Assignment 1: Test Your Neural Network on a Toy Dataset

Use this notebook to build your neural network by first implementing the following functions in the python files under assignment2/algorithms,

1. linear.py
2. relu.py
3. softmax.py
4. loss_func.py

First you will be testing your 2 layer neural network implementation on a toy dataset.

47
```
import sys
sys.path.append("../")
# Setup
import matplotlib.pyplot as plt
import numpy as np

from layers.sequential import Sequential
from layers.linear import Linear
from layers.relu import ReLU
from layers.softmax import Softmax
from layers.loss_func import CrossEntropyLoss
from utils.optimizer import SGD

%matplotlib inline
plt.rcParams['figure.figsize'] = (10.0, 8.0)  # set default size of plots

# For auto-reloading external modules
# See http://stackoverflow.com/questions/1907993/autoreload-of-modules-in-ipython
%load_ext autoreload
%autoreload 2
```

```
The autoreload extension is already loaded. To reload it, use:
  %reload_ext autoreload
```

We will use the class Sequential as implemented in the file assignment2/layers/sequential.py to build a layer by layer model of our neural network. Below we initialize the toy model and the toy random data that you will use to develop your implementation.

48
```
# Create a small net and some toy data to check your implementations.
# Note that we set the random seed for repeatable experiments.

input_size = 4
hidden_size = 10
num_classes = 3 # Output
num_inputs = 10 # N
```

```
def init_toy_model():
    np.random.seed(0)
    l1 = Linear(input_size, hidden_size)
    l2 = Linear(hidden_size, num_classes)

    r1 = ReLU()
    softmax = Softmax()
    return Sequential([l1, r1, l2, softmax])

def init_toy_data():
    np.random.seed(0)
    X = 10 * np.random.randn(num_inputs, input_size)
    y = np.random.randint(num_classes, size=num_inputs)
    #y = np.array([0, 1, 2, 2, 1])
    return X, y

net = init_toy_model()
X, y = init_toy_data()
```

Forward Pass: Compute Scores (20%)

Implement the forward functions in Linear, Relu and Softmax layers and get the output by passing our toy data X The output must match the given output scores

49

```
scores = net.forward(X)
print('Your scores:')
print(scores)
print()
print('correct scores:')
correct_scores = np.asarray([[0.33333514, 0.33333826, 0.33332661],
 [0.3333351,  0.33333828, 0.33332661],
 [0.3333351,  0.33333828, 0.33332662],
 [0.3333351,  0.33333828, 0.33332662],
 [0.33333509, 0.33333829, 0.33332662],
 [0.33333508, 0.33333829, 0.33332662],
 [0.33333511, 0.33333828, 0.33332661],
 [0.33333512, 0.33333827, 0.33332661],
 [0.33333508, 0.33333829, 0.33332662],
 [0.33333511, 0.33333828, 0.33332662]])
print(correct_scores)

# The difference should be very small. We get < 1e-7
print('Difference between your scores and correct scores:')
print(np.sum(np.abs(scores - correct_scores)))
```

```
Your scores:
[[0.33333514 0.33333826 0.33332661]
 [0.3333351  0.33333828 0.33332661]
 [0.3333351  0.33333828 0.33332662]
 [0.3333351  0.33333828 0.33332662]
 [0.33333509 0.33333829 0.33332662]
 [0.33333508 0.33333829 0.33332662]
 [0.33333511 0.33333828 0.33332661]
 [0.33333512 0.33333827 0.33332661]
 [0.33333508 0.33333829 0.33332662]
 [0.33333511 0.33333828 0.33332662]]

correct scores:
[[0.33333514 0.33333826 0.33332661]
 [0.3333351  0.33333828 0.33332661]
```

```
 [0.3333351  0.33333828 0.33332662]
 [0.3333351  0.33333828 0.33332662]
 [0.33333509 0.33333829 0.33332662]
 [0.33333508 0.33333829 0.33332662]
 [0.33333511 0.33333828 0.33332661]
 [0.33333512 0.33333827 0.33332661]
 [0.33333508 0.33333829 0.33332662]
 [0.33333511 0.33333828 0.33332662]]
Difference between your scores and correct scores:
8.799388540037256e-08
```

Forward Pass: Compute loss given the output scores from the previous step (10%)

Implement the forward function in the loss_func.py file, and output the loss value. The loss value must match the given loss value.

50

```
Loss = CrossEntropyLoss()
loss = Loss.forward(scores,y)
correct_loss = 1.098612723362578
print(loss)
# should be very small, we get < 1e-12
print('Difference between your loss and correct loss:')
print(np.sum(np.abs(loss - correct_loss)))
```

```
1.098612723362578
Difference between your loss and correct loss:
0.0
```

Backward Pass (40%)

Implement the rest of the functions in the given files. Specifically, implement the backward function in all the 4 files as mentioned in the files. Note: No backward function in the softmax file, the gradient for softmax is jointly calculated with the cross entropy loss in the loss_func.backward function.

You will use the chain rule to calculate gradient individually for each layer. You can assume that this calculated gradeint then is passed to the next layers in a reversed manner due to the Sequential implementation. So all you need to worry about is implementing the gradient for the current layer and multiply it will the incoming gradient (passed to the backward function as dout) to calculate the total gradient for the parameters of that layer.

51

```
# No need to edit anything in this block ( 20% of the above 40% )
net.backward(Loss.backward())

gradients = []
for module in net._modules:
    for para, grad in zip(module.parameters, module.grads):
        assert grad is not None, "No Gradient"
        #Print gradients of the linear layer
        print(grad.shape)
        gradients.append(grad)

# Check shapes of your gradient. Note that only the linear layer has parameters
```

```
#(4, 10) -> Layer 1 W
#(10,)   -> Layer 1 b
#(10, 3) -> Layer 2 W
#(3,)    -> Layer 2 b
```

```
(4, 10)
(10,)
(10, 3)
(3,)
```

52

```
# No need to edit anything in this block ( 20% of the above 40% )
# Now we check the values for these gradients. Here are the values for these gradients, Below we calculat
# difference, you must get difference < 1e-10
grad_w1 = np.array([[-6.24320917e-05,  3.41037180e-06, -1.69125969e-05,
         2.41514079e-05,  3.88697976e-06,  7.63842314e-05,
        -8.88925758e-05,  3.34909890e-05, -1.42758303e-05,
        -4.74748560e-06],
       [-7.16182867e-05,  4.63270039e-06, -2.20344270e-05,
        -2.72027034e-06,  6.52903437e-07,  8.97294847e-05,
        -1.05981609e-04,  4.15825391e-05, -2.12210745e-05,
         3.06061658e-05],
       [-1.69074923e-05, -8.83185056e-06,  3.10730840e-05,
         1.23010428e-05,  5.25830316e-05, -7.82980115e-06,
         3.02117990e-05, -3.37645284e-05,  6.17276346e-05,
        -1.10735656e-05],
       [-4.35902272e-05,  3.71512704e-06, -1.66837877e-05,
         2.54069557e-06, -4.33258099e-06,  5.72310022e-05,
        -6.94881762e-05,  2.92408329e-05, -1.89369767e-05,
         2.01692516e-05]])
grad_b1 = np.array([-2.27150209e-06,  5.14674340e-07, -2.04284403e-06,  6.08849787e-07, -1.92177796e-06,
                    3.92085824e-06, -5.40772636e-06,  2.93354593e-06, -3.14568138e-06,  5.27501592e-11])

grad_w2 = np.array([[ 1.28932983e-04,  1.19946731e-04, -2.48879714e-04],
       [ 1.08784150e-04,  1.55140199e-04, -2.63924349e-04],
       [ 6.96017544e-05,  1.42748410e-04, -2.12350164e-04],
       [ 9.92512487e-05,  1.73257611e-04, -2.72508860e-04],
       [ 2.05484895e-05,  4.96161144e-05, -7.01646039e-05],
       [ 8.20539510e-05,  9.37063861e-05, -1.75760337e-04],
       [ 2.45831715e-05,  8.74369112e-05, -1.12020083e-04],
       [ 1.34073379e-04,  1.86253064e-04, -3.20326443e-04],
       [ 8.86473128e-05,  2.35554414e-04, -3.24201726e-04],
       [ 3.57433149e-05,  1.91164061e-04, -2.26907376e-04]])

grad_b2 = np.array([-0.1666649 ,  0.13333828,  0.03332662])

difference = np.sum(np.abs(gradients[0]-grad_w1)) + np.sum(np.abs(gradients[1]-grad_b1)) + np.sum(np.abs(
             np.sum(np.abs(gradients[3]-grad_b2))
print("Difference in Gradient values", difference)
```

```
Difference in Gradient values 7.70191643436727e-09
```

Train the complete network on the toy data.

To train the network we will use stochastic gradient descent (SGD), we have implemented the optimizer for you. You do not implement any more functions in the python files. Below we implement the training procedure, you should get yourself familiar with the training process. Specifically looking at which functions to call and when.

Once you have implemented the method and tested various parts in the above blocks, run the code below to train a two-layer network on toy data. You should see your the training loss decrease to

53
```
# Training Procedure
# Initialize the optimizer. DO NOT change any of the hyper-parameters here or above.
# We have implemented the SGD optimizer class for you here, which visits each layer sequentially to
# get the gradients and optimize the respective parameters.
# You should work with the given parameters and only edit your implementation in the .py files

epochs = 1000
optim = SGD(net, lr=0.1, weight_decay=0.00001)

epoch_loss = []
for epoch in range(epochs):
            # Get output scores from the network
            output_x = net(X)
            # Calculate the loss for these output scores, given the true labels
            loss = Loss.forward(output_x, y)
            # Initialize your gradients to None in each epoch
            optim.zero_grad()
            # Make a backward pass to update the internal gradients in the layers
            net.backward(Loss.backward())
            # call the step function in the optimizer to update the values of the params with the gradien
            optim.step()
            # Append the loss at each iteration
            epoch_loss.append(loss)

            print("Epoch Loss: {:3f}".format(epoch_loss[-1]))
```

```
Epoch Loss: 1.098613
Epoch Loss: 1.094024
Epoch Loss: 1.089737
Epoch Loss: 1.085733
Epoch Loss: 1.081994
Epoch Loss: 1.078505
Epoch Loss: 1.075248
Epoch Loss: 1.072208
Epoch Loss: 1.069372
Epoch Loss: 1.066726
Epoch Loss: 1.064257
Epoch Loss: 1.061952
Epoch Loss: 1.059799
Epoch Loss: 1.057785
Epoch Loss: 1.055899
Epoch Loss: 1.054124
Epoch Loss: 1.052445
Epoch Loss: 1.050839
Epoch Loss: 1.049277
Epoch Loss: 1.047713
Epoch Loss: 1.046081
Epoch Loss: 1.044279
Epoch Loss: 1.042147
Epoch Loss: 1.039435
Epoch Loss: 1.035774
Epoch Loss: 1.030637
Epoch Loss: 1.023365
Epoch Loss: 1.013337
Epoch Loss: 1.000400
Epoch Loss: 0.985502
Epoch Loss: 0.970780
Epoch Loss: 0.958224
```

```
Epoch Loss: 0.948146
Epoch Loss: 0.939687
Epoch Loss: 0.932411
Epoch Loss: 0.926597
Epoch Loss: 0.920919
Epoch Loss: 0.914630
Epoch Loss: 0.908446
Epoch Loss: 0.902708
Epoch Loss: 0.895794
Epoch Loss: 0.889273
Epoch Loss: 0.882132
Epoch Loss: 0.875647
Epoch Loss: 0.870536
Epoch Loss: 0.861775
Epoch Loss: 0.855117
Epoch Loss: 0.848626
Epoch Loss: 0.839904
Epoch Loss: 0.832706
Epoch Loss: 0.827365
Epoch Loss: 0.815973
Epoch Loss: 0.810844
Epoch Loss: 0.802194
Epoch Loss: 0.790601
Epoch Loss: 0.783502
Epoch Loss: 0.770632
Epoch Loss: 0.760160
Epoch Loss: 0.749472
Epoch Loss: 0.739295
Epoch Loss: 0.732478
Epoch Loss: 0.719142
Epoch Loss: 0.713844
Epoch Loss: 0.700037
Epoch Loss: 0.693629
Epoch Loss: 0.689958
Epoch Loss: 0.674298
Epoch Loss: 0.659023
Epoch Loss: 0.647852
Epoch Loss: 0.638275
Epoch Loss: 0.628526
Epoch Loss: 0.622772
Epoch Loss: 0.617140
Epoch Loss: 0.608889
Epoch Loss: 0.601315
Epoch Loss: 0.590965
Epoch Loss: 0.588483
Epoch Loss: 0.580111
Epoch Loss: 0.580237
Epoch Loss: 0.579590
Epoch Loss: 0.575265
Epoch Loss: 0.583167
Epoch Loss: 0.569299
Epoch Loss: 0.567051
Epoch Loss: 0.558750
Epoch Loss: 0.570455
Epoch Loss: 0.550303
Epoch Loss: 0.556647
Epoch Loss: 0.528476
Epoch Loss: 0.523165
Epoch Loss: 0.503276
Epoch Loss: 0.508981
Epoch Loss: 0.493973
Epoch Loss: 0.512211
Epoch Loss: 0.477031
Epoch Loss: 0.483331
Epoch Loss: 0.456475
```

```
Epoch Loss: 0.472357
Epoch Loss: 0.443569
Epoch Loss: 0.454687
Epoch Loss: 0.429413
Epoch Loss: 0.441666
Epoch Loss: 0.413838
Epoch Loss: 0.443494
Epoch Loss: 0.407146
Epoch Loss: 0.444991
Epoch Loss: 0.403453
Epoch Loss: 0.436572
Epoch Loss: 0.397764
Epoch Loss: 0.439797
Epoch Loss: 0.422846
Epoch Loss: 0.526893
Epoch Loss: 0.553876
Epoch Loss: 0.728167
Epoch Loss: 0.582263
Epoch Loss: 0.547702
Epoch Loss: 0.416240
Epoch Loss: 0.453693
Epoch Loss: 0.381732
Epoch Loss: 0.362521
Epoch Loss: 0.324897
Epoch Loss: 0.338047
Epoch Loss: 0.342623
Epoch Loss: 0.326819
Epoch Loss: 0.392589
Epoch Loss: 0.292861
Epoch Loss: 0.314854
Epoch Loss: 0.235608
Epoch Loss: 0.227234
Epoch Loss: 0.207276
Epoch Loss: 0.220121
Epoch Loss: 0.208397
Epoch Loss: 0.249017
Epoch Loss: 0.226502
Epoch Loss: 0.284894
Epoch Loss: 0.233826
Epoch Loss: 0.259128
Epoch Loss: 0.195131
Epoch Loss: 0.177242
Epoch Loss: 0.156995
Epoch Loss: 0.147287
Epoch Loss: 0.142642
Epoch Loss: 0.139515
Epoch Loss: 0.136126
Epoch Loss: 0.135978
Epoch Loss: 0.130053
Epoch Loss: 0.127937
Epoch Loss: 0.122477
Epoch Loss: 0.120641
Epoch Loss: 0.118350
Epoch Loss: 0.117382
Epoch Loss: 0.112601
Epoch Loss: 0.111264
Epoch Loss: 0.106784
Epoch Loss: 0.105601
Epoch Loss: 0.105261
Epoch Loss: 0.103690
Epoch Loss: 0.099740
Epoch Loss: 0.097608
Epoch Loss: 0.094972
Epoch Loss: 0.094597
Epoch Loss: 0.093439
```

```
Epoch Loss: 0.092069
Epoch Loss: 0.090610
Epoch Loss: 0.087682
Epoch Loss: 0.085767
Epoch Loss: 0.084664
Epoch Loss: 0.083181
Epoch Loss: 0.082531
Epoch Loss: 0.081955
Epoch Loss: 0.079550
Epoch Loss: 0.078298
Epoch Loss: 0.077149
Epoch Loss: 0.075943
Epoch Loss: 0.074927
Epoch Loss: 0.074027
Epoch Loss: 0.074400
Epoch Loss: 0.072842
Epoch Loss: 0.071362
Epoch Loss: 0.070140
Epoch Loss: 0.069109
Epoch Loss: 0.068189
Epoch Loss: 0.067438
Epoch Loss: 0.066542
Epoch Loss: 0.065878
Epoch Loss: 0.065619
Epoch Loss: 0.065236
Epoch Loss: 0.064023
Epoch Loss: 0.063155
Epoch Loss: 0.062142
Epoch Loss: 0.061411
Epoch Loss: 0.060913
Epoch Loss: 0.060049
Epoch Loss: 0.059368
Epoch Loss: 0.059265
Epoch Loss: 0.058653
Epoch Loss: 0.057884
Epoch Loss: 0.057724
Epoch Loss: 0.056640
Epoch Loss: 0.055911
Epoch Loss: 0.055581
Epoch Loss: 0.054974
Epoch Loss: 0.054336
Epoch Loss: 0.053731
Epoch Loss: 0.053237
Epoch Loss: 0.052698
Epoch Loss: 0.052610
Epoch Loss: 0.052149
Epoch Loss: 0.051605
Epoch Loss: 0.051018
Epoch Loss: 0.050427
Epoch Loss: 0.050176
Epoch Loss: 0.050089
Epoch Loss: 0.049610
Epoch Loss: 0.048936
Epoch Loss: 0.048407
Epoch Loss: 0.048042
Epoch Loss: 0.047693
Epoch Loss: 0.047232
Epoch Loss: 0.047006
Epoch Loss: 0.046590
Epoch Loss: 0.046342
Epoch Loss: 0.045836
Epoch Loss: 0.045452
Epoch Loss: 0.045112
Epoch Loss: 0.044795
Epoch Loss: 0.044464
```

```
Epoch Loss: 0.044038
Epoch Loss: 0.043724
Epoch Loss: 0.043653
Epoch Loss: 0.043454
Epoch Loss: 0.042933
Epoch Loss: 0.042565
Epoch Loss: 0.042280
Epoch Loss: 0.041901
Epoch Loss: 0.041709
Epoch Loss: 0.041528
Epoch Loss: 0.041169
Epoch Loss: 0.040963
Epoch Loss: 0.040683
Epoch Loss: 0.040507
Epoch Loss: 0.040178
Epoch Loss: 0.039961
Epoch Loss: 0.039609
Epoch Loss: 0.039391
Epoch Loss: 0.039159
Epoch Loss: 0.038825
Epoch Loss: 0.038546
Epoch Loss: 0.038335
Epoch Loss: 0.038039
Epoch Loss: 0.037768
Epoch Loss: 0.037602
Epoch Loss: 0.037523
Epoch Loss: 0.037224
Epoch Loss: 0.036954
Epoch Loss: 0.036836
Epoch Loss: 0.036784
Epoch Loss: 0.036457
Epoch Loss: 0.036186
Epoch Loss: 0.035966
Epoch Loss: 0.035764
Epoch Loss: 0.035549
Epoch Loss: 0.035444
Epoch Loss: 0.035394
Epoch Loss: 0.035143
Epoch Loss: 0.034874
Epoch Loss: 0.034649
Epoch Loss: 0.034474
Epoch Loss: 0.034245
Epoch Loss: 0.034032
Epoch Loss: 0.033865
Epoch Loss: 0.033680
Epoch Loss: 0.033470
Epoch Loss: 0.033298
Epoch Loss: 0.033177
Epoch Loss: 0.032968
Epoch Loss: 0.032779
Epoch Loss: 0.032638
Epoch Loss: 0.032442
Epoch Loss: 0.032259
Epoch Loss: 0.032117
Epoch Loss: 0.032015
Epoch Loss: 0.032165
Epoch Loss: 0.031888
Epoch Loss: 0.031721
Epoch Loss: 0.031566
Epoch Loss: 0.031528
Epoch Loss: 0.031292
Epoch Loss: 0.031094
Epoch Loss: 0.030956
Epoch Loss: 0.030888
Epoch Loss: 0.030697
```

```
Epoch Loss: 0.030521
Epoch Loss: 0.030383
Epoch Loss: 0.030241
Epoch Loss: 0.030111
Epoch Loss: 0.029963
Epoch Loss: 0.029830
Epoch Loss: 0.029668
Epoch Loss: 0.029528
Epoch Loss: 0.029408
Epoch Loss: 0.029255
Epoch Loss: 0.029114
Epoch Loss: 0.029005
Epoch Loss: 0.028858
Epoch Loss: 0.028718
Epoch Loss: 0.028612
Epoch Loss: 0.028480
Epoch Loss: 0.028342
Epoch Loss: 0.028228
Epoch Loss: 0.028112
Epoch Loss: 0.028042
Epoch Loss: 0.028009
Epoch Loss: 0.027869
Epoch Loss: 0.027722
Epoch Loss: 0.027666
Epoch Loss: 0.027744
Epoch Loss: 0.027542
Epoch Loss: 0.027385
Epoch Loss: 0.027244
Epoch Loss: 0.027115
Epoch Loss: 0.027015
Epoch Loss: 0.026908
Epoch Loss: 0.026801
Epoch Loss: 0.026707
Epoch Loss: 0.026660
Epoch Loss: 0.026527
Epoch Loss: 0.026404
Epoch Loss: 0.026311
Epoch Loss: 0.026192
Epoch Loss: 0.026078
Epoch Loss: 0.025984
Epoch Loss: 0.025883
Epoch Loss: 0.025771
Epoch Loss: 0.025673
Epoch Loss: 0.025583
Epoch Loss: 0.025474
Epoch Loss: 0.025383
Epoch Loss: 0.025316
Epoch Loss: 0.025206
Epoch Loss: 0.025106
Epoch Loss: 0.025025
Epoch Loss: 0.024921
Epoch Loss: 0.024822
Epoch Loss: 0.024753
Epoch Loss: 0.024686
Epoch Loss: 0.024655
Epoch Loss: 0.024557
Epoch Loss: 0.024513
Epoch Loss: 0.024400
Epoch Loss: 0.024295
Epoch Loss: 0.024202
Epoch Loss: 0.024120
Epoch Loss: 0.024023
Epoch Loss: 0.023949
Epoch Loss: 0.023882
Epoch Loss: 0.023873
```

```
Epoch Loss: 0.023803
Epoch Loss: 0.023696
Epoch Loss: 0.023608
Epoch Loss: 0.023513
Epoch Loss: 0.023421
Epoch Loss: 0.023337
Epoch Loss: 0.023263
Epoch Loss: 0.023175
Epoch Loss: 0.023089
Epoch Loss: 0.023058
Epoch Loss: 0.022986
Epoch Loss: 0.022898
Epoch Loss: 0.022820
Epoch Loss: 0.022764
Epoch Loss: 0.022686
Epoch Loss: 0.022602
Epoch Loss: 0.022536
Epoch Loss: 0.022452
Epoch Loss: 0.022371
Epoch Loss: 0.022310
Epoch Loss: 0.022230
Epoch Loss: 0.022152
Epoch Loss: 0.022091
Epoch Loss: 0.022017
Epoch Loss: 0.021940
Epoch Loss: 0.021877
Epoch Loss: 0.021808
Epoch Loss: 0.021733
Epoch Loss: 0.021676
Epoch Loss: 0.021642
Epoch Loss: 0.021625
Epoch Loss: 0.021543
Epoch Loss: 0.021474
Epoch Loss: 0.021424
Epoch Loss: 0.021348
Epoch Loss: 0.021282
Epoch Loss: 0.021211
Epoch Loss: 0.021137
Epoch Loss: 0.021074
Epoch Loss: 0.021011
Epoch Loss: 0.020940
Epoch Loss: 0.020878
Epoch Loss: 0.020819
Epoch Loss: 0.020751
Epoch Loss: 0.020688
Epoch Loss: 0.020633
Epoch Loss: 0.020566
Epoch Loss: 0.020504
Epoch Loss: 0.020451
Epoch Loss: 0.020385
Epoch Loss: 0.020313
Epoch Loss: 0.020251
Epoch Loss: 0.020197
Epoch Loss: 0.020169
Epoch Loss: 0.020110
Epoch Loss: 0.020046
Epoch Loss: 0.019971
Epoch Loss: 0.019898
Epoch Loss: 0.019833
Epoch Loss: 0.019767
Epoch Loss: 0.019698
Epoch Loss: 0.019641
Epoch Loss: 0.019610
Epoch Loss: 0.019540
Epoch Loss: 0.019472
```

```
Epoch Loss: 0.019407
Epoch Loss: 0.019350
Epoch Loss: 0.019300
Epoch Loss: 0.019235
Epoch Loss: 0.019172
Epoch Loss: 0.019116
Epoch Loss: 0.019053
Epoch Loss: 0.018993
Epoch Loss: 0.018957
Epoch Loss: 0.018905
Epoch Loss: 0.018855
Epoch Loss: 0.018792
Epoch Loss: 0.018782
Epoch Loss: 0.018733
Epoch Loss: 0.018664
Epoch Loss: 0.018605
Epoch Loss: 0.018545
Epoch Loss: 0.018484
Epoch Loss: 0.018426
Epoch Loss: 0.018374
Epoch Loss: 0.018317
Epoch Loss: 0.018260
Epoch Loss: 0.018212
Epoch Loss: 0.018156
Epoch Loss: 0.018101
Epoch Loss: 0.018054
Epoch Loss: 0.018001
Epoch Loss: 0.017947
Epoch Loss: 0.017900
Epoch Loss: 0.017862
Epoch Loss: 0.017808
Epoch Loss: 0.017761
Epoch Loss: 0.017711
Epoch Loss: 0.017659
Epoch Loss: 0.017613
Epoch Loss: 0.017564
Epoch Loss: 0.017513
Epoch Loss: 0.017468
Epoch Loss: 0.017421
Epoch Loss: 0.017371
Epoch Loss: 0.017327
Epoch Loss: 0.017281
Epoch Loss: 0.017232
Epoch Loss: 0.017188
Epoch Loss: 0.017143
Epoch Loss: 0.017096
Epoch Loss: 0.017052
Epoch Loss: 0.017009
Epoch Loss: 0.016962
Epoch Loss: 0.016919
Epoch Loss: 0.016877
Epoch Loss: 0.016831
Epoch Loss: 0.016825
Epoch Loss: 0.016779
Epoch Loss: 0.016733
Epoch Loss: 0.016688
Epoch Loss: 0.016648
Epoch Loss: 0.016615
Epoch Loss: 0.016571
Epoch Loss: 0.016538
Epoch Loss: 0.016511
Epoch Loss: 0.016489
Epoch Loss: 0.016449
Epoch Loss: 0.016401
Epoch Loss: 0.016354
```

```
Epoch Loss: 0.016312
Epoch Loss: 0.016269
Epoch Loss: 0.016226
Epoch Loss: 0.016184
Epoch Loss: 0.016159
Epoch Loss: 0.016143
Epoch Loss: 0.016097
Epoch Loss: 0.016054
Epoch Loss: 0.016014
Epoch Loss: 0.015972
Epoch Loss: 0.015933
Epoch Loss: 0.015906
Epoch Loss: 0.015866
Epoch Loss: 0.015825
Epoch Loss: 0.015785
Epoch Loss: 0.015748
Epoch Loss: 0.015709
Epoch Loss: 0.015670
Epoch Loss: 0.015635
Epoch Loss: 0.015597
Epoch Loss: 0.015559
Epoch Loss: 0.015523
Epoch Loss: 0.015487
Epoch Loss: 0.015450
Epoch Loss: 0.015415
Epoch Loss: 0.015380
Epoch Loss: 0.015343
Epoch Loss: 0.015308
Epoch Loss: 0.015275
Epoch Loss: 0.015238
Epoch Loss: 0.015203
Epoch Loss: 0.015171
Epoch Loss: 0.015135
Epoch Loss: 0.015100
Epoch Loss: 0.015069
Epoch Loss: 0.015042
Epoch Loss: 0.015007
Epoch Loss: 0.014976
Epoch Loss: 0.014941
Epoch Loss: 0.014907
Epoch Loss: 0.014876
Epoch Loss: 0.014842
Epoch Loss: 0.014809
Epoch Loss: 0.014777
Epoch Loss: 0.014745
Epoch Loss: 0.014712
Epoch Loss: 0.014680
Epoch Loss: 0.014649
Epoch Loss: 0.014617
Epoch Loss: 0.014585
Epoch Loss: 0.014555
Epoch Loss: 0.014523
Epoch Loss: 0.014495
Epoch Loss: 0.014474
Epoch Loss: 0.014441
Epoch Loss: 0.014409
Epoch Loss: 0.014380
Epoch Loss: 0.014349
Epoch Loss: 0.014318
Epoch Loss: 0.014288
Epoch Loss: 0.014259
Epoch Loss: 0.014228
Epoch Loss: 0.014198
Epoch Loss: 0.014170
Epoch Loss: 0.014139
```

```
Epoch Loss: 0.014110
Epoch Loss: 0.014082
Epoch Loss: 0.014052
Epoch Loss: 0.014022
Epoch Loss: 0.014014
Epoch Loss: 0.014001
Epoch Loss: 0.013977
Epoch Loss: 0.013963
Epoch Loss: 0.013930
Epoch Loss: 0.013898
Epoch Loss: 0.013888
Epoch Loss: 0.013871
Epoch Loss: 0.013836
Epoch Loss: 0.013803
Epoch Loss: 0.013771
Epoch Loss: 0.013741
Epoch Loss: 0.013710
Epoch Loss: 0.013680
Epoch Loss: 0.013653
Epoch Loss: 0.013631
Epoch Loss: 0.013602
Epoch Loss: 0.013579
Epoch Loss: 0.013554
Epoch Loss: 0.013525
Epoch Loss: 0.013498
Epoch Loss: 0.013470
Epoch Loss: 0.013442
Epoch Loss: 0.013414
Epoch Loss: 0.013388
Epoch Loss: 0.013362
Epoch Loss: 0.013334
Epoch Loss: 0.013308
Epoch Loss: 0.013283
Epoch Loss: 0.013257
Epoch Loss: 0.013230
Epoch Loss: 0.013204
Epoch Loss: 0.013180
Epoch Loss: 0.013154
Epoch Loss: 0.013128
Epoch Loss: 0.013104
Epoch Loss: 0.013079
Epoch Loss: 0.013053
Epoch Loss: 0.013029
Epoch Loss: 0.013005
Epoch Loss: 0.012980
Epoch Loss: 0.012956
Epoch Loss: 0.012941
Epoch Loss: 0.012916
Epoch Loss: 0.012891
Epoch Loss: 0.012867
Epoch Loss: 0.012843
Epoch Loss: 0.012818
Epoch Loss: 0.012797
Epoch Loss: 0.012779
Epoch Loss: 0.012754
Epoch Loss: 0.012730
Epoch Loss: 0.012707
Epoch Loss: 0.012682
Epoch Loss: 0.012658
Epoch Loss: 0.012636
Epoch Loss: 0.012612
Epoch Loss: 0.012589
Epoch Loss: 0.012567
Epoch Loss: 0.012544
Epoch Loss: 0.012520
```

```
Epoch Loss: 0.012498
Epoch Loss: 0.012476
Epoch Loss: 0.012453
Epoch Loss: 0.012431
Epoch Loss: 0.012409
Epoch Loss: 0.012386
Epoch Loss: 0.012364
Epoch Loss: 0.012343
Epoch Loss: 0.012320
Epoch Loss: 0.012298
Epoch Loss: 0.012277
Epoch Loss: 0.012255
Epoch Loss: 0.012233
Epoch Loss: 0.012212
Epoch Loss: 0.012191
Epoch Loss: 0.012169
Epoch Loss: 0.012148
Epoch Loss: 0.012127
Epoch Loss: 0.012106
Epoch Loss: 0.012090
Epoch Loss: 0.012080
Epoch Loss: 0.012064
Epoch Loss: 0.012042
Epoch Loss: 0.012024
Epoch Loss: 0.012014
Epoch Loss: 0.011991
Epoch Loss: 0.011968
Epoch Loss: 0.011947
Epoch Loss: 0.011928
Epoch Loss: 0.011928
Epoch Loss: 0.011904
Epoch Loss: 0.011881
Epoch Loss: 0.011859
Epoch Loss: 0.011837
Epoch Loss: 0.011815
Epoch Loss: 0.011794
Epoch Loss: 0.011773
Epoch Loss: 0.011752
Epoch Loss: 0.011732
Epoch Loss: 0.011711
Epoch Loss: 0.011691
Epoch Loss: 0.011670
Epoch Loss: 0.011651
Epoch Loss: 0.011631
Epoch Loss: 0.011611
Epoch Loss: 0.011597
Epoch Loss: 0.011578
Epoch Loss: 0.011559
Epoch Loss: 0.011544
Epoch Loss: 0.011529
Epoch Loss: 0.011510
Epoch Loss: 0.011490
Epoch Loss: 0.011471
Epoch Loss: 0.011451
Epoch Loss: 0.011433
Epoch Loss: 0.011414
Epoch Loss: 0.011395
Epoch Loss: 0.011377
Epoch Loss: 0.011359
Epoch Loss: 0.011340
Epoch Loss: 0.011322
Epoch Loss: 0.011304
Epoch Loss: 0.011286
Epoch Loss: 0.011267
Epoch Loss: 0.011250
```

```
Epoch Loss: 0.011232
Epoch Loss: 0.011214
Epoch Loss: 0.011197
Epoch Loss: 0.011180
Epoch Loss: 0.011162
Epoch Loss: 0.011149
Epoch Loss: 0.011132
Epoch Loss: 0.011114
Epoch Loss: 0.011097
Epoch Loss: 0.011080
Epoch Loss: 0.011062
Epoch Loss: 0.011045
Epoch Loss: 0.011029
Epoch Loss: 0.011011
Epoch Loss: 0.010994
Epoch Loss: 0.010977
Epoch Loss: 0.010961
Epoch Loss: 0.010943
Epoch Loss: 0.010927
Epoch Loss: 0.010910
Epoch Loss: 0.010893
Epoch Loss: 0.010877
Epoch Loss: 0.010861
Epoch Loss: 0.010844
Epoch Loss: 0.010828
Epoch Loss: 0.010812
Epoch Loss: 0.010795
Epoch Loss: 0.010779
Epoch Loss: 0.010763
Epoch Loss: 0.010747
Epoch Loss: 0.010731
Epoch Loss: 0.010716
Epoch Loss: 0.010700
Epoch Loss: 0.010684
Epoch Loss: 0.010676
Epoch Loss: 0.010663
Epoch Loss: 0.010647
Epoch Loss: 0.010632
Epoch Loss: 0.010616
Epoch Loss: 0.010600
Epoch Loss: 0.010584
Epoch Loss: 0.010569
Epoch Loss: 0.010553
Epoch Loss: 0.010538
Epoch Loss: 0.010523
Epoch Loss: 0.010510
Epoch Loss: 0.010497
Epoch Loss: 0.010482
Epoch Loss: 0.010467
Epoch Loss: 0.010451
Epoch Loss: 0.010436
Epoch Loss: 0.010421
Epoch Loss: 0.010406
Epoch Loss: 0.010391
Epoch Loss: 0.010378
Epoch Loss: 0.010369
Epoch Loss: 0.010354
Epoch Loss: 0.010338
Epoch Loss: 0.010324
Epoch Loss: 0.010309
Epoch Loss: 0.010294
Epoch Loss: 0.010280
Epoch Loss: 0.010272
Epoch Loss: 0.010257
Epoch Loss: 0.010242
```

```
Epoch Loss: 0.010227
Epoch Loss: 0.010213
Epoch Loss: 0.010198
Epoch Loss: 0.010183
Epoch Loss: 0.010169
Epoch Loss: 0.010159
Epoch Loss: 0.010145
Epoch Loss: 0.010131
Epoch Loss: 0.010116
Epoch Loss: 0.010102
Epoch Loss: 0.010088
Epoch Loss: 0.010080
Epoch Loss: 0.010072
Epoch Loss: 0.010057
Epoch Loss: 0.010043
Epoch Loss: 0.010028
Epoch Loss: 0.010013
Epoch Loss: 0.009999
Epoch Loss: 0.009985
Epoch Loss: 0.009971
Epoch Loss: 0.009957
Epoch Loss: 0.009943
Epoch Loss: 0.009929
Epoch Loss: 0.009915
Epoch Loss: 0.009902
Epoch Loss: 0.009888
Epoch Loss: 0.009874
Epoch Loss: 0.009861
Epoch Loss: 0.009847
Epoch Loss: 0.009834
Epoch Loss: 0.009821
Epoch Loss: 0.009807
Epoch Loss: 0.009794
Epoch Loss: 0.009781
Epoch Loss: 0.009768
Epoch Loss: 0.009755
Epoch Loss: 0.009741
Epoch Loss: 0.009728
Epoch Loss: 0.009716
Epoch Loss: 0.009702
Epoch Loss: 0.009690
Epoch Loss: 0.009681
Epoch Loss: 0.009668
Epoch Loss: 0.009654
Epoch Loss: 0.009642
Epoch Loss: 0.009629
Epoch Loss: 0.009616
Epoch Loss: 0.009603
Epoch Loss: 0.009591
Epoch Loss: 0.009578
Epoch Loss: 0.009565
Epoch Loss: 0.009553
Epoch Loss: 0.009540
Epoch Loss: 0.009528
Epoch Loss: 0.009516
Epoch Loss: 0.009503
Epoch Loss: 0.009491
Epoch Loss: 0.009478
Epoch Loss: 0.009468
Epoch Loss: 0.009462
Epoch Loss: 0.009450
Epoch Loss: 0.009437
Epoch Loss: 0.009425
Epoch Loss: 0.009412
Epoch Loss: 0.009400
```

```
Epoch Loss: 0.009388
Epoch Loss: 0.009375
Epoch Loss: 0.009363
Epoch Loss: 0.009351
Epoch Loss: 0.009340
Epoch Loss: 0.009331
Epoch Loss: 0.009319
Epoch Loss: 0.009307
Epoch Loss: 0.009295
Epoch Loss: 0.009289
Epoch Loss: 0.009277
Epoch Loss: 0.009264
Epoch Loss: 0.009252
Epoch Loss: 0.009240
Epoch Loss: 0.009228
Epoch Loss: 0.009216
Epoch Loss: 0.009205
Epoch Loss: 0.009193
Epoch Loss: 0.009181
Epoch Loss: 0.009169
Epoch Loss: 0.009158
Epoch Loss: 0.009146
Epoch Loss: 0.009134
Epoch Loss: 0.009123
Epoch Loss: 0.009111
Epoch Loss: 0.009100
Epoch Loss: 0.009089
Epoch Loss: 0.009077
Epoch Loss: 0.009066
Epoch Loss: 0.009055
Epoch Loss: 0.009044
Epoch Loss: 0.009032
Epoch Loss: 0.009021
Epoch Loss: 0.009011
Epoch Loss: 0.009004
Epoch Loss: 0.008992
Epoch Loss: 0.008981
Epoch Loss: 0.008970
Epoch Loss: 0.008959
Epoch Loss: 0.008948
Epoch Loss: 0.008937
Epoch Loss: 0.008926
Epoch Loss: 0.008914
Epoch Loss: 0.008904
Epoch Loss: 0.008895
Epoch Loss: 0.008885
Epoch Loss: 0.008874
Epoch Loss: 0.008863
Epoch Loss: 0.008852
Epoch Loss: 0.008841
Epoch Loss: 0.008831
Epoch Loss: 0.008820
Epoch Loss: 0.008809
Epoch Loss: 0.008799
Epoch Loss: 0.008788
Epoch Loss: 0.008777
Epoch Loss: 0.008767
Epoch Loss: 0.008756
Epoch Loss: 0.008745
Epoch Loss: 0.008735
Epoch Loss: 0.008725
Epoch Loss: 0.008714
Epoch Loss: 0.008704
Epoch Loss: 0.008694
Epoch Loss: 0.008683
```

```
Epoch Loss: 0.008673
Epoch Loss: 0.008663
Epoch Loss: 0.008652
Epoch Loss: 0.008642
Epoch Loss: 0.008632
Epoch Loss: 0.008622
Epoch Loss: 0.008611
Epoch Loss: 0.008601
Epoch Loss: 0.008591
Epoch Loss: 0.008581
Epoch Loss: 0.008571
Epoch Loss: 0.008564
Epoch Loss: 0.008553
Epoch Loss: 0.008543
Epoch Loss: 0.008534
Epoch Loss: 0.008523
Epoch Loss: 0.008513
Epoch Loss: 0.008504
Epoch Loss: 0.008494
Epoch Loss: 0.008483
Epoch Loss: 0.008474
Epoch Loss: 0.008464
Epoch Loss: 0.008454
Epoch Loss: 0.008448
Epoch Loss: 0.008441
Epoch Loss: 0.008431
Epoch Loss: 0.008422
Epoch Loss: 0.008413
Epoch Loss: 0.008403
Epoch Loss: 0.008393
Epoch Loss: 0.008383
Epoch Loss: 0.008373
Epoch Loss: 0.008368
Epoch Loss: 0.008364
Epoch Loss: 0.008358
Epoch Loss: 0.008348
Epoch Loss: 0.008337
Epoch Loss: 0.008327
Epoch Loss: 0.008317
Epoch Loss: 0.008307
Epoch Loss: 0.008297
Epoch Loss: 0.008287
Epoch Loss: 0.008277
Epoch Loss: 0.008267
Epoch Loss: 0.008258
Epoch Loss: 0.008248
Epoch Loss: 0.008238
Epoch Loss: 0.008229
Epoch Loss: 0.008219
Epoch Loss: 0.008210
Epoch Loss: 0.008201
Epoch Loss: 0.008194
Epoch Loss: 0.008185
Epoch Loss: 0.008175
Epoch Loss: 0.008166
Epoch Loss: 0.008157
Epoch Loss: 0.008147
Epoch Loss: 0.008138
Epoch Loss: 0.008129
Epoch Loss: 0.008120
Epoch Loss: 0.008110
Epoch Loss: 0.008102
Epoch Loss: 0.008092
Epoch Loss: 0.008083
Epoch Loss: 0.008074
```

```
Epoch Loss: 0.008065
Epoch Loss: 0.008056
Epoch Loss: 0.008047
Epoch Loss: 0.008039
Epoch Loss: 0.008030
Epoch Loss: 0.008021
Epoch Loss: 0.008012
Epoch Loss: 0.008003
Epoch Loss: 0.007994
Epoch Loss: 0.007985
Epoch Loss: 0.007977
Epoch Loss: 0.007968
Epoch Loss: 0.007959
Epoch Loss: 0.007951
Epoch Loss: 0.007943
Epoch Loss: 0.007935
Epoch Loss: 0.007927
Epoch Loss: 0.007918
Epoch Loss: 0.007909
Epoch Loss: 0.007901
Epoch Loss: 0.007892
Epoch Loss: 0.007884
Epoch Loss: 0.007875
Epoch Loss: 0.007867
Epoch Loss: 0.007858
Epoch Loss: 0.007850
Epoch Loss: 0.007844
Epoch Loss: 0.007836
Epoch Loss: 0.007828
Epoch Loss: 0.007820
Epoch Loss: 0.007811
Epoch Loss: 0.007802
Epoch Loss: 0.007794
Epoch Loss: 0.007786
Epoch Loss: 0.007777
Epoch Loss: 0.007769
Epoch Loss: 0.007761
Epoch Loss: 0.007752
Epoch Loss: 0.007744
Epoch Loss: 0.007736
Epoch Loss: 0.007728
Epoch Loss: 0.007719
Epoch Loss: 0.007711
Epoch Loss: 0.007703
Epoch Loss: 0.007695
Epoch Loss: 0.007687
Epoch Loss: 0.007679
Epoch Loss: 0.007671
Epoch Loss: 0.007662
Epoch Loss: 0.007655
Epoch Loss: 0.007646
Epoch Loss: 0.007638
Epoch Loss: 0.007630
Epoch Loss: 0.007622
Epoch Loss: 0.007614
Epoch Loss: 0.007606
Epoch Loss: 0.007601
Epoch Loss: 0.007593
```

55

```
# Test your predictions. The predictions must match the labels
print(net.predict(X))
print(y)
```

```
[2 1 0 1 2 0 0 2 0 0]
```

```
[2 1 0 1 2 0 0 2 0 0]
```

```
# You should be able to achieve a training loss of less than 0.02 (10%)
print("Final training loss", epoch_loss[-1])
```

56
```
# Plot the training loss curve. The loss in the curve should be decreasing (20%)
plt.plot(epoch_loss)
plt.title('Loss history')
plt.xlabel('Iteration')
plt.ylabel('Loss')
```

56
```
Text(0, 0.5, 'Loss')
```