
tensor([[70.4382]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[70.43817]
[70.]
epoch: 11 loss: 0.19199416041374207
tensor([[167.1103]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')

```



```

[167.11034]
[165.]
epoch: 12 loss: 4.453519344329834
tensor([[201.5913]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[201.59132]
[199.]
epoch: 13 loss: 6.714959144592285
tensor([[108.6776]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[108.67755]
[109.]
epoch: 14 loss: 0.10397318005561829
tensor([[199.9024]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[199.90242]
[203.]
epoch: 15 loss: 9.595001220703125
tensor([[30.6661]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[30.666124]
[29.]
epoch: 16 loss: 2.775970220565796
tensor([[47.8179]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[47.81793]
[46.]
epoch: 17 loss: 3.304863452911377
tensor([[89.1672]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[89.16715]
[88.]
epoch: 18 loss: 1.362244725227356
tensor([[44.2049]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[44.204926]
[43.]
epoch: 19 loss: 1.4518455266952515
tensor([[191.9709]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[191.97086]
[194.]
epoch: 20 loss: 4.11742639541626
tensor([[126.0115]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[126.01154]
[126.]
epoch: 21 loss: 0.00013324717292562127

```

```

tensor([[7.1599]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[7.1598873]
[6.]
epoch: 22 loss: 1.3453385829925537
tensor([[150.9184]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[150.91844]
[148.]
epoch: 23 loss: 8.517302513122559
tensor([[171.3178]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[171.31781]
[170.]
epoch: 24 loss: 1.7366234064102173
tensor([[81.9138]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[81.91385]
[82.]
epoch: 25 loss: 0.007422016002237797
tensor([[104.6268]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[104.626785]
[106.]
epoch: 26 loss: 1.8857187032699585
tensor([[4.5772]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[4.577235]
[3.]
epoch: 27 loss: 2.4876708984375
tensor([[160.1872]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[160.18716]
[163.]
epoch: 28 loss: 7.912044525146484
tensor([[60.2329]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[60.23294]
[62.]
epoch: 29 loss: 3.1224987506866455
tensor([[13.6296]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[13.629551]
[16.]
epoch: 30 loss: 5.619028568267822
tensor([[118.5965]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[118.596466]

```

```

[120.]
epoch: 31 loss: 1.969907522201538
tensor([[47.8605]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[47.860523]
[47.]
epoch: 32 loss: 0.7405002117156982
tensor([[126.8126]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[126.812614]
[124.]
epoch: 33 loss: 7.910799980163574
tensor([[196.3555]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[196.35547]
[193.]
epoch: 34 loss: 11.259170532226562
tensor([[49.7074]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[49.707355]
[49.]
epoch: 35 loss: 0.5003517866134644
tensor([[12.7556]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[12.755614]
[14.]
epoch: 36 loss: 1.5484957695007324
tensor([[3.1139]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[3.1138754]
[2.]
epoch: 37 loss: 1.2407183647155762
tensor([[11.3155]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[11.315545]
[12.]
epoch: 38 loss: 0.4684785306453705
tensor([[101.9654]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[101.9654]
[107.]
epoch: 39 loss: 25.347190856933594
tensor([[87.7012]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[87.70116]
[91.]
epoch: 40 loss: 10.882368087768555
tensor([[87.6547]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],

```

```

device='cuda:0')
[87.6547]
[90.]
epoch: 41 loss: 5.500426292419434
tensor([[39.0403]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[39.040264]
[39.]
epoch: 42 loss: 0.0016212001210078597
tensor([[138.1223]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[138.1223]
[138.]
epoch: 43 loss: 0.014957092702388763
tensor([[9.3226]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[9.322636]
[8.]
epoch: 44 loss: 1.7493650913238525
tensor([[173.9858]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[173.98584]
[171.]
epoch: 45 loss: 8.915239334106445
tensor([[13.8043]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[13.8043165]
[11.]
epoch: 46 loss: 7.864191055297852
tensor([[168.9564]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[168.9564]
[167.]
epoch: 47 loss: 3.8275229930877686
tensor([[19.1714]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[19.171381]
[21.]
epoch: 48 loss: 3.3438475131988525
tensor([[117.5370]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[117.53703]
[119.]
epoch: 49 loss: 2.1402721405029297
tensor([[123.5957]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[123.59574]
[125.]

```

```

epoch: 50 loss: 1.9719425439834595
tensor([[55.5196]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[55.519558]
[53.]
epoch: 51 loss: 6.348172187805176
tensor([[162.3445]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[162.34451]
[166.]
epoch: 52 loss: 13.36258602142334
tensor([[184.6528]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[184.6528]
[189.]
epoch: 53 loss: 18.898134231567383
tensor([[55.1697]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[55.169743]
[51.]
epoch: 54 loss: 17.38675308227539
tensor([[81.8651]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[81.8651]
[76.]
epoch: 55 loss: 34.39936447143555
tensor([[148.3129]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[148.31293]
[141.]
epoch: 56 loss: 53.478904724121094
tensor([[2.7190]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[2.719009]
[0.]
epoch: 57 loss: 7.393009662628174
tensor([[178.2133]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[178.21333]
[179.]
epoch: 58 loss: 0.618844747543335
tensor([[29.0449]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[29.044853]
[30.]
epoch: 59 loss: 0.9123054146766663
tensor([[94.4491]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')

```

```

[94.44907]
[96.]
epoch: 60 loss: 2.405372142791748
tensor([[61.1370]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[61.137012]
[65.]
epoch: 61 loss: 14.922672271728516
tensor([[90.9815]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[90.98154]
[94.]
epoch: 62 loss: 9.111119270324707
tensor([[84.7050]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[84.70502]
[87.]
epoch: 63 loss: 5.266946792602539
tensor([[114.6659]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[114.66593]
[117.]
epoch: 64 loss: 5.447875022888184
tensor([[55.9075]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[55.907475]
[56.]
epoch: 65 loss: 0.008560964837670326
tensor([[205.2389]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[205.23894]
[201.]
epoch: 66 loss: 17.968589782714844
tensor([[112.1324]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[112.132385]
[111.]
epoch: 67 loss: 1.2822964191436768
tensor([[ -0.6269]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[ -0.6269166]
[1.]
epoch: 68 loss: 2.646857738494873
tensor([[32.5980]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[32.597977]
[33.]
epoch: 69 loss: 0.16162274777889252

```

```

tensor([[127.9414]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[127.94141]
[130.]
epoch: 70 loss: 4.2378082275390625
tensor([[191.5490]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[191.54897]
[196.]
epoch: 71 loss: 19.811708450317383
tensor([[161.1760]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[161.176]
[159.]
epoch: 72 loss: 4.734953880310059
tensor([[124.9225]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[124.92252]
[122.]
epoch: 73 loss: 8.541143417358398
tensor([[114.3243]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[114.32433]
[112.]
epoch: 74 loss: 5.402524948120117
tensor([[73.2195]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[73.2195]
[71.]
epoch: 75 loss: 4.9261698722839355
tensor([[84.5148]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[84.51477]
[83.]
epoch: 76 loss: 2.294529676437378
tensor([[23.1468]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[23.146793]
[25.]
epoch: 77 loss: 3.4343748092651367
tensor([[11.4836]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[11.483577]
[17.]
epoch: 78 loss: 30.430925369262695
tensor([[132.6666]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[132.66663]

```

```

[136.]
epoch: 79 loss: 11.111382484436035
tensor([[166.9829]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[166.98293]
[172.]
epoch: 80 loss: 25.171037673950195
tensor([[16.2105]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[16.210503]
[19.]
epoch: 81 loss: 7.7812957763671875
tensor([[18.4001]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[18.40006]
[20.]
epoch: 82 loss: 2.5598058700561523
tensor([[99.1978]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[99.19781]
[92.]
epoch: 83 loss: 51.80842971801758
tensor([[60.6873]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[60.687305]
[55.]
epoch: 84 loss: 32.34544372558594
tensor([[213.3515]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[213.35147]
[204.]
epoch: 85 loss: 87.45001220703125
tensor([[188.3227]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[188.32272]
[190.]
epoch: 86 loss: 2.8132567405700684
tensor([[36.8785]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[36.878544]
[38.]
epoch: 87 loss: 1.2576638460159302
tensor([[94.0089]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[94.00892]
[101.]
epoch: 88 loss: 48.87521743774414
tensor([[136.1406]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],

```



```

device='cuda:0')
[136.14061]
[149.]
epoch: 89 loss: 165.36392211914062
tensor([[81.3643]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[81.36425]
[86.]
epoch: 90 loss: 21.490177154541016
tensor([[60.3594]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[60.359383]
[61.]
epoch: 91 loss: 0.41039061546325684
tensor([[79.3919]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[79.39186]
[77.]
epoch: 92 loss: 5.720998764038086
tensor([[100.5620]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[100.562035]
[95.]
epoch: 93 loss: 30.936229705810547
tensor([[182.2621]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[182.26208]
[175.]
epoch: 94 loss: 52.737876892089844
tensor([[118.3537]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[118.353745]
[115.]
epoch: 95 loss: 11.247602462768555
tensor([[115.6509]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[115.65092]
[116.]
epoch: 96 loss: 0.12185890227556229
tensor([[150.9125]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[150.91249]
[155.]
epoch: 97 loss: 16.707731246948242
tensor([[18.5152]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[18.515202]
[15.]

```

```

epoch: 98 loss: 12.35664176940918
tensor([[66.1121]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[66.11207]
[69.]
epoch: 99 loss: 8.340149879455566
tensor([[34.3385]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[34.33854]
[34.]
epoch: 100 loss: 0.11460873484611511
tensor([[69.9646]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[69.96458]
[74.]
epoch: 101 loss: 16.28464126586914
tensor([[106.6510]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[106.650955]
[113.]
epoch: 102 loss: 40.31037139892578
tensor([[97.7907]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[97.79074]
[100.]
epoch: 103 loss: 4.880825519561768
tensor([[99.5687]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[99.568695]
[98.]
epoch: 104 loss: 2.4608042240142822
tensor([[13.3601]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[13.360121]
[4.]
epoch: 105 loss: 87.61186218261719
tensor([[60.7907]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[60.79071]
[57.]
epoch: 106 loss: 14.369485855102539
tensor([[21.0715]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[21.07152]
[18.]
epoch: 107 loss: 9.434234619140625
tensor([[194.2508]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')

```

```

[194.25076]
[195.]
epoch: 108 loss: 0.5613561868667603
tensor([[22.4789]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[22.478916]
[22.]
epoch: 109 loss: 0.2293606996536255
tensor([[181.4772]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[181.47719]
[180.]
epoch: 110 loss: 2.182084798812866
tensor([[83.2778]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[83.27782]
[80.]
epoch: 111 loss: 10.7440824508667
tensor([[100.4684]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[100.46835]
[97.]
epoch: 112 loss: 12.029474258422852
tensor([[173.8451]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[173.84512]
[177.]
epoch: 113 loss: 9.9532470703125
tensor([[11.9299]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[11.9298525]
[13.]
epoch: 114 loss: 1.1452157497406006
tensor([[168.8524]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[168.85243]
[173.]
epoch: 115 loss: 17.20231819152832
tensor([[156.2229]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[156.2229]
[153.]
epoch: 116 loss: 10.387086868286133
tensor([[139.1961]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[139.19614]
[137.]
epoch: 117 loss: 4.823015213012695

```

```

tensor([[168.9992]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[168.99924]
[169.]
epoch: 118 loss: 5.820766091346741e-07
tensor([[83.1041]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[83.10412]
[85.]
epoch: 119 loss: 3.594367265701294
tensor([[139.9363]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[139.93633]
[140.]
epoch: 120 loss: 0.004054496064782143
tensor([[108.2688]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[108.26877]
[110.]
epoch: 121 loss: 2.9971630573272705
tensor([[161.7370]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[161.737]
[160.]
epoch: 122 loss: 3.017167329788208
tensor([[78.5835]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[78.58354]
[81.]
epoch: 123 loss: 5.839270114898682
tensor([[167.9347]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[167.9347]
[168.]
epoch: 124 loss: 0.004265084862709045
tensor([[46.5649]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[46.564934]
[50.]
epoch: 125 loss: 11.79967975616455
tensor([[113.7747]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[113.774704]
[114.]
epoch: 126 loss: 0.050758298486471176
tensor([[174.7766]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[174.77661]

```

```

[174.]
epoch: 127 loss: 0.6031251549720764
tensor([[106.0151]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[106.0151]
[104.]
epoch: 128 loss: 4.060622215270996
tensor([[42.8000]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[42.799957]
[45.]
epoch: 129 loss: 4.840188026428223
tensor([[6.1672]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[6.1672325]
[9.]
epoch: 130 loss: 8.024571418762207
tensor([[24.9786]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[24.978556]
[27.]
epoch: 131 loss: 4.086236953735352
tensor([[186.2980]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[186.29799]
[185.]
epoch: 132 loss: 1.6847751140594482
tensor([[84.9644]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[84.96443]
[84.]
epoch: 133 loss: 0.9301286339759827
tensor([[175.3215]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[175.32153]
[176.]
epoch: 134 loss: 0.4603171944618225
tensor([[20.1224]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[20.122366]
[23.]
epoch: 135 loss: 8.280777931213379
tensor([[36.8810]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[36.88097]
[37.]
epoch: 136 loss: 0.014168140478432178
tensor([[130.2526]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],

```

```

device='cuda:0')
[130.25261]
[129.]
epoch: 137 loss: 1.5690299272537231
tensor([[136.5496]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[136.54956]
[135.]
epoch: 138 loss: 2.4011378288269043
tensor([[89.5947]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[89.59475]
[89.]
epoch: 139 loss: 0.35372692346572876
tensor([[36.1271]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[36.12707]
[36.]
epoch: 140 loss: 0.016147136688232422
tensor([[182.6822]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[182.68222]
[191.]
epoch: 141 loss: 69.18545532226562
tensor([[95.0565]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[95.05652]
[93.]
epoch: 142 loss: 4.229268550872803
tensor([[198.8171]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[198.81714]
[197.]
epoch: 143 loss: 3.301992893218994
tensor([[30.5246]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[30.52464]
[28.]
epoch: 144 loss: 6.373802661895752
tensor([[152.0390]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[152.039]
[144.]
epoch: 145 loss: 64.62554168701172
tensor([[159.0585]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[159.05853]
[154.]

```

```

epoch: 146 loss: 25.58875274658203
tensor([[154.3404]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[154.34044]
[156.]
epoch: 147 loss: 2.754143238067627
tensor([[57.9808]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[57.980755]
[60.]
epoch: 148 loss: 4.077351093292236
tensor([[123.8819]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[123.88191]
[133.]
epoch: 149 loss: 83.1395263671875
tensor([[101.3371]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[101.33712]
[108.]
epoch: 150 loss: 44.39397048950195
tensor([[22.3316]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[22.331594]
[24.]
epoch: 151 loss: 2.7835769653320312
tensor([[68.0871]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[68.08712]
[68.]
epoch: 152 loss: 0.007589904125779867
tensor([[65.5670]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[65.566986]
[64.]
epoch: 153 loss: 2.4554452896118164
tensor([[160.8839]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[160.88385]
[157.]
epoch: 154 loss: 15.084291458129883
tensor([[197.7345]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[197.73447]
[198.]
epoch: 155 loss: 0.07050801068544388
tensor([[178.1360]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')

```

```

[178.13599]
[178.]
epoch: 156 loss: 0.018492281436920166
tensor([[61.5639]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[61.563885]
[58.]
epoch: 157 loss: 12.701274871826172
tensor([[190.1304]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[190.13042]
[192.]
epoch: 158 loss: 3.4953410625457764
tensor([[165.5350]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[165.53502]
[164.]
epoch: 159 loss: 2.356283187866211
tensor([[125.1466]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[125.146645]
[123.]
epoch: 160 loss: 4.6080827713012695
tensor([[80.8300]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[80.829956]
[79.]
epoch: 161 loss: 3.3487391471862793
tensor([[180.0069]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[180.00693]
[184.]
epoch: 162 loss: 15.94462776184082
tensor([[149.4306]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[149.43056]
[151.]
epoch: 163 loss: 2.4631505012512207
tensor([[182.4271]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[182.42714]
[183.]
epoch: 164 loss: 0.32816940546035767
tensor([[68.7772]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[68.77717]
[67.]
epoch: 165 loss: 3.158327102661133

```



```

tensor([[27.1219]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[27.12193]
[26.]
epoch: 166 loss: 1.2587250471115112
tensor([[33.8732]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[33.873158]
[32.]
epoch: 167 loss: 3.508718967437744
tensor([[121.8488]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[121.84885]
[118.]
epoch: 168 loss: 14.813618659973145
tensor([[102.8061]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[102.80605]
[99.]
epoch: 169 loss: 14.486041069030762
tensor([[53.6441]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[53.64406]
[54.]
epoch: 170 loss: 0.12669454514980316
tensor([[140.7755]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[140.77545]
[142.]
epoch: 171 loss: 1.499518632888794
tensor([[40.6940]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[40.69398]
[44.]
epoch: 172 loss: 10.929760932922363
tensor([[99.0836]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[99.08365]
[102.]
epoch: 173 loss: 8.505105018615723
tensor([[10.2524]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[10.252377]
[10.]
epoch: 174 loss: 0.06369392573833466
tensor([[5.8551]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[5.8550878]

```

```

[5.]
epoch: 175 loss: 0.7311750650405884
tensor([[196.8459]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[196.84595]
[200.]
epoch: 176 loss: 9.94804859161377
tensor([[104.6622]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[104.662155]
[105.]
epoch: 177 loss: 0.1141391396522522
tensor([[141.5668]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[141.5668]
[139.]
epoch: 178 loss: 6.588477611541748
tensor([[77.5855]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[77.58553]
[78.]
epoch: 179 loss: 0.17178277671337128
tensor([[168.9330]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[168.93304]
[162.]
epoch: 180 loss: 48.06710433959961
tensor([[49.5420]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[49.541965]
[52.]
epoch: 181 loss: 6.041933536529541
tensor([[180.5858]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[180.58585]
[182.]
epoch: 182 loss: 1.9998316764831543
tensor([[31.8688]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[31.868828]
[35.]
epoch: 183 loss: 9.804239273071289
tensor([[44.3698]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[44.369827]
[48.]
epoch: 184 loss: 13.178153991699219
tensor([[70.6518]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],

```

```

device='cuda:0')
[70.65184]
[75.]
epoch: 185 loss: 18.906494140625
tensor([[100.5045]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[100.50446]
[103.]
epoch: 186 loss: 6.227704048156738
For epoch 19 the test accuracy over the whole test set is 0 %
tensor([[65.3655]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[65.365456]
[67.]
epoch: 0 loss: 2.6717352867126465
tensor([[28.6461]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[28.646126]
[28.]
epoch: 1 loss: 0.4174785315990448
tensor([[168.5345]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[168.53448]
[165.]
epoch: 2 loss: 12.492583274841309
tensor([[123.8463]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[123.84632]
[122.]
epoch: 3 loss: 3.4089016914367676
tensor([[39.1597]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[39.15966]
[37.]
epoch: 4 loss: 4.664132595062256
tensor([[16.8375]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[16.837513]
[16.]
epoch: 5 loss: 0.7014279961585999
tensor([[83.5380]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[83.53804]
[83.]
epoch: 6 loss: 0.28948721289634705
tensor([[100.7710]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[100.77104]

```

```

[101.]
epoch: 7 loss: 0.052421826869249344
tensor([[187.1293]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[187.12929]
[194.]
epoch: 8 loss: 47.206687927246094
tensor([[100.6499]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[100.649895]
[100.]
epoch: 9 loss: 0.42236313223838806
tensor([[20.4315]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[20.431458]
[21.]
epoch: 10 loss: 0.3232405483722687
tensor([[112.1164]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[112.116356]
[107.]
epoch: 11 loss: 26.17709732055664
tensor([[33.7973]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[33.79729]
[30.]
epoch: 12 loss: 14.419417381286621
tensor([[209.6087]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[209.60866]
[204.]
epoch: 13 loss: 31.457042694091797
tensor([[22.8905]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[22.8905]
[23.]
epoch: 14 loss: 0.011990443803369999
tensor([[76.5767]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[76.57674]
[77.]
epoch: 15 loss: 0.17915202677249908
tensor([[25.8501]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[25.850101]
[26.]
epoch: 16 loss: 0.022469568997621536
tensor([[57.8014]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],

```

```

device='cuda:0')
[57.8014]
[60.]
epoch: 17 loss: 4.833845138549805
tensor([[2.5191]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[2.5191205]
[0.]
epoch: 18 loss: 6.345967769622803
tensor([[125.8285]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[125.828545]
[141.]
epoch: 19 loss: 230.17306518554688
tensor([[70.3814]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[70.38141]
[74.]
epoch: 20 loss: 13.094202995300293
tensor([[77.6562]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[77.65623]
[78.]
epoch: 21 loss: 0.1181797981262207
tensor([[5.1424]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[5.142351]
[3.]
epoch: 22 loss: 4.589668273925781
tensor([[59.8492]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[59.84921]
[54.]
epoch: 23 loss: 34.2132453918457
tensor([[30.2815]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[30.281519]
[27.]
epoch: 24 loss: 10.768366813659668
tensor([[126.4196]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[126.41958]
[118.]
epoch: 25 loss: 70.8893051147461
tensor([[112.5836]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[112.58363]
[105.]

```

```

epoch: 26 loss: 57.511497497558594
tensor([[24.4306]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[24.430567]
[25.]
epoch: 27 loss: 0.32425418496131897
tensor([[13.8996]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[13.899624]
[20.]
epoch: 28 loss: 37.2145881652832
tensor([[118.0486]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[118.04857]
[123.]
epoch: 29 loss: 24.516672134399414
tensor([[87.5391]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[87.53914]
[92.]
epoch: 30 loss: 19.899282455444336
tensor([[157.9062]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[157.90619]
[164.]
epoch: 31 loss: 37.1345329284668
tensor([[191.4679]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[191.46793]
[192.]
epoch: 32 loss: 0.28310272097587585
tensor([[125.8285]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[125.828545]
[124.]
epoch: 33 loss: 3.3435754776000977
tensor([[201.5192]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[201.51917]
[193.]
epoch: 34 loss: 72.576171875
tensor([[75.6546]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[75.65464]
[75.]
epoch: 35 loss: 0.4285537898540497
tensor([[142.9653]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')

```

```

[142.96532]
[144.]
epoch: 36 loss: 1.070569396018982
tensor([[9.2971]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[9.297065]
[10.]
epoch: 37 loss: 0.4941179156303406
tensor([[160.6449]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[160.64488]
[168.]
epoch: 38 loss: 54.09775924682617
tensor([[181.7600]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[181.76]
[186.]
epoch: 39 loss: 17.977645874023438
tensor([[169.8775]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[169.87746]
[169.]
epoch: 40 loss: 0.7699301838874817
tensor([[82.1660]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[82.166]
[80.]
epoch: 41 loss: 4.69155740737915
tensor([[5.9018]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[5.901782]
[4.]
epoch: 42 loss: 3.616774797439575
tensor([[88.9829]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[88.98288]
[84.]
epoch: 43 loss: 24.829090118408203
tensor([[157.8439]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[157.84395]
[146.]
epoch: 44 loss: 140.27911376953125
tensor([[157.0137]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[157.01367]
[152.]
epoch: 45 loss: 25.136905670166016

```

```

tensor([[60.2438]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[60.24375]
[65.]
epoch: 46 loss: 22.621898651123047
tensor([[147.6095]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[147.60953]
[159.]
epoch: 47 loss: 129.74285888671875
tensor([[0.7987]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[0.79872864]
[1.]
epoch: 48 loss: 0.040510158985853195
tensor([[149.4536]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[149.45357]
[160.]
epoch: 49 loss: 111.22724151611328
tensor([[50.8187]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[50.818665]
[55.]
epoch: 50 loss: 17.483566284179688
tensor([[91.6597]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[91.6597]
[93.]
epoch: 51 loss: 1.7964081764221191
tensor([[113.9268]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[113.926765]
[111.]
epoch: 52 loss: 8.565956115722656
tensor([[37.2941]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[37.294052]
[36.]
epoch: 53 loss: 1.674570918083191
tensor([[190.4495]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[190.44948]
[170.]
epoch: 54 loss: 418.18115234375
tensor([[153.6906]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[153.69058]

```



```

[148.]
epoch: 55 loss: 32.38272476196289
tensor([[87.9674]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[87.967354]
[95.]
epoch: 56 loss: 49.45811080932617
tensor([[62.1659]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[62.165882]
[71.]
epoch: 57 loss: 78.04164123535156
tensor([[52.8863]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[52.88634]
[62.]
epoch: 58 loss: 83.05877685546875
tensor([[14.4028]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[14.402766]
[18.]
epoch: 59 loss: 12.940091133117676
tensor([[105.6507]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[105.6507]
[126.]
epoch: 60 loss: 414.0938720703125
tensor([[137.0168]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[137.01677]
[153.]
epoch: 61 loss: 255.46365356445312
tensor([[11.9276]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[11.927584]
[5.]
epoch: 62 loss: 47.991416931152344
tensor([[23.7331]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[23.733145]
[15.]
epoch: 63 loss: 76.26781463623047
tensor([[170.6777]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[170.6777]
[142.]
epoch: 64 loss: 822.4107055664062
tensor([[48.8950]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],

```

```

device='cuda:0')
[48.894955]
[38.]
epoch: 65 loss: 118.70003509521484
tensor([[92.7810]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[92.781]
[87.]
epoch: 66 loss: 33.41994094848633
tensor([[4.9877]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[4.987707]
[8.]
epoch: 67 loss: 9.073907852172852
tensor([[34.2190]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[34.219044]
[39.]
epoch: 68 loss: 22.857542037963867
tensor([[134.5010]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[134.50099]
[172.]
epoch: 69 loss: 1406.1756591796875
tensor([[67.9374]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[67.93744]
[76.]
epoch: 70 loss: 65.00489044189453
tensor([[133.4185]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[133.41847]
[139.]
epoch: 71 loss: 31.153451919555664
tensor([[201.2850]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[201.28503]
[190.]
epoch: 72 loss: 127.35199737548828
tensor([[53.3588]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[53.358795]
[47.]
epoch: 73 loss: 40.43427658081055
tensor([[8.6498]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[8.649751]
[12.]

```

```

epoch: 74 loss: 11.224170684814453
tensor([[133.3441]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[133.34406]
[116.]
epoch: 75 loss: 300.8162536621094
tensor([[167.1266]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[167.12659]
[150.]
epoch: 76 loss: 293.3199768066406
tensor([[32.4264]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[32.426407]
[34.]
epoch: 77 loss: 2.4761953353881836
tensor([[27.0119]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[27.011862]
[32.]
epoch: 78 loss: 24.88152313232422
tensor([[120.3639]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[120.36391]
[143.]
epoch: 79 loss: 512.3927001953125
tensor([[21.9534]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[21.953373]
[29.]
epoch: 80 loss: 49.65495300292969
tensor([[0.7034]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[0.703351]
[6.]
epoch: 81 loss: 28.05449104309082
tensor([[112.6312]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[112.63116]
[132.]
epoch: 82 loss: 375.1520690917969
tensor([[144.8706]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[144.87062]
[157.]
epoch: 83 loss: 147.12184143066406
tensor([[82.7397]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')

```

```

[82.739746]
[82.]
epoch: 84 loss: 0.5472242832183838
tensor([[219.1312]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[219.13116]
[203.]
epoch: 85 loss: 260.2144775390625
tensor([[213.3517]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[213.35165]
[200.]
epoch: 86 loss: 178.26666259765625
tensor([[89.2093]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[89.20926]
[86.]
epoch: 87 loss: 10.29934310913086
tensor([[79.5927]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[79.59266]
[81.]
epoch: 88 loss: 1.9806087017059326
tensor([[167.4421]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[167.44214]
[195.]
epoch: 89 loss: 759.4357299804688
tensor([[92.4186]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[92.41856]
[96.]
epoch: 90 loss: 12.826739311218262
tensor([[137.2076]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[137.2076]
[137.]
epoch: 91 loss: 0.04309602826833725
tensor([[179.7013]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[179.70126]
[175.]
epoch: 92 loss: 22.101877212524414
tensor([[74.4304]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[74.43044]
[64.]
epoch: 93 loss: 108.79413604736328

```

```

tensor([[151.1122]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[151.11224]
[140.]
epoch: 94 loss: 123.4819564819336
tensor([[153.5659]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[153.56592]
[147.]
epoch: 95 loss: 43.11127853393555
tensor([[136.2025]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[136.20253]
[136.]
epoch: 96 loss: 0.04101836308836937
tensor([[171.4314]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[171.43141]
[183.]
epoch: 97 loss: 133.8322296142578
tensor([[8.9371]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[8.937118]
[7.]
epoch: 98 loss: 3.752424478530884
tensor([[177.7630]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[177.76305]
[191.]
epoch: 99 loss: 175.21694946289062
tensor([[183.5532]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[183.55319]
[184.]
epoch: 100 loss: 0.1996372640132904
tensor([[60.0690]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[60.068954]
[58.]
epoch: 101 loss: 4.280572414398193
tensor([[24.4321]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[24.432089]
[24.]
epoch: 102 loss: 0.18670077621936798
tensor([[202.1129]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[202.11293]

```

```

[180.]
epoch: 103 loss: 488.981689453125
tensor([[107.6137]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[107.613655]
[103.]
epoch: 104 loss: 21.285812377929688
tensor([[13.6736]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[13.673629]
[17.]
epoch: 105 loss: 11.06474494934082
tensor([[155.9324]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[155.93236]
[163.]
epoch: 106 loss: 49.95156478881836
tensor([[35.5882]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[35.588207]
[45.]
epoch: 107 loss: 88.58184051513672
tensor([[159.9455]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[159.94554]
[176.]
epoch: 108 loss: 257.7456359863281
tensor([[81.8422]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[81.842155]
[90.]
epoch: 109 loss: 66.55043029785156
tensor([[170.5013]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[170.50128]
[167.]
epoch: 110 loss: 12.258974075317383
tensor([[143.9880]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[143.98796]
[138.]
epoch: 111 loss: 35.855674743652344
tensor([[142.1312]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[142.13123]
[134.]
epoch: 112 loss: 66.11682891845703
tensor([[42.0408]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],

```

```

device='cuda:0')
[42.040764]
[43.]
epoch: 113 loss: 0.9201340079307556
tensor([[198.6247]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[198.6247]
[196.]
epoch: 114 loss: 6.8890228271484375
tensor([[86.3569]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[86.356926]
[88.]
epoch: 115 loss: 2.6996922492980957
tensor([[82.1149]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[82.11485]
[85.]
epoch: 116 loss: 8.324073791503906
tensor([[32.6442]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[32.64421]
[33.]
epoch: 117 loss: 0.1265859454870224
tensor([[103.0730]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[103.07296]
[110.]
epoch: 118 loss: 47.983882904052734
tensor([[88.4801]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[88.48007]
[91.]
epoch: 119 loss: 6.350037097930908
tensor([[127.2072]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[127.20724]
[130.]
epoch: 120 loss: 7.799523830413818
tensor([[58.8494]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[58.849384]
[56.]
epoch: 121 loss: 8.118990898132324
tensor([[71.9203]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[71.920265]
[68.]

```

```

epoch: 122 loss: 15.368478775024414
tensor([[119.5246]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[119.52458]
[119.]
epoch: 123 loss: 0.2751861810684204
tensor([[113.0050]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[113.00503]
[112.]
epoch: 124 loss: 1.0100808143615723
tensor([[49.3151]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[49.31511]
[46.]
epoch: 125 loss: 10.989949226379395
tensor([[14.7854]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[14.785356]
[9.]
epoch: 126 loss: 33.470340728759766
tensor([[170.6245]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[170.62454]
[166.]
epoch: 127 loss: 21.386390686035156
tensor([[51.3451]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[51.345146]
[50.]
epoch: 128 loss: 1.8094182014465332
tensor([[13.6595]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[13.659475]
[11.]
epoch: 129 loss: 7.072809219360352
tensor([[98.6038]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[98.60379]
[104.]
epoch: 130 loss: 29.11907958984375
tensor([[163.6210]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[163.62105]
[171.]
epoch: 131 loss: 54.44893264770508
tensor([[152.7325]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')

```



```

[152.7325]
[156.]
epoch: 132 loss: 10.676568031311035
tensor([[12.4134]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[12.41335]
[13.]
epoch: 133 loss: 0.3441581130027771
tensor([[114.1257]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[114.12574]
[114.]
epoch: 134 loss: 0.015810560435056686
tensor([[140.8716]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[140.87164]
[135.]
epoch: 135 loss: 34.476192474365234
tensor([[112.0909]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[112.09092]
[108.]
epoch: 136 loss: 16.73562240600586
tensor([[204.1675]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[204.16748]
[199.]
epoch: 137 loss: 26.70285415649414
tensor([[70.4527]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[70.4527]
[70.]
epoch: 138 loss: 0.2049352526664734
tensor([[153.2956]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[153.29564]
[154.]
epoch: 139 loss: 0.49612435698509216
tensor([[120.1823]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[120.18229]
[125.]
epoch: 140 loss: 23.210338592529297
tensor([[49.8973]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[49.89731]
[53.]
epoch: 141 loss: 9.62669563293457

```

```

tensor([[44.3103]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[44.3103]
[48.]
epoch: 142 loss: 13.61389446258545
tensor([[167.6789]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[167.67892]
[174.]
epoch: 143 loss: 39.95599365234375
tensor([[174.3449]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[174.34486]
[177.]
epoch: 144 loss: 7.049747943878174
tensor([[51.6696]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[51.669605]
[51.]
epoch: 145 loss: 0.44837120175361633
tensor([[72.5979]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[72.597885]
[69.]
epoch: 146 loss: 12.944777488708496
tensor([[65.6119]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[65.611916]
[61.]
epoch: 147 loss: 21.269765853881836
tensor([[201.6713]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[201.67126]
[189.]
epoch: 148 loss: 160.56094360351562
tensor([[190.3838]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[190.38379]
[185.]
epoch: 149 loss: 28.985183715820312
tensor([[54.4473]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[54.44732]
[57.]
epoch: 150 loss: 6.516180038452148
tensor([[167.4468]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[167.44676]

```

```

[182.]
epoch: 151 loss: 211.79673767089844
tensor([[164.9852]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[164.98524]
[178.]
epoch: 152 loss: 169.38385009765625
tensor([[9.5300]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[9.530015]
[14.]
epoch: 153 loss: 19.98076629638672
tensor([[99.8420]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[99.84205]
[98.]
epoch: 154 loss: 3.3931431770324707
tensor([[44.9844]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[44.98439]
[44.]
epoch: 155 loss: 0.9690241813659668
tensor([[190.0511]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[190.05113]
[173.]
epoch: 156 loss: 290.7411193847656
tensor([[93.2411]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[93.24115]
[89.]
epoch: 157 loss: 17.98735237121582
tensor([[197.0812]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[197.08122]
[198.]
epoch: 158 loss: 0.84415203332901
tensor([[13.9594]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[13.95938]
[19.]
epoch: 159 loss: 25.407848358154297
tensor([[187.7909]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[187.79091]
[201.]
epoch: 160 loss: 174.48008728027344
tensor([[90.5244]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],

```

```

device='cuda:0')
[90.52436]
[94.]
epoch: 161 loss: 12.080068588256836
tensor([[129.3123]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[129.31227]
[129.]
epoch: 162 loss: 0.09751325100660324
tensor([[137.2384]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[137.23836]
[133.]
epoch: 163 loss: 17.963674545288086
tensor([[52.3323]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[52.332336]
[52.]
epoch: 164 loss: 0.11044749617576599
tensor([[210.8807]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[210.88072]
[205.]
epoch: 165 loss: 34.58289337158203
tensor([[34.7514]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[34.75138]
[35.]
epoch: 166 loss: 0.06181144714355469
tensor([[99.6596]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[99.65955]
[99.]
epoch: 167 loss: 0.4350108504295349
tensor([[7.1558]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[7.155799]
[2.]
epoch: 168 loss: 26.58226203918457
tensor([[179.9912]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[179.99123]
[179.]
epoch: 169 loss: 0.9825294017791748
tensor([[103.8432]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[103.84319]
[109.]

```

```

epoch: 170 loss: 26.5926570892334
tensor([[92.8281]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[92.828094]
[97.]
epoch: 171 loss: 17.404796600341797
tensor([[109.4752]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[109.47522]
[115.]
epoch: 172 loss: 30.523197174072266
tensor([[149.1661]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[149.16609]
[149.]
epoch: 173 loss: 0.027586525306105614
tensor([[20.0695]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[20.069464]
[22.]
epoch: 174 loss: 3.7269701957702637
tensor([[155.7250]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[155.72496]
[151.]
epoch: 175 loss: 22.32525062561035
tensor([[200.7370]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[200.73698]
[197.]
epoch: 176 loss: 13.965051651000977
tensor([[102.9263]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[102.926346]
[102.]
epoch: 177 loss: 0.8581165671348572
tensor([[117.6018]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[117.60177]
[120.]
epoch: 178 loss: 5.751514434814453
tensor([[48.1434]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[48.143394]
[49.]
epoch: 179 loss: 0.7337730526924133
tensor([[78.3112]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')

```

```

[78.311226]
[79.]
epoch: 180 loss: 0.4744097590446472
tensor([[141.0018]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[141.00185]
[145.]
epoch: 181 loss: 15.985233306884766
tensor([[162.4083]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[162.40833]
[162.]
epoch: 182 loss: 0.16672946512699127
tensor([[114.3280]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[114.328]
[117.]
epoch: 183 loss: 7.139568328857422
tensor([[112.1358]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[112.1358]
[113.]
epoch: 184 loss: 0.7468360662460327
tensor([[158.5200]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[158.52003]
[155.]
epoch: 185 loss: 12.390645027160645
tensor([[108.7466]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[108.74664]
[106.]
epoch: 186 loss: 7.544048309326172
For epoch 20 the test accuracy over the whole test set is 0 %
tensor([[139.6105]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[139.61055]
[140.]
epoch: 0 loss: 0.15167136490345
tensor([[14.1267]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[14.126655]
[10.]
epoch: 1 loss: 17.029277801513672
tensor([[77.2515]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[77.25146]
[75.]

```

```

epoch: 2 loss: 5.069059371948242
tensor([[34.7843]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[34.78433]
[32.]
epoch: 3 loss: 7.7524847984313965
tensor([[105.0325]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[105.03247]
[107.]
epoch: 4 loss: 3.871171474456787
tensor([[103.9967]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[103.99672]
[106.]
epoch: 5 loss: 4.013133525848389
tensor([[83.1259]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[83.1259]
[84.]
epoch: 6 loss: 0.7640503644943237
tensor([[130.7016]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[130.70157]
[135.]
epoch: 7 loss: 18.476512908935547
tensor([[184.0794]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[184.07936]
[191.]
epoch: 8 loss: 47.89524459838867
tensor([[94.1752]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[94.175156]
[90.]
epoch: 9 loss: 17.43192481994629
tensor([[148.9522]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[148.95216]
[143.]
epoch: 10 loss: 35.428253173828125
tensor([[147.6119]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[147.61192]
[142.]
epoch: 11 loss: 31.493682861328125
tensor([[68.7833]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')

```

```

[68.783295]
[65.]
epoch: 12 loss: 14.313318252563477
tensor([[82.7031]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[82.70306]
[80.]
epoch: 13 loss: 7.306513786315918
tensor([[87.3903]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[87.390335]
[87.]
epoch: 14 loss: 0.15236148238182068
tensor([[192.5395]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[192.53954]
[203.]
epoch: 15 loss: 109.42131805419922
tensor([[179.4050]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[179.40503]
[179.]
epoch: 16 loss: 0.16404873132705688
tensor([[123.4346]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[123.43457]
[122.]
epoch: 17 loss: 2.0579919815063477
tensor([[159.0884]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[159.08844]
[153.]
epoch: 18 loss: 37.06909942626953
tensor([[199.1886]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[199.18858]
[195.]
epoch: 19 loss: 17.544231414794922
tensor([[46.4696]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[46.469555]
[50.]
epoch: 20 loss: 12.464042663574219
tensor([[144.0934]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[144.09341]
[148.]
epoch: 21 loss: 15.261411666870117

```



```

tensor([[9.1511]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[9.151114]
[12.]
epoch: 22 loss: 8.116148948669434
tensor([[159.4967]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[159.49675]
[160.]
epoch: 23 loss: 0.25326067209243774
tensor([[105.7977]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[105.79765]
[112.]
epoch: 24 loss: 38.4691047668457
tensor([[82.2089]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[82.208885]
[86.]
epoch: 25 loss: 14.372551918029785
tensor([[8.4846]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[8.484605]
[9.]
epoch: 26 loss: 0.26563218235969543
tensor([[34.3544]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[34.354416]
[36.]
epoch: 27 loss: 2.707947015762329
tensor([[175.9807]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[175.98073]
[165.]
epoch: 28 loss: 120.5763931274414
tensor([[67.2830]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[67.28303]
[68.]
epoch: 29 loss: 0.5140493512153625
tensor([[154.9654]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[154.96541]
[154.]
epoch: 30 loss: 0.932013213634491
tensor([[18.2780]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[18.27799]

```

```

[22.]
epoch: 31 loss: 13.85335636138916
tensor([[179.8733]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[179.87334]
[185.]
epoch: 32 loss: 26.28267478942871
tensor([[44.2874]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[44.287384]
[47.]
epoch: 33 loss: 7.358285427093506
tensor([[128.9170]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[128.91699]
[133.]
epoch: 34 loss: 16.67095184326172
tensor([[55.3131]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[55.313072]
[55.]
epoch: 35 loss: 0.09801420569419861
tensor([[76.8857]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[76.88565]
[76.]
epoch: 36 loss: 0.7843770384788513
tensor([[103.6255]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[103.62554]
[101.]
epoch: 37 loss: 6.893469333648682
tensor([[173.6760]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[173.67601]
[169.]
epoch: 38 loss: 21.865070343017578
tensor([[161.9953]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[161.9953]
[159.]
epoch: 39 loss: 8.971823692321777
tensor([[170.5522]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[170.55223]
[173.]
epoch: 40 loss: 5.991573810577393
tensor([[164.2999]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],

```

```

device='cuda:0')
[164.29994]
[167.]
epoch: 41 loss: 7.290313243865967
tensor([[57.8831]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[57.88312]
[57.]
epoch: 42 loss: 0.7799035906791687
tensor([[181.0138]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[181.0138]
[184.]
epoch: 43 loss: 8.917427062988281
tensor([[91.9908]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[91.990845]
[92.]
epoch: 44 loss: 8.381903171539307e-05
tensor([[163.8405]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[163.84048]
[164.]
epoch: 45 loss: 0.025445155799388885
tensor([[110.0754]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[110.075356]
[110.]
epoch: 46 loss: 0.00567845581099391
tensor([[40.2783]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[40.27827]
[38.]
epoch: 47 loss: 5.190517425537109
tensor([[66.2810]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[66.280975]
[64.]
epoch: 48 loss: 5.202848434448242
tensor([[147.3297]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[147.32974]
[147.]
epoch: 49 loss: 0.10873007029294968
tensor([[174.8992]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[174.89923]
[171.]

```

```

epoch: 50 loss: 15.204002380371094
tensor([[78.8311]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[78.83114]
[78.]
epoch: 51 loss: 0.6907913684844971
tensor([[146.7072]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[146.70721]
[151.]
epoch: 52 loss: 18.428009033203125
tensor([[156.0365]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[156.03647]
[157.]
epoch: 53 loss: 0.9283929467201233
tensor([[35.4672]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[35.46722]
[34.]
epoch: 54 loss: 2.152735471725464
tensor([[96.5929]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[96.59287]
[97.]
epoch: 55 loss: 0.16575270891189575
tensor([[8.0267]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[8.026725]
[1.]
epoch: 56 loss: 49.37486267089844
tensor([[68.8344]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[68.83435]
[70.]
epoch: 57 loss: 1.3587385416030884
tensor([[199.5351]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[199.5351]
[204.]
epoch: 58 loss: 19.935375213623047
tensor([[73.7279]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[73.727905]
[74.]
epoch: 59 loss: 0.07403554022312164
tensor([[83.0058]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')

```

```

[83.005775]
[81.]
epoch: 60 loss: 4.023135185241699
tensor([[143.0559]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[143.0559]
[139.]
epoch: 61 loss: 16.450267791748047
tensor([[116.7254]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[116.72539]
[114.]
epoch: 62 loss: 7.427737236022949
tensor([[133.3862]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[133.38615]
[130.]
epoch: 63 loss: 11.466039657592773
tensor([[43.2650]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[43.265045]
[45.]
epoch: 64 loss: 3.01006817817688
tensor([[125.4530]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[125.45301]
[126.]
epoch: 65 loss: 0.2991974353790283
tensor([[114.0849]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[114.084885]
[118.]
epoch: 66 loss: 15.32812786102295
tensor([[58.6787]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[58.67875]
[62.]
epoch: 67 loss: 11.030707359313965
tensor([[93.2409]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[93.24089]
[95.]
epoch: 68 loss: 3.094466209411621
tensor([[102.0801]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[102.08014]
[103.]
epoch: 69 loss: 0.8461439609527588

```

```

tensor([[28.0108]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[28.010845]
[28.]
epoch: 70 loss: 0.0001176180230686441
tensor([[105.2906]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[105.29059]
[104.]
epoch: 71 loss: 1.6656184196472168
tensor([[42.9680]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[42.968033]
[44.]
epoch: 72 loss: 1.0649561882019043
tensor([[67.2095]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[67.20948]
[67.]
epoch: 73 loss: 0.043881990015506744
tensor([[118.0279]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[118.027885]
[116.]
epoch: 74 loss: 4.112319469451904
tensor([[83.7331]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[83.73312]
[82.]
epoch: 75 loss: 3.003718137741089
tensor([[200.4811]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[200.48106]
[193.]
epoch: 76 loss: 55.96631622314453
tensor([[59.7627]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[59.762672]
[61.]
epoch: 77 loss: 1.5309795141220093
tensor([[45.3076]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[45.30764]
[48.]
epoch: 78 loss: 7.248802185058594
tensor([[169.3112]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[169.31125]

```

```

[177.]
epoch: 79 loss: 59.11689376831055
tensor([[190.0876]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[190.08755]
[199.]
epoch: 80 loss: 79.43167877197266
tensor([[44.5171]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[44.51706]
[46.]
epoch: 81 loss: 2.199113130569458
tensor([[185.3973]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[185.39728]
[178.]
epoch: 82 loss: 54.71971893310547
tensor([[54.1641]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[54.164104]
[52.]
epoch: 83 loss: 4.683348178863525
tensor([[97.3881]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[97.388145]
[94.]
epoch: 84 loss: 11.47952938079834
tensor([[16.2215]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[16.221498]
[21.]
epoch: 85 loss: 22.834075927734375
tensor([[179.5115]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[179.51146]
[172.]
epoch: 86 loss: 56.422019958496094
tensor([[22.5433]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[22.543322]
[24.]
epoch: 87 loss: 2.1219120025634766
tensor([[139.5904]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[139.5904]
[145.]
epoch: 88 loss: 29.26384735107422
tensor([[109.4821]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],

```

```

device='cuda:0')
[109.48214]
[115.]
epoch: 89 loss: 30.4467830657959
tensor([[198.3723]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[198.37231]
[205.]
epoch: 90 loss: 43.92621612548828
tensor([[112.2956]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[112.295555]
[113.]
epoch: 91 loss: 0.49624258279800415
tensor([[109.2837]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[109.283745]
[105.]
epoch: 92 loss: 18.3504695892334
tensor([[24.9460]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[24.945972]
[23.]
epoch: 93 loss: 3.786808729171753
tensor([[160.0770]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[160.07703]
[150.]
epoch: 94 loss: 101.54646301269531
tensor([[209.3274]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[209.32742]
[197.]
epoch: 95 loss: 151.96536254882812
tensor([[36.8037]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[36.803745]
[37.]
epoch: 96 loss: 0.03851591795682907
tensor([[182.3404]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[182.34041]
[196.]
epoch: 97 loss: 186.58444213867188
tensor([[117.5177]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[117.51774]
[129.]

```



```

epoch: 98 loss: 131.8423309326172
tensor([[188.6899]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[188.68993]
[198.]
epoch: 99 loss: 86.67747497558594
tensor([[49.6937]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[49.69367]
[49.]
epoch: 100 loss: 0.48117581009864807
tensor([[178.3466]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[178.34662]
[168.]
epoch: 101 loss: 107.05252075195312
tensor([[116.2855]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[116.28551]
[111.]
epoch: 102 loss: 27.936586380004883
tensor([[29.7006]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[29.700563]
[26.]
epoch: 103 loss: 13.694169998168945
tensor([[162.6842]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[162.68417]
[155.]
epoch: 104 loss: 59.04652404785156
tensor([[118.4768]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[118.47676]
[119.]
epoch: 105 loss: 0.2737791836261749
tensor([[8.6790]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[8.678994]
[8.]
epoch: 106 loss: 0.4610331058502197
tensor([[5.6943]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[5.694289]
[3.]
epoch: 107 loss: 7.259194374084473
tensor([[15.1180]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')

```

```

[15.117995]
[18.]
epoch: 108 loss: 8.305951118469238
tensor([[121.8713]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[121.87132]
[134.]
epoch: 109 loss: 147.1048126220703
tensor([[75.3763]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[75.376335]
[83.]
epoch: 110 loss: 58.12026596069336
tensor([[16.0184]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[16.018398]
[20.]
epoch: 111 loss: 15.85315227508545
tensor([[56.1700]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[56.16999]
[58.]
epoch: 112 loss: 3.3489346504211426
tensor([[6.8593]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[6.8592615]
[4.]
epoch: 113 loss: 8.175375938415527
tensor([[198.3144]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[198.3144]
[190.]
epoch: 114 loss: 69.1293716430664
tensor([[178.8068]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[178.8068]
[175.]
epoch: 115 loss: 14.491674423217773
tensor([[28.2771]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[28.27715]
[27.]
epoch: 116 loss: 1.6311100721359253
tensor([[188.9356]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[188.93558]
[194.]
epoch: 117 loss: 25.64837646484375

```

```

tensor([[53.2952]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[53.29523]
[53.]
epoch: 118 loss: 0.08716126531362534
tensor([[88.7391]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[88.73907]
[91.]
epoch: 119 loss: 5.111817836761475
tensor([[77.7663]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[77.76635]
[79.]
epoch: 120 loss: 1.521892786026001
tensor([[104.1432]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[104.14321]
[108.]
epoch: 121 loss: 14.874818801879883
tensor([[31.9269]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[31.926891]
[30.]
epoch: 122 loss: 3.7129101753234863
tensor([[144.3392]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[144.33923]
[149.]
epoch: 123 loss: 21.722745895385742
tensor([[185.8337]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[185.83368]
[183.]
epoch: 124 loss: 8.02973747253418
tensor([[175.5601]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[175.56013]
[174.]
epoch: 125 loss: 2.434020757675171
tensor([[157.8282]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[157.82825]
[156.]
epoch: 126 loss: 3.342487335205078
tensor([[137.8979]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[137.8979]

```

```

[136.]
epoch: 127 loss: 3.6020374298095703
tensor([[144.5781]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[144.57814]
[146.]
epoch: 128 loss: 2.0216851234436035
tensor([[96.5776]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[96.57755]
[93.]
epoch: 129 loss: 12.798884391784668
tensor([[17.2734]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[17.273445]
[16.]
epoch: 130 loss: 1.6216624975204468
tensor([[72.0443]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[72.04435]
[69.]
epoch: 131 loss: 9.268064498901367
tensor([[144.5746]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[144.57458]
[152.]
epoch: 132 loss: 55.13678741455078
tensor([[12.6903]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[12.690334]
[11.]
epoch: 133 loss: 2.8572301864624023
tensor([[79.6171]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[79.61706]
[77.]
epoch: 134 loss: 6.848991394042969
tensor([[188.3700]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[188.37004]
[192.]
epoch: 135 loss: 13.176603317260742
tensor([[188.9976]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[188.9976]
[189.]
epoch: 136 loss: 5.7390425354242325e-06
tensor([[12.5673]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],

```

```

device='cuda:0')
[12.5672655]
[13.]
epoch: 137 loss: 0.18725913763046265
tensor([[6.4739]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[6.4739037]
[7.]
epoch: 138 loss: 0.27677735686302185
tensor([[38.5957]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[38.595695]
[35.]
epoch: 139 loss: 12.929025650024414
tensor([[192.3025]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[192.30249]
[182.]
epoch: 140 loss: 106.14130401611328
tensor([[40.4154]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[40.415432]
[39.]
epoch: 141 loss: 2.0034477710723877
tensor([[62.3341]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[62.334095]
[60.]
epoch: 142 loss: 5.447999477386475
tensor([[191.5161]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[191.51613]
[200.]
epoch: 143 loss: 71.97607421875
tensor([[96.6571]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[96.65708]
[99.]
epoch: 144 loss: 5.489266395568848
tensor([[0.0743]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[0.07431046]
[0.]
epoch: 145 loss: 0.005522044375538826
tensor([[183.6205]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[183.6205]
[186.]

```

```

epoch: 146 loss: 5.662026405334473
tensor([[27.9874]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[27.987385]
[29.]
epoch: 147 loss: 1.025389552116394
tensor([[110.4756]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[110.475624]
[109.]
epoch: 148 loss: 2.17746639251709
tensor([[99.0375]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[99.03754]
[96.]
epoch: 149 loss: 9.226628303527832
tensor([[32.1442]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[32.144222]
[33.]
epoch: 150 loss: 0.732355535030365
tensor([[119.2042]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[119.204155]
[117.]
epoch: 151 loss: 4.858299255371094
tensor([[142.4544]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[142.45438]
[141.]
epoch: 152 loss: 2.1152102947235107
tensor([[100.7582]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[100.75824]
[100.]
epoch: 153 loss: 0.5749275088310242
tensor([[167.9411]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[167.9411]
[166.]
epoch: 154 loss: 3.7678732872009277
tensor([[39.1215]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[39.12147]
[43.]
epoch: 155 loss: 15.042984008789062
tensor([[21.6533]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')

```

```

[21.653296]
[25.]
epoch: 156 loss: 11.200430870056152
tensor([[9.4428]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[9.442779]
[15.]
epoch: 157 loss: 30.882709503173828
tensor([[201.9872]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[201.98715]
[201.]
epoch: 158 loss: 0.9744692444801331
tensor([[133.0616]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[133.06158]
[138.]
epoch: 159 loss: 24.38794708251953
tensor([[99.9709]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[99.97089]
[102.]
epoch: 160 loss: 4.117302894592285
tensor([[85.1099]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[85.109856]
[85.]
epoch: 161 loss: 0.012068264186382294
tensor([[14.6618]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[14.661786]
[14.]
epoch: 162 loss: 0.43796080350875854
tensor([[147.7604]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[147.76038]
[144.]
epoch: 163 loss: 14.140427589416504
tensor([[53.4514]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[53.451397]
[51.]
epoch: 164 loss: 6.009346961975098
tensor([[18.7559]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[18.755882]
[19.]
epoch: 165 loss: 0.05959346890449524

```

```

tensor([[54.4772]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[54.477226]
[54.]
epoch: 166 loss: 0.2277449071407318
tensor([[55.3393]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[55.339287]
[56.]
epoch: 167 loss: 0.43654191493988037
tensor([[174.5040]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[174.504]
[180.]
epoch: 168 loss: 30.206039428710938
tensor([[7.9030]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[7.9029527]
[5.]
epoch: 169 loss: 8.42713451385498
tensor([[71.7459]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[71.74588]
[71.]
epoch: 170 loss: 0.5563371777534485
tensor([[8.8254]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[8.825373]
[6.]
epoch: 171 loss: 7.982730865478516
tensor([[90.7662]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[90.76624]
[88.]
epoch: 172 loss: 7.652100086212158
tensor([[122.8891]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[122.8891]
[120.]
epoch: 173 loss: 8.346893310546875
tensor([[175.3501]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[175.35005]
[170.]
epoch: 174 loss: 28.62305450439453
tensor([[136.8109]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[136.81091]

```



```

[137.]
epoch: 175 loss: 0.03575386106967926
tensor([[13.8382]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[13.838195]
[17.]
epoch: 176 loss: 9.9970121383667
tensor([[93.6363]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[93.63632]
[98.]
epoch: 177 loss: 19.041685104370117
tensor([[83.7037]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[83.70367]
[89.]
epoch: 178 loss: 28.0511474609375
tensor([[128.1382]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[128.13823]
[132.]
epoch: 179 loss: 14.913272857666016
tensor([[122.7543]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[122.75427]
[125.]
epoch: 180 loss: 5.043292045593262
tensor([[162.4183]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[162.41826]
[162.]
epoch: 181 loss: 0.1749403178691864
tensor([[126.3133]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[126.31333]
[124.]
epoch: 182 loss: 5.351502895355225
tensor([[180.4239]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[180.42392]
[176.]
epoch: 183 loss: 19.57106590270996
tensor([[124.9258]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[124.925835]
[123.]
epoch: 184 loss: 3.708839178085327
tensor([[162.4155]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],

```

```

device='cuda:0')
[162.41553]
[163.]
epoch: 185 loss: 0.34160828590393066
tensor([[3.1910]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[3.191016]
[2.]
epoch: 186 loss: 1.4185190200805664
For epoch 21 the test accuracy over the whole test set is 0 %
tensor([[179.2885]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[179.2885]
[182.]
epoch: 0 loss: 7.352243423461914
tensor([[67.3073]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[67.30727]
[68.]
epoch: 1 loss: 0.47988006472587585
tensor([[77.2423]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[77.242294]
[78.]
epoch: 2 loss: 0.57411789894104
tensor([[104.7383]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[104.73834]
[106.]
epoch: 3 loss: 1.591780185699463
tensor([[81.2474]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[81.24735]
[81.]
epoch: 4 loss: 0.06118330731987953
tensor([[96.8725]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[96.872505]
[96.]
epoch: 5 loss: 0.7612652778625488
tensor([[26.0241]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[26.024094]
[25.]
epoch: 6 loss: 1.0487678050994873
tensor([[118.8545]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[118.85447]

```

```

[120.]
epoch: 7 loss: 1.3122406005859375
tensor([[31.5322]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[31.532248]
[30.]
epoch: 8 loss: 2.347782611846924
tensor([[199.5289]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[199.5289]
[196.]
epoch: 9 loss: 12.453136444091797
tensor([[21.4550]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[21.454971]
[22.]
epoch: 10 loss: 0.29705625772476196
tensor([[154.4246]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[154.42456]
[157.]
epoch: 11 loss: 6.632888317108154
tensor([[73.7599]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[73.75993]
[74.]
epoch: 12 loss: 0.05763193964958191
tensor([[61.3325]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[61.33252]
[60.]
epoch: 13 loss: 1.7756083011627197
tensor([[81.6055]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[81.605515]
[82.]
epoch: 14 loss: 0.1556187868118286
tensor([[5.9405]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[5.940528]
[6.]
epoch: 15 loss: 0.003536928677931428
tensor([[123.3875]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[123.38753]
[125.]
epoch: 16 loss: 2.600067615509033
tensor([[6.2346]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],

```

```

device='cuda:0')
[6.234605]
[7.]
epoch: 17 loss: 0.5858297348022461
tensor([[5.2487]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[5.2486944]
[5.]
epoch: 18 loss: 0.06184891611337662
tensor([[116.8508]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[116.85075]
[119.]
epoch: 19 loss: 4.619259357452393
tensor([[136.0918]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[136.09183]
[137.]
epoch: 20 loss: 0.8247774839401245
tensor([[193.8174]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[193.81735]
[192.]
epoch: 21 loss: 3.302769422531128
tensor([[48.7642]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[48.764225]
[47.]
epoch: 22 loss: 3.112489938735962
tensor([[12.6870]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[12.686986]
[12.]
epoch: 23 loss: 0.4719497263431549
tensor([[19.5025]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[19.502455]
[21.]
epoch: 24 loss: 2.2426416873931885
tensor([[166.3986]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[166.39862]
[167.]
epoch: 25 loss: 0.36165717244148254
tensor([[204.6460]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[204.64601]
[205.]

```

```

epoch: 26 loss: 0.12530796229839325
tensor([[131.9999]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[131.99991]
[129.]
epoch: 27 loss: 8.99945068359375
tensor([[123.4267]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[123.42668]
[122.]
epoch: 28 loss: 2.0354201793670654
tensor([[93.4907]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[93.49068]
[92.]
epoch: 29 loss: 2.2221176624298096
tensor([[14.4616]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[14.461573]
[15.]
epoch: 30 loss: 0.28990402817726135
tensor([[146.9694]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[146.96942]
[152.]
epoch: 31 loss: 25.306720733642578
tensor([[86.0271]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[86.02709]
[86.]
epoch: 32 loss: 0.0007339753792621195
tensor([[48.5757]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[48.57574]
[48.]
epoch: 33 loss: 0.33147749304771423
tensor([[145.5895]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[145.58951]
[148.]
epoch: 34 loss: 5.810471534729004
tensor([[12.9310]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[12.9310255]
[11.]
epoch: 35 loss: 3.7288594245910645
tensor([[113.3504]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')

```

```

[113.350365]
[112.]
epoch: 36 loss: 1.8234847784042358
tensor([[29.1544]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[29.154448]
[28.]
epoch: 37 loss: 1.3327491283416748
tensor([[11.7824]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[11.782379]
[10.]
epoch: 38 loss: 3.176875352859497
tensor([[7.5966]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[7.5965548]
[8.]
epoch: 39 loss: 0.16276806592941284
tensor([[51.7136]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[51.713608]
[50.]
epoch: 40 loss: 2.9364516735076904
tensor([[52.3013]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[52.30127]
[51.]
epoch: 41 loss: 1.6933023929595947
tensor([[119.8153]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[119.81532]
[117.]
epoch: 42 loss: 7.9260430335998535
tensor([[90.9943]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[90.994286]
[89.]
epoch: 43 loss: 3.977174997329712
tensor([[169.2106]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[169.2106]
[168.]
epoch: 44 loss: 1.4655567407608032
tensor([[137.6926]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[137.69258]
[139.]
epoch: 45 loss: 1.7093440294265747

```

```

tensor([[52.1578]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[52.157757]
[55.]
epoch: 46 loss: 8.078346252441406
tensor([[122.3490]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[122.348976]
[124.]
epoch: 47 loss: 2.7258799076080322
tensor([[107.1249]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[107.12489]
[111.]
epoch: 48 loss: 15.01645278930664
tensor([[134.5023]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[134.50235]
[135.]
epoch: 49 loss: 0.24765567481517792
tensor([[155.3062]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[155.30617]
[155.]
epoch: 50 loss: 0.0937386006116867
tensor([[113.3036]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[113.30363]
[114.]
epoch: 51 loss: 0.48493534326553345
tensor([[13.9754]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[13.975444]
[18.]
epoch: 52 loss: 16.197052001953125
tensor([[33.1011]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[33.10109]
[35.]
epoch: 53 loss: 3.605861186981201
tensor([[149.8108]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[149.81084]
[145.]
epoch: 54 loss: 23.144149780273438
tensor([[34.4178]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[34.41775]

```

```

[36.]
epoch: 55 loss: 2.5035109519958496
tensor([[176.1142]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[176.11423]
[172.]
epoch: 56 loss: 16.92686653137207
tensor([[105.2878]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[105.287766]
[104.]
epoch: 57 loss: 1.6583399772644043
tensor([[4.2509]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[4.2509227]
[4.]
epoch: 58 loss: 0.06296218931674957
tensor([[195.8121]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[195.81207]
[200.]
epoch: 59 loss: 17.538734436035156
tensor([[144.1768]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[144.1768]
[149.]
epoch: 60 loss: 23.26322364807129
tensor([[194.2930]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[194.29305]
[195.]
epoch: 61 loss: 0.49978530406951904
tensor([[173.5068]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[173.5068]
[174.]
epoch: 62 loss: 0.24324089288711548
tensor([[101.3967]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[101.39668]
[98.]
epoch: 63 loss: 11.537453651428223
tensor([[3.4562]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[3.4562068]
[1.]
epoch: 64 loss: 6.032951831817627
tensor([[79.8715]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],

```



```

device='cuda:0')
[79.87149]
[77.]
epoch: 65 loss: 8.245457649230957
tensor([[82.5324]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[82.532394]
[80.]
epoch: 66 loss: 6.413021564483643
tensor([[15.9953]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[15.995292]
[17.]
epoch: 67 loss: 1.0094387531280518
tensor([[89.7939]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[89.793945]
[88.]
epoch: 68 loss: 3.2182397842407227
tensor([[64.5438]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[64.54378]
[65.]
epoch: 69 loss: 0.20813900232315063
tensor([[109.5573]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[109.5573]
[109.]
epoch: 70 loss: 0.31057965755462646
tensor([[92.3104]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[92.31036]
[93.]
epoch: 71 loss: 0.47559812664985657
tensor([[114.3268]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[114.326775]
[116.]
epoch: 72 loss: 2.7996833324432373
tensor([[92.5770]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[92.57699]
[94.]
epoch: 73 loss: 2.0249624252319336
tensor([[170.8309]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[170.8309]
[170.]

```

```

epoch: 74 loss: 0.6903982758522034
tensor([[147.2607]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[147.2607]
[147.]
epoch: 75 loss: 0.06796261668205261
tensor([[51.6243]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[51.624264]
[54.]
epoch: 76 loss: 5.64412260055542
tensor([[190.9951]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[190.99506]
[186.]
epoch: 77 loss: 24.950586318969727
tensor([[25.2201]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[25.220121]
[27.]
epoch: 78 loss: 3.1679677963256836
tensor([[38.5878]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[38.587788]
[43.]
epoch: 79 loss: 19.467618942260742
tensor([[165.9991]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[165.99905]
[173.]
epoch: 80 loss: 49.01324462890625
tensor([[104.9580]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[104.95802]
[108.]
epoch: 81 loss: 9.253623962402344
tensor([[160.8410]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[160.84097]
[159.]
epoch: 82 loss: 3.389181137084961
tensor([[10.2924]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[10.29236]
[9.]
epoch: 83 loss: 1.6701951026916504
tensor([[135.0226]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')

```

```

[135.02255]
[130.]
epoch: 84 loss: 25.22603416442871
tensor([[105.8027]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[105.802734]
[103.]
epoch: 85 loss: 7.855319976806641
tensor([[194.8731]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[194.87314]
[190.]
epoch: 86 loss: 23.747478485107422
tensor([[100.6235]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[100.62347]
[102.]
epoch: 87 loss: 1.8948445320129395
tensor([[164.5646]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[164.56459]
[171.]
epoch: 88 loss: 41.41449737548828
tensor([[54.8680]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[54.867954]
[56.]
epoch: 89 loss: 1.2815275192260742
tensor([[80.8232]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[80.823204]
[84.]
epoch: 90 loss: 10.092032432556152
tensor([[170.7379]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[170.73787]
[175.]
epoch: 91 loss: 18.16575813293457
tensor([[31.2234]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[31.223415]
[29.]
epoch: 92 loss: 4.943575859069824
tensor([[59.8563]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[59.85632]
[57.]
epoch: 93 loss: 8.158560752868652

```

```

tensor([[40.0124]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[40.012398]
[38.]
epoch: 94 loss: 4.049744606018066
tensor([[142.1879]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[142.1879]
[136.]
epoch: 95 loss: 38.29006576538086
tensor([[168.7346]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[168.73462]
[163.]
epoch: 96 loss: 32.88585662841797
tensor([[44.0863]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[44.086296]
[44.]
epoch: 97 loss: 0.007447013631463051
tensor([[60.9428]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[60.942772]
[62.]
epoch: 98 loss: 1.1177312135696411
tensor([[153.5508]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[153.55084]
[160.]
epoch: 99 loss: 41.591636657714844
tensor([[71.8877]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[71.88768]
[76.]
epoch: 100 loss: 16.911174774169922
tensor([[148.0716]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[148.07156]
[154.]
epoch: 101 loss: 35.146358489990234
tensor([[12.1741]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[12.174135]
[13.]
epoch: 102 loss: 0.6820526719093323
tensor([[69.8167]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[69.81673]

```

```

[69.]
epoch: 103 loss: 0.6670424938201904
tensor([[87.9497]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[87.94969]
[85.]
epoch: 104 loss: 8.700681686401367
tensor([[213.6676]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[213.66765]
[201.]
epoch: 105 loss: 160.4693145751953
tensor([[61.9295]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[61.929535]
[61.]
epoch: 106 loss: 0.8640351295471191
tensor([[23.4097]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[23.409737]
[24.]
epoch: 107 loss: 0.34841084480285645
tensor([[90.1296]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[90.1296]
[97.]
epoch: 108 loss: 47.202388763427734
tensor([[90.4758]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[90.475815]
[100.]
epoch: 109 loss: 90.7101058959961
tensor([[49.9993]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[49.999332]
[53.]
epoch: 110 loss: 9.004005432128906
tensor([[160.1624]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[160.16235]
[176.]
epoch: 111 loss: 250.83103942871094
tensor([[114.4928]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[114.49283]
[118.]
epoch: 112 loss: 12.300252914428711
tensor([[57.6450]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],

```

```

device='cuda:0')
[57.645042]
[52.]
epoch: 113 loss: 31.866504669189453
tensor([[190.8385]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[190.83855]
[177.]
epoch: 114 loss: 191.50537109375
tensor([[52.6878]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[52.6878]
[46.]
epoch: 115 loss: 44.726688385009766
tensor([[103.4983]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[103.49831]
[95.]
epoch: 116 loss: 72.22120666503906
tensor([[15.7147]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[15.714682]
[14.]
epoch: 117 loss: 2.9401330947875977
tensor([[80.0781]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[80.0781]
[79.]
epoch: 118 loss: 1.162304162979126
tensor([[17.6714]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[17.67137]
[20.]
epoch: 119 loss: 5.422519683837891
tensor([[179.2611]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[179.26106]
[194.]
epoch: 120 loss: 217.2362823486328
tensor([[182.4027]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[182.40274]
[191.]
epoch: 121 loss: 73.91287231445312
tensor([[108.0015]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[108.001526]
[105.]

```

```

epoch: 122 loss: 9.009157180786133
tensor([[16.7619]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[16.761873]
[16.]
epoch: 123 loss: 0.5804508328437805
tensor([[215.1333]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[215.13332]
[193.]
epoch: 124 loss: 489.8836669921875
tensor([[158.4507]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[158.45067]
[151.]
epoch: 125 loss: 55.51245880126953
tensor([[29.8858]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[29.885788]
[33.]
epoch: 126 loss: 9.69831657409668
tensor([[3.1593]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[3.1592603]
[3.]
epoch: 127 loss: 0.02536383457481861
tensor([[24.9230]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[24.922955]
[32.]
epoch: 128 loss: 50.084571838378906
tensor([[152.1398]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[152.1398]
[185.]
epoch: 129 loss: 1079.792724609375
tensor([[52.2398]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[52.239758]
[64.]
epoch: 130 loss: 138.3032989501953
tensor([[189.6221]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[189.62212]
[204.]
epoch: 131 loss: 206.72354125976562
tensor([[112.3750]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')

```

```

[112.37499]
[107.]
epoch: 132 loss: 28.89054298400879
tensor([[232.9227]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[232.9227]
[203.]
epoch: 133 loss: 895.367919921875
tensor([[195.1565]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[195.15648]
[178.]
epoch: 134 loss: 294.3447570800781
tensor([[111.0054]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[111.0054]
[113.]
epoch: 135 loss: 3.9784226417541504
tensor([[160.1264]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[160.12643]
[179.]
epoch: 136 loss: 356.21148681640625
tensor([[127.0495]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[127.049515]
[146.]
epoch: 137 loss: 359.1208801269531
tensor([[93.1692]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[93.16921]
[101.]
epoch: 138 loss: 61.32123565673828
tensor([[120.2717]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[120.27169]
[126.]
epoch: 139 loss: 32.81352996826172
tensor([[69.7036]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[69.70363]
[67.]
epoch: 140 loss: 7.30960750579834
tensor([[166.4248]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[166.42477]
[164.]
epoch: 141 loss: 5.87952995300293

```



```

tensor([[100.4557]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[100.4557]
[91.]
epoch: 142 loss: 89.41033172607422
tensor([[32.1482]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[32.14817]
[26.]
epoch: 143 loss: 37.79999923706055
tensor([[151.0359]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[151.03595]
[143.]
epoch: 144 loss: 64.57648468017578
tensor([[107.0965]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[107.09653]
[99.]
epoch: 145 loss: 65.55374908447266
tensor([[140.1800]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[140.17998]
[142.]
epoch: 146 loss: 3.3124821186065674
tensor([[46.9580]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[46.95801]
[45.]
epoch: 147 loss: 3.8338096141815186
tensor([[57.8021]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[57.802116]
[58.]
epoch: 148 loss: 0.0391579233109951
tensor([[102.2535]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[102.25348]
[110.]
epoch: 149 loss: 60.00858688354492
tensor([[161.4481]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[161.44806]
[180.]
epoch: 150 loss: 344.17449951171875
tensor([[150.1170]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[150.11703]

```

```

[156.]
epoch: 151 loss: 34.60927963256836
tensor([[23.5735]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[23.573496]
[23.]
epoch: 152 loss: 0.32889750599861145
tensor([[178.9807]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[178.98071]
[166.]
epoch: 153 loss: 168.4989013671875
tensor([[211.5022]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[211.5022]
[198.]
epoch: 154 loss: 182.309326171875
tensor([[141.9977]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[141.99771]
[134.]
epoch: 155 loss: 63.963382720947266
tensor([[135.6421]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[135.64212]
[138.]
epoch: 156 loss: 5.559596538543701
tensor([[130.6173]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[130.61734]
[141.]
epoch: 157 loss: 107.79962921142578
tensor([[78.0128]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[78.01285]
[83.]
epoch: 158 loss: 24.871686935424805
tensor([[139.3935]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[139.39351]
[153.]
epoch: 159 loss: 185.1365966796875
tensor([[127.5857]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[127.58572]
[132.]
epoch: 160 loss: 19.48590087890625
tensor([[185.1444]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],

```

```

device='cuda:0')
[185.1444]
[199.]
epoch: 161 loss: 191.97781372070312
tensor([[129.1726]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[129.1726]
[115.]
epoch: 162 loss: 200.86280822753906
tensor([[89.0852]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[89.085236]
[75.]
epoch: 163 loss: 198.39385986328125
tensor([[57.4845]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[57.484547]
[49.]
epoch: 164 loss: 71.98753356933594
tensor([[207.6232]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[207.62318]
[183.]
epoch: 165 loss: 606.3012084960938
tensor([[188.1876]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[188.18758]
[184.]
epoch: 166 loss: 17.535795211791992
tensor([[125.4079]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[125.40785]
[140.]
epoch: 167 loss: 212.93077087402344
tensor([[10.9263]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[10.926254]
[19.]
epoch: 168 loss: 65.18537139892578
tensor([[66.8402]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[66.84017]
[87.]
epoch: 169 loss: 406.4186706542969
tensor([[53.2212]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[53.221237]
[70.]

```

```

epoch: 170 loss: 281.5268859863281
tensor([[27.7117]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[27.7117]
[34.]
epoch: 171 loss: 39.5427131652832
tensor([[129.0929]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[129.0929]
[150.]
epoch: 172 loss: 437.1070251464844
tensor([[84.6625]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[84.66249]
[90.]
epoch: 173 loss: 28.489004135131836
tensor([[40.7188]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[40.71881]
[37.]
epoch: 174 loss: 13.82955551147461
tensor([[148.7288]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[148.72876]
[133.]
epoch: 175 loss: 247.39389038085938
tensor([[138.8008]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[138.80075]
[123.]
epoch: 176 loss: 249.66372680664062
tensor([[171.9209]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[171.92094]
[165.]
epoch: 177 loss: 47.89946746826172
tensor([[14.0631]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[14.063114]
[2.]
epoch: 178 loss: 145.5187225341797
tensor([[154.9527]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[154.95268]
[169.]
epoch: 179 loss: 197.32713317871094
tensor([[169.6387]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')

```

```

[169.63869]
[189.]
epoch: 180 loss: 374.8604431152344
tensor([[10.9133]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[10.913319]
[0.]
epoch: 181 loss: 119.10052490234375
tensor([[147.7757]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[147.77573]
[144.]
epoch: 182 loss: 14.256109237670898
tensor([[76.4580]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[76.45798]
[71.]
epoch: 183 loss: 29.78951644897461
tensor([[173.6656]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[173.66559]
[162.]
epoch: 184 loss: 136.08595275878906
tensor([[38.0493]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[38.049297]
[39.]
epoch: 185 loss: 0.9038355350494385
tensor([[198.3631]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[198.36311]
[197.]
epoch: 186 loss: 1.858078122138977
For epoch 22 the test accuracy over the whole test set is 0 %
tensor([[157.4037]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[157.40372]
[156.]
epoch: 0 loss: 1.970421552658081
tensor([[190.9191]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[190.91905]
[200.]
epoch: 1 loss: 82.46361541748047
tensor([[122.9064]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[122.9064]
[126.]

```

```

epoch: 2 loss: 9.570344924926758
tensor([[128.3499]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[128.34991]
[129.]
epoch: 3 loss: 0.4226110875606537
tensor([[173.0966]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[173.0966]
[169.]
epoch: 4 loss: 16.78215980529785
tensor([[91.1072]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[91.107216]
[93.]
epoch: 5 loss: 3.582631826400757
tensor([[13.0961]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[13.096051]
[22.]
epoch: 6 loss: 79.28030395507812
tensor([[154.5593]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[154.55928]
[148.]
epoch: 7 loss: 43.0241584777832
tensor([[8.0967]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[8.096723]
[13.]
epoch: 8 loss: 24.042129516601562
tensor([[191.2207]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[191.22075]
[185.]
epoch: 9 loss: 38.697715759277344
tensor([[46.6601]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[46.66007]
[54.]
epoch: 10 loss: 53.874595642089844
tensor([[66.9550]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[66.95497]
[75.]
epoch: 11 loss: 64.72248840332031
tensor([[9.7764]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')

```

```

[9.776443]
[12.]
epoch: 12 loss: 4.944207668304443
tensor([[9.7643]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[9.764251]
[10.]
epoch: 13 loss: 0.055577706545591354
tensor([[187.4080]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[187.40804]
[195.]
epoch: 14 loss: 57.6379280090332
tensor([[186.8269]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[186.8269]
[190.]
epoch: 15 loss: 10.068536758422852
tensor([[166.3659]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[166.3659]
[165.]
epoch: 16 loss: 1.8656985759735107
tensor([[171.2756]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[171.27556]
[166.]
epoch: 17 loss: 27.83151626586914
tensor([[197.3392]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[197.3392]
[189.]
epoch: 18 loss: 69.54230499267578
tensor([[34.7625]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[34.762486]
[32.]
epoch: 19 loss: 7.631326198577881
tensor([[174.7541]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[174.75409]
[176.]
epoch: 20 loss: 1.552293300628662
tensor([[46.8869]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[46.886852]
[47.]
epoch: 21 loss: 0.012802409939467907

```

```

tensor([[192.1582]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[192.15822]
[205.]
epoch: 22 loss: 164.91136169433594
tensor([[19.1902]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[19.190184]
[17.]
epoch: 23 loss: 4.796904563903809
tensor([[44.6135]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[44.613525]
[43.]
epoch: 24 loss: 2.603464126586914
tensor([[38.4258]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[38.42576]
[35.]
epoch: 25 loss: 11.735820770263672
tensor([[123.8154]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[123.81544]
[119.]
epoch: 26 loss: 23.18843650817871
tensor([[27.2287]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[27.228743]
[24.]
epoch: 27 loss: 10.424778938293457
tensor([[112.5478]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[112.54781]
[107.]
epoch: 28 loss: 30.778234481811523
tensor([[207.2550]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[207.25496]
[197.]
epoch: 29 loss: 105.1641845703125
tensor([[35.3580]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[35.35797]
[37.]
epoch: 30 loss: 2.696258544921875
tensor([[40.9170]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[40.916985]

```



```

[45.]
epoch: 31 loss: 16.6710147857666
tensor([[59.1746]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[59.1746]
[65.]
epoch: 32 loss: 33.935298919677734
tensor([[23.3728]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[23.372768]
[26.]
epoch: 33 loss: 6.902345657348633
tensor([[72.3837]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[72.38372]
[82.]
epoch: 34 loss: 92.47283172607422
tensor([[119.5250]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[119.52498]
[135.]
epoch: 35 loss: 239.47628784179688
tensor([[48.1093]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[48.109283]
[51.]
epoch: 36 loss: 8.356242179870605
tensor([[159.5518]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[159.55179]
[170.]
epoch: 37 loss: 109.16513061523438
tensor([[46.9950]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[46.994972]
[44.]
epoch: 38 loss: 8.96985912322998
tensor([[118.9015]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[118.90145]
[114.]
epoch: 39 loss: 24.02422332763672
tensor([[163.3724]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[163.37242]
[150.]
epoch: 40 loss: 178.8216552734375
tensor([[23.6323]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],

```

```

device='cuda:0')
[23.632267]
[18.]
epoch: 41 loss: 31.722431182861328
tensor([[189.0097]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[189.00967]
[179.]
epoch: 42 loss: 100.19357299804688
tensor([[179.6254]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[179.62544]
[178.]
epoch: 43 loss: 2.6420633792877197
tensor([[19.7419]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[19.741901]
[19.]
epoch: 44 loss: 0.5504176616668701
tensor([[51.5484]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[51.54844]
[49.]
epoch: 45 loss: 6.494541645050049
tensor([[90.3697]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[90.36965]
[97.]
epoch: 46 loss: 43.961517333984375
tensor([[173.7191]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[173.71912]
[201.]
epoch: 47 loss: 744.2466430664062
tensor([[86.9974]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[86.99742]
[89.]
epoch: 48 loss: 4.010321617126465
tensor([[154.1746]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[154.17459]
[155.]
epoch: 49 loss: 0.6812999248504639
tensor([[111.3752]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[111.37519]
[104.]

```

```
epoch: 50 loss: 54.39344024658203
tensor([[207.9189]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[207.91893]
[191.]
epoch: 51 loss: 286.25018310546875
tensor([[175.1524]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[175.15236]
[167.]
epoch: 52 loss: 66.4609603881836
tensor([[138.8442]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[138.8442]
[138.]
epoch: 53 loss: 0.7126609683036804
tensor([[80.7497]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[80.74974]
[80.]
epoch: 54 loss: 0.562110960483551
tensor([[89.4346]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[89.43458]
[95.]
epoch: 55 loss: 30.973922729492188
tensor([[135.0410]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[135.04105]
[152.]
epoch: 56 loss: 287.6061096191406
tensor([[156.0757]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[156.07568]
[174.]
epoch: 57 loss: 321.2811279296875
tensor([[164.6438]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[164.64377]
[173.]
epoch: 58 loss: 69.82660675048828
tensor([[77.2852]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[77.28516]
[68.]
epoch: 59 loss: 86.21412658691406
tensor([[83.4782]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
```

```

[83.47822]
[70.]
epoch: 60 loss: 181.66236877441406
tensor([[159.8593]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[159.85928]
[140.]
epoch: 61 loss: 394.3911437988281
tensor([[195.3600]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[195.35997]
[175.]
epoch: 62 loss: 414.52838134765625
tensor([[3.9657]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[3.9656985]
[5.]
epoch: 63 loss: 1.069779634475708
tensor([[132.5888]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[132.58879]
[142.]
epoch: 64 loss: 88.57085418701172
tensor([[88.5124]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[88.51237]
[99.]
epoch: 65 loss: 109.99044036865234
tensor([[97.9200]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[97.92002]
[116.]
epoch: 66 loss: 326.8856506347656
tensor([[116.4434]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[116.44339]
[132.]
epoch: 67 loss: 242.00811767578125
tensor([[122.1658]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[122.16585]
[133.]
epoch: 68 loss: 117.37885284423828
tensor([[76.2873]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[76.287315]
[78.]
epoch: 69 loss: 2.93328857421875

```

```

tensor([[175.7373]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[175.73732]
[171.]
epoch: 70 loss: 22.44219970703125
tensor([[85.6905]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[85.69054]
[79.]
epoch: 71 loss: 44.76327896118164
tensor([[59.1619]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[59.16189]
[53.]
epoch: 72 loss: 37.96886444091797
tensor([[93.1023]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[93.10231]
[83.]
epoch: 73 loss: 102.05667114257812
tensor([[209.2933]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[209.29326]
[193.]
epoch: 74 loss: 265.47027587890625
tensor([[98.0414]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[98.041374]
[94.]
epoch: 75 loss: 16.332704544067383
tensor([[94.5348]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[94.53475]
[96.]
epoch: 76 loss: 2.146951913833618
tensor([[74.7231]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[74.72305]
[81.]
epoch: 77 loss: 39.400062561035156
tensor([[133.2815]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[133.28152]
[149.]
epoch: 78 loss: 247.07046508789062
tensor([[89.0560]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[89.055984]

```

```

[98.]
epoch: 79 loss: 79.99541473388672
tensor([[3.4367]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[3.436733]
[3.]
epoch: 80 loss: 0.19073571264743805
tensor([[101.2160]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[101.216]
[108.]
epoch: 81 loss: 46.02260971069336
tensor([[116.5992]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[116.59923]
[120.]
epoch: 82 loss: 11.565250396728516
tensor([[32.1596]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[32.15963]
[33.]
epoch: 83 loss: 0.7062220573425293
tensor([[39.3796]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[39.37961]
[38.]
epoch: 84 loss: 1.9033186435699463
tensor([[177.2708]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[177.27081]
[159.]
epoch: 85 loss: 333.8226013183594
tensor([[40.0620]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[40.06198]
[39.]
epoch: 86 loss: 1.1278040409088135
tensor([[4.3505]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[4.3504705]
[0.]
epoch: 87 loss: 18.926593780517578
tensor([[121.0178]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[121.01776]
[124.]
epoch: 88 loss: 8.89374828338623
tensor([[136.5078]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],

```

```

device='cuda:0')
[136.50778]
[147.]
epoch: 89 loss: 110.08663940429688
tensor([[83.0566]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[83.05664]
[84.]
epoch: 90 loss: 0.8899269104003906
tensor([[132.2038]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[132.20375]
[137.]
epoch: 91 loss: 23.00400733947754
tensor([[140.6507]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[140.6507]
[144.]
epoch: 92 loss: 11.217838287353516
tensor([[51.1861]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[51.186115]
[48.]
epoch: 93 loss: 10.151330947875977
tensor([[60.9514]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[60.95141]
[55.]
epoch: 94 loss: 35.41926193237305
tensor([[181.7031]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[181.70314]
[194.]
epoch: 95 loss: 151.21275329589844
tensor([[196.5057]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[196.50572]
[186.]
epoch: 96 loss: 110.37019348144531
tensor([[177.3846]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[177.3846]
[160.]
epoch: 97 loss: 302.2242431640625
tensor([[68.7943]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[68.79434]
[64.]

```

```

epoch: 98 loss: 22.985715866088867
tensor([[185.4083]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[185.40826]
[192.]
epoch: 99 loss: 43.45098114013672
tensor([[89.9820]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[89.98202]
[90.]
epoch: 100 loss: 0.0003233696916140616
tensor([[100.0203]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[100.02028]
[101.]
epoch: 101 loss: 0.9598533511161804
tensor([[99.0953]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[99.09528]
[103.]
epoch: 102 loss: 15.246810913085938
tensor([[185.5303]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[185.5303]
[204.]
epoch: 103 loss: 341.1296691894531
tensor([[146.1765]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[146.17645]
[146.]
epoch: 104 loss: 0.03113553300499916
tensor([[137.9509]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[137.95085]
[130.]
epoch: 105 loss: 63.21603775024414
tensor([[122.0208]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[122.02075]
[115.]
epoch: 106 loss: 49.290958404541016
tensor([[184.3009]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[184.30086]
[177.]
epoch: 107 loss: 53.302520751953125
tensor([[157.4205]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')

```



```

[157.42053]
[151.]
epoch: 108 loss: 41.22323226928711
tensor([[115.6970]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[115.69701]
[117.]
epoch: 109 loss: 1.6977927684783936
tensor([[117.5410]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[117.54098]
[122.]
epoch: 110 loss: 19.88288116455078
tensor([[174.2756]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[174.27556]
[183.]
epoch: 111 loss: 76.11588287353516
tensor([[8.6499]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[8.649911]
[11.]
epoch: 112 loss: 5.522918701171875
tensor([[94.4188]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[94.41882]
[100.]
epoch: 113 loss: 31.149534225463867
tensor([[79.5898]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[79.58976]
[85.]
epoch: 114 loss: 29.27069854736328
tensor([[26.3696]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[26.369642]
[30.]
epoch: 115 loss: 13.179497718811035
tensor([[107.8449]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[107.84489]
[102.]
epoch: 116 loss: 34.16270065307617
tensor([[116.5129]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[116.512886]
[110.]
epoch: 117 loss: 42.417686462402344

```

```

tensor([[9.3044]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[9.304355]
[8.]
epoch: 118 loss: 1.701341152191162
tensor([[189.1723]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[189.17232]
[180.]
epoch: 119 loss: 84.13140869140625
tensor([[118.0717]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[118.07173]
[118.]
epoch: 120 loss: 0.0051454175263643265
tensor([[30.1084]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[30.10835]
[36.]
epoch: 121 loss: 34.71153259277344
tensor([[84.6919]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[84.69192]
[91.]
epoch: 122 loss: 39.79190444946289
tensor([[13.3662]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[13.366203]
[15.]
epoch: 123 loss: 2.6692917346954346
tensor([[62.8945]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[62.894493]
[71.]
epoch: 124 loss: 65.6992416381836
tensor([[9.4626]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[9.462564]
[6.]
epoch: 125 loss: 11.989352226257324
tensor([[197.6632]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[197.66318]
[199.]
epoch: 126 loss: 1.7870944738388062
tensor([[66.4567]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[66.45674]

```

```

[69.]
epoch: 127 loss: 6.468164443969727
tensor([[18.7924]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[18.792381]
[16.]
epoch: 128 loss: 7.797393321990967
tensor([[89.3213]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[89.32127]
[86.]
epoch: 129 loss: 11.03085994720459
tensor([[13.9396]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[13.939555]
[9.]
epoch: 130 loss: 24.399206161499023
tensor([[168.6301]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[168.6301]
[168.]
epoch: 131 loss: 0.3970215320587158
tensor([[77.1197]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[77.119705]
[76.]
epoch: 132 loss: 1.2537397146224976
tensor([[137.6180]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[137.618]
[136.]
epoch: 133 loss: 2.6179118156433105
tensor([[151.2422]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[151.24225]
[153.]
epoch: 134 loss: 3.0896902084350586
tensor([[192.3992]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[192.39919]
[196.]
epoch: 135 loss: 12.965867042541504
tensor([[88.0884]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[88.08842]
[88.]
epoch: 136 loss: 0.007817575708031654
tensor([[3.6661]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],

```

```

device='cuda:0')
[3.666122]
[7.]
epoch: 137 loss: 11.11474323272705
tensor([[28.2110]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[28.211002]
[27.]
epoch: 138 loss: 1.466526746749878
tensor([[2.1949]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[2.1949]
[4.]
epoch: 139 loss: 3.2583858966827393
tensor([[46.1606]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[46.16065]
[46.]
epoch: 140 loss: 0.02580789104104042
tensor([[145.7329]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[145.73291]
[141.]
epoch: 141 loss: 22.40043830871582
tensor([[173.0665]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[173.06647]
[172.]
epoch: 142 loss: 1.137352466583252
tensor([[157.7858]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[157.78578]
[157.]
epoch: 143 loss: 0.6174531579017639
tensor([[91.6047]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[91.6047]
[92.]
epoch: 144 loss: 0.156263530254364
tensor([[54.1205]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[54.120502]
[56.]
epoch: 145 loss: 3.532510995864868
tensor([[34.4865]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[34.486458]
[34.]

```

```

epoch: 146 loss: 0.23664121329784393
tensor([[129.2643]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[129.26427]
[134.]
epoch: 147 loss: 22.427167892456055
tensor([[121.3617]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[121.361664]
[125.]
epoch: 148 loss: 13.237489700317383
tensor([[71.9520]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[71.95201]
[74.]
epoch: 149 loss: 4.194258689880371
tensor([[160.0215]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[160.02155]
[164.]
epoch: 150 loss: 15.82810115814209
tensor([[51.7272]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[51.727154]
[50.]
epoch: 151 loss: 2.983060121536255
tensor([[193.1370]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[193.137]
[184.]
epoch: 152 loss: 83.48464965820312
tensor([[17.9793]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[17.97935]
[20.]
epoch: 153 loss: 4.083029747009277
tensor([[91.2076]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[91.207634]
[87.]
epoch: 154 loss: 17.70418357849121
tensor([[63.4770]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[63.47704]
[62.]
epoch: 155 loss: 2.181645154953003
tensor([[26.2977]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')

```

```

[26.297718]
[25.]
epoch: 156 loss: 1.6840721368789673
tensor([[50.3815]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[50.38151]
[52.]
epoch: 157 loss: 2.619504451751709
tensor([[152.3495]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[152.34947]
[163.]
epoch: 158 loss: 113.43374633789062
tensor([[197.6882]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[197.68825]
[203.]
epoch: 159 loss: 28.214712142944336
tensor([[201.1626]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[201.16258]
[198.]
epoch: 160 loss: 10.001927375793457
tensor([[4.0309]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[4.030921]
[2.]
epoch: 161 loss: 4.124639987945557
tensor([[112.4186]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[112.41859]
[106.]
epoch: 162 loss: 41.19825744628906
tensor([[130.5897]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[130.58972]
[123.]
epoch: 163 loss: 57.60387420654297
tensor([[23.4525]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[23.45252]
[23.]
epoch: 164 loss: 0.20477469265460968
tensor([[57.5211]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[57.52113]
[57.]
epoch: 165 loss: 0.27157607674598694

```

```

tensor([[59.4993]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[59.499287]
[60.]
epoch: 166 loss: 0.25071385502815247
tensor([[152.8553]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[152.85529]
[154.]
epoch: 167 loss: 1.310370922088623
tensor([[177.0236]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[177.0236]
[182.]
epoch: 168 loss: 24.764503479003906
tensor([[108.7342]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[108.73416]
[112.]
epoch: 169 loss: 10.665701866149902
tensor([[146.0313]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[146.03131]
[145.]
epoch: 170 loss: 1.0636024475097656
tensor([[65.9008]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[65.90084]
[67.]
epoch: 171 loss: 1.2081509828567505
tensor([[12.8837]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[12.88369]
[14.]
epoch: 172 loss: 1.2461482286453247
tensor([[57.7899]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[57.789852]
[58.]
epoch: 173 loss: 0.044162120670080185
tensor([[108.3210]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[108.32101]
[105.]
epoch: 174 loss: 11.029086112976074
tensor([[149.5222]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[149.52219]

```

```

[143.]
epoch: 175 loss: 42.53891372680664
tensor([[76.9344]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[76.934364]
[77.]
epoch: 176 loss: 0.004308042582124472
tensor([[2.1352]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[2.1352391]
[1.]
epoch: 177 loss: 1.2887678146362305
tensor([[17.2318]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[17.23184]
[21.]
epoch: 178 loss: 14.199028968811035
tensor([[108.9241]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[108.92407]
[113.]
epoch: 179 loss: 16.613187789916992
tensor([[29.6488]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[29.64882]
[29.]
epoch: 180 loss: 0.42096853256225586
tensor([[57.8462]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[57.846245]
[61.]
epoch: 181 loss: 9.946171760559082
tensor([[28.8390]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[28.83903]
[28.]
epoch: 182 loss: 0.703970193862915
tensor([[137.3267]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[137.32666]
[139.]
epoch: 183 loss: 2.8000662326812744
tensor([[160.5381]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[160.53812]
[162.]
epoch: 184 loss: 2.13710355758667
tensor([[108.0987]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],

```



```

device='cuda:0')
[108.09875]
[109.]
epoch: 185 loss: 0.8122565150260925
tensor([[110.5164]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[110.516426]
[111.]
epoch: 186 loss: 0.23384372889995575
For epoch 23 the test accuracy over the whole test set is 0 %
tensor([[161.6848]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[161.68483]
[157.]
epoch: 0 loss: 21.947629928588867
tensor([[110.6167]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[110.616745]
[110.]
epoch: 1 loss: 0.38037440180778503
tensor([[178.3293]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[178.32927]
[177.]
epoch: 2 loss: 1.76695716381073
tensor([[16.9659]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[16.965855]
[17.]
epoch: 3 loss: 0.0011659052688628435
tensor([[4.3538]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[4.353818]
[2.]
epoch: 4 loss: 5.540458679199219
tensor([[41.7586]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[41.75859]
[43.]
epoch: 5 loss: 1.5410970449447632
tensor([[173.9259]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[173.92595]
[178.]
epoch: 6 loss: 16.597890853881836
tensor([[98.2253]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[98.225296]

```

```

[98.]
epoch: 7 loss: 0.050758298486471176
tensor([[54.6424]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[54.642403]
[55.]
epoch: 8 loss: 0.12787586450576782
tensor([[64.6938]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[64.6938]
[65.]
epoch: 9 loss: 0.09375728666782379
tensor([[39.4468]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[39.44685]
[39.]
epoch: 10 loss: 0.19967477023601532
tensor([[143.2692]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[143.2692]
[142.]
epoch: 11 loss: 1.610857367515564
tensor([[193.3997]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[193.39969]
[190.]
epoch: 12 loss: 11.557883262634277
tensor([[151.7441]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[151.74406]
[151.]
epoch: 13 loss: 0.5536317229270935
tensor([[206.6554]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[206.65538]
[205.]
epoch: 14 loss: 2.740283727645874
tensor([[17.6268]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[17.626799]
[20.]
epoch: 15 loss: 5.632084846496582
tensor([[12.4322]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[12.4321995]
[12.]
epoch: 16 loss: 0.1867963820695877
tensor([[97.3669]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],

```

```

device='cuda:0')
[97.36692]
[102.]
epoch: 17 loss: 21.465425491333008
tensor([[81.8631]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[81.86307]
[86.]
epoch: 18 loss: 17.11421012878418
tensor([[112.8256]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[112.82563]
[119.]
epoch: 19 loss: 38.122840881347656
tensor([[6.7555]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[6.755548]
[5.]
epoch: 20 loss: 3.081948757171631
tensor([[168.0849]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[168.08488]
[173.]
epoch: 21 loss: 24.15835952758789
tensor([[70.7825]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[70.782455]
[68.]
epoch: 22 loss: 7.742058277130127
tensor([[73.3661]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[73.3661]
[69.]
epoch: 23 loss: 19.06279945373535
tensor([[122.1032]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[122.10317]
[116.]
epoch: 24 loss: 37.24871063232422
tensor([[97.6585]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[97.658485]
[91.]
epoch: 25 loss: 44.335426330566406
tensor([[182.7078]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[182.70782]
[179.]

```

```

epoch: 26 loss: 13.747963905334473
tensor([[78.0480]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[78.04799]
[77.]
epoch: 27 loss: 1.098280668258667
tensor([[52.2000]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[52.19996]
[53.]
epoch: 28 loss: 0.6400659084320068
tensor([[123.2711]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[123.27111]
[126.]
epoch: 29 loss: 7.446837902069092
tensor([[176.1268]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[176.12679]
[184.]
epoch: 30 loss: 61.987510681152344
tensor([[135.9491]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[135.94908]
[141.]
epoch: 31 loss: 25.511777877807617
tensor([[46.1507]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[46.15075]
[48.]
epoch: 32 loss: 3.4197285175323486
tensor([[141.9264]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[141.92639]
[140.]
epoch: 33 loss: 3.710984706878662
tensor([[161.3511]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[161.35112]
[156.]
epoch: 34 loss: 28.634485244750977
tensor([[94.3125]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[94.312485]
[92.]
epoch: 35 loss: 5.347585678100586
tensor([[187.9802]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')

```

```

[187.98021]
[183.]
epoch: 36 loss: 24.8024845123291
tensor([[22.6276]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[22.627575]
[24.]
epoch: 37 loss: 1.8835506439208984
tensor([[106.9291]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[106.929085]
[109.]
epoch: 38 loss: 4.288690090179443
tensor([[4.7943]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[4.794321]
[7.]
epoch: 39 loss: 4.865019798278809
tensor([[33.2349]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[33.234943]
[35.]
epoch: 40 loss: 3.115424871444702
tensor([[92.0727]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[92.07274]
[96.]
epoch: 41 loss: 15.423381805419922
tensor([[31.5576]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[31.55755]
[32.]
epoch: 42 loss: 0.19576162099838257
tensor([[139.8833]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[139.88333]
[145.]
epoch: 43 loss: 26.1802978515625
tensor([[1.7035]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[1.7034979]
[1.]
epoch: 44 loss: 0.49490928649902344
tensor([[72.6707]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[72.67071]
[75.]
epoch: 45 loss: 5.425602436065674

```

```

tensor([[134.3500]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[134.34998]
[132.]
epoch: 46 loss: 5.522385120391846
tensor([[94.5882]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[94.58815]
[94.]
epoch: 47 loss: 0.3459204435348511
tensor([[145.5280]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[145.52805]
[144.]
epoch: 48 loss: 2.334923505783081
tensor([[162.7327]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[162.73267]
[163.]
epoch: 49 loss: 0.07146745920181274
tensor([[166.3592]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[166.35922]
[167.]
epoch: 50 loss: 0.41059592366218567
tensor([[127.1843]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[127.18427]
[124.]
epoch: 51 loss: 10.139593124389648
tensor([[64.5022]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[64.50223]
[64.]
epoch: 52 loss: 0.25223276019096375
tensor([[8.8246]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[8.824633]
[8.]
epoch: 53 loss: 0.6800190210342407
tensor([[191.3302]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[191.33023]
[192.]
epoch: 54 loss: 0.4485909044742584
tensor([[102.5039]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[102.50391]

```

```

[101.]
epoch: 55 loss: 2.2617340087890625
tensor([[149.2184]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[149.21838]
[149.]
epoch: 56 loss: 0.047691479325294495
tensor([[57.1718]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[57.1718]
[58.]
epoch: 57 loss: 0.6859173774719238
tensor([[92.0104]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[92.01039]
[93.]
epoch: 58 loss: 0.979325532913208
tensor([[145.5568]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[145.55676]
[147.]
epoch: 59 loss: 2.0829339027404785
tensor([[55.5713]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[55.5713]
[56.]
epoch: 60 loss: 0.18378326296806335
tensor([[200.1173]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[200.11734]
[200.]
epoch: 61 loss: 0.0137686962261796
tensor([[115.3114]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[115.31141]
[117.]
epoch: 62 loss: 2.85133957862854
tensor([[112.6490]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[112.649]
[114.]
epoch: 63 loss: 1.8251954317092896
tensor([[173.5254]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[173.52538]
[175.]
epoch: 64 loss: 2.174517869949341
tensor([[187.8417]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],

```

```

device='cuda:0')
[187.84169]
[186.]
epoch: 65 loss: 3.391822338104248
tensor([[200.5642]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[200.56416]
[196.]
epoch: 66 loss: 20.831584930419922
tensor([[88.9480]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[88.948006]
[87.]
epoch: 67 loss: 3.7947261333465576
tensor([[51.1293]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[51.129333]
[51.]
epoch: 68 loss: 0.016727153211832047
tensor([[29.6523]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[29.652294]
[27.]
epoch: 69 loss: 7.034664154052734
tensor([[43.3943]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[43.394257]
[44.]
epoch: 70 loss: 0.36692509055137634
tensor([[78.2127]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[78.21273]
[81.]
epoch: 71 loss: 7.768871784210205
tensor([[45.4201]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[45.420086]
[47.]
epoch: 72 loss: 2.496128559112549
tensor([[143.2902]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[143.2902]
[150.]
epoch: 73 loss: 45.021324157714844
tensor([[86.9168]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[86.91678]
[89.]

```



```

epoch: 74 loss: 4.339811325073242
tensor([[155.3907]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[155.39075]
[160.]
epoch: 75 loss: 21.24521255493164
tensor([[127.9035]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[127.90349]
[125.]
epoch: 76 loss: 8.430243492126465
tensor([[93.9520]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[93.95202]
[90.]
epoch: 77 loss: 15.618452072143555
tensor([[88.2578]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[88.25779]
[84.]
epoch: 78 loss: 18.128772735595703
tensor([[136.3340]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[136.33398]
[130.]
epoch: 79 loss: 40.11935806274414
tensor([[196.6923]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[196.69229]
[194.]
epoch: 80 loss: 7.248432159423828
tensor([[46.2375]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[46.23746]
[46.]
epoch: 81 loss: 0.0563877709209919
tensor([[28.2160]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[28.215996]
[26.]
epoch: 82 loss: 4.910637378692627
tensor([[4.8977]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[4.8976803]
[4.]
epoch: 83 loss: 0.8058298826217651
tensor([[157.8980]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')

```

```

[157.89804]
[170.]
epoch: 84 loss: 146.45741271972656
tensor([[124.8370]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[124.83703]
[129.]
epoch: 85 loss: 17.330331802368164
tensor([[83.4559]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[83.45588]
[85.]
epoch: 86 loss: 2.3843090534210205
tensor([[111.6512]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[111.651245]
[112.]
epoch: 87 loss: 0.12162996828556061
tensor([[196.8436]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[196.84358]
[193.]
epoch: 88 loss: 14.773123741149902
tensor([[204.6762]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[204.67624]
[197.]
epoch: 89 loss: 58.924644470214844
tensor([[206.3202]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[206.32024]
[201.]
epoch: 90 loss: 28.304912567138672
tensor([[195.3322]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[195.3322]
[198.]
epoch: 91 loss: 7.117161750793457
tensor([[35.1610]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[35.161003]
[36.]
epoch: 92 loss: 0.7039157748222351
tensor([[33.6028]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[33.602764]
[34.]
epoch: 93 loss: 0.15779633820056915

```

```

tensor([[163.4329]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[163.43286]
[180.]
epoch: 94 loss: 274.4700927734375
tensor([[101.9385]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[101.93849]
[107.]
epoch: 95 loss: 25.618865966796875
tensor([[12.3002]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[12.300247]
[13.]
epoch: 96 loss: 0.489654004573822
tensor([[190.2167]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[190.2167]
[189.]
epoch: 97 loss: 1.4803718328475952
tensor([[73.6673]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[73.667274]
[70.]
epoch: 98 loss: 13.448902130126953
tensor([[83.4939]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[83.49387]
[78.]
epoch: 99 loss: 30.182647705078125
tensor([[112.2070]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[112.20701]
[103.]
epoch: 100 loss: 84.76900482177734
tensor([[70.7445]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[70.74449]
[67.]
epoch: 101 loss: 14.021217346191406
tensor([[162.2457]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[162.2457]
[153.]
epoch: 102 loss: 85.48291015625
tensor([[182.3006]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[182.30057]

```

```

[182.]
epoch: 103 loss: 0.09034089744091034
tensor([[157.3962]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[157.39624]
[162.]
epoch: 104 loss: 21.194604873657227
tensor([[97.3441]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[97.3441]
[104.]
epoch: 105 loss: 44.30099105834961
tensor([[72.2128]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[72.21284]
[80.]
epoch: 106 loss: 60.6399040222168
tensor([[48.5727]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[48.572727]
[54.]
epoch: 107 loss: 29.455289840698242
tensor([[34.1323]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[34.132347]
[37.]
epoch: 108 loss: 8.223433494567871
tensor([[20.8681]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[20.868067]
[23.]
epoch: 109 loss: 4.545139312744141
tensor([[103.0569]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[103.056854]
[106.]
epoch: 110 loss: 8.66210651397705
tensor([[158.6226]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[158.62257]
[159.]
epoch: 111 loss: 0.1424504965543747
tensor([[205.8576]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[205.85756]
[204.]
epoch: 112 loss: 3.450526237487793
tensor([[32.6812]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],

```

```

device='cuda:0')
[32.681183]
[29.]
epoch: 113 loss: 13.551107406616211
tensor([[22.6805]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[22.680502]
[22.]
epoch: 114 loss: 0.463082879781723
tensor([[193.4575]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[193.45747]
[191.]
epoch: 115 loss: 6.039177417755127
tensor([[62.5209]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[62.520885]
[61.]
epoch: 116 loss: 2.3130927085876465
tensor([[74.4291]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[74.42909]
[74.]
epoch: 117 loss: 0.1841202974319458
tensor([[124.9060]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[124.90598]
[123.]
epoch: 118 loss: 3.6327710151672363
tensor([[88.3235]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[88.32351]
[88.]
epoch: 119 loss: 0.10465821623802185
tensor([[15.1276]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[15.127601]
[15.]
epoch: 120 loss: 0.01628193072974682
tensor([[108.5621]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[108.56206]
[111.]
epoch: 121 loss: 5.943563461303711
tensor([[145.7553]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[145.75534]
[148.]

```

```
epoch: 122 loss: 5.038496017456055
tensor([[48.5014]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[48.501373]
[49.]
epoch: 123 loss: 0.24862860143184662
tensor([[99.4075]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[99.40753]
[99.]
epoch: 124 loss: 0.16608211398124695
tensor([[164.2112]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[164.21117]
[168.]
epoch: 125 loss: 14.355259895324707
tensor([[135.1448]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[135.14479]
[133.]
epoch: 126 loss: 4.60012674331665
tensor([[139.8665]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[139.86652]
[138.]
epoch: 127 loss: 3.483882427215576
tensor([[5.3124]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[5.312385]
[3.]
epoch: 128 loss: 5.3471245765686035
tensor([[79.6988]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[79.698814]
[79.]
epoch: 129 loss: 0.4883415400981903
tensor([[137.7668]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[137.76682]
[135.]
epoch: 130 loss: 7.655266284942627
tensor([[81.9278]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[81.92778]
[82.]
epoch: 131 loss: 0.005215706303715706
tensor([[196.4951]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
```

```

[196.49507]
[199.]
epoch: 132 loss: 6.274667263031006
tensor([[100.2345]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[100.23454]
[100.]
epoch: 133 loss: 0.0550103485584259
tensor([[107.9960]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[107.99596]
[108.]
epoch: 134 loss: 1.6350531950592995e-05
tensor([[50.0860]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[50.085957]
[50.]
epoch: 135 loss: 0.007388532627373934
tensor([[170.6497]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[170.64966]
[172.]
epoch: 136 loss: 1.823422908782959
tensor([[114.8409]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[114.84092]
[115.]
epoch: 137 loss: 0.02530660666525364
tensor([[105.9776]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[105.977646]
[105.]
epoch: 138 loss: 0.9557914733886719
tensor([[35.1538]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[35.153797]
[33.]
epoch: 139 loss: 4.6388421058654785
tensor([[166.7818]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[166.78178]
[166.]
epoch: 140 loss: 0.6111863255500793
tensor([[154.7676]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[154.76758]
[152.]
epoch: 141 loss: 7.659488677978516

```

```

tensor([[[-0.3696]]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[-0.36955172]
[0.]
epoch: 142 loss: 0.13656847178936005
tensor([[137.1635]]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[137.16345]
[137.]
epoch: 143 loss: 0.026716604828834534
tensor([[119.9357]]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[119.93572]
[122.]
epoch: 144 loss: 4.261242389678955
tensor([[9.5525]]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[9.552491]
[11.]
epoch: 145 loss: 2.0952818393707275
tensor([[23.9572]]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[23.957214]
[25.]
epoch: 146 loss: 1.0874018669128418
tensor([[16.7863]]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[16.78634]
[21.]
epoch: 147 loss: 17.754924774169922
tensor([[170.3140]]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[170.31398]
[174.]
epoch: 148 loss: 13.586742401123047
tensor([[74.1112]]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[74.11116]
[76.]
epoch: 149 loss: 3.5677154064178467
tensor([[29.8618]]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[29.861845]
[28.]
epoch: 150 loss: 3.4664669036865234
tensor([[137.7113]]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[137.71129]

```



```

[134.]
epoch: 151 loss: 13.773661613464355
tensor([[38.8158]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[38.815834]
[38.]
epoch: 152 loss: 0.6655851602554321
tensor([[96.5016]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[96.50157]
[95.]
epoch: 153 loss: 2.2547173500061035
tensor([[171.0213]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[171.02127]
[169.]
epoch: 154 loss: 4.085535526275635
tensor([[11.3651]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[11.365066]
[9.]
epoch: 155 loss: 5.59353494644165
tensor([[82.0054]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[82.00537]
[83.]
epoch: 156 loss: 0.9892866611480713
tensor([[68.9763]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[68.97633]
[71.]
epoch: 157 loss: 4.095225811004639
tensor([[115.6657]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[115.665665]
[118.]
epoch: 158 loss: 5.449121475219727
tensor([[17.9824]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[17.98243]
[18.]
epoch: 159 loss: 0.00030872231582179666
tensor([[162.5172]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[162.51718]
[164.]
epoch: 160 loss: 2.1987509727478027
tensor([[59.4991]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],

```

```

device='cuda:0')
[59.49915]
[60.]
epoch: 161 loss: 0.2508513927459717
tensor([[147.8985]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[147.89845]
[146.]
epoch: 162 loss: 3.6041228771209717
tensor([[31.9314]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[31.931406]
[30.]
epoch: 163 loss: 3.7303292751312256
tensor([[12.4960]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[12.496034]
[10.]
epoch: 164 loss: 6.230184078216553
tensor([[195.0985]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[195.09853]
[195.]
epoch: 165 loss: 0.009707372635602951
tensor([[17.5525]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[17.552498]
[19.]
epoch: 166 loss: 2.0952625274658203
tensor([[112.5195]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[112.51949]
[113.]
epoch: 167 loss: 0.23088687658309937
tensor([[61.3794]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[61.379353]
[62.]
epoch: 168 loss: 0.3852032423019409
tensor([[139.7088]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[139.70877]
[139.]
epoch: 169 loss: 0.5023559927940369
tensor([[155.5505]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[155.55054]
[154.]

```

```

epoch: 170 loss: 2.404165267944336
tensor([[163.8284]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[163.82835]
[165.]
epoch: 171 loss: 1.3727545738220215
tensor([[170.6365]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[170.63649]
[171.]
epoch: 172 loss: 0.13213962316513062
tensor([[97.6792]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[97.67916]
[97.]
epoch: 173 loss: 0.46125975251197815
tensor([[43.6668]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[43.666836]
[45.]
epoch: 174 loss: 1.7773268222808838
tensor([[174.9804]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[174.98036]
[176.]
epoch: 175 loss: 1.0396617650985718
tensor([[14.7819]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[14.781903]
[16.]
epoch: 176 loss: 1.4837596416473389
tensor([[136.9832]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[136.9832]
[136.]
epoch: 177 loss: 0.9666823744773865
tensor([[13.4475]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[13.447526]
[14.]
epoch: 178 loss: 0.3052275478839874
tensor([[56.7900]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[56.78996]
[57.]
epoch: 179 loss: 0.044117242097854614
tensor([[156.1081]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')

```

```

[156.10812]
[155.]
epoch: 180 loss: 1.2279382944107056
tensor([[184.5835]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[184.58347]
[185.]
epoch: 181 loss: 0.1735009253025055
tensor([[119.3930]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[119.392975]
[120.]
epoch: 182 loss: 0.36847952008247375
tensor([[51.6448]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[51.644836]
[52.]
epoch: 183 loss: 0.12614116072654724
tensor([[6.7011]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[6.701138]
[6.]
epoch: 184 loss: 0.491594523191452
tensor([[202.5998]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[202.59976]
[203.]
epoch: 185 loss: 0.16019049286842346
tensor([[143.5558]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[143.55582]
[143.]
epoch: 186 loss: 0.3089321553707123
For epoch 24 the test accuracy over the whole test set is 0 %
tensor([[49.1041]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[49.104137]
[49.]
epoch: 0 loss: 0.01084460224956274
tensor([[66.6037]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[66.6037]
[67.]
epoch: 1 loss: 0.15705469250679016
tensor([[74.8948]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[74.894844]
[75.]

```

```

epoch: 2 loss: 0.011057772673666477
tensor([[37.7170]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[37.716995]
[37.]
epoch: 3 loss: 0.5140821933746338
tensor([[18.6902]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[18.690226]
[20.]
epoch: 4 loss: 1.7155089378356934
tensor([[179.2919]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[179.2919]
[179.]
epoch: 5 loss: 0.08520597964525223
tensor([[18.2429]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[18.24286]
[18.]
epoch: 6 loss: 0.05898136645555496
tensor([[163.0584]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[163.05841]
[162.]
epoch: 7 loss: 1.1202330589294434
tensor([[23.2390]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[23.238998]
[23.]
epoch: 8 loss: 0.05712024122476578
tensor([[84.0029]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[84.0029]
[84.]
epoch: 9 loss: 8.405186235904694e-06
tensor([[152.2846]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[152.28459]
[152.]
epoch: 10 loss: 0.08099242299795151
tensor([[137.2563]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[137.25626]
[138.]
epoch: 11 loss: 0.5531550049781799
tensor([[31.1174]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')

```

```

[31.117365]
[29.]
epoch: 12 loss: 4.48323392868042
tensor([[50.4267]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[50.42669]
[51.]
epoch: 13 loss: 0.32868534326553345
tensor([[35.8113]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[35.8113]
[35.]
epoch: 14 loss: 0.6582050323486328
tensor([[14.8503]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[14.85028]
[15.]
epoch: 15 loss: 0.022416135296225548
tensor([[78.3188]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[78.31885]
[80.]
epoch: 16 loss: 2.826273202896118
tensor([[160.4788]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[160.47884]
[163.]
epoch: 17 loss: 6.35626745223999
tensor([[91.4718]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[91.4718]
[92.]
epoch: 18 loss: 0.2789933681488037
tensor([[139.4153]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[139.41528]
[139.]
epoch: 19 loss: 0.17246013879776
tensor([[14.6146]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[14.614615]
[14.]
epoch: 20 loss: 0.3777521252632141
tensor([[80.0877]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[80.087715]
[79.]
epoch: 21 loss: 1.183124303817749

```

```

tensor([[191.3256]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[191.3256]
[190.]
epoch: 22 loss: 1.7572346925735474
tensor([[24.4844]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[24.484364]
[24.]
epoch: 23 loss: 0.23460805416107178
tensor([[178.5313]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[178.53134]
[176.]
epoch: 24 loss: 6.407690048217773
tensor([[48.1398]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[48.139824]
[48.]
epoch: 25 loss: 0.019550727680325508
tensor([[63.1364]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[63.136444]
[64.]
epoch: 26 loss: 0.7457287907600403
tensor([[3.5792]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[3.579182]
[3.]
epoch: 27 loss: 0.3354516923427582
tensor([[67.2005]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[67.20048]
[69.]
epoch: 28 loss: 3.238280773162842
tensor([[161.1814]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[161.18143]
[165.]
epoch: 29 loss: 14.581500053405762
tensor([[99.7241]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[99.724144]
[100.]
epoch: 30 loss: 0.0760965421795845
tensor([[159.9427]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[159.94272]

```

```

[160.]
epoch: 31 loss: 0.003281169570982456
tensor([[89.0469]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[89.0469]
[88.]
epoch: 32 loss: 1.0959951877593994
tensor([[101.9272]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[101.92718]
[99.]
epoch: 33 loss: 8.568367958068848
tensor([[160.4444]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[160.4444]
[159.]
epoch: 34 loss: 2.086282730102539
tensor([[151.3945]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[151.39453]
[150.]
epoch: 35 loss: 1.9447174072265625
tensor([[14.9997]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[14.999692]
[17.]
epoch: 36 loss: 4.001232147216797
tensor([[153.6004]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[153.60042]
[155.]
epoch: 37 loss: 1.9588295221328735
tensor([[102.0410]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[102.04098]
[103.]
epoch: 38 loss: 0.9197242259979248
tensor([[91.6825]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[91.68251]
[93.]
epoch: 39 loss: 1.7357789278030396
tensor([[75.7272]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[75.72723]
[77.]
epoch: 40 loss: 1.6199530363082886
tensor([[59.8555]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],

```



```

device='cuda:0')
[59.855488]
[60.]
epoch: 41 loss: 0.020883768796920776
tensor([[38.5405]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[38.54045]
[39.]
epoch: 42 loss: 0.21118523180484772
tensor([[186.2971]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[186.29712]
[189.]
epoch: 43 loss: 7.305564880371094
tensor([[106.4970]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[106.49696]
[105.]
epoch: 44 loss: 2.2408998012542725
tensor([[117.8797]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[117.879745]
[116.]
epoch: 45 loss: 3.533442974090576
tensor([[147.0534]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[147.05339]
[144.]
epoch: 46 loss: 9.323193550109863
tensor([[111.7673]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[111.76725]
[110.]
epoch: 47 loss: 3.1231727600097656
tensor([[157.5798]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[157.57983]
[156.]
epoch: 48 loss: 2.495875358581543
tensor([[200.2985]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[200.29851]
[201.]
epoch: 49 loss: 0.49209144711494446
tensor([[190.2731]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[190.27309]
[194.]

```

```

epoch: 50 loss: 13.889883995056152
tensor([[198.1937]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[198.19368]
[200.]
epoch: 51 loss: 3.2627925872802734
tensor([[12.5452]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[12.545189]
[10.]
epoch: 52 loss: 6.4779863357543945
tensor([[18.4341]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[18.434105]
[19.]
epoch: 53 loss: 0.32023724913597107
tensor([[105.6233]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[105.62325]
[104.]
epoch: 54 loss: 2.6349499225616455
tensor([[189.0777]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[189.07773]
[185.]
epoch: 55 loss: 16.62786865234375
tensor([[207.0635]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[207.06352]
[205.]
epoch: 56 loss: 4.258124351501465
tensor([[164.6383]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[164.6383]
[166.]
epoch: 57 loss: 1.854211449623108
tensor([[79.7961]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[79.796135]
[82.]
epoch: 58 loss: 4.857021331787109
tensor([[108.9691]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[108.969124]
[113.]
epoch: 59 loss: 16.247962951660156
tensor([[91.3539]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')

```

```

[91.353905]
[94.]
epoch: 60 loss: 7.001820087432861
tensor([[110.9687]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[110.96871]
[114.]
epoch: 61 loss: 9.188708305358887
tensor([[85.9805]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[85.980545]
[86.]
epoch: 62 loss: 0.0003784953150898218
tensor([[31.6110]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[31.611021]
[30.]
epoch: 63 loss: 2.595388889312744
tensor([[35.0609]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[35.060875]
[33.]
epoch: 64 loss: 4.24720573425293
tensor([[143.2511]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[143.25105]
[140.]
epoch: 65 loss: 10.569344520568848
tensor([[15.4141]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[15.414114]
[16.]
epoch: 66 loss: 0.3432624042034149
tensor([[100.9054]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[100.90538]
[97.]
epoch: 67 loss: 15.251995086669922
tensor([[199.4963]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[199.4963]
[197.]
epoch: 68 loss: 6.231550693511963
tensor([[190.4612]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[190.46115]
[193.]
epoch: 69 loss: 6.445753574371338

```

```

tensor([[168.4218]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[168.42178]
[170.]
epoch: 70 loss: 2.490767478942871
tensor([[149.2543]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[149.25427]
[151.]
epoch: 71 loss: 3.047564744949341
tensor([[145.3403]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[145.34029]
[146.]
epoch: 72 loss: 0.43522223830223083
tensor([[181.1397]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[181.13968]
[182.]
epoch: 73 loss: 0.7401522994041443
tensor([[17.9075]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[17.907537]
[21.]
epoch: 74 loss: 9.563324928283691
tensor([[97.8278]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[97.82783]
[96.]
epoch: 75 loss: 3.3409531116485596
tensor([[137.0023]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[137.00229]
[134.]
epoch: 76 loss: 9.013737678527832
tensor([[27.5425]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[27.542479]
[27.]
epoch: 77 loss: 0.294283002614975
tensor([[173.7244]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[173.72438]
[173.]
epoch: 78 loss: 0.5247271060943604
tensor([[121.7417]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[121.74167]

```

```

[122.]
epoch: 79 loss: 0.06673505902290344
tensor([[203.9252]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[203.92522]
[204.]
epoch: 80 loss: 0.005592545494437218
tensor([[175.5201]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[175.52008]
[177.]
epoch: 81 loss: 2.1901614665985107
tensor([[36.5703]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[36.570312]
[38.]
epoch: 82 loss: 2.04400634765625
tensor([[6.9621]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[6.9621153]
[8.]
epoch: 83 loss: 1.077204704284668
tensor([[28.5501]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[28.550127]
[28.]
epoch: 84 loss: 0.30263975262641907
tensor([[190.9377]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[190.93774]
[192.]
epoch: 85 loss: 1.128387451171875
tensor([[5.9702]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[5.9701753]
[5.]
epoch: 86 loss: 0.9412400722503662
tensor([[5.3487]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[5.3486977]
[4.]
epoch: 87 loss: 1.8189853429794312
tensor([[1.9909]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[1.990887]
[0.]
epoch: 88 loss: 3.9636311531066895
tensor([[65.0211]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],

```

```

device='cuda:0')
[65.0211]
[65.]
epoch: 89 loss: 0.00044533261097967625
tensor([[33.1555]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[33.15551]
[32.]
epoch: 90 loss: 1.3352032899856567
tensor([[35.8096]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[35.80958]
[36.]
epoch: 91 loss: 0.036259107291698456
tensor([[188.5260]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[188.52599]
[184.]
epoch: 92 loss: 20.484546661376953
tensor([[197.3733]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[197.37334]
[196.]
epoch: 93 loss: 1.8860539197921753
tensor([[54.6018]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[54.60181]
[56.]
epoch: 94 loss: 1.954934000968933
tensor([[130.0141]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[130.01414]
[133.]
epoch: 95 loss: 8.91533088684082
tensor([[95.3111]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[95.31105]
[98.]
epoch: 96 loss: 7.230449676513672
tensor([[169.6680]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[169.66798]
[175.]
epoch: 97 loss: 28.430395126342773
tensor([[81.4283]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[81.428314]
[83.]

```

```

epoch: 98 loss: 2.47019624710083
tensor([[130.7307]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[130.73074]
[130.]
epoch: 99 loss: 0.5339859127998352
tensor([[68.4983]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[68.49831]
[68.]
epoch: 100 loss: 0.2483091503381729
tensor([[139.5868]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[139.58679]
[135.]
epoch: 101 loss: 21.038660049438477
tensor([[176.3914]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[176.3914]
[171.]
epoch: 102 loss: 29.067228317260742
tensor([[42.0977]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[42.09774]
[43.]
epoch: 103 loss: 0.8140727877616882
tensor([[49.3327]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[49.33274]
[50.]
epoch: 104 loss: 0.44523486495018005
tensor([[107.0720]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[107.072]
[109.]
epoch: 105 loss: 3.717189311981201
tensor([[138.4053]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[138.40535]
[143.]
epoch: 106 loss: 21.110811233520508
tensor([[120.4328]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[120.43284]
[123.]
epoch: 107 loss: 6.59031867980957
tensor([[70.3892]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')

```

```

[70.38919]
[71.]
epoch: 108 loss: 0.3730880320072174
tensor([[[-1.9605]]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[-1.9604639]
[2.]
epoch: 109 loss: 15.685275077819824
tensor([[194.8672]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[194.86722]
[195.]
epoch: 110 loss: 0.017631053924560547
tensor([[60.2516]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[60.251583]
[58.]
epoch: 111 loss: 5.069626331329346
tensor([[122.5988]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[122.598785]
[120.]
epoch: 112 loss: 6.753685474395752
tensor([[12.5457]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[12.545708]
[12.]
epoch: 113 loss: 0.29779690504074097
tensor([[121.5074]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[121.50738]
[119.]
epoch: 114 loss: 6.286942481994629
tensor([[156.4309]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[156.43092]
[154.]
epoch: 115 loss: 5.909389019012451
tensor([[183.7245]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[183.72447]
[186.]
epoch: 116 loss: 5.178027629852295
tensor([[139.2981]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[139.2981]
[141.]
epoch: 117 loss: 2.8964781761169434

```



```

tensor([[149.3373]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[149.33728]
[149.]
epoch: 118 loss: 0.11375798285007477
tensor([[187.6373]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[187.63731]
[191.]
epoch: 119 loss: 11.307658195495605
tensor([[45.0030]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[45.002995]
[44.]
epoch: 120 loss: 1.0059980154037476
tensor([[56.2626]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[56.262623]
[54.]
epoch: 121 loss: 5.119462013244629
tensor([[55.2949]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[55.294888]
[53.]
epoch: 122 loss: 5.266509056091309
tensor([[206.5380]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[206.53796]
[203.]
epoch: 123 loss: 12.51718807220459
tensor([[118.7464]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[118.74639]
[118.]
epoch: 124 loss: 0.5570999383926392
tensor([[85.5453]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[85.5453]
[85.]
epoch: 125 loss: 0.2973557412624359
tensor([[9.7665]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[9.766505]
[9.]
epoch: 126 loss: 0.587530255317688
tensor([[111.6715]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[111.671524]

```

```

[115.]
epoch: 127 loss: 11.078752517700195
tensor([[58.9061]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[58.906086]
[61.]
epoch: 128 loss: 4.384476184844971
tensor([[167.8788]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[167.87877]
[172.]
epoch: 129 loss: 16.98454475402832
tensor([[87.9151]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[87.915054]
[89.]
epoch: 130 loss: 1.1771070957183838
tensor([[127.6280]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[127.62802]
[126.]
epoch: 131 loss: 2.6504530906677246
tensor([[181.0046]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[181.00458]
[180.]
epoch: 132 loss: 1.009176254272461
tensor([[150.9408]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[150.94083]
[148.]
epoch: 133 loss: 8.648460388183594
tensor([[168.6334]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[168.63338]
[167.]
epoch: 134 loss: 2.6679205894470215
tensor([[13.1082]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[13.108169]
[13.]
epoch: 135 loss: 0.011700446717441082
tensor([[107.5319]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[107.531876]
[108.]
epoch: 136 loss: 0.2191404402256012
tensor([[11.8590]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],

```

```

device='cuda:0')
[11.858951]
[11.]
epoch: 137 loss: 0.7377961874008179
tensor([[49.7262]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[49.726246]
[52.]
epoch: 138 loss: 5.169957637786865
tensor([[60.1704]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[60.17037]
[62.]
epoch: 139 loss: 3.347552537918091
tensor([[122.0136]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[122.01359]
[124.]
epoch: 140 loss: 3.9458327293395996
tensor([[76.3195]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[76.31954]
[78.]
epoch: 141 loss: 2.823939323425293
tensor([[171.8879]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[171.8879]
[174.]
epoch: 142 loss: 4.46099328994751
tensor([[107.6441]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[107.644104]
[107.]
epoch: 143 loss: 0.4148699641227722
tensor([[169.3563]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[169.35626]
[168.]
epoch: 144 loss: 1.8394471406936646
tensor([[171.1518]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[171.15184]
[169.]
epoch: 145 loss: 4.630416393280029
tensor([[21.1647]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[21.16471]
[22.]

```

```

epoch: 146 loss: 0.6977109313011169
tensor([[138.1524]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[138.15236]
[136.]
epoch: 147 loss: 4.6326494216918945
tensor([[1.6294]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[1.6294026]
[1.]
epoch: 148 loss: 0.3961476683616638
tensor([[70.6625]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[70.662506]
[70.]
epoch: 149 loss: 0.43891432881355286
tensor([[128.8637]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[128.86368]
[129.]
epoch: 150 loss: 0.018583692610263824
tensor([[25.2487]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[25.248716]
[25.]
epoch: 151 loss: 0.061859823763370514
tensor([[26.7242]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[26.724247]
[26.]
epoch: 152 loss: 0.5245336890220642
tensor([[72.8462]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[72.84624]
[74.]
epoch: 153 loss: 1.3311686515808105
tensor([[133.4718]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[133.47177]
[137.]
epoch: 154 loss: 12.44839859008789
tensor([[138.2081]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[138.20811]
[142.]
epoch: 155 loss: 14.378395080566406
tensor([[198.7534]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')

```

```

[198.75337]
[199.]
epoch: 156 loss: 0.060825277119874954
tensor([[6.5307]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[6.530681]
[6.]
epoch: 157 loss: 0.2816224694252014
tensor([[96.1615]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[96.16145]
[90.]
epoch: 158 loss: 37.963504791259766
tensor([[96.6602]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[96.660164]
[91.]
epoch: 159 loss: 32.03745651245117
tensor([[46.5529]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[46.552948]
[45.]
epoch: 160 loss: 2.4116475582122803
tensor([[82.5862]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[82.58618]
[81.]
epoch: 161 loss: 2.515972137451172
tensor([[132.5604]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[132.56042]
[132.]
epoch: 162 loss: 0.31407594680786133
tensor([[159.3938]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[159.3938]
[164.]
epoch: 163 loss: 21.21708869934082
tensor([[53.4658]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[53.465828]
[55.]
epoch: 164 loss: 2.3536839485168457
tensor([[99.9595]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[99.95951]
[101.]
epoch: 165 loss: 1.0826177597045898

```

```

tensor([[114.8136]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[114.81363]
[117.]
epoch: 166 loss: 4.78021764755249
tensor([[102.1281]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[102.1281]
[102.]
epoch: 167 loss: 0.016408978030085564
tensor([[158.9163]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[158.91626]
[157.]
epoch: 168 loss: 3.672051429748535
tensor([[112.0178]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[112.017845]
[112.]
epoch: 169 loss: 0.000318449514452368
tensor([[200.5734]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[200.57338]
[198.]
epoch: 170 loss: 6.622282028198242
tensor([[94.4292]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[94.42921]
[95.]
epoch: 171 loss: 0.32580482959747314
tensor([[105.2797]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[105.27969]
[106.]
epoch: 172 loss: 0.5188413262367249
tensor([[146.9306]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[146.93065]
[147.]
epoch: 173 loss: 0.004809588193893433
tensor([[85.2159]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[85.21592]
[87.]
epoch: 174 loss: 3.182943344116211
tensor([[73.7718]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[73.771835]

```

```

[76.]
epoch: 175 loss: 4.964717864990234
tensor([[152.8672]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[152.86716]
[153.]
epoch: 176 loss: 0.017647266387939453
tensor([[32.7117]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[32.71173]
[34.]
epoch: 177 loss: 1.6596370935440063
tensor([[110.6796]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[110.67963]
[111.]
epoch: 178 loss: 0.10263919830322266
tensor([[182.8192]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[182.81921]
[183.]
epoch: 179 loss: 0.03268362581729889
tensor([[44.5169]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[44.516865]
[46.]
epoch: 180 loss: 2.199690103530884
tensor([[4.1754]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[4.1754103]
[7.]
epoch: 181 loss: 7.978307247161865
tensor([[148.8360]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[148.83603]
[145.]
epoch: 182 loss: 14.715119361877441
tensor([[57.9905]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[57.990524]
[57.]
epoch: 183 loss: 0.9811383485794067
tensor([[46.8881]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[46.88814]
[47.]
epoch: 184 loss: 0.012512294575572014
tensor([[126.9938]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],

```

```

device='cuda:0')
[126.99377]
[125.]
epoch: 185 loss: 3.9751060009002686
tensor([[174.9346]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[174.93457]
[178.]
epoch: 186 loss: 9.396859169006348
For epoch 25 the test accuracy over the whole test set is 0 %
tensor([[137.4321]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[137.43211]
[138.]
epoch: 0 loss: 0.3224949240684509
tensor([[9.4387]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[9.438675]
[9.]
epoch: 1 loss: 0.19243569672107697
tensor([[163.0950]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[163.095]
[164.]
epoch: 2 loss: 0.8190227746963501
tensor([[56.1743]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[56.17434]
[55.]
epoch: 3 loss: 1.3790727853775024
tensor([[47.2886]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[47.28855]
[47.]
epoch: 4 loss: 0.0832618698477745
tensor([[2.5045]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[2.5044765]
[0.]
epoch: 5 loss: 6.272402763366699
tensor([[61.3173]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[61.317318]
[60.]
epoch: 6 loss: 1.7353266477584839
tensor([[165.7271]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[165.72705]

```



```

[165.]
epoch: 7 loss: 0.5286028385162354
tensor([[107.1779]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[107.17786]
[106.]
epoch: 8 loss: 1.387345790863037
tensor([[126.4386]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[126.438614]
[125.]
epoch: 9 loss: 2.0696098804473877
tensor([[35.3478]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[35.347797]
[36.]
epoch: 10 loss: 0.4253682494163513
tensor([[50.6348]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[50.634796]
[52.]
epoch: 11 loss: 1.8637815713882446
tensor([[157.4794]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[157.47942]
[160.]
epoch: 12 loss: 6.353344440460205
tensor([[175.7684]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[175.76843]
[179.]
epoch: 13 loss: 10.44302749633789
tensor([[150.5091]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[150.5091]
[150.]
epoch: 14 loss: 0.25917693972587585
tensor([[47.9666]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[47.966602]
[48.]
epoch: 15 loss: 0.0011154046515002847
tensor([[84.5650]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[84.565025]
[83.]
epoch: 16 loss: 2.4493043422698975
tensor([[30.8662]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],

```

```

device='cuda:0')
[30.86621]
[30.]
epoch: 17 loss: 0.7503213882446289
tensor([[131.1149]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[131.1149]
[126.]
epoch: 18 loss: 26.162187576293945
tensor([[102.0086]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[102.00856]
[99.]
epoch: 19 loss: 9.051434516906738
tensor([[15.8280]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[15.828005]
[18.]
epoch: 20 loss: 4.717563152313232
tensor([[83.7339]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[83.73392]
[85.]
epoch: 21 loss: 1.6029655933380127
tensor([[99.6823]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[99.68229]
[101.]
epoch: 22 loss: 1.7363619804382324
tensor([[191.6278]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[191.62779]
[194.]
epoch: 23 loss: 5.627368927001953
tensor([[14.3759]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[14.375863]
[17.]
epoch: 24 loss: 6.886094570159912
tensor([[78.9534]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[78.95341]
[81.]
epoch: 25 loss: 4.188541889190674
tensor([[1.4514]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[1.4514196]
[1.]

```

```

epoch: 26 loss: 0.2037796527147293
tensor([[68.7256]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[68.725624]
[70.]
epoch: 27 loss: 1.6240339279174805
tensor([[13.2758]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[13.275826]
[12.]
epoch: 28 loss: 1.6277331113815308
tensor([[25.5075]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[25.507542]
[25.]
epoch: 29 loss: 0.25759851932525635
tensor([[33.6419]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[33.641876]
[32.]
epoch: 30 loss: 2.6957576274871826
tensor([[57.2119]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[57.211887]
[56.]
epoch: 31 loss: 1.4686709642410278
tensor([[134.0772]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[134.07718]
[130.]
epoch: 32 loss: 16.623388290405273
tensor([[205.1493]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[205.14929]
[198.]
epoch: 33 loss: 51.11237716674805
tensor([[194.3547]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[194.35468]
[195.]
epoch: 34 loss: 0.41644397377967834
tensor([[128.3416]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[128.34157]
[134.]
epoch: 35 loss: 32.017852783203125
tensor([[26.6431]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')

```

```

[26.64309]
[28.]
epoch: 36 loss: 1.841206669807434
tensor([[52.2366]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[52.236633]
[57.]
epoch: 37 loss: 22.68966293334961
tensor([[95.4183]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[95.41829]
[103.]
epoch: 38 loss: 57.482337951660156
tensor([[74.8868]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[74.88678]
[80.]
epoch: 39 loss: 26.145021438598633
tensor([[111.9163]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[111.91631]
[117.]
epoch: 40 loss: 25.8438720703125
tensor([[136.2894]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[136.28943]
[137.]
epoch: 41 loss: 0.5049115419387817
tensor([[179.8977]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[179.89769]
[176.]
epoch: 42 loss: 15.191986083984375
tensor([[21.8520]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[21.851963]
[21.]
epoch: 43 loss: 0.7258410453796387
tensor([[96.8439]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[96.84386]
[91.]
epoch: 44 loss: 34.15066146850586
tensor([[151.1502]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[151.15018]
[145.]
epoch: 45 loss: 37.824676513671875

```

```

tensor([[137.0293]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[137.02931]
[133.]
epoch: 46 loss: 16.235355377197266
tensor([[198.5608]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[198.5608]
[205.]
epoch: 47 loss: 41.46321487426758
tensor([[3.3684]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[3.3684065]
[2.]
epoch: 48 loss: 1.8725364208221436
tensor([[39.5025]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[39.5025]
[37.]
epoch: 49 loss: 6.2624993324279785
tensor([[196.2667]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[196.26672]
[203.]
epoch: 50 loss: 45.33700942993164
tensor([[99.9867]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[99.98673]
[97.]
epoch: 51 loss: 8.920571327209473
tensor([[14.1207]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[14.120683]
[11.]
epoch: 52 loss: 9.73866081237793
tensor([[95.4473]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[95.447296]
[92.]
epoch: 53 loss: 11.883851051330566
tensor([[35.9775]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[35.977463]
[33.]
epoch: 54 loss: 8.86528491973877
tensor([[125.2926]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[125.29259]

```

```

[122.]
epoch: 55 loss: 10.841131210327148
tensor([[101.3715]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[101.37147]
[98.]
epoch: 56 loss: 11.366793632507324
tensor([[62.5568]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[62.55682]
[62.]
epoch: 57 loss: 0.3100484311580658
tensor([[48.7220]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[48.72201]
[50.]
epoch: 58 loss: 1.6332544088363647
tensor([[97.2289]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[97.22891]
[100.]
epoch: 59 loss: 7.678926944732666
tensor([[109.0261]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[109.026054]
[113.]
epoch: 60 loss: 15.792243957519531
tensor([[126.2461]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[126.24613]
[129.]
epoch: 61 loss: 7.583789348602295
tensor([[139.7542]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[139.7542]
[142.]
epoch: 62 loss: 5.04363489151001
tensor([[115.7008]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[115.7008]
[116.]
epoch: 63 loss: 0.08952181786298752
tensor([[84.3806]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[84.38061]
[84.]
epoch: 64 loss: 0.14486214518547058
tensor([[127.1879]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],

```

```

device='cuda:0')
[127.187935]
[123.]
epoch: 65 loss: 17.538799285888672
tensor([[200.8846]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[200.88458]
[191.]
epoch: 66 loss: 97.70497131347656
tensor([[9.7342]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[9.734151]
[16.]
epoch: 67 loss: 39.2608642578125
tensor([[54.1193]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[54.11931]
[58.]
epoch: 68 loss: 15.05976676940918
tensor([[175.1156]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[175.11557]
[182.]
epoch: 69 loss: 47.3953742980957
tensor([[64.0270]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[64.02698]
[69.]
epoch: 70 loss: 24.730953216552734
tensor([[134.7413]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[134.74132]
[141.]
epoch: 71 loss: 39.171104431152344
tensor([[59.7930]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[59.79305]
[61.]
epoch: 72 loss: 1.4567310810089111
tensor([[120.3079]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[120.30788]
[120.]
epoch: 73 loss: 0.09478799253702164
tensor([[213.2281]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[213.22815]
[204.]

```

```

epoch: 74 loss: 85.15874481201172
tensor([[110.3017]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[110.30169]
[108.]
epoch: 75 loss: 5.2977728843688965
tensor([[168.4763]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[168.47632]
[169.]
epoch: 76 loss: 0.27424246072769165
tensor([[179.3102]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[179.31017]
[183.]
epoch: 77 loss: 13.614879608154297
tensor([[180.8070]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[180.80696]
[186.]
epoch: 78 loss: 26.967653274536133
tensor([[45.4077]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[45.407692]
[45.]
epoch: 79 loss: 0.16621273756027222
tensor([[154.4122]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[154.41225]
[153.]
epoch: 80 loss: 1.9944407939910889
tensor([[144.4597]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[144.45969]
[143.]
epoch: 81 loss: 2.1306841373443604
tensor([[105.3804]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[105.38045]
[102.]
epoch: 82 loss: 11.427424430847168
tensor([[174.1615]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[174.16148]
[173.]
epoch: 83 loss: 1.3490445613861084
tensor([[17.0164]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')

```



```

[17.016401]
[14.]
epoch: 84 loss: 9.098676681518555
tensor([[171.1407]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[171.14069]
[172.]
epoch: 85 loss: 0.7384204864501953
tensor([[52.2301]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[52.230053]
[51.]
epoch: 86 loss: 1.5130302906036377
tensor([[116.5723]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[116.57226]
[119.]
epoch: 87 loss: 5.8939313888549805
tensor([[8.8943]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[8.894328]
[7.]
epoch: 88 loss: 3.5884790420532227
tensor([[147.3861]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[147.3861]
[149.]
epoch: 89 loss: 2.6046953201293945
tensor([[79.4230]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[79.423004]
[79.]
epoch: 90 loss: 0.17893251776695251
tensor([[30.6215]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[30.62151]
[29.]
epoch: 91 loss: 2.629293203353882
tensor([[35.8011]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[35.80108]
[35.]
epoch: 92 loss: 0.6417272090911865
tensor([[164.6286]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[164.62862]
[166.]
epoch: 93 loss: 1.8806931972503662

```

```

tensor([[175.3847]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[175.3847]
[175.]
epoch: 94 loss: 0.14799761772155762
tensor([[89.3227]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[89.322685]
[89.]
epoch: 95 loss: 0.10412576794624329
tensor([[13.4321]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[13.43214]
[13.]
epoch: 96 loss: 0.186745285987854
tensor([[163.8675]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[163.86751]
[162.]
epoch: 97 loss: 3.4875857830047607
tensor([[178.0520]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[178.05202]
[177.]
epoch: 98 loss: 1.1067402362823486
tensor([[167.6670]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[167.66704]
[168.]
epoch: 99 loss: 0.11086371541023254
tensor([[137.8058]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[137.80583]
[139.]
epoch: 100 loss: 1.4260374307632446
tensor([[63.5924]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[63.59244]
[64.]
epoch: 101 loss: 0.16610388457775116
tensor([[131.8696]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[131.86957]
[132.]
epoch: 102 loss: 0.017012540251016617
tensor([[182.7523]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[182.7523]

```

```

[184.]
epoch: 103 loss: 1.556745171546936
tensor([[189.6946]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[189.69461]
[193.]
epoch: 104 loss: 10.925599098205566
tensor([[110.2938]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[110.29378]
[109.]
epoch: 105 loss: 1.6738600730895996
tensor([[149.4645]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[149.46446]
[146.]
epoch: 106 loss: 12.002498626708984
tensor([[4.5048]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[4.504753]
[6.]
epoch: 107 loss: 2.2357633113861084
tensor([[94.0553]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[94.055275]
[93.]
epoch: 108 loss: 1.113605260848999
tensor([[5.6787]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[5.6787434]
[8.]
epoch: 109 loss: 5.388232231140137
tensor([[72.0608]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[72.060814]
[71.]
epoch: 110 loss: 1.125326156616211
tensor([[203.5398]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[203.53981]
[199.]
epoch: 111 loss: 20.60987663269043
tensor([[151.9646]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[151.96461]
[152.]
epoch: 112 loss: 0.0012521075550466776
tensor([[191.3543]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],

```

```

device='cuda:0')
[191.35432]
[196.]
epoch: 113 loss: 21.58230209350586
tensor([[74.6464]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[74.64644]
[77.]
epoch: 114 loss: 5.539251327514648
tensor([[74.0637]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[74.0637]
[76.]
epoch: 115 loss: 3.7492661476135254
tensor([[177.1592]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[177.15921]
[178.]
epoch: 116 loss: 0.706927478313446
tensor([[167.4155]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[167.4155]
[167.]
epoch: 117 loss: 0.17263761162757874
tensor([[6.0978]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[6.097837]
[4.]
epoch: 118 loss: 4.4009199142456055
tensor([[48.4245]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[48.424534]
[49.]
epoch: 119 loss: 0.3311612904071808
tensor([[116.3366]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[116.336555]
[115.]
epoch: 120 loss: 1.7863805294036865
tensor([[204.7690]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[204.76895]
[200.]
epoch: 121 loss: 22.742897033691406
tensor([[54.5373]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[54.5373]
[54.]

```

```

epoch: 122 loss: 0.2886914014816284
tensor([[6.1164]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[6.1164474]
[5.]
epoch: 123 loss: 1.2464549541473389
tensor([[110.0157]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[110.01574]
[114.]
epoch: 124 loss: 15.874332427978516
tensor([[63.3346]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[63.334583]
[65.]
epoch: 125 loss: 2.7736127376556396
tensor([[167.0215]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[167.02155]
[171.]
epoch: 126 loss: 15.82810115814209
tensor([[66.1064]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[66.10636]
[67.]
epoch: 127 loss: 0.7985899448394775
tensor([[123.4450]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[123.44504]
[124.]
epoch: 128 loss: 0.30798301100730896
tensor([[87.0127]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[87.01273]
[86.]
epoch: 129 loss: 1.0256290435791016
tensor([[191.0614]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[191.0614]
[190.]
epoch: 130 loss: 1.126572847366333
tensor([[38.3115]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[38.311504]
[38.]
epoch: 131 loss: 0.09703496843576431
tensor([[39.3289]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')

```

```

[39.328938]
[39.]
epoch: 132 loss: 0.10819990187883377
tensor([[46.8690]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[46.868977]
[46.]
epoch: 133 loss: 0.7551203370094299
tensor([[162.1202]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[162.12018]
[157.]
epoch: 134 loss: 26.216224670410156
tensor([[111.6785]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[111.67852]
[110.]
epoch: 135 loss: 2.817430019378662
tensor([[186.5317]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[186.5317]
[189.]
epoch: 136 loss: 6.092541694641113
tensor([[152.6996]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[152.69955]
[154.]
epoch: 137 loss: 1.691158652305603
tensor([[177.2666]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[177.26657]
[180.]
epoch: 138 loss: 7.4716339111328125
tensor([[103.8905]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[103.890526]
[104.]
epoch: 139 loss: 0.011984596960246563
tensor([[150.8601]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[150.86014]
[151.]
epoch: 140 loss: 0.019561395049095154
tensor([[21.4877]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[21.487726]
[23.]
epoch: 141 loss: 2.2869720458984375

```

```

tensor([[75.9534]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[75.9534]
[74.]
epoch: 142 loss: 3.815770149230957
tensor([[199.2314]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[199.2314]
[197.]
epoch: 143 loss: 4.979144096374512
tensor([[69.9116]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[69.91164]
[68.]
epoch: 144 loss: 3.654353618621826
tensor([[136.2596]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[136.25955]
[135.]
epoch: 145 loss: 1.5864711999893188
tensor([[13.5754]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[13.575376]
[15.]
epoch: 146 loss: 2.029554843902588
tensor([[160.6969]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[160.69688]
[163.]
epoch: 147 loss: 5.304342746734619
tensor([[115.3595]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[115.35952]
[118.]
epoch: 148 loss: 6.972135066986084
tensor([[77.3305]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[77.330475]
[78.]
epoch: 149 loss: 0.44826391339302063
tensor([[53.4914]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[53.491405]
[53.]
epoch: 150 loss: 0.24147935211658478
tensor([[190.7739]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[190.77393]

```

```

[192.]
epoch: 151 loss: 1.5032579898834229
tensor([[90.4488]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[90.44882]
[90.]
epoch: 152 loss: 0.2014412134885788
tensor([[177.9887]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[177.98868]
[174.]
epoch: 153 loss: 15.909551620483398
tensor([[157.8984]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[157.89838]
[156.]
epoch: 154 loss: 3.6038331985473633
tensor([[147.4861]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[147.48605]
[148.]
epoch: 155 loss: 0.2641410529613495
tensor([[26.4607]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[26.460678]
[27.]
epoch: 156 loss: 0.2908681035041809
tensor([[40.6966]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[40.696564]
[43.]
epoch: 157 loss: 5.305818557739258
tensor([[91.9362]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[91.93618]
[95.]
epoch: 158 loss: 9.386992454528809
tensor([[22.8846]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[22.884647]
[24.]
epoch: 159 loss: 1.2440115213394165
tensor([[151.3186]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[151.31856]
[155.]
epoch: 160 loss: 13.553016662597656
tensor([[137.7429]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],

```



```

device='cuda:0')
[137.74289]
[140.]
epoch: 161 loss: 5.094548225402832
tensor([[111.7234]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[111.723434]
[111.]
epoch: 162 loss: 0.5233573913574219
tensor([[84.9485]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[84.948456]
[82.]
epoch: 163 loss: 8.693391799926758
tensor([[163.2242]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[163.22421]
[159.]
epoch: 164 loss: 17.84397315979004
tensor([[107.6764]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[107.67639]
[105.]
epoch: 165 loss: 7.163072109222412
tensor([[137.5019]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[137.50186]
[136.]
epoch: 166 loss: 2.2555882930755615
tensor([[93.4504]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[93.4504]
[94.]
epoch: 167 loss: 0.30205872654914856
tensor([[43.5247]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[43.5247]
[44.]
epoch: 168 loss: 0.22590993344783783
tensor([[85.1533]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[85.15326]
[87.]
epoch: 169 loss: 3.4104514122009277
tensor([[26.6723]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[26.672344]
[26.]

```

```

epoch: 170 loss: 0.4520467221736908
tensor([[20.9808]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[20.98081]
[22.]
epoch: 171 loss: 1.0387479066848755
tensor([[84.4435]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[84.44346]
[88.]
epoch: 172 loss: 12.64898681640625
tensor([[12.4699]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[12.469879]
[10.]
epoch: 173 loss: 6.1003031730651855
tensor([[34.5660]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[34.566]
[34.]
epoch: 174 loss: 0.32035812735557556
tensor([[140.4124]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[140.4124]
[147.]
epoch: 175 loss: 43.396484375
tensor([[4.6012]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[4.6012397]
[3.]
epoch: 176 loss: 2.5639684200286865
tensor([[18.7229]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[18.722914]
[20.]
epoch: 177 loss: 1.6309492588043213
tensor([[185.1763]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[185.17627]
[185.]
epoch: 178 loss: 0.031070947647094727
tensor([[109.7110]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[109.71103]
[107.]
epoch: 179 loss: 7.3496785163879395
tensor([[114.9917]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')

```

```

[114.99171]
[112.]
epoch: 180 loss: 8.950309753417969
tensor([[99.3895]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[99.38954]
[96.]
epoch: 181 loss: 11.488992691040039
tensor([[18.1064]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[18.106426]
[19.]
epoch: 182 loss: 0.7984740734100342
tensor([[200.5388]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[200.5388]
[201.]
epoch: 183 loss: 0.21270258724689484
tensor([[143.6180]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[143.61795]
[144.]
epoch: 184 loss: 0.14596186578273773
tensor([[169.5823]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[169.5823]
[170.]
epoch: 185 loss: 0.17446835339069366
tensor([[76.0346]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[76.03459]
[75.]
epoch: 186 loss: 1.0703799724578857
For epoch 26 the test accuracy over the whole test set is 0 %
tensor([[95.6495]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[95.64945]
[96.]
epoch: 0 loss: 0.12288375198841095
tensor([[167.1968]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[167.19676]
[169.]
epoch: 1 loss: 3.251667022705078
tensor([[9.8996]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[9.89964]
[10.]

```

```

epoch: 2 loss: 0.010072113014757633
tensor([[185.1662]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[185.16623]
[186.]
epoch: 3 loss: 0.6951736807823181
tensor([[110.7970]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[110.796974]
[111.]
epoch: 4 loss: 0.04121948406100273
tensor([[11.0383]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[11.038265]
[11.]
epoch: 5 loss: 0.0014642276801168919
tensor([[35.8136]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[35.81361]
[34.]
epoch: 6 loss: 3.2891814708709717
tensor([[94.3034]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[94.30344]
[93.]
epoch: 7 loss: 1.6989461183547974
tensor([[151.1725]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[151.17255]
[148.]
epoch: 8 loss: 10.06505012512207
tensor([[2.4185]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[2.4185193]
[4.]
epoch: 9 loss: 2.5010812282562256
tensor([[130.3523]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[130.35231]
[130.]
epoch: 10 loss: 0.12412246316671371
tensor([[75.6148]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[75.614784]
[76.]
epoch: 11 loss: 0.14839118719100952
tensor([[ -0.2077]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')

```

```

[-0.20767158]
[1.]
epoch: 12 loss: 1.4584708213806152
tensor([[32.5154]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[32.51545]
[33.]
epoch: 13 loss: 0.2347891628742218
tensor([[59.2538]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[59.25385]
[60.]
epoch: 14 loss: 0.5567412972450256
tensor([[40.8835]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[40.883465]
[43.]
epoch: 15 loss: 4.4797210693359375
tensor([[17.2112]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[17.211205]
[20.]
epoch: 16 loss: 7.777379989624023
tensor([[122.8060]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[122.80603]
[126.]
epoch: 17 loss: 10.20144271850586
tensor([[76.4464]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[76.44638]
[78.]
epoch: 18 loss: 2.4137332439422607
tensor([[30.3116]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[30.31165]
[29.]
epoch: 19 loss: 1.720423936843872
tensor([[150.4603]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[150.46033]
[150.]
epoch: 20 loss: 0.21190108358860016
tensor([[16.4911]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[16.491055]
[15.]
epoch: 21 loss: 2.2232437133789062

```

```

tensor([[11.6605]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[11.66047]
[9.]
epoch: 22 loss: 7.078100681304932
tensor([[75.7257]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[75.72568]
[74.]
epoch: 23 loss: 2.9779627323150635
tensor([[14.7914]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[14.79143]
[13.]
epoch: 24 loss: 3.2092230319976807
tensor([[20.1799]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[20.179865]
[21.]
epoch: 25 loss: 0.6726216077804565
tensor([[110.8071]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[110.8071]
[110.]
epoch: 26 loss: 0.6514078378677368
tensor([[201.6459]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[201.64587]
[199.]
epoch: 27 loss: 7.000649452209473
tensor([[1.1570]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[1.1569982]
[2.]
epoch: 28 loss: 0.7106521129608154
tensor([[127.8585]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[127.858536]
[129.]
epoch: 29 loss: 1.3029406070709229
tensor([[51.7468]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[51.746845]
[52.]
epoch: 30 loss: 0.06408733129501343
tensor([[88.9824]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[88.9824]

```

```

[92.]
epoch: 31 loss: 9.105916023254395
tensor([[57.9595]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[57.95954]
[58.]
epoch: 32 loss: 0.001636904664337635
tensor([[86.2286]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[86.228584]
[88.]
epoch: 33 loss: 3.137913703918457
tensor([[101.7229]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[101.72288]
[104.]
epoch: 34 loss: 5.185286998748779
tensor([[30.7795]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[30.779505]
[30.]
epoch: 35 loss: 0.6076276898384094
tensor([[115.2999]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[115.29985]
[116.]
epoch: 36 loss: 0.4902093708515167
tensor([[134.0389]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[134.0389]
[132.]
epoch: 37 loss: 4.1570916175842285
tensor([[117.6109]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[117.61094]
[117.]
epoch: 38 loss: 0.3732464909553528
tensor([[183.0659]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[183.06587]
[182.]
epoch: 39 loss: 1.1360834836959839
tensor([[76.8322]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[76.83218]
[75.]
epoch: 40 loss: 3.356869697570801
tensor([[36.7192]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],

```

```

device='cuda:0')
[36.719227]
[36.]
epoch: 41 loss: 0.5172872543334961
tensor([[165.7515]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[165.75154]
[168.]
epoch: 42 loss: 5.055567264556885
tensor([[192.1832]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[192.1832]
[197.]
epoch: 43 loss: 23.20159149169922
tensor([[203.2294]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[203.22942]
[204.]
epoch: 44 loss: 0.593799889087677
tensor([[97.3400]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[97.33998]
[94.]
epoch: 45 loss: 11.155473709106445
tensor([[169.1514]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[169.15135]
[162.]
epoch: 46 loss: 51.1418342590332
tensor([[72.3540]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[72.35405]
[69.]
epoch: 47 loss: 11.249649047851562
tensor([[118.7910]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[118.79104]
[118.]
epoch: 48 loss: 0.6257419586181641
tensor([[43.3303]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[43.330338]
[44.]
epoch: 49 loss: 0.4484478235244751
tensor([[36.8610]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[36.860992]
[39.]

```



```

epoch: 50 loss: 4.575353145599365
tensor([[55.8667]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[55.86672]
[57.]
epoch: 51 loss: 1.2843273878097534
tensor([[73.6888]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[73.68882]
[77.]
epoch: 52 loss: 10.963913917541504
tensor([[100.4258]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[100.425766]
[106.]
epoch: 53 loss: 31.072084426879883
tensor([[2.5960]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[2.5960472]
[5.]
epoch: 54 loss: 5.778989315032959
tensor([[153.0869]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[153.0869]
[157.]
epoch: 55 loss: 15.312360763549805
tensor([[161.5739]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[161.57391]
[163.]
epoch: 56 loss: 2.033722400665283
tensor([[101.5727]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[101.57272]
[99.]
epoch: 57 loss: 6.618905544281006
tensor([[125.5026]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[125.502625]
[122.]
epoch: 58 loss: 12.268378257751465
tensor([[34.8868]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[34.886826]
[32.]
epoch: 59 loss: 8.333762168884277
tensor([[53.2894]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')

```

```

[53.289448]
[50.]
epoch: 60 loss: 10.820466995239258
tensor([[82.5904]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[82.59036]
[79.]
epoch: 61 loss: 12.890703201293945
tensor([[91.0884]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[91.088394]
[89.]
epoch: 62 loss: 4.361390113830566
tensor([[54.3002]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[54.300167]
[53.]
epoch: 63 loss: 1.690434455871582
tensor([[165.3765]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[165.3765]
[165.]
epoch: 64 loss: 0.14174875617027283
tensor([[14.4724]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[14.472381]
[16.]
epoch: 65 loss: 2.333621025085449
tensor([[103.2330]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[103.23302]
[107.]
epoch: 66 loss: 14.190160751342773
tensor([[122.1765]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[122.17648]
[125.]
epoch: 67 loss: 7.97224760055542
tensor([[175.4999]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[175.49992]
[176.]
epoch: 68 loss: 0.2500762939453125
tensor([[109.2707]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[109.27068]
[112.]
epoch: 69 loss: 7.449169635772705

```

```

tensor([[141.2950]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[141.29503]
[141.]
epoch: 70 loss: 0.0870419293642044
tensor([[207.0164]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[207.01645]
[203.]
epoch: 71 loss: 16.13186264038086
tensor([[35.2782]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[35.27824]
[38.]
epoch: 72 loss: 7.407976150512695
tensor([[138.9866]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[138.98663]
[138.]
epoch: 73 loss: 0.9734452962875366
tensor([[179.5149]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[179.51494]
[178.]
epoch: 74 loss: 2.2950382232666016
tensor([[26.4031]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[26.403084]
[28.]
epoch: 75 loss: 2.5501413345336914
tensor([[20.4523]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[20.452318]
[23.]
epoch: 76 loss: 6.490682601928711
tensor([[32.7170]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[32.71699]
[35.]
epoch: 77 loss: 5.212128162384033
tensor([[119.2368]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[119.236824]
[123.]
epoch: 78 loss: 14.161493301391602
tensor([[132.3731]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[132.37311]

```

```

[135.]
epoch: 79 loss: 6.900562286376953
tensor([[23.5717]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[23.571651]
[24.]
epoch: 80 loss: 0.18348246812820435
tensor([[200.2598]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[200.25983]
[201.]
epoch: 81 loss: 0.5478565692901611
tensor([[10.2092]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[10.209236]
[8.]
epoch: 82 loss: 4.8807244300842285
tensor([[188.9339]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[188.9339]
[184.]
epoch: 83 loss: 24.343358993530273
tensor([[107.2569]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[107.256905]
[105.]
epoch: 84 loss: 5.093618392944336
tensor([[179.0499]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[179.04991]
[179.]
epoch: 85 loss: 0.0024911577347666025
tensor([[158.4927]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[158.49268]
[160.]
epoch: 86 loss: 2.272026300430298
tensor([[137.1649]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[137.16486]
[139.]
epoch: 87 loss: 3.367753744125366
tensor([[83.2781]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[83.27809]
[83.]
epoch: 88 loss: 0.07733484357595444
tensor([[8.3603]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],

```

```

device='cuda:0')
[8.360264]
[6.]
epoch: 89 loss: 5.570845127105713
tensor([[184.7875]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[184.78752]
[191.]
epoch: 90 loss: 38.59489059448242
tensor([[155.9992]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[155.99919]
[155.]
epoch: 91 loss: 0.9983832240104675
tensor([[166.9670]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[166.96703]
[164.]
epoch: 92 loss: 8.803241729736328
tensor([[54.4970]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[54.49703]
[51.]
epoch: 93 loss: 12.229207038879395
tensor([[106.3908]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[106.39077]
[102.]
epoch: 94 loss: 19.278860092163086
tensor([[146.5463]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[146.5463]
[144.]
epoch: 95 loss: 6.483619213104248
tensor([[146.5875]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[146.58751]
[146.]
epoch: 96 loss: 0.34516701102256775
tensor([[150.0105]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[150.01045]
[152.]
epoch: 97 loss: 3.9583001136779785
tensor([[12.4469]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[12.446921]
[14.]

```

```

epoch: 98 loss: 2.412053346633911
tensor([[55.5542]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[55.55418]
[56.]
epoch: 99 loss: 0.19875533878803253
tensor([[167.8199]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[167.81995]
[174.]
epoch: 100 loss: 38.19306564331055
tensor([[14.5833]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[14.583335]
[17.]
epoch: 101 loss: 5.840270042419434
tensor([[46.6753]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[46.675255]
[47.]
epoch: 102 loss: 0.10545942932367325
tensor([[171.9617]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[171.96175]
[170.]
epoch: 103 loss: 3.8484482765197754
tensor([[1.8695]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[1.8694799]
[0.]
epoch: 104 loss: 3.494955062866211
tensor([[46.3913]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[46.391293]
[46.]
epoch: 105 loss: 0.15310987830162048
tensor([[28.6696]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[28.669613]
[27.]
epoch: 106 loss: 2.787607192993164
tensor([[148.6504]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[148.65044]
[145.]
epoch: 107 loss: 13.325685501098633
tensor([[182.9776]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')

```

```

[182.97762]
[180.]
epoch: 108 loss: 8.866192817687988
tensor([[202.0560]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[202.05597]
[205.]
epoch: 109 loss: 8.667317390441895
tensor([[171.5101]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[171.5101]
[172.]
epoch: 110 loss: 0.24000072479248047
tensor([[131.9640]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[131.964]
[133.]
epoch: 111 loss: 1.0732866525650024
tensor([[87.0960]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[87.09605]
[90.]
epoch: 112 loss: 8.43294620513916
tensor([[26.2054]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[26.205381]
[26.]
epoch: 113 loss: 0.04218151792883873
tensor([[24.7698]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[24.769808]
[25.]
epoch: 114 loss: 0.0529884435236454
tensor([[70.2626]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[70.2626]
[71.]
epoch: 115 loss: 0.5437532067298889
tensor([[185.8024]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[185.80241]
[183.]
epoch: 116 loss: 7.8535237312316895
tensor([[54.0155]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[54.015503]
[54.]
epoch: 117 loss: 0.00024034082889556885

```

```

tensor([[177.3915]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[177.3915]
[177.]
epoch: 118 loss: 0.15326814353466034
tensor([[117.7511]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[117.75105]
[120.]
epoch: 119 loss: 5.05776309967041
tensor([[174.5196]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[174.51956]
[175.]
epoch: 120 loss: 0.230820894241333
tensor([[36.1291]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[36.129074]
[37.]
epoch: 121 loss: 0.7585119009017944
tensor([[100.3328]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[100.3328]
[101.]
epoch: 122 loss: 0.4451534152030945
tensor([[156.1233]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[156.12332]
[156.]
epoch: 123 loss: 0.015208200551569462
tensor([[134.9834]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[134.98343]
[134.]
epoch: 124 loss: 0.9671325087547302
tensor([[83.7953]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[83.79527]
[84.]
epoch: 125 loss: 0.041913215070962906
tensor([[147.5522]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[147.55217]
[147.]
epoch: 126 loss: 0.3048914968967438
tensor([[68.0705]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[68.07053]

```



```

[68.]
epoch: 127 loss: 0.004975010175257921
tensor([[153.2717]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[153.27171]
[153.]
epoch: 128 loss: 0.07382809370756149
tensor([[118.2535]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[118.25349]
[119.]
epoch: 129 loss: 0.5572822093963623
tensor([[198.5460]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[198.54604]
[200.]
epoch: 130 loss: 2.1140120029449463
tensor([[159.1023]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[159.10234]
[159.]
epoch: 131 loss: 0.010473618283867836
tensor([[85.9312]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[85.9312]
[85.]
epoch: 132 loss: 0.867129921913147
tensor([[104.0506]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[104.05063]
[103.]
epoch: 133 loss: 1.103820562362671
tensor([[18.0422]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[18.042208]
[18.]
epoch: 134 loss: 0.0017814914463087916
tensor([[82.9087]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[82.90867]
[82.]
epoch: 135 loss: 0.8256784677505493
tensor([[186.8860]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[186.88599]
[185.]
epoch: 136 loss: 3.5569443702697754
tensor([[79.7818]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],

```

```

device='cuda:0')
[79.78183]
[80.]
epoch: 137 loss: 0.0475982204079628
tensor([[111.7915]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[111.79151]
[113.]
epoch: 138 loss: 1.4604443311691284
tensor([[140.6973]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[140.69728]
[143.]
epoch: 139 loss: 5.302515506744385
tensor([[123.6004]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[123.60037]
[124.]
epoch: 140 loss: 0.15970228612422943
tensor([[136.8098]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[136.80975]
[137.]
epoch: 141 loss: 0.036193761974573135
tensor([[114.8290]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[114.82902]
[115.]
epoch: 142 loss: 0.02923496812582016
tensor([[21.6188]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[21.618793]
[22.]
epoch: 143 loss: 0.145318403840065
tensor([[86.8081]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[86.80807]
[86.]
epoch: 144 loss: 0.6529728174209595
tensor([[95.3678]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[95.36783]
[95.]
epoch: 145 loss: 0.1352977156639099
tensor([[197.2561]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[197.25615]
[196.]

```

```

epoch: 146 loss: 1.5779110193252563
tensor([[81.8173]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[81.81726]
[81.]
epoch: 147 loss: 0.6679151058197021
tensor([[18.0872]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[18.087246]
[19.]
epoch: 148 loss: 0.8331199884414673
tensor([[197.1989]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[197.19885]
[198.]
epoch: 149 loss: 0.6418372392654419
tensor([[61.7809]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[61.78095]
[61.]
epoch: 150 loss: 0.609880805015564
tensor([[193.6746]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[193.67459]
[195.]
epoch: 151 loss: 1.7567088603973389
tensor([[90.2914]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[90.291374]
[91.]
epoch: 152 loss: 0.5021505355834961
tensor([[98.7964]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[98.79643]
[98.]
epoch: 153 loss: 0.6343047022819519
tensor([[13.2698]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[13.2698345]
[12.]
epoch: 154 loss: 1.6124796867370605
tensor([[67.4397]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[67.43972]
[67.]
epoch: 155 loss: 0.1933538168668747
tensor([[190.9011]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')

```

```

[190.90114]
[190.]
epoch: 156 loss: 0.8120502233505249
tensor([[114.5388]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[114.5388]
[114.]
epoch: 157 loss: 0.29030877351760864
tensor([[44.6302]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[44.63018]
[45.]
epoch: 158 loss: 0.13676656782627106
tensor([[136.6551]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[136.65508]
[136.]
epoch: 159 loss: 0.4291233420372009
tensor([[61.9824]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[61.9824]
[62.]
epoch: 160 loss: 0.000309795665089041
tensor([[108.0506]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[108.0506]
[109.]
epoch: 161 loss: 0.9013639092445374
tensor([[106.6737]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[106.67373]
[108.]
epoch: 162 loss: 1.7589949369430542
tensor([[138.5679]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[138.56787]
[140.]
epoch: 163 loss: 2.0509932041168213
tensor([[193.6966]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[193.6966]
[194.]
epoch: 164 loss: 0.0920550525188446
tensor([[99.5351]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[99.53514]
[100.]
epoch: 165 loss: 0.21609389781951904

```

```

tensor([[143.0994]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[143.09944]
[142.]
epoch: 166 loss: 1.2087717056274414
tensor([[193.8277]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[193.82765]
[192.]
epoch: 167 loss: 3.3403117656707764
tensor([[5.7761]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[5.7761035]
[7.]
epoch: 168 loss: 1.497922658920288
tensor([[168.0253]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[168.0253]
[167.]
epoch: 169 loss: 1.0512381792068481
tensor([[149.6536]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[149.65363]
[149.]
epoch: 170 loss: 0.4272262752056122
tensor([[170.4845]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[170.48454]
[171.]
epoch: 171 loss: 0.26569607853889465
tensor([[54.9306]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[54.93058]
[55.]
epoch: 172 loss: 0.004819117020815611
tensor([[162.0837]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[162.08366]
[166.]
epoch: 173 loss: 15.337688446044922
tensor([[69.0267]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[69.026665]
[70.]
epoch: 174 loss: 0.9473815560340881
tensor([[48.0610]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[48.06098]

```

```

[48.]
epoch: 175 loss: 0.0037187738344073296
tensor([[189.0344]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[189.03441]
[189.]
epoch: 176 loss: 0.0011839496437460184
tensor([[156.6529]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[156.65294]
[154.]
epoch: 177 loss: 7.038084506988525
tensor([[65.0878]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[65.08784]
[64.]
epoch: 178 loss: 1.1833897829055786
tensor([[4.6042]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[4.604185]
[3.]
epoch: 179 loss: 2.5734097957611084
tensor([[97.7664]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[97.76637]
[97.]
epoch: 180 loss: 0.5873270630836487
tensor([[65.2499]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[65.24987]
[65.]
epoch: 181 loss: 0.06243516877293587
tensor([[48.5888]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[48.588825]
[49.]
epoch: 182 loss: 0.1690647006034851
tensor([[189.3456]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[189.34558]
[193.]
epoch: 183 loss: 13.354778289794922
tensor([[173.2693]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[173.26935]
[173.]
epoch: 184 loss: 0.07254841923713684
tensor([[152.1678]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],

```

```

device='cuda:0')
[152.16785]
[151.]
epoch: 185 loss: 1.363865852355957
tensor([[88.0926]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[88.09263]
[87.]
epoch: 186 loss: 1.19383704662323
For epoch 27 the test accuracy over the whole test set is 0 %
tensor([[62.3735]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[62.373478]
[61.]
epoch: 0 loss: 1.8864415884017944
tensor([[13.5459]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[13.545898]
[14.]
epoch: 1 loss: 0.2062082290649414
tensor([[138.5346]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[138.53458]
[136.]
epoch: 2 loss: 6.42407751083374
tensor([[94.2715]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[94.27152]
[95.]
epoch: 3 loss: 0.53067946434021
tensor([[32.4977]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[32.49775]
[32.]
epoch: 4 loss: 0.24775439500808716
tensor([[141.9629]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[141.96292]
[143.]
epoch: 5 loss: 1.0755325555801392
tensor([[159.5690]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[159.56898]
[160.]
epoch: 6 loss: 0.18577393889427185
tensor([[0.1050]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[0.10499646]

```

```

[1.]
epoch: 7 loss: 0.8010313510894775
tensor([[149.6481]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[149.64807]
[151.]
epoch: 8 loss: 1.8277112245559692
tensor([[0.3558]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[0.35578156]
[0.]
epoch: 9 loss: 0.12658052146434784
tensor([[57.2849]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[57.284943]
[58.]
epoch: 10 loss: 0.5113070607185364
tensor([[182.0633]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[182.06332]
[180.]
epoch: 11 loss: 4.25730562210083
tensor([[193.5119]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[193.51187]
[193.]
epoch: 12 loss: 0.2620122730731964
tensor([[66.3627]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[66.36274]
[68.]
epoch: 13 loss: 2.68062162399292
tensor([[78.1009]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[78.10093]
[80.]
epoch: 14 loss: 3.6064696311950684
tensor([[75.6195]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[75.61947]
[77.]
epoch: 15 loss: 1.9058667421340942
tensor([[97.6656]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[97.66557]
[98.]
epoch: 16 loss: 0.11184133589267731
tensor([[48.3354]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],

```



```

device='cuda:0')
[48.335392]
[49.]
epoch: 17 loss: 0.44170379638671875
tensor([[87.2937]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[87.29373]
[87.]
epoch: 18 loss: 0.08627830445766449
tensor([[3.5704]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[3.570439]
[3.]
epoch: 19 loss: 0.32540076971054077
tensor([[88.9848]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[88.984764]
[88.]
epoch: 20 loss: 0.9697603583335876
tensor([[198.7163]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[198.71631]
[196.]
epoch: 21 loss: 7.378332138061523
tensor([[204.5125]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[204.51248]
[204.]
epoch: 22 loss: 0.26263749599456787
tensor([[4.5238]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[4.523767]
[5.]
epoch: 23 loss: 0.22679787874221802
tensor([[79.7580]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[79.75798]
[81.]
epoch: 24 loss: 1.5426127910614014
tensor([[5.6897]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[5.6896515]
[7.]
epoch: 25 loss: 1.7170132398605347
tensor([[125.9053]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[125.9053]
[129.]

```

```

epoch: 26 loss: 9.577190399169922
tensor([[167.2593]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[167.2593]
[171.]
epoch: 27 loss: 13.992892265319824
tensor([[116.1312]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[116.1312]
[117.]
epoch: 28 loss: 0.7548087239265442
tensor([[90.6977]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[90.69767]
[90.]
epoch: 29 loss: 0.48674339056015015
tensor([[12.9577]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[12.957677]
[10.]
epoch: 30 loss: 8.747852325439453
tensor([[53.8814]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[53.88144]
[51.]
epoch: 31 loss: 8.302691459655762
tensor([[87.3812]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[87.38124]
[83.]
epoch: 32 loss: 19.195270538330078
tensor([[29.5776]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[29.57764]
[27.]
epoch: 33 loss: 6.644230842590332
tensor([[20.2316]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[20.231619]
[21.]
epoch: 34 loss: 0.590409517288208
tensor([[77.1598]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[77.159836]
[75.]
epoch: 35 loss: 4.664890766143799
tensor([[99.0445]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')

```

```

[99.04454]
[97.]
epoch: 36 loss: 4.180145263671875
tensor([[86.0152]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[86.01516]
[86.]
epoch: 37 loss: 0.0002298136823810637
tensor([[184.8422]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[184.84224]
[185.]
epoch: 38 loss: 0.0248884130269289
tensor([[113.5129]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[113.51294]
[115.]
epoch: 39 loss: 2.2113490104675293
tensor([[173.6152]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[173.61516]
[175.]
epoch: 40 loss: 1.9177871942520142
tensor([[30.8478]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[30.84778]
[33.]
epoch: 41 loss: 4.632050037384033
tensor([[151.1968]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[151.1968]
[152.]
epoch: 42 loss: 0.645142138004303
tensor([[173.0767]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[173.07669]
[173.]
epoch: 43 loss: 0.005881459452211857
tensor([[164.5890]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[164.58902]
[164.]
epoch: 44 loss: 0.34694430232048035
tensor([[204.5383]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[204.53835]
[205.]
epoch: 45 loss: 0.21312503516674042

```

```

tensor([[204.0161]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[204.0161]
[203.]
epoch: 46 loss: 1.0324552059173584
tensor([[105.6656]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[105.66562]
[106.]
epoch: 47 loss: 0.11181072145700455
tensor([[54.3302]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[54.330166]
[56.]
epoch: 48 loss: 2.7883460521698
tensor([[34.2070]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[34.207016]
[37.]
epoch: 49 loss: 7.800759792327881
tensor([[52.6681]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[52.668102]
[54.]
epoch: 50 loss: 1.773951530456543
tensor([[24.9271]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[24.927107]
[26.]
epoch: 51 loss: 1.151099681854248
tensor([[151.4283]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[151.42825]
[147.]
epoch: 52 loss: 19.609426498413086
tensor([[29.4127]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[29.412706]
[30.]
epoch: 53 loss: 0.3449138104915619
tensor([[118.5925]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[118.592514]
[118.]
epoch: 54 loss: 0.35107287764549255
tensor([[123.3404]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[123.340355]

```

```

[123.]
epoch: 55 loss: 0.11584147065877914
tensor([[75.5718]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[75.571846]
[78.]
epoch: 56 loss: 5.895931720733643
tensor([[51.8321]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[51.8321]
[53.]
epoch: 57 loss: 1.3639906644821167
tensor([[108.7364]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[108.73635]
[111.]
epoch: 58 loss: 5.1241068840026855
tensor([[44.7075]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[44.707455]
[46.]
epoch: 59 loss: 1.6706733703613281
tensor([[2.8460]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[2.8459904]
[4.]
epoch: 60 loss: 1.3317381143569946
tensor([[133.2291]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[133.2291]
[132.]
epoch: 61 loss: 1.5106756687164307
tensor([[153.7523]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[153.75229]
[154.]
epoch: 62 loss: 0.06136082857847214
tensor([[181.4054]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[181.4054]
[182.]
epoch: 63 loss: 0.35355448722839355
tensor([[122.7046]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[122.70458]
[122.]
epoch: 64 loss: 0.49643608927726746
tensor([[62.3743]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],

```

```

device='cuda:0')
[62.374275]
[60.]
epoch: 65 loss: 5.637182712554932
tensor([[161.9094]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[161.90941]
[162.]
epoch: 66 loss: 0.008206807076931
tensor([[155.7530]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[155.75298]
[156.]
epoch: 67 loss: 0.061021123081445694
tensor([[182.2229]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[182.22295]
[183.]
epoch: 68 loss: 0.6038126349449158
tensor([[92.7323]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[92.73226]
[93.]
epoch: 69 loss: 0.07168381661176682
tensor([[20.1961]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[20.19611]
[20.]
epoch: 70 loss: 0.03845904394984245
tensor([[170.0293]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[170.02927]
[169.]
epoch: 71 loss: 1.0593892335891724
tensor([[109.2892]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[109.28923]
[108.]
epoch: 72 loss: 1.6621148586273193
tensor([[17.5992]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[17.599247]
[16.]
epoch: 73 loss: 2.557590961456299
tensor([[32.4185]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[32.418465]
[29.]

```

```

epoch: 74 loss: 11.685900688171387
tensor([[174.5591]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[174.5591]
[177.]
epoch: 75 loss: 5.958005905151367
tensor([[137.9121]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[137.91211]
[140.]
epoch: 76 loss: 4.359287261962891
tensor([[6.1566]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[6.15656]
[6.]
epoch: 77 loss: 0.02451101690530777
tensor([[22.3475]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[22.347488]
[22.]
epoch: 78 loss: 0.12074819207191467
tensor([[181.4238]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[181.4238]
[179.]
epoch: 79 loss: 5.874794960021973
tensor([[108.0810]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[108.08098]
[107.]
epoch: 80 loss: 1.1685142517089844
tensor([[159.5161]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[159.51608]
[159.]
epoch: 81 loss: 0.2663414180278778
tensor([[147.1885]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[147.18846]
[148.]
epoch: 82 loss: 0.658595085144043
tensor([[13.1644]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[13.164416]
[15.]
epoch: 83 loss: 3.3693673610687256
tensor([[43.6023]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')

```

```

[43.602264]
[44.]
epoch: 84 loss: 0.15819360315799713
tensor([[164.1226]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[164.12259]
[166.]
epoch: 85 loss: 3.5246715545654297
tensor([[113.3046]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[113.30465]
[114.]
epoch: 86 loss: 0.4835125207901001
tensor([[190.1099]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[190.10994]
[191.]
epoch: 87 loss: 0.7922075390815735
tensor([[140.6304]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[140.63039]
[139.]
epoch: 88 loss: 2.6581597328186035
tensor([[43.4973]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[43.497265]
[43.]
epoch: 89 loss: 0.2472723424434662
tensor([[17.8233]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[17.823307]
[19.]
epoch: 90 loss: 1.3846063613891602
tensor([[94.5694]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[94.56941]
[94.]
epoch: 91 loss: 0.32423028349876404
tensor([[114.1894]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[114.189445]
[112.]
epoch: 92 loss: 4.793671607971191
tensor([[138.1594]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[138.15944]
[135.]
epoch: 93 loss: 9.9820556640625

```



```

tensor([[196.6434]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[196.64342]
[197.]
epoch: 94 loss: 0.12715117633342743
tensor([[83.3166]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[83.31656]
[85.]
epoch: 95 loss: 2.8339741230010986
tensor([[155.4020]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[155.40198]
[157.]
epoch: 96 loss: 2.553675889968872
tensor([[37.0187]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[37.018684]
[38.]
epoch: 97 loss: 0.962980329990387
tensor([[50.6859]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[50.685932]
[52.]
epoch: 98 loss: 1.7267743349075317
tensor([[153.1997]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[153.19968]
[155.]
epoch: 99 loss: 3.2411646842956543
tensor([[10.3840]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[10.384046]
[9.]
epoch: 100 loss: 1.9155821800231934
tensor([[189.6309]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[189.6309]
[190.]
epoch: 101 loss: 0.13623100519180298
tensor([[110.3778]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[110.37783]
[110.]
epoch: 102 loss: 0.1427558958530426
tensor([[8.1082]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[8.108179]

```

```

[8.]
epoch: 103 loss: 0.01170271635055542
tensor([[58.8022]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[58.802242]
[57.]
epoch: 104 loss: 3.248077154159546
tensor([[147.8528]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[147.8528]
[145.]
epoch: 105 loss: 8.138459205627441
tensor([[135.5349]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[135.53494]
[133.]
epoch: 106 loss: 6.425934314727783
tensor([[194.5224]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[194.52235]
[194.]
epoch: 107 loss: 0.272853821516037
tensor([[87.5796]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[87.57961]
[91.]
epoch: 108 loss: 11.69904899597168
tensor([[62.8331]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[62.83306]
[65.]
epoch: 109 loss: 4.695623874664307
tensor([[66.4882]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[66.4882]
[69.]
epoch: 110 loss: 6.309152603149414
tensor([[167.5203]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[167.52032]
[170.]
epoch: 111 loss: 6.148789405822754
tensor([[195.8286]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[195.82863]
[198.]
epoch: 112 loss: 4.7148542404174805
tensor([[102.0826]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],

```

```

device='cuda:0')
[102.08264]
[102.]
epoch: 113 loss: 0.0068296343088150024
tensor([[90.0870]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[90.087036]
[89.]
epoch: 114 loss: 1.181647539138794
tensor([[123.5680]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[123.56797]
[120.]
epoch: 115 loss: 12.730411529541016
tensor([[142.6349]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[142.63492]
[138.]
epoch: 116 loss: 21.482467651367188
tensor([[12.8715]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[12.87147]
[11.]
epoch: 117 loss: 3.50240159034729
tensor([[177.6205]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[177.62054]
[174.]
epoch: 118 loss: 13.108342170715332
tensor([[36.4666]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[36.4666]
[36.]
epoch: 119 loss: 0.21771417558193207
tensor([[94.2387]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[94.238716]
[99.]
epoch: 120 loss: 22.669824600219727
tensor([[190.7182]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[190.71825]
[199.]
epoch: 121 loss: 68.58744049072266
tensor([[121.4739]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[121.47391]
[125.]

```

```

epoch: 122 loss: 12.433328628540039
tensor([[23.0490]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[23.048985]
[23.]
epoch: 123 loss: 0.0023994839284569025
tensor([[37.5872]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[37.58723]
[35.]
epoch: 124 loss: 6.69376277923584
tensor([[141.7993]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[141.79932]
[134.]
epoch: 125 loss: 60.82933807373047
tensor([[15.2111]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[15.211105]
[17.]
epoch: 126 loss: 3.200144052505493
tensor([[108.5651]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[108.565094]
[104.]
epoch: 127 loss: 20.840084075927734
tensor([[170.4369]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[170.43689]
[163.]
epoch: 128 loss: 55.30732727050781
tensor([[65.8781]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[65.87807]
[67.]
epoch: 129 loss: 1.2587336301803589
tensor([[172.4844]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[172.48444]
[176.]
epoch: 130 loss: 12.359189987182617
tensor([[59.3152]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[59.315247]
[64.]
epoch: 131 loss: 21.946914672851562
tensor([[159.0850]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')

```

```

[159.08496]
[168.]
epoch: 132 loss: 79.47792053222656
tensor([[138.7782]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[138.77817]
[146.]
epoch: 133 loss: 52.15486145019531
tensor([[98.3646]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[98.364555]
[100.]
epoch: 134 loss: 2.6746792793273926
tensor([[12.1266]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[12.12661]
[12.]
epoch: 135 loss: 0.016030041500926018
tensor([[38.1660]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[38.165997]
[34.]
epoch: 136 loss: 17.355527877807617
tensor([[121.8659]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[121.8659]
[109.]
epoch: 137 loss: 165.53134155273438
tensor([[202.7944]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[202.79445]
[178.]
epoch: 138 loss: 614.7647094726562
tensor([[113.7438]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[113.7438]
[116.]
epoch: 139 loss: 5.090450763702393
tensor([[80.6669]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[80.66695]
[92.]
epoch: 140 loss: 128.4381103515625
tensor([[108.7961]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[108.79614]
[130.]
epoch: 141 loss: 449.60357666015625

```

```

tensor([[173.3826]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[173.38263]
[200.]
epoch: 142 loss: 708.4844360351562
tensor([[94.4400]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[94.440025]
[103.]
epoch: 143 loss: 73.27316284179688
tensor([[144.7411]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[144.74109]
[141.]
epoch: 144 loss: 13.995745658874512
tensor([[31.4613]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[31.461267]
[28.]
epoch: 145 loss: 11.980372428894043
tensor([[226.3510]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[226.35103]
[186.]
epoch: 146 loss: 1628.2054443359375
tensor([[90.8960]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[90.89596]
[84.]
epoch: 147 loss: 47.554237365722656
tensor([[22.3152]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[22.315216]
[25.]
epoch: 148 loss: 7.208064556121826
tensor([[108.2393]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[108.23928]
[124.]
epoch: 149 loss: 248.4002685546875
tensor([[45.9147]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[45.914692]
[55.]
epoch: 150 loss: 82.5428237915039
tensor([[129.3886]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[129.38863]

```

```

[167.]
epoch: 151 loss: 1414.615478515625
tensor([[127.7162]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[127.716156]
[153.]
epoch: 152 loss: 639.2727661132812
tensor([[1.4672]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[1.467214]
[2.]
epoch: 153 loss: 0.2838609218597412
tensor([[149.3071]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[149.30711]
[142.]
epoch: 154 loss: 53.3939094543457
tensor([[83.8384]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[83.83838]
[70.]
epoch: 155 loss: 191.500732421875
tensor([[97.2862]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[97.28623]
[79.]
epoch: 156 loss: 334.38629150390625
tensor([[163.7056]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[163.70558]
[137.]
epoch: 157 loss: 713.1881103515625
tensor([[78.6201]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[78.62008]
[71.]
epoch: 158 loss: 58.06560516357422
tensor([[189.9779]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[189.9779]
[189.]
epoch: 159 loss: 0.9562987089157104
tensor([[93.7338]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[93.733765]
[101.]
epoch: 160 loss: 52.79817581176758
tensor([[133.7588]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],

```

```

device='cuda:0')
[133.75883]
[150.]
epoch: 161 loss: 263.77545166015625
tensor([[62.3241]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[62.324112]
[76.]
epoch: 162 loss: 187.0299072265625
tensor([[68.4391]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[68.4391]
[82.]
epoch: 163 loss: 183.89794921875
tensor([[84.7220]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[84.72202]
[96.]
epoch: 164 loss: 127.19276428222656
tensor([[198.4851]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[198.48506]
[201.]
epoch: 165 loss: 6.324914932250977
tensor([[199.5756]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[199.57558]
[195.]
epoch: 166 loss: 20.935903549194336
tensor([[17.0928]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[17.092836]
[18.]
epoch: 167 loss: 0.8229458332061768
tensor([[178.0825]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[178.08246]
[165.]
epoch: 168 loss: 171.15072631835938
tensor([[191.3179]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[191.31793]
[184.]
epoch: 169 loss: 53.55213165283203
tensor([[128.4729]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[128.47289]
[126.]

```



```

epoch: 170 loss: 6.115160942077637
tensor([[145.5645]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[145.5645]
[149.]
epoch: 171 loss: 11.802667617797852
tensor([[41.9360]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[41.93595]
[45.]
epoch: 172 loss: 9.388395309448242
tensor([[13.1381]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[13.138074]
[13.]
epoch: 173 loss: 0.01906440779566765
tensor([[46.1412]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[46.141205]
[50.]
epoch: 174 loss: 14.890299797058105
tensor([[92.9897]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[92.9897]
[105.]
epoch: 175 loss: 144.24729919433594
tensor([[45.3570]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[45.356956]
[48.]
epoch: 176 loss: 6.9856791496276855
tensor([[37.9549]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[37.954937]
[39.]
epoch: 177 loss: 1.0921567678451538
tensor([[59.8992]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[59.899197]
[62.]
epoch: 178 loss: 4.413374900817871
tensor([[154.4687]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[154.46872]
[172.]
epoch: 179 loss: 307.3457946777344
tensor([[111.9126]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')

```

```

[111.91262]
[113.]
epoch: 180 loss: 1.182394027709961
tensor([[188.4568]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[188.4568]
[192.]
epoch: 181 loss: 12.55424976348877
tensor([[85.5913]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[85.59127]
[74.]
epoch: 182 loss: 134.3575439453125
tensor([[161.0401]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[161.04012]
[144.]
epoch: 183 loss: 290.36553955078125
tensor([[32.1568]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[32.15685]
[24.]
epoch: 184 loss: 66.53418731689453
tensor([[57.4620]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[57.462044]
[47.]
epoch: 185 loss: 109.45436096191406
tensor([[127.9754]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[127.97544]
[119.]
epoch: 186 loss: 80.55854034423828
For epoch 28 the test accuracy over the whole test set is 0 %
tensor([[132.9674]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[132.96738]
[130.]
epoch: 0 loss: 8.80532455444336
tensor([[186.0848]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[186.08481]
[196.]
epoch: 1 loss: 98.31102752685547
tensor([[127.5030]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[127.502975]
[135.]

```

```

epoch: 2 loss: 56.20537567138672
tensor([[124.4698]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[124.46981]
[129.]
epoch: 3 loss: 20.52261734008789
tensor([[9.0532]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[9.053191]
[8.]
epoch: 4 loss: 1.1092116832733154
tensor([[17.0536]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[17.053572]
[20.]
epoch: 5 loss: 8.681439399719238
tensor([[139.8341]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[139.8341]
[142.]
epoch: 6 loss: 4.691094875335693
tensor([[20.1276]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[20.127562]
[23.]
epoch: 7 loss: 8.25090217590332
tensor([[183.6576]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[183.65761]
[178.]
epoch: 8 loss: 32.00852966308594
tensor([[67.8628]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[67.86282]
[69.]
epoch: 9 loss: 1.29317045211792
tensor([[169.6411]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[169.64107]
[162.]
epoch: 10 loss: 58.38591384887695
tensor([[29.7633]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[29.76332]
[33.]
epoch: 11 loss: 10.476091384887695
tensor([[131.4792]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')

```

```

[131.47923]
[133.]
epoch: 12 loss: 2.3127329349517822
tensor([[111.7919]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[111.791885]
[115.]
epoch: 13 loss: 10.291999816894531
tensor([[113.1177]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[113.11773]
[117.]
epoch: 14 loss: 15.072026252746582
tensor([[106.6048]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[106.60475]
[110.]
epoch: 15 loss: 11.527711868286133
tensor([[203.2536]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[203.25362]
[203.]
epoch: 16 loss: 0.06432124227285385
tensor([[34.0756]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[34.07565]
[38.]
epoch: 17 loss: 15.400528907775879
tensor([[207.6415]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[207.64148]
[205.]
epoch: 18 loss: 6.977414131164551
tensor([[159.7344]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[159.73442]
[157.]
epoch: 19 loss: 7.477056980133057
tensor([[86.9739]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[86.97391]
[88.]
epoch: 20 loss: 1.0528658628463745
tensor([[59.2242]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[59.224182]
[60.]
epoch: 21 loss: 0.6018933653831482

```

```

tensor([[24.9967]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[24.996698]
[26.]
epoch: 22 loss: 1.006614089012146
tensor([[145.5056]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[145.50563]
[148.]
epoch: 23 loss: 6.221879005432129
tensor([[57.8051]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[57.805054]
[58.]
epoch: 24 loss: 0.03800405561923981
tensor([[154.0977]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[154.09767]
[154.]
epoch: 25 loss: 0.00953972339630127
tensor([[18.4292]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[18.429169]
[19.]
epoch: 26 loss: 0.32584837079048157
tensor([[90.3159]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[90.31595]
[90.]
epoch: 27 loss: 0.09982344508171082
tensor([[62.5630]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[62.56304]
[61.]
epoch: 28 loss: 2.4430992603302
tensor([[77.8683]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[77.868286]
[77.]
epoch: 29 loss: 0.7539207935333252
tensor([[5.1575]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[5.157455]
[2.]
epoch: 30 loss: 9.969521522521973
tensor([[9.0203]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[9.020349]

```

```

[6.]
epoch: 31 loss: 9.122505187988281
tensor([[33.3432]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[33.343185]
[32.]
epoch: 32 loss: 1.8041471242904663
tensor([[117.7451]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[117.74511]
[118.]
epoch: 33 loss: 0.06496913731098175
tensor([[93.8397]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[93.83969]
[94.]
epoch: 34 loss: 0.02569892257452011
tensor([[105.5703]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[105.57031]
[105.]
epoch: 35 loss: 0.32525634765625
tensor([[122.4262]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[122.42618]
[123.]
epoch: 36 loss: 0.32927170395851135
tensor([[171.7188]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[171.71877]
[180.]
epoch: 37 loss: 68.57884979248047
tensor([[169.8921]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[169.89209]
[169.]
epoch: 38 loss: 0.7958242893218994
tensor([[116.8030]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[116.80305]
[112.]
epoch: 39 loss: 23.06926155090332
tensor([[131.6594]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[131.65938]
[125.]
epoch: 40 loss: 44.34731674194336
tensor([[197.0699]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],

```

```

device='cuda:0')
[197.06995]
[194.]
epoch: 41 loss: 9.424570083618164
tensor([[122.7168]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[122.71681]
[122.]
epoch: 42 loss: 0.5138196349143982
tensor([[153.0786]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[153.07864]
[155.]
epoch: 43 loss: 3.6916096210479736
tensor([[95.8514]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[95.85143]
[97.]
epoch: 44 loss: 1.3192065954208374
tensor([[183.2771]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[183.27708]
[193.]
epoch: 45 loss: 94.53508758544922
tensor([[43.7432]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[43.743187]
[46.]
epoch: 46 loss: 5.093204975128174
tensor([[23.3775]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[23.377504]
[25.]
epoch: 47 loss: 2.6324920654296875
tensor([[4.4625]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[4.4625063]
[9.]
epoch: 48 loss: 20.588850021362305
tensor([[184.3419]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[184.34189]
[176.]
epoch: 49 loss: 69.58710479736328
tensor([[105.5177]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[105.51772]
[100.]

```

```

epoch: 50 loss: 30.445268630981445
tensor([[173.2023]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[173.20233]
[168.]
epoch: 51 loss: 27.064252853393555
tensor([[194.5511]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[194.55109]
[195.]
epoch: 52 loss: 0.20152339339256287
tensor([[139.7756]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[139.77557]
[146.]
epoch: 53 loss: 38.74348068237305
tensor([[8.8142]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[8.8141985]
[10.]
epoch: 54 loss: 1.4061251878738403
tensor([[34.7058]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[34.705814]
[37.]
epoch: 55 loss: 5.263287544250488
tensor([[41.9857]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[41.9857]
[45.]
epoch: 56 loss: 9.08601188659668
tensor([[189.5173]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[189.5173]
[198.]
epoch: 57 loss: 71.95613861083984
tensor([[127.5765]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[127.57649]
[132.]
epoch: 58 loss: 19.567420959472656
tensor([[20.1955]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[20.195541]
[21.]
epoch: 59 loss: 0.6471536755561829
tensor([[60.5551]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')

```



```

[60.555145]
[57.]
epoch: 60 loss: 12.639058113098145
tensor([[202.4169]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[202.41695]
[192.]
epoch: 61 loss: 108.51277160644531
tensor([[17.3382]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[17.33815]
[14.]
epoch: 62 loss: 11.143245697021484
tensor([[145.5104]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[145.5104]
[144.]
epoch: 63 loss: 2.281327724456787
tensor([[52.0707]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[52.0707]
[50.]
epoch: 64 loss: 4.287805080413818
tensor([[55.2400]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[55.240013]
[54.]
epoch: 65 loss: 1.5376325845718384
tensor([[14.4566]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[14.456574]
[13.]
epoch: 66 loss: 2.1216092109680176
tensor([[79.1465]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[79.14654]
[83.]
epoch: 67 loss: 14.849170684814453
tensor([[4.9483]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[4.9483366]
[3.]
epoch: 68 loss: 3.796015501022339
tensor([[62.1124]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],
device='cuda:0')
[62.112408]
[67.]
epoch: 69 loss: 23.888559341430664

```

```

tensor([[72.4331]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[72.43311]
[78.]
epoch: 70 loss: 30.990230560302734
tensor([[91.3209]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[91.3209]
[96.]
epoch: 71 loss: 21.89397621154785
tensor([[48.5589]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[48.558895]
[49.]
epoch: 72 loss: 0.1945735216140747
tensor([[90.7881]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[90.78815]
[93.]
epoch: 73 loss: 4.892293930053711
tensor([[1.0058]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],
device='cuda:0')
[1.0058167]
[0.]
epoch: 74 loss: 1.011667251586914
tensor([[83.9455]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[83.94554]
[82.]
epoch: 75 loss: 3.7851312160491943
tensor([[147.6478]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[147.64775]
[147.]
epoch: 76 loss: 0.41958117485046387
tensor([[115.9580]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[115.958]
[113.]
epoch: 77 loss: 8.749765396118164
tensor([[188.3006]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[188.30064]
[190.]
epoch: 78 loss: 2.8878111839294434
tensor([[103.8131]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')
[103.8131]

```

```

[98.]
epoch: 79 loss: 33.79216384887695
tensor([[88.9725]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[88.97252]
[85.]
epoch: 80 loss: 15.780906677246094
tensor([[178.8454]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[178.84538]
[183.]
epoch: 81 loss: 17.260845184326172
tensor([[90.7589]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[90.75891]
[89.]
epoch: 82 loss: 3.093768358230591
tensor([[72.4885]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[72.48849]
[71.]
epoch: 83 loss: 2.2155942916870117
tensor([[187.4991]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[187.49911]
[191.]
epoch: 84 loss: 12.256196022033691
tensor([[66.3095]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[66.309494]
[64.]
epoch: 85 loss: 5.3337626457214355
tensor([[207.3549]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[207.35492]
[204.]
epoch: 86 loss: 11.255484580993652
tensor([[162.3729]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[162.37292]
[163.]
epoch: 87 loss: 0.39322328567504883
tensor([[171.1535]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[171.1535]
[173.]
epoch: 88 loss: 3.4095497131347656
tensor([[158.2298]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],

```

```

device='cuda:0')
[158.22978]
[159.]
epoch: 89 loss: 0.5932356119155884
tensor([[67.8003]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')
[67.80035]
[68.]
epoch: 90 loss: 0.03986157104372978
tensor([[78.7904]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],
device='cuda:0')
[78.79037]
[79.]
epoch: 91 loss: 0.04394594207406044
tensor([[175.4396]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[175.43964]
[175.]
epoch: 92 loss: 0.19328001141548157
tensor([[91.7866]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[91.786644]
[92.]
epoch: 93 loss: 0.04552078992128372
tensor([[ -0.9208]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[ -0.92078394]
[4.]
epoch: 94 loss: 24.214115142822266
tensor([[104.6132]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[104.6132]
[103.]
epoch: 95 loss: 2.602405548095703
tensor([[206.0878]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[206.08783]
[200.]
epoch: 96 loss: 37.061668395996094
tensor([[11.0268]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[11.0268]
[15.]
epoch: 97 loss: 15.786316871643066
tensor([[102.9398]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[102.93976]
[107.]

```

```

epoch: 98 loss: 16.485563278198242
tensor([[103.0617]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[103.06172]
[108.]
epoch: 99 loss: 24.3865909576416
tensor([[42.9042]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[42.904205]
[44.]
epoch: 100 loss: 1.2007659673690796
tensor([[160.0617]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[160.06174]
[166.]
epoch: 101 loss: 35.26296615600586
tensor([[148.1686]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[148.16858]
[150.]
epoch: 102 loss: 3.354102611541748
tensor([[138.6787]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[138.67868]
[141.]
epoch: 103 loss: 5.388524532318115
tensor([[6.1465]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[6.1465464]
[1.]
epoch: 104 loss: 26.486940383911133
tensor([[21.2589]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[21.25888]
[18.]
epoch: 105 loss: 10.6203031539917
tensor([[19.6289]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[19.62889]
[16.]
epoch: 106 loss: 13.168835639953613
tensor([[33.0008]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[33.00075]
[28.]
epoch: 107 loss: 25.00751495361328
tensor([[60.0608]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')

```

```

[60.060776]
[53.]
epoch: 108 loss: 49.85455322265625
tensor([[149.7368]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[149.7368]
[138.]
epoch: 109 loss: 137.75250244140625
tensor([[109.8632]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[109.86319]
[101.]
epoch: 110 loss: 78.5561294555664
tensor([[143.9872]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[143.98724]
[143.]
epoch: 111 loss: 0.9746500253677368
tensor([[75.5609]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[75.56093]
[80.]
epoch: 112 loss: 19.70535659790039
tensor([[106.3382]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[106.33819]
[114.]
epoch: 113 loss: 58.70336151123047
tensor([[133.8934]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[133.89339]
[145.]
epoch: 114 loss: 123.35685729980469
tensor([[119.3335]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[119.33354]
[126.]
epoch: 115 loss: 44.44166564941406
tensor([[170.8335]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[170.8335]
[177.]
epoch: 116 loss: 38.02577209472656
tensor([[45.0707]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[45.070686]
[48.]
epoch: 117 loss: 8.580878257751465

```

```

tensor([[27.2900]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[27.290041]
[30.]
epoch: 118 loss: 7.343877792358398
tensor([[171.1694]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[171.16936]
[153.]
epoch: 119 loss: 330.12554931640625
tensor([[85.0288]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[85.02876]
[81.]
epoch: 120 loss: 16.23093032836914
tensor([[89.4879]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')
[89.48789]
[87.]
epoch: 121 loss: 6.1896071434021
tensor([[161.5910]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[161.59102]
[156.]
epoch: 122 loss: 31.259489059448242
tensor([[134.2072]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[134.20717]
[139.]
epoch: 123 loss: 22.971233367919922
tensor([[77.7221]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[77.72208]
[84.]
epoch: 124 loss: 39.412322998046875
tensor([[152.6822]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[152.68217]
[160.]
epoch: 125 loss: 53.550567626953125
tensor([[32.8188]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[32.818775]
[39.]
epoch: 126 loss: 38.20753860473633
tensor([[30.2943]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[30.29426]

```

```

[35.]
epoch: 127 loss: 22.14398956298828
tensor([[171.9966]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[171.9966]
[174.]
epoch: 128 loss: 4.013622283935547
tensor([[19.1429]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[19.142912]
[24.]
epoch: 129 loss: 23.591304779052734
tensor([[168.2606]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[168.26056]
[164.]
epoch: 130 loss: 18.152362823486328
tensor([[173.2598]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[173.2598]
[167.]
epoch: 131 loss: 39.1850471496582
tensor([[111.6737]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[111.6737]
[109.]
epoch: 132 loss: 7.14866304397583
tensor([[116.4764]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[116.47639]
[116.]
epoch: 133 loss: 0.22694459557533264
tensor([[9.3654]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[9.365356]
[12.]
epoch: 134 loss: 6.941346645355225
tensor([[171.6115]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[171.61153]
[182.]
epoch: 135 loss: 107.92037963867188
tensor([[176.6092]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[176.60924]
[184.]
epoch: 136 loss: 54.62336730957031
tensor([[31.5883]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],

```



```

device='cuda:0')
[31.588327]
[29.]
epoch: 137 loss: 6.699438571929932
tensor([[16.6569]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[16.65691]
[17.]
epoch: 138 loss: 0.11771078407764435
tensor([[94.8123]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[94.81231]
[86.]
epoch: 139 loss: 77.65679168701172
tensor([[149.4751]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[149.47514]
[137.]
epoch: 140 loss: 155.6291961669922
tensor([[183.8890]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[183.88896]
[170.]
epoch: 141 loss: 192.90325927734375
tensor([[13.3459]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[13.345891]
[11.]
epoch: 142 loss: 5.503204345703125
tensor([[6.1994]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[6.1993594]
[5.]
epoch: 143 loss: 1.4384629726409912
tensor([[105.9857]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[105.985695]
[119.]
epoch: 144 loss: 169.37213134765625
tensor([[152.0335]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[152.03351]
[172.]
epoch: 145 loss: 398.6607971191406
tensor([[66.6152]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[66.615166]
[75.]

```

```

epoch: 146 loss: 70.3054428100586
tensor([[98.4134]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[98.41336]
[104.]
epoch: 147 loss: 31.210540771484375
tensor([[148.6470]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[148.64703]
[149.]
epoch: 148 loss: 0.12458521127700806
tensor([[211.8190]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[211.81898]
[201.]
epoch: 149 loss: 117.0504379272461
tensor([[55.0027]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[55.00272]
[52.]
epoch: 150 loss: 9.016326904296875
tensor([[130.2274]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[130.22739]
[124.]
epoch: 151 loss: 38.78034210205078
tensor([[49.2659]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[49.26589]
[47.]
epoch: 152 loss: 5.134249210357666
tensor([[120.7589]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[120.75894]
[120.]
epoch: 153 loss: 0.5759924054145813
tensor([[183.7924]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[183.79239]
[189.]
epoch: 154 loss: 27.119213104248047
tensor([[169.6692]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[169.66922]
[171.]
epoch: 155 loss: 1.7709754705429077
tensor([[135.9962]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')

```

```

[135.99622]
[140.]
epoch: 156 loss: 16.030288696289062
tensor([[74.0718]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[74.07181]
[74.]
epoch: 157 loss: 0.00515636894851923
tensor([[97.4634]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[97.463394]
[95.]
epoch: 158 loss: 6.068310737609863
tensor([[36.4137]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[36.41365]
[34.]
epoch: 159 loss: 5.825708866119385
tensor([[104.4514]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[104.45136]
[99.]
epoch: 160 loss: 29.71735382080078
tensor([[168.0739]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[168.07394]
[165.]
epoch: 161 loss: 9.449131965637207
tensor([[51.3908]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],
device='cuda:0')
[51.39075]
[51.]
epoch: 162 loss: 0.1526862531900406
tensor([[89.9659]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[89.96592]
[91.]
epoch: 163 loss: 1.0693224668502808
tensor([[6.8160]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')
[6.8159885]
[7.]
epoch: 164 loss: 0.0338602177798748
tensor([[147.3017]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[147.30167]
[151.]
epoch: 165 loss: 13.677672386169434

```

```

tensor([[54.8823]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[54.88228]
[56.]
epoch: 166 loss: 1.2493014335632324
tensor([[62.7384]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[62.7384]
[65.]
epoch: 167 loss: 5.114836692810059
tensor([[41.3330]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[41.333023]
[43.]
epoch: 168 loss: 2.7788121700286865
tensor([[102.1408]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[102.14078]
[102.]
epoch: 169 loss: 0.019818328320980072
tensor([[34.8867]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[34.886677]
[36.]
epoch: 170 loss: 1.2394886016845703
tensor([[62.0036]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[62.00356]
[62.]
epoch: 171 loss: 1.2667282135225832e-05
tensor([[153.4979]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[153.49791]
[152.]
epoch: 172 loss: 2.2437329292297363
tensor([[184.2829]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[184.28293]
[185.]
epoch: 173 loss: 0.5141915678977966
tensor([[107.7595]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[107.75947]
[106.]
epoch: 174 loss: 3.0957279205322266
tensor([[134.9202]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[134.92015]

```

```

[136.]
epoch: 175 loss: 1.1660743951797485
tensor([[56.3014]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[56.30138]
[55.]
epoch: 176 loss: 1.6935902833938599
tensor([[185.0804]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[185.0804]
[186.]
epoch: 177 loss: 0.8456668257713318
tensor([[202.2339]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[202.23389]
[199.]
epoch: 178 loss: 10.458023071289062
tensor([[75.7247]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[75.72473]
[76.]
epoch: 179 loss: 0.07577277719974518
tensor([[68.4334]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[68.433426]
[70.]
epoch: 180 loss: 2.4541544914245605
tensor([[19.9253]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[19.92533]
[22.]
epoch: 181 loss: 4.304258823394775
tensor([[131.3759]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[131.37587]
[134.]
epoch: 182 loss: 6.886059761047363
tensor([[175.8056]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[175.80562]
[179.]
epoch: 183 loss: 10.20407485961914
tensor([[27.0180]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[27.017992]
[27.]
epoch: 184 loss: 0.00032371276756748557
tensor([[198.4994]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],

```

```

device='cuda:0')
[198.49936]
[197.]
epoch: 185 loss: 2.248077869415283
tensor([[114.0160]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[114.01598]
[111.]
epoch: 186 loss: 9.09615707397461
For epoch 29 the test accuracy over the whole test set is 0 %
tensor([[92.2652]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([91.],
device='cuda:0')
[92.26516]
[91.]
epoch: 0 loss: 1.6006288528442383
tensor([[108.7019]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([106.],
device='cuda:0')
[108.701935]
[106.]
epoch: 1 loss: 7.300451755523682
tensor([[153.4136]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([152.],
device='cuda:0')
[153.41357]
[152.]
epoch: 2 loss: 1.9981920719146729
tensor([[197.3570]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([203.],
device='cuda:0')
[197.357]
[203.]
epoch: 3 loss: 31.843509674072266
tensor([[61.1089]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([60.],
device='cuda:0')
[61.1089]
[60.]
epoch: 4 loss: 1.229655146598816
tensor([[26.4669]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([26.],
device='cuda:0')
[26.466866]
[26.]
epoch: 5 loss: 0.21796342730522156
tensor([[12.6978]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([11.],
device='cuda:0')
[12.697802]
[11.]
epoch: 6 loss: 2.8825302124023438
tensor([[70.1634]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([70.],
device='cuda:0')
[70.16343]

```

```

[70.]
epoch: 7 loss: 0.026709122583270073
tensor([[167.8820]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([160.],
device='cuda:0')
[167.882]
[160.]
epoch: 8 loss: 62.12598419189453
tensor([[20.6184]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([22.],
device='cuda:0')
[20.618357]
[22.]
epoch: 9 loss: 1.9089381694793701
tensor([[120.9014]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([122.],
device='cuda:0')
[120.90142]
[122.]
epoch: 10 loss: 1.2068767547607422
tensor([[55.9562]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([58.],
device='cuda:0')
[55.95616]
[58.]
epoch: 11 loss: 4.177275657653809
tensor([[139.2360]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([145.],
device='cuda:0')
[139.23596]
[145.]
epoch: 12 loss: 33.22413635253906
tensor([[150.8167]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([154.],
device='cuda:0')
[150.8167]
[154.]
epoch: 13 loss: 10.1334228515625
tensor([[196.5964]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([196.],
device='cuda:0')
[196.59639]
[196.]
epoch: 14 loss: 0.35568076372146606
tensor([[183.1710]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([182.],
device='cuda:0')
[183.17102]
[182.]
epoch: 15 loss: 1.3712890148162842
tensor([[7.3267]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([8.],
device='cuda:0')
[7.3266764]
[8.]
epoch: 16 loss: 0.45336470007896423
tensor([[65.8638]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([67.],

```

```

device='cuda:0')
[65.863815]
[67.]
epoch: 17 loss: 1.290915608406067
tensor([[103.3201]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([102.],
device='cuda:0')
[103.32011]
[102.]
epoch: 18 loss: 1.7426811456680298
tensor([[119.0866]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([117.],
device='cuda:0')
[119.086586]
[117.]
epoch: 19 loss: 4.353841304779053
tensor([[207.8043]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([201.],
device='cuda:0')
[207.80429]
[201.]
epoch: 20 loss: 46.29837417602539
tensor([[160.1776]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([162.],
device='cuda:0')
[160.17764]
[162.]
epoch: 21 loss: 3.320985794067383
tensor([[152.9473]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([157.],
device='cuda:0')
[152.94727]
[157.]
epoch: 22 loss: 16.42465591430664
tensor([[133.2296]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([141.],
device='cuda:0')
[133.22957]
[141.]
epoch: 23 loss: 60.37960433959961
tensor([[162.5991]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([166.],
device='cuda:0')
[162.59906]
[166.]
epoch: 24 loss: 11.56639289855957
tensor([[132.9809]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([134.],
device='cuda:0')
[132.9809]
[134.]
epoch: 25 loss: 1.0385730266571045
tensor([[3.2073]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([1.],
device='cuda:0')
[3.207334]
[1.]

```



```

epoch: 26 loss: 4.872323513031006
tensor([[94.7676]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([92.],
device='cuda:0')
[94.76758]
[92.]
epoch: 27 loss: 7.659488677978516
tensor([[65.3872]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([61.],
device='cuda:0')
[65.38718]
[61.]
epoch: 28 loss: 19.247318267822266
tensor([[58.4892]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([54.],
device='cuda:0')
[58.489227]
[54.]
epoch: 29 loss: 20.153162002563477
tensor([[32.9658]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([30.],
device='cuda:0')
[32.96582]
[30.]
epoch: 30 loss: 8.796090126037598
tensor([[213.5792]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([199.],
device='cuda:0')
[213.57921]
[199.]
epoch: 31 loss: 212.55331420898438
tensor([[38.0379]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([38.],
device='cuda:0')
[38.037937]
[38.]
epoch: 32 loss: 0.0014392284210771322
tensor([[185.7006]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([194.],
device='cuda:0')
[185.70062]
[194.]
epoch: 33 loss: 68.87966918945312
tensor([[122.6246]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([129.],
device='cuda:0')
[122.62458]
[129.]
epoch: 34 loss: 40.64597702026367
tensor([[78.3613]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([85.],
device='cuda:0')
[78.36131]
[85.]
epoch: 35 loss: 44.072166442871094
tensor([[4.4595]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([7.],
device='cuda:0')

```

```

[4.459504]
[7.]
epoch: 36 loss: 6.4541192054748535
tensor([[119.0214]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([123.],
device='cuda:0')
[119.021385]
[123.]
epoch: 37 loss: 15.829376220703125
tensor([[108.0755]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([111.],
device='cuda:0')
[108.07552]
[111.]
epoch: 38 loss: 8.552563667297363
tensor([[5.4939]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([6.],
device='cuda:0')
[5.493864]
[6.]
epoch: 39 loss: 0.25617358088493347
tensor([[164.3648]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([163.],
device='cuda:0')
[164.3648]
[163.]
epoch: 40 loss: 1.8626985549926758
tensor([[27.8570]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([28.],
device='cuda:0')
[27.856958]
[28.]
epoch: 41 loss: 0.02046090178191662
tensor([[59.1693]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([57.],
device='cuda:0')
[59.169323]
[57.]
epoch: 42 loss: 4.705962181091309
tensor([[147.1799]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([142.],
device='cuda:0')
[147.17995]
[142.]
epoch: 43 loss: 26.831850051879883
tensor([[85.0220]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([82.],
device='cuda:0')
[85.02203]
[82.]
epoch: 44 loss: 9.13268756866455
tensor([[130.3271]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([124.],
device='cuda:0')
[130.32707]
[124.]
epoch: 45 loss: 40.03184127807617

```

```

tensor([[136.6762]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([137.],
device='cuda:0')
[136.6762]
[137.]
epoch: 46 loss: 0.10485082119703293
tensor([[131.9489]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([133.],
device='cuda:0')
[131.94893]
[133.]
epoch: 47 loss: 1.104750633239746
tensor([[143.3552]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([148.],
device='cuda:0')
[143.3552]
[148.]
epoch: 48 loss: 21.574222564697266
tensor([[21.8987]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([25.],
device='cuda:0')
[21.898746]
[25.]
epoch: 49 loss: 9.617773056030273
tensor([[174.0141]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([186.],
device='cuda:0')
[174.01413]
[186.]
epoch: 50 loss: 143.6610870361328
tensor([[162.5697]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([164.],
device='cuda:0')
[162.56973]
[164.]
epoch: 51 loss: 2.0456645488739014
tensor([[188.0863]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([185.],
device='cuda:0')
[188.08633]
[185.]
epoch: 52 loss: 9.525459289550781
tensor([[108.8960]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([103.],
device='cuda:0')
[108.89596]
[103.]
epoch: 53 loss: 34.76232147216797
tensor([[213.2953]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([204.],
device='cuda:0')
[213.2953]
[204.]
epoch: 54 loss: 86.40266418457031
tensor([[14.4860]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([14.],
device='cuda:0')
[14.485981]

```

```

[14.]
epoch: 55 loss: 0.23617751896381378
tensor([[112.4583]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([112.],
device='cuda:0')
[112.45833]
[112.]
epoch: 56 loss: 0.21006478369235992
tensor([[32.8211]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([33.],
device='cuda:0')
[32.821068]
[33.]
epoch: 57 loss: 0.032016728073358536
tensor([[137.6366]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([144.],
device='cuda:0')
[137.63664]
[144.]
epoch: 58 loss: 40.4923210144043
tensor([[102.0694]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([107.],
device='cuda:0')
[102.069435]
[107.]
epoch: 59 loss: 24.310470581054688
tensor([[112.9384]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([116.],
device='cuda:0')
[112.93837]
[116.]
epoch: 60 loss: 9.373579978942871
tensor([[83.4504]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([86.],
device='cuda:0')
[83.45038]
[86.]
epoch: 61 loss: 6.500570297241211
tensor([[16.2076]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([17.],
device='cuda:0')
[16.20758]
[17.]
epoch: 62 loss: 0.6279285550117493
tensor([[170.6899]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([167.],
device='cuda:0')
[170.68994]
[167.]
epoch: 63 loss: 13.615667343139648
tensor([[184.0516]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([183.],
device='cuda:0')
[184.05157]
[183.]
epoch: 64 loss: 1.1058093309402466
tensor([[51.8574]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([51.],

```

```

device='cuda:0')
[51.85743]
[51.]
epoch: 65 loss: 0.7351853847503662
tensor([[153.4116]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([151.],
device='cuda:0')
[153.41156]
[151.]
epoch: 66 loss: 5.815621852874756
tensor([[57.4850]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([56.],
device='cuda:0')
[57.484978]
[56.]
epoch: 67 loss: 2.2051589488983154
tensor([[185.9624]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([191.],
device='cuda:0')
[185.9624]
[191.]
epoch: 68 loss: 25.377389907836914
tensor([[53.5540]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([52.],
device='cuda:0')
[53.554005]
[52.]
epoch: 69 loss: 2.414930582046509
tensor([[43.6952]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([43.],
device='cuda:0')
[43.695152]
[43.]
epoch: 70 loss: 0.483236700296402
tensor([[83.6532]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([83.],
device='cuda:0')
[83.65322]
[83.]
epoch: 71 loss: 0.42669785022735596
tensor([[153.9742]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([150.],
device='cuda:0')
[153.97417]
[150.]
epoch: 72 loss: 15.794002532958984
tensor([[95.7824]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([95.],
device='cuda:0')
[95.78239]
[95.]
epoch: 73 loss: 0.6121290922164917
tensor([[173.7696]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([178.],
device='cuda:0')
[173.76964]
[178.]

```

```

epoch: 74 loss: 17.89596176147461
tensor([[133.5527]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([130.],
device='cuda:0')
[133.55275]
[130.]
epoch: 75 loss: 12.622030258178711
tensor([[178.7237]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([195.],
device='cuda:0')
[178.72368]
[195.]
epoch: 76 loss: 264.91864013671875
tensor([[96.9510]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([93.],
device='cuda:0')
[96.95099]
[93.]
epoch: 77 loss: 15.610312461853027
tensor([[76.9010]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([74.],
device='cuda:0')
[76.901024]
[74.]
epoch: 78 loss: 8.415939331054688
tensor([[120.3425]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([109.],
device='cuda:0')
[120.342476]
[109.]
epoch: 79 loss: 128.65176391601562
tensor([[23.4246]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([24.],
device='cuda:0')
[23.424593]
[24.]
epoch: 80 loss: 0.3310932517051697
tensor([[6.4612]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([4.],
device='cuda:0')
[6.461185]
[4.]
epoch: 81 loss: 6.057431697845459
tensor([[183.7102]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([169.],
device='cuda:0')
[183.71024]
[169.]
epoch: 82 loss: 216.3910369873047
tensor([[7.7826]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([9.],
device='cuda:0')
[7.782585]
[9.]
epoch: 83 loss: 1.482098937034607
tensor([[86.4615]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([98.],
device='cuda:0')

```

```

[86.461494]
[98.]
epoch: 84 loss: 133.13711547851562
tensor([[190.5576]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([192.],
device='cuda:0')
[190.55762]
[192.]
epoch: 85 loss: 2.08046817779541
tensor([[55.1725]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([45.],
device='cuda:0')
[55.172493]
[45.]
epoch: 86 loss: 103.4796142578125
tensor([[34.5250]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([27.],
device='cuda:0')
[34.524982]
[27.]
epoch: 87 loss: 56.625362396240234
tensor([[102.7735]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([139.],
device='cuda:0')
[102.773544]
[139.]
epoch: 88 loss: 1312.3560791015625
tensor([[138.4568]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([159.],
device='cuda:0')
[138.45683]
[159.]
epoch: 89 loss: 422.021728515625
tensor([[115.8446]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([108.],
device='cuda:0')
[115.84461]
[108.]
epoch: 90 loss: 61.537940979003906
tensor([[35.2789]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([16.],
device='cuda:0')
[35.278896]
[16.]
epoch: 91 loss: 371.67584228515625
tensor([[154.5471]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([120.],
device='cuda:0')
[154.54712]
[120.]
epoch: 92 loss: 1193.50341796875
tensor([[40.8637]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([48.],
device='cuda:0')
[40.863728]
[48.]
epoch: 93 loss: 50.92638397216797

```

```

tensor([[35.1566]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([53.],
device='cuda:0')
[35.156574]
[53.]
epoch: 94 loss: 318.3878479003906
tensor([[184.2363]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([179.],
device='cuda:0')
[184.2363]
[179.]
epoch: 95 loss: 27.418811798095703
tensor([[29.9319]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([47.],
device='cuda:0')
[29.931866]
[47.]
epoch: 96 loss: 291.3211975097656
tensor([[5.2674]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([12.],
device='cuda:0')
[5.267396]
[12.]
epoch: 97 loss: 45.32795715332031
tensor([[168.7785]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([174.],
device='cuda:0')
[168.77855]
[174.]
epoch: 98 loss: 27.263547897338867
tensor([[79.2864]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([84.],
device='cuda:0')
[79.28644]
[84.]
epoch: 99 loss: 22.217666625976562
tensor([[98.1713]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([97.],
device='cuda:0')
[98.171326]
[97.]
epoch: 100 loss: 1.3720039129257202
tensor([[202.1584]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([189.],
device='cuda:0')
[202.15839]
[189.]
epoch: 101 loss: 173.14312744140625
tensor([[154.9704]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([153.],
device='cuda:0')
[154.97035]
[153.]
epoch: 102 loss: 3.8822877407073975
tensor([[18.9501]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([2.],
device='cuda:0')
[18.95014]

```



```

[2.]
epoch: 103 loss: 287.3072509765625
tensor([[152.1311]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([173.],
device='cuda:0')
[152.13112]
[173.]
epoch: 104 loss: 435.51019287109375
tensor([[179.7964]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([197.],
device='cuda:0')
[179.79639]
[197.]
epoch: 105 loss: 295.9643249511719
tensor([[203.4756]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([200.],
device='cuda:0')
[203.47563]
[200.]
epoch: 106 loss: 12.080016136169434
tensor([[189.0519]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([172.],
device='cuda:0')
[189.05191]
[172.]
epoch: 107 loss: 290.76763916015625
tensor([[56.0896]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([49.],
device='cuda:0')
[56.08962]
[49.]
epoch: 108 loss: 50.262691497802734
tensor([[111.6829]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([110.],
device='cuda:0')
[111.68286]
[110.]
epoch: 109 loss: 2.832022190093994
tensor([[23.5679]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([10.],
device='cuda:0')
[23.567942]
[10.]
epoch: 110 loss: 184.0890350341797
tensor([[77.9666]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([81.],
device='cuda:0')
[77.966606]
[81.]
epoch: 111 loss: 9.201478004455566
tensor([[106.0559]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([114.],
device='cuda:0')
[106.055855]
[114.]
epoch: 112 loss: 63.10944366455078
tensor([[73.4169]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([79.],

```

```

device='cuda:0')
[73.41688]
[79.]
epoch: 113 loss: 31.171253204345703
tensor([[111.5531]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([118.],
device='cuda:0')
[111.5531]
[118.]
epoch: 114 loss: 41.5625114440918
tensor([[37.2933]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([36.],
device='cuda:0')
[37.293262]
[36.]
epoch: 115 loss: 1.67252779006958
tensor([[179.6427]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([170.],
device='cuda:0')
[179.64268]
[170.]
epoch: 116 loss: 92.98136901855469
tensor([[77.0346]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([78.],
device='cuda:0')
[77.0346]
[78.]
epoch: 117 loss: 0.9319984912872314
tensor([[60.6763]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([62.],
device='cuda:0')
[60.676296]
[62.]
epoch: 118 loss: 1.7521916627883911
tensor([[181.2235]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([180.],
device='cuda:0')
[181.22354]
[180.]
epoch: 119 loss: 1.4970532655715942
tensor([[201.5725]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([198.],
device='cuda:0')
[201.57248]
[198.]
epoch: 120 loss: 12.76260757446289
tensor([[171.4728]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([176.],
device='cuda:0')
[171.47278]
[176.]
epoch: 121 loss: 20.495737075805664
tensor([[110.3600]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([119.],
device='cuda:0')
[110.35998]
[119.]

```

```

epoch: 122 loss: 74.64998626708984
tensor([[19.0738]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([20.],
device='cuda:0')
[19.073776]
[20.]
epoch: 123 loss: 0.8578904271125793
tensor([[132.1970]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([135.],
device='cuda:0')
[132.197]
[135.]
epoch: 124 loss: 7.85677433013916
tensor([[89.4223]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([90.],
device='cuda:0')
[89.42228]
[90.]
epoch: 125 loss: 0.3337611258029938
tensor([[174.4213]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([168.],
device='cuda:0')
[174.42134]
[168.]
epoch: 126 loss: 41.233619689941406
tensor([[151.9421]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([147.],
device='cuda:0')
[151.94211]
[147.]
epoch: 127 loss: 24.4244327545166
tensor([[114.5067]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([115.],
device='cuda:0')
[114.506676]
[115.]
epoch: 128 loss: 0.24336884915828705
tensor([[69.4699]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([71.],
device='cuda:0')
[69.46991]
[71.]
epoch: 129 loss: 2.3411765098571777
tensor([[22.6141]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([23.],
device='cuda:0')
[22.614115]
[23.]
epoch: 130 loss: 0.14890742301940918
tensor([[61.7988]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([64.],
device='cuda:0')
[61.79879]
[64.]
epoch: 131 loss: 4.845325469970703
tensor([[83.6784]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([87.],
device='cuda:0')

```

```

[83.678375]
[87.]
epoch: 132 loss: 11.033190727233887
tensor([[140.2144]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([146.],
device='cuda:0')
[140.21445]
[146.]
epoch: 133 loss: 33.47262191772461
tensor([[139.0251]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([143.],
device='cuda:0')
[139.02515]
[143.]
epoch: 134 loss: 15.799460411071777
tensor([[82.0998]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([80.],
device='cuda:0')
[82.0998]
[80.]
epoch: 135 loss: 4.409160614013672
tensor([[18.5446]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([13.],
device='cuda:0')
[18.544592]
[13.]
epoch: 136 loss: 30.74250030517578
tensor([[93.7765]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([88.],
device='cuda:0')
[93.776505]
[88.]
epoch: 137 loss: 33.368003845214844
tensor([[107.3261]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([101.],
device='cuda:0')
[107.32606]
[101.]
epoch: 138 loss: 40.01900100708008
tensor([[59.4088]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([55.],
device='cuda:0')
[59.40884]
[55.]
epoch: 139 loss: 19.4378719329834
tensor([[71.0957]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([69.],
device='cuda:0')
[71.09575]
[69.]
epoch: 140 loss: 4.392163276672363
tensor([[99.1109]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([100.],
device='cuda:0')
[99.11095]
[100.]
epoch: 141 loss: 0.7904158234596252

```

```

tensor([[183.8057]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([190.],
device='cuda:0')
[183.80566]
[190.]
epoch: 142 loss: 38.36979675292969
tensor([[124.6566]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([126.],
device='cuda:0')
[124.65658]
[126.]
epoch: 143 loss: 1.804782509803772
tensor([[31.7715]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([32.],
device='cuda:0')
[31.77151]
[32.]
epoch: 144 loss: 0.052208058536052704
tensor([[17.1631]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([18.],
device='cuda:0')
[17.163132]
[18.]
epoch: 145 loss: 0.7003485560417175
tensor([[97.3914]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([96.],
device='cuda:0')
[97.39136]
[96.]
epoch: 146 loss: 1.935875415802002
tensor([[7.8053]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([3.],
device='cuda:0')
[7.8053226]
[3.]
epoch: 147 loss: 23.09112548828125
tensor([[161.6209]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([156.],
device='cuda:0')
[161.62093]
[156.]
epoch: 148 loss: 31.59480857849121
tensor([[73.7072]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([75.],
device='cuda:0')
[73.707245]
[75.]
epoch: 149 loss: 1.6712157726287842
tensor([[205.4413]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([205.],
device='cuda:0')
[205.4413]
[205.]
epoch: 150 loss: 0.19474519789218903
tensor([[90.3706]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([94.],
device='cuda:0')
[90.37061]

```

```

[94.]
epoch: 151 loss: 13.172449111938477
tensor([[73.4198]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([77.],
device='cuda:0')
[73.41977]
[77.]
epoch: 152 loss: 12.818052291870117
tensor([[131.0890]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([132.],
device='cuda:0')
[131.08896]
[132.]
epoch: 153 loss: 0.8299961686134338
tensor([[182.9231]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([184.],
device='cuda:0')
[182.92314]
[184.]
epoch: 154 loss: 1.1596242189407349
tensor([[102.7837]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([105.],
device='cuda:0')
[102.78371]
[105.]
epoch: 155 loss: 4.911956310272217
tensor([[167.1394]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([165.],
device='cuda:0')
[167.1394]
[165.]
epoch: 156 loss: 4.577050685882568
tensor([[104.2412]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([104.],
device='cuda:0')
[104.24119]
[104.]
epoch: 157 loss: 0.058171674609184265
tensor([[138.4449]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([136.],
device='cuda:0')
[138.44493]
[136.]
epoch: 158 loss: 5.977687835693359
tensor([[14.9809]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([21.],
device='cuda:0')
[14.980917]
[21.]
epoch: 159 loss: 36.2293586730957
tensor([[127.5515]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([125.],
device='cuda:0')
[127.55154]
[125.]
epoch: 160 loss: 6.51033878326416
tensor([[2.5751]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([0.],

```

```

device='cuda:0')
[2.5751355]
[0.]
epoch: 161 loss: 6.631322860717773
tensor([[31.7909]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([34.],
device='cuda:0')
[31.790874]
[34.]
epoch: 162 loss: 4.88023567199707
tensor([[176.4801]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([177.],
device='cuda:0')
[176.4801]
[177.]
epoch: 163 loss: 0.27029335498809814
tensor([[11.8384]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([15.],
device='cuda:0')
[11.838362]
[15.]
epoch: 164 loss: 9.995956420898438
tensor([[174.3790]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([175.],
device='cuda:0')
[174.37904]
[175.]
epoch: 165 loss: 0.3855868875980377
tensor([[43.4007]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([46.],
device='cuda:0')
[43.400665]
[46.]
epoch: 166 loss: 6.756540775299072
tensor([[33.9796]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([35.],
device='cuda:0')
[33.97962]
[35.]
epoch: 167 loss: 1.0411715507507324
tensor([[88.6367]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([89.],
device='cuda:0')
[88.63666]
[89.]
epoch: 168 loss: 0.1320176124572754
tensor([[65.8481]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([65.],
device='cuda:0')
[65.84805]
[65.]
epoch: 169 loss: 0.7191938757896423
tensor([[195.7339]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([193.],
device='cuda:0')
[195.7339]
[193.]

```

```

epoch: 170 loss: 7.474219799041748
tensor([[29.9399]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([29.],
device='cuda:0')
[29.939878]
[29.]
epoch: 171 loss: 0.8833715319633484
tensor([[8.1462]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([5.],
device='cuda:0')
[8.146207]
[5.]
epoch: 172 loss: 9.89861774444458
tensor([[75.9489]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([76.],
device='cuda:0')
[75.94887]
[76.]
epoch: 173 loss: 0.0026145020965486765
tensor([[167.5378]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([171.],
device='cuda:0')
[167.53775]
[171.]
epoch: 174 loss: 11.987173080444336
tensor([[43.6159]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([44.],
device='cuda:0')
[43.615932]
[44.]
epoch: 175 loss: 0.14750787615776062
tensor([[111.2243]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([113.],
device='cuda:0')
[111.22427]
[113.]
epoch: 176 loss: 3.1532039642333984
tensor([[17.9326]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([19.],
device='cuda:0')
[17.932636]
[19.]
epoch: 177 loss: 1.1392652988433838
tensor([[39.9639]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([39.],
device='cuda:0')
[39.96389]
[39.]
epoch: 178 loss: 0.929084062576294
tensor([[51.8451]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([50.],
device='cuda:0')
[51.84505]
[50.]
epoch: 179 loss: 3.404212474822998
tensor([[70.4745]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([68.],
device='cuda:0')

```



```

[70.4745]
[68.]
epoch: 180 loss: 6.123162746429443
tensor([[38.6917]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([37.],
device='cuda:0')
[38.69171]
[37.]
epoch: 181 loss: 2.8618874549865723
tensor([[101.5712]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([99.],
device='cuda:0')
[101.57116]
[99.]
epoch: 182 loss: 6.610860347747803
tensor([[156.1131]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([155.],
device='cuda:0')
[156.11308]
[155.]
epoch: 183 loss: 1.2389534711837769
tensor([[149.5119]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([149.],
device='cuda:0')
[149.51189]
[149.]
epoch: 184 loss: 0.26202788949012756
tensor([[137.1286]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([140.],
device='cuda:0')
[137.12862]
[140.]
epoch: 185 loss: 8.244844436645508
tensor([[134.4577]], device='cuda:0', grad_fn=<AddmmBackward0>) tensor([138.],
device='cuda:0')
[134.45772]
[138.]
epoch: 186 loss: 12.547762870788574
For epoch 30 the test accuracy over the whole test set is 0 %
Finished Training
Finished Training
(tensor([[39.0987]], device='cuda:0'), tensor([40.]))

```

```

[ ]: print(predict(net, device))

(tensor([[159.0677]], device='cuda:0'), tensor([161.]))

```

7 save the Net

```

[ ]: torch.save(net, 'net.pkl')

```

8 Load the Net

```
[ ]: def restore_net():  
    net = torch.load('net.pkl').cuda()  
    print(predict(net, device))
```