

DATA_598_HW_1

January 14, 2022

1

Homework 1: Multilayer Perceptrons and Stochastic Gradient Descent

Apoorv Sharma

<center> DATA 598 (Winter 2022), University of Washington </center>

In this exercise, we will experiment with the number of hidden units in a multilayer perceptron (MLP) with a single hidden layer. The number of hidden units is also referred to the width of the hidden layer. Here are the details: * The setup is identical to the demo/lab and you may reuse that code. Take the FashionMNIST dataset and randomly subsample 10% of its training set to work with. As a test set, we will use the full test set of FashionMNIST. * Define a MLP with one hidden layer of width $h = 32$. Find the divergent learning rate for this model and use a fixed learning rate of $\frac{1}{2}$, as we discussed in class. * Train the model for 100 passes over the data. * Repeat this procedure for widths $h = 8, 16, 128, 512$ with the same learning rate $\frac{1}{2}$ as above (i.e., you do not need to find the divergent learning rate of each model for this exercise).

```
[60]: import numpy as np
import torch
from torchvision.datasets import FashionMNIST
from torch.nn.functional import cross_entropy
import time
import copy
import os

from pprint import pprint as pp

import matplotlib.pyplot as plt
%matplotlib inline

torch.manual_seed(0)
np.random.seed(1)
```

```
[2]: MLP_Q1_RESULTS_FILE = 'models/hw1_q1_fashion_mlp_summary.pt'
MLP_Q2_RESULTS_FILE = 'models/hw1_q2_fashion_mlp_summary.pt'
```

Download the FashionMNIST dataset and subsample 10% of it

```
[3]: # download dataset (~117M in size)
train_dataset = FashionMNIST('./data/train', train=True, download=True)
X_train = train_dataset.data # torch tensor of type uint8
y_train = train_dataset.targets # torch tensor of type Long

test_dataset = FashionMNIST('./data/test', train=False, download=True)
X_test = test_dataset.data
y_test = test_dataset.targets

# choose a subsample of 10% of the data:
idxs_train = torch.from_numpy(
    np.random.choice(X_train.shape[0], replace=False, size=X_train.shape[0]//
↳10))
X_train, y_train = X_train[idxs_train], y_train[idxs_train]

print(f'X_train.shape = {X_train.shape}')
print(f'n_train: {X_train.shape[0]}, n_test: {X_test.shape[0]}')
print(f'Image size: {X_train.shape[1:]}')
```

```
X_train.shape = torch.Size([6000, 28, 28])
n_train: 6000, n_test: 10000
Image size: torch.Size([28, 28])
```

```
[4]: # Normalize dataset: pixel values lie between 0 and 255
# Normalize them so the pixelwise mean is zero and standard deviation is 1

X_train = X_train.float() # convert to float32
X_train = X_train.view(-1, 784) # flatten into a (n, d) shape
mean, std = X_train.mean(axis=0), X_train.std(axis=0)
X_train = (X_train - mean[None, :]) / (std[None, :] + 1e-6) # avoid divide by 0
↳zero

X_test = X_test.float()
X_test = X_test.view(-1, 784)
X_test = (X_test - mean[None, :]) / (std[None, :] + 1e-6)

n_class = np.unique(y_train).shape[0]
```

```
[5]: # Some utility functions to compute the objective and the accuracy

def mlp(X, ws, bs):
    hidden = X # Shape: (n, d_0)
    for w, b in zip(ws[:-1], bs[:-1]):
        hidden = torch.matmul(hidden, w) + b[None, :] # Shape: (n, d_{j-1}) * 1
↳(d_{j-1}, d_j) = (n, d_j)
        hidden = torch.nn.functional.relu(hidden)
    return torch.matmul(hidden, ws[-1]) + bs[-1][None, :]
```

```

def compute_objective(ws, bs, X, y, reg_param):
    """ Compute the multinomial logistic loss.
        ws is a list of tensors of consistent shapes,
        X of shape (n, d) and y of shape (n,)
    """
    score = mlp(X, ws, bs) # Shape: (n, K)
    # PyTorch's function cross_entropy computes the multinomial logistic loss
    return (
        cross_entropy(input=score, target=y, reduction='mean')
        + 0.5 * reg_param * sum([torch.norm(w)**2 for w in ws])
    )

@torch.no_grad()
def compute_accuracy(ws, bs, X, y):
    """ Compute the classification accuracy
        ws is a list of tensors of consistent shapes
        X of shape (n, d) and y of shape (n,)
    """
    score = mlp(X, ws, bs) # shape: (n, K)
    predictions = torch.argmax(score, axis=1) # class with highest score is
    ↪ predicted
    return (predictions == y).sum() * 1.0 / y.shape[0]

@torch.no_grad()
def compute_logs(ws, bs, reg_param, verbose=False):
    train_loss = compute_objective(ws, bs, X_train, y_train, reg_param)
    test_loss = compute_objective(ws, bs, X_test, y_test, reg_param)
    train_accuracy = compute_accuracy(ws, bs, X_train, y_train)
    test_accuracy = compute_accuracy(ws, bs, X_test, y_test)
    if verbose:
        print(('Train Loss = {:.3f}, Train Accuracy = {:.3f}, ' +
              'Test Loss = {:.3f}, Test Accuracy = {:.3f}').format(
                train_loss.item(), train_accuracy.item(),
                test_loss.item(), test_accuracy.item()))
    )
    return (train_loss, train_accuracy, test_loss, test_accuracy)

```

```

[6]: def initialize_ws(hidden_size, n_class):
    ws = [1e-6 * torch.randn(784, hidden_size, requires_grad=True),
          1e-6 * torch.randn(hidden_size, n_class, requires_grad=True)]
    return ws

def initialize_bs(hidden_size, n_class):
    bs = [torch.zeros(hidden_size, requires_grad=True),
          torch.zeros(n_class, requires_grad=True)]
    return bs

```

```
[7]: def sgd_one_pass(ws, bs, X, y, reg_param, learning_rate, verbose=False):
    num_examples = X.shape[0]
    average_loss = 0.0
    for i in range(num_examples):
        idx = np.random.choice(X.shape[0])
        # compute the objective.
        # Note: This function requires X to be of shape (n,d). In this case,
        ↪n=1
        objective = compute_objective(ws, bs, X[idx:idx+1], y[idx:idx+1],
        ↪reg_param)
        average_loss = 0.99 * average_loss + 0.01 * objective.item()
        if verbose and (i+1) % 100 == 0:
            print(average_loss)

        # compute the gradient using automatic differentiation
        all_parameters = [*ws, *bs]
        gradients = torch.autograd.grad(outputs=objective,
        ↪inputs=all_parameters)

        # perform SGD update. IMPORTANT: Make the update inplace!
        with torch.no_grad():
            for (w, g) in zip(all_parameters, gradients):
                w -= learning_rate * g
    return ws, bs
```

```
[8]: # reg_param = 0.1 / X_train.shape[0]
    reg_param = 0
```

Find the divergent learning rate

```
[9]: learning_rate = 0.01 / 2

hidden_size = 32
ws = initialize_ws(hidden_size, n_class)
bs = initialize_bs(hidden_size, n_class)

_ = compute_logs(ws, bs, reg_param, verbose=True)

ws, bs = sgd_one_pass(ws, bs, X_train, y_train, reg_param, learning_rate,
    ↪verbose=True)
_ = compute_logs(ws, bs, reg_param, verbose=True)
```

Train Loss = 2.303, Train Accuracy = 0.094, Test Loss = 2.303, Test Accuracy = 0.096
 1.4596734945436483
 1.9945874147627942
 2.1907702811847978

2.262258175462675
2.288168216893089
2.2950623880166683
2.298219839564293
2.302896123190581
2.30257610207377
2.2915321294114945
2.280852721102953
2.126430165573341
1.9287736125690487
1.8493492605486324
1.7123184657147137
1.5586019151842219
1.4214616694003144
1.390881504237791
1.3405923644022517
1.1929667180198427
1.1020713150444366
1.070320012709279
0.9729690050926664
0.9245819557951745
0.8588912656283177
0.9412617632724704
0.9733787434576562
0.901405160316614
0.9182102662294258
0.7444109468615022
0.8170800067544532
0.7932568922523444
0.8063622863686092
0.741420326250601
0.7817283040031301
0.6811794254503886
0.694335086149683
0.6539764663752585
0.7827460163532026
0.734852175898242
0.5769021581364675
0.609761253720165
0.574642390944477
0.6409038684058812
0.5790878158547016
0.6099915783967417
0.6605070878151497
0.6063592603720036
0.651159947928013
0.6548265432749135
0.6603105942833384

```
0.5986338487307346
0.5504842544433707
0.5827600440769217
0.5459764514576187
0.495702489723848
0.5256435893719871
0.46258766276508123
0.5573034576365449
0.5060381139626582
Train Loss = 0.594, Train Accuracy = 0.770, Test Loss = 0.639, Test Accuracy = 0.758
```

Run the data for 100 passes over the data

```
[10]: logs = []
      num_passes = 100

      if not os.path.exists(MLP_Q1_RESULTS_FILE):
          ws = initialize_ws(hidden_size, n_class)
          bs = initialize_bs(hidden_size, n_class)
          logs.append(compute_logs(ws, bs, reg_param, verbose=True))

      for j in range(num_passes):
          print(j+1, end=', ')
          ws, bs = sgd_one_pass(ws, bs, X_train, y_train, reg_param,
          ↪learning_rate, verbose=False)
          logs.append(compute_logs(ws, bs, reg_param, verbose=True))
```

```
Train Loss = 2.303, Train Accuracy = 0.052, Test Loss = 2.303, Test Accuracy = 0.051
1, Train Loss = 0.580, Train Accuracy = 0.787, Test Loss = 0.615, Test Accuracy = 0.776
2, Train Loss = 0.423, Train Accuracy = 0.848, Test Loss = 0.532, Test Accuracy = 0.820
3, Train Loss = 0.412, Train Accuracy = 0.849, Test Loss = 0.578, Test Accuracy = 0.808
4, Train Loss = 0.417, Train Accuracy = 0.852, Test Loss = 0.623, Test Accuracy = 0.795
5, Train Loss = 0.354, Train Accuracy = 0.869, Test Loss = 0.618, Test Accuracy = 0.800
6, Train Loss = 0.311, Train Accuracy = 0.887, Test Loss = 0.610, Test Accuracy = 0.816
7, Train Loss = 0.288, Train Accuracy = 0.901, Test Loss = 0.585, Test Accuracy = 0.819
8, Train Loss = 0.265, Train Accuracy = 0.909, Test Loss = 0.611, Test Accuracy = 0.809
9, Train Loss = 0.268, Train Accuracy = 0.901, Test Loss = 0.663, Test Accuracy = 0.810
```

10, Train Loss = 0.207, Train Accuracy = 0.923, Test Loss = 0.629, Test Accuracy = 0.828
11, Train Loss = 0.212, Train Accuracy = 0.923, Test Loss = 0.675, Test Accuracy = 0.820
12, Train Loss = 0.229, Train Accuracy = 0.915, Test Loss = 0.714, Test Accuracy = 0.816
13, Train Loss = 0.207, Train Accuracy = 0.922, Test Loss = 0.754, Test Accuracy = 0.809
14, Train Loss = 0.153, Train Accuracy = 0.947, Test Loss = 0.701, Test Accuracy = 0.828
15, Train Loss = 0.141, Train Accuracy = 0.954, Test Loss = 0.722, Test Accuracy = 0.834
16, Train Loss = 0.138, Train Accuracy = 0.950, Test Loss = 0.758, Test Accuracy = 0.825
17, Train Loss = 0.138, Train Accuracy = 0.949, Test Loss = 0.803, Test Accuracy = 0.821
18, Train Loss = 0.160, Train Accuracy = 0.938, Test Loss = 0.873, Test Accuracy = 0.811
19, Train Loss = 0.296, Train Accuracy = 0.918, Test Loss = 1.058, Test Accuracy = 0.791
20, Train Loss = 0.107, Train Accuracy = 0.959, Test Loss = 0.891, Test Accuracy = 0.818
21, Train Loss = 0.121, Train Accuracy = 0.954, Test Loss = 0.971, Test Accuracy = 0.808
22, Train Loss = 0.090, Train Accuracy = 0.968, Test Loss = 0.943, Test Accuracy = 0.818
23, Train Loss = 0.138, Train Accuracy = 0.950, Test Loss = 1.027, Test Accuracy = 0.807
24, Train Loss = 0.114, Train Accuracy = 0.958, Test Loss = 1.025, Test Accuracy = 0.817
25, Train Loss = 0.076, Train Accuracy = 0.972, Test Loss = 1.014, Test Accuracy = 0.820
26, Train Loss = 0.078, Train Accuracy = 0.973, Test Loss = 1.023, Test Accuracy = 0.818
27, Train Loss = 0.057, Train Accuracy = 0.980, Test Loss = 1.052, Test Accuracy = 0.820
28, Train Loss = 0.081, Train Accuracy = 0.971, Test Loss = 1.088, Test Accuracy = 0.814
29, Train Loss = 0.078, Train Accuracy = 0.971, Test Loss = 1.156, Test Accuracy = 0.815
30, Train Loss = 0.044, Train Accuracy = 0.985, Test Loss = 1.103, Test Accuracy = 0.824
31, Train Loss = 0.125, Train Accuracy = 0.960, Test Loss = 1.245, Test Accuracy = 0.812
32, Train Loss = 0.096, Train Accuracy = 0.969, Test Loss = 1.233, Test Accuracy = 0.808
33, Train Loss = 0.050, Train Accuracy = 0.982, Test Loss = 1.192, Test Accuracy = 0.823

34, Train Loss = 0.038, Train Accuracy = 0.987, Test Loss = 1.191, Test Accuracy = 0.820
35, Train Loss = 0.063, Train Accuracy = 0.977, Test Loss = 1.235, Test Accuracy = 0.816
36, Train Loss = 0.058, Train Accuracy = 0.980, Test Loss = 1.263, Test Accuracy = 0.811
37, Train Loss = 0.019, Train Accuracy = 0.995, Test Loss = 1.185, Test Accuracy = 0.828
38, Train Loss = 0.018, Train Accuracy = 0.996, Test Loss = 1.242, Test Accuracy = 0.825
39, Train Loss = 0.021, Train Accuracy = 0.993, Test Loss = 1.289, Test Accuracy = 0.823
40, Train Loss = 0.023, Train Accuracy = 0.993, Test Loss = 1.280, Test Accuracy = 0.824
41, Train Loss = 0.011, Train Accuracy = 0.998, Test Loss = 1.279, Test Accuracy = 0.828
42, Train Loss = 0.011, Train Accuracy = 0.998, Test Loss = 1.291, Test Accuracy = 0.827
43, Train Loss = 0.009, Train Accuracy = 0.998, Test Loss = 1.314, Test Accuracy = 0.826
44, Train Loss = 0.006, Train Accuracy = 0.999, Test Loss = 1.306, Test Accuracy = 0.827
45, Train Loss = 0.013, Train Accuracy = 0.996, Test Loss = 1.356, Test Accuracy = 0.824
46, Train Loss = 0.005, Train Accuracy = 0.999, Test Loss = 1.335, Test Accuracy = 0.829
47, Train Loss = 0.008, Train Accuracy = 0.998, Test Loss = 1.365, Test Accuracy = 0.827
48, Train Loss = 0.012, Train Accuracy = 0.997, Test Loss = 1.396, Test Accuracy = 0.825
49, Train Loss = 0.007, Train Accuracy = 0.998, Test Loss = 1.380, Test Accuracy = 0.827
50, Train Loss = 0.005, Train Accuracy = 0.999, Test Loss = 1.397, Test Accuracy = 0.826
51, Train Loss = 0.003, Train Accuracy = 1.000, Test Loss = 1.393, Test Accuracy = 0.824
52, Train Loss = 0.003, Train Accuracy = 1.000, Test Loss = 1.402, Test Accuracy = 0.827
53, Train Loss = 0.003, Train Accuracy = 1.000, Test Loss = 1.413, Test Accuracy = 0.827
54, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.414, Test Accuracy = 0.828
55, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.426, Test Accuracy = 0.827
56, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.429, Test Accuracy = 0.829
57, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.438, Test Accuracy = 0.828

58, Train Loss = 0.003, Train Accuracy = 1.000, Test Loss = 1.454, Test Accuracy = 0.826
59, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.449, Test Accuracy = 0.828
60, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.457, Test Accuracy = 0.827
61, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.458, Test Accuracy = 0.827
62, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.464, Test Accuracy = 0.827
63, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.471, Test Accuracy = 0.827
64, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.476, Test Accuracy = 0.828
65, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.480, Test Accuracy = 0.829
66, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.483, Test Accuracy = 0.828
67, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.489, Test Accuracy = 0.828
68, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.494, Test Accuracy = 0.828
69, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.499, Test Accuracy = 0.828
70, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.507, Test Accuracy = 0.827
71, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.513, Test Accuracy = 0.829
72, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.512, Test Accuracy = 0.828
73, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.515, Test Accuracy = 0.828
74, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.521, Test Accuracy = 0.828
75, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.523, Test Accuracy = 0.829
76, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.525, Test Accuracy = 0.829
77, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.529, Test Accuracy = 0.828
78, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.535, Test Accuracy = 0.828
79, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.536, Test Accuracy = 0.829
80, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.541, Test Accuracy = 0.828
81, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.543, Test Accuracy = 0.829

```

82, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.547, Test Accuracy
= 0.829
83, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.551, Test Accuracy
= 0.828
84, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.552, Test Accuracy
= 0.828
85, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.557, Test Accuracy
= 0.829
86, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.557, Test Accuracy
= 0.829
87, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.561, Test Accuracy
= 0.829
88, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.563, Test Accuracy
= 0.829
89, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.568, Test Accuracy
= 0.828
90, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.570, Test Accuracy
= 0.828
91, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.572, Test Accuracy
= 0.828
92, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.576, Test Accuracy
= 0.829
93, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.578, Test Accuracy
= 0.828
94, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.582, Test Accuracy
= 0.828
95, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.583, Test Accuracy
= 0.828
96, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.587, Test Accuracy
= 0.829
97, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.590, Test Accuracy
= 0.828
98, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.593, Test Accuracy
= 0.828
99, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.594, Test Accuracy
= 0.828
100, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.595, Test
Accuracy = 0.828

```

1.1 1. Performance vs. Width

```

[11]: num_passes = 100
      hidden_sizes = [8, 16, 32, 128, 512]

      if not os.path.exists(MLP_Q1_RESULTS_FILE):
          hidden_size_logs = {}
          for hidden_size in hidden_sizes:

```

```

print(f'Hidden Size: {hidden_size}')
hidden_size_logs[hidden_size] = []

ws = initialize_ws(hidden_size, n_class)
bs = initialize_bs(hidden_size, n_class)
hidden_size_logs[hidden_size].append(compute_logs(ws, bs, reg_param,
↪verbose=True))

for j in range(num_passes):
    print(f'\t{j+1}', end=', ')
    ws, bs = sgd_one_pass(ws, bs, X_train, y_train, reg_param,
↪learning_rate, verbose=False)
    hidden_size_logs[hidden_size].append(compute_logs(ws, bs,
↪reg_param, verbose=True))

print()

```

Hidden Size: 8

Train Loss = 2.303, Train Accuracy = 0.098, Test Loss = 2.303, Test Accuracy = 0.099

1, Train Loss = 0.637, Train Accuracy = 0.763, Test Loss = 0.673, Test Accuracy = 0.754

2, Train Loss = 0.491, Train Accuracy = 0.822, Test Loss = 0.593, Test Accuracy = 0.791

3, Train Loss = 0.469, Train Accuracy = 0.830, Test Loss = 0.580, Test Accuracy = 0.799

4, Train Loss = 0.507, Train Accuracy = 0.808, Test Loss = 0.664, Test Accuracy = 0.770

5, Train Loss = 0.388, Train Accuracy = 0.862, Test Loss = 0.592, Test Accuracy = 0.810

6, Train Loss = 0.456, Train Accuracy = 0.830, Test Loss = 0.667, Test Accuracy = 0.783

7, Train Loss = 0.360, Train Accuracy = 0.864, Test Loss = 0.607, Test Accuracy = 0.805

8, Train Loss = 0.377, Train Accuracy = 0.862, Test Loss = 0.664, Test Accuracy = 0.786

9, Train Loss = 0.380, Train Accuracy = 0.869, Test Loss = 0.700, Test Accuracy = 0.797

10, Train Loss = 0.353, Train Accuracy = 0.869, Test Loss = 0.680, Test Accuracy = 0.790

11, Train Loss = 0.308, Train Accuracy = 0.892, Test Loss = 0.638, Test Accuracy = 0.810

12, Train Loss = 0.287, Train Accuracy = 0.894, Test Loss = 0.667, Test Accuracy = 0.801

13, Train Loss = 0.339, Train Accuracy = 0.877, Test Loss = 0.753, Test Accuracy = 0.794

14, Train Loss = 0.286, Train Accuracy = 0.894, Test Loss = 0.695, Test

Accuracy = 0.813
 15, Train Loss = 0.300, Train Accuracy = 0.897, Test Loss = 0.737, Test Accuracy = 0.799
 16, Train Loss = 0.272, Train Accuracy = 0.897, Test Loss = 0.732, Test Accuracy = 0.802
 17, Train Loss = 0.300, Train Accuracy = 0.887, Test Loss = 0.782, Test Accuracy = 0.796
 18, Train Loss = 0.307, Train Accuracy = 0.884, Test Loss = 0.793, Test Accuracy = 0.789
 19, Train Loss = 0.254, Train Accuracy = 0.900, Test Loss = 0.782, Test Accuracy = 0.800
 20, Train Loss = 0.239, Train Accuracy = 0.911, Test Loss = 0.791, Test Accuracy = 0.810
 21, Train Loss = 0.260, Train Accuracy = 0.903, Test Loss = 0.834, Test Accuracy = 0.795
 22, Train Loss = 0.228, Train Accuracy = 0.914, Test Loss = 0.812, Test Accuracy = 0.799
 23, Train Loss = 0.214, Train Accuracy = 0.920, Test Loss = 0.788, Test Accuracy = 0.811
 24, Train Loss = 0.206, Train Accuracy = 0.924, Test Loss = 0.819, Test Accuracy = 0.805
 25, Train Loss = 0.250, Train Accuracy = 0.910, Test Loss = 0.883, Test Accuracy = 0.791
 26, Train Loss = 0.248, Train Accuracy = 0.911, Test Loss = 0.906, Test Accuracy = 0.795
 27, Train Loss = 0.227, Train Accuracy = 0.914, Test Loss = 0.882, Test Accuracy = 0.797
 28, Train Loss = 0.248, Train Accuracy = 0.903, Test Loss = 0.911, Test Accuracy = 0.790
 29, Train Loss = 0.207, Train Accuracy = 0.919, Test Loss = 0.917, Test Accuracy = 0.790
 30, Train Loss = 0.213, Train Accuracy = 0.920, Test Loss = 0.947, Test Accuracy = 0.797
 31, Train Loss = 0.201, Train Accuracy = 0.922, Test Loss = 0.945, Test Accuracy = 0.798
 32, Train Loss = 0.268, Train Accuracy = 0.902, Test Loss = 0.998, Test Accuracy = 0.777
 33, Train Loss = 0.209, Train Accuracy = 0.919, Test Loss = 0.957, Test Accuracy = 0.788
 34, Train Loss = 0.193, Train Accuracy = 0.928, Test Loss = 0.980, Test Accuracy = 0.795
 35, Train Loss = 0.209, Train Accuracy = 0.924, Test Loss = 0.977, Test Accuracy = 0.797
 36, Train Loss = 0.179, Train Accuracy = 0.938, Test Loss = 0.974, Test Accuracy = 0.792
 37, Train Loss = 0.171, Train Accuracy = 0.939, Test Loss = 1.024, Test Accuracy = 0.795
 38, Train Loss = 0.237, Train Accuracy = 0.918, Test Loss = 1.110, Test

Accuracy = 0.791
 39, Train Loss = 0.228, Train Accuracy = 0.917, Test Loss = 1.095, Test Accuracy = 0.793
 40, Train Loss = 0.200, Train Accuracy = 0.928, Test Loss = 1.063, Test Accuracy = 0.794
 41, Train Loss = 0.211, Train Accuracy = 0.920, Test Loss = 1.119, Test Accuracy = 0.785
 42, Train Loss = 0.253, Train Accuracy = 0.904, Test Loss = 1.165, Test Accuracy = 0.767
 43, Train Loss = 0.160, Train Accuracy = 0.939, Test Loss = 1.072, Test Accuracy = 0.793
 44, Train Loss = 0.204, Train Accuracy = 0.922, Test Loss = 1.134, Test Accuracy = 0.789
 45, Train Loss = 0.148, Train Accuracy = 0.945, Test Loss = 1.122, Test Accuracy = 0.797
 46, Train Loss = 0.160, Train Accuracy = 0.939, Test Loss = 1.207, Test Accuracy = 0.792
 47, Train Loss = 0.190, Train Accuracy = 0.930, Test Loss = 1.218, Test Accuracy = 0.789
 48, Train Loss = 0.203, Train Accuracy = 0.928, Test Loss = 1.232, Test Accuracy = 0.792
 49, Train Loss = 0.150, Train Accuracy = 0.943, Test Loss = 1.189, Test Accuracy = 0.786
 50, Train Loss = 0.178, Train Accuracy = 0.929, Test Loss = 1.231, Test Accuracy = 0.787
 51, Train Loss = 0.145, Train Accuracy = 0.945, Test Loss = 1.212, Test Accuracy = 0.797
 52, Train Loss = 0.152, Train Accuracy = 0.943, Test Loss = 1.256, Test Accuracy = 0.786
 53, Train Loss = 0.162, Train Accuracy = 0.944, Test Loss = 1.269, Test Accuracy = 0.786
 54, Train Loss = 0.148, Train Accuracy = 0.943, Test Loss = 1.280, Test Accuracy = 0.787
 55, Train Loss = 0.161, Train Accuracy = 0.945, Test Loss = 1.325, Test Accuracy = 0.792
 56, Train Loss = 0.156, Train Accuracy = 0.941, Test Loss = 1.317, Test Accuracy = 0.784
 57, Train Loss = 0.180, Train Accuracy = 0.929, Test Loss = 1.374, Test Accuracy = 0.779
 58, Train Loss = 0.147, Train Accuracy = 0.947, Test Loss = 1.373, Test Accuracy = 0.784
 59, Train Loss = 0.207, Train Accuracy = 0.927, Test Loss = 1.446, Test Accuracy = 0.781
 60, Train Loss = 0.167, Train Accuracy = 0.939, Test Loss = 1.407, Test Accuracy = 0.781
 61, Train Loss = 0.148, Train Accuracy = 0.945, Test Loss = 1.383, Test Accuracy = 0.793
 62, Train Loss = 0.115, Train Accuracy = 0.960, Test Loss = 1.381, Test

Accuracy = 0.786
 63, Train Loss = 0.115, Train Accuracy = 0.956, Test Loss = 1.401, Test Accuracy = 0.790
 64, Train Loss = 0.151, Train Accuracy = 0.945, Test Loss = 1.472, Test Accuracy = 0.785
 65, Train Loss = 0.118, Train Accuracy = 0.956, Test Loss = 1.442, Test Accuracy = 0.784
 66, Train Loss = 0.215, Train Accuracy = 0.936, Test Loss = 1.753, Test Accuracy = 0.774
 67, Train Loss = 0.189, Train Accuracy = 0.926, Test Loss = 1.648, Test Accuracy = 0.775
 68, Train Loss = 0.157, Train Accuracy = 0.950, Test Loss = 1.719, Test Accuracy = 0.788
 69, Train Loss = 0.146, Train Accuracy = 0.947, Test Loss = 1.702, Test Accuracy = 0.785
 70, Train Loss = 0.131, Train Accuracy = 0.951, Test Loss = 1.665, Test Accuracy = 0.792
 71, Train Loss = 0.105, Train Accuracy = 0.962, Test Loss = 1.667, Test Accuracy = 0.794
 72, Train Loss = 0.106, Train Accuracy = 0.961, Test Loss = 1.645, Test Accuracy = 0.784
 73, Train Loss = 0.190, Train Accuracy = 0.928, Test Loss = 1.780, Test Accuracy = 0.780
 74, Train Loss = 0.168, Train Accuracy = 0.939, Test Loss = 1.755, Test Accuracy = 0.777
 75, Train Loss = 0.128, Train Accuracy = 0.950, Test Loss = 1.738, Test Accuracy = 0.779
 76, Train Loss = 0.129, Train Accuracy = 0.949, Test Loss = 1.736, Test Accuracy = 0.784
 77, Train Loss = 0.112, Train Accuracy = 0.959, Test Loss = 1.785, Test Accuracy = 0.790
 78, Train Loss = 0.103, Train Accuracy = 0.962, Test Loss = 1.756, Test Accuracy = 0.781
 79, Train Loss = 0.134, Train Accuracy = 0.950, Test Loss = 1.836, Test Accuracy = 0.785
 80, Train Loss = 0.131, Train Accuracy = 0.951, Test Loss = 1.815, Test Accuracy = 0.785
 81, Train Loss = 0.117, Train Accuracy = 0.956, Test Loss = 1.885, Test Accuracy = 0.783
 82, Train Loss = 0.112, Train Accuracy = 0.957, Test Loss = 1.791, Test Accuracy = 0.789
 83, Train Loss = 0.103, Train Accuracy = 0.960, Test Loss = 1.836, Test Accuracy = 0.786
 84, Train Loss = 0.107, Train Accuracy = 0.959, Test Loss = 1.857, Test Accuracy = 0.786
 85, Train Loss = 0.151, Train Accuracy = 0.946, Test Loss = 1.893, Test Accuracy = 0.781
 86, Train Loss = 0.113, Train Accuracy = 0.954, Test Loss = 1.846, Test

Accuracy = 0.780
 87, Train Loss = 0.113, Train Accuracy = 0.959, Test Loss = 1.850, Test Accuracy = 0.783
 88, Train Loss = 0.101, Train Accuracy = 0.961, Test Loss = 1.887, Test Accuracy = 0.776
 89, Train Loss = 0.128, Train Accuracy = 0.953, Test Loss = 1.960, Test Accuracy = 0.783
 90, Train Loss = 0.186, Train Accuracy = 0.940, Test Loss = 2.064, Test Accuracy = 0.777
 91, Train Loss = 0.113, Train Accuracy = 0.954, Test Loss = 1.926, Test Accuracy = 0.777
 92, Train Loss = 0.089, Train Accuracy = 0.965, Test Loss = 1.951, Test Accuracy = 0.782
 93, Train Loss = 0.108, Train Accuracy = 0.960, Test Loss = 1.953, Test Accuracy = 0.785
 94, Train Loss = 0.157, Train Accuracy = 0.941, Test Loss = 2.102, Test Accuracy = 0.777
 95, Train Loss = 0.104, Train Accuracy = 0.960, Test Loss = 2.033, Test Accuracy = 0.779
 96, Train Loss = 0.193, Train Accuracy = 0.939, Test Loss = 2.269, Test Accuracy = 0.768
 97, Train Loss = 0.147, Train Accuracy = 0.951, Test Loss = 2.081, Test Accuracy = 0.780
 98, Train Loss = 0.111, Train Accuracy = 0.957, Test Loss = 2.105, Test Accuracy = 0.782
 99, Train Loss = 0.118, Train Accuracy = 0.956, Test Loss = 2.144, Test Accuracy = 0.772
 100, Train Loss = 0.106, Train Accuracy = 0.959, Test Loss = 2.095, Test Accuracy = 0.781

Hidden Size: 16

Train Loss = 2.303, Train Accuracy = 0.106, Test Loss = 2.303, Test Accuracy = 0.108

1, Train Loss = 0.572, Train Accuracy = 0.779, Test Loss = 0.617, Test Accuracy = 0.769
 2, Train Loss = 0.479, Train Accuracy = 0.834, Test Loss = 0.589, Test Accuracy = 0.805
 3, Train Loss = 0.407, Train Accuracy = 0.859, Test Loss = 0.575, Test Accuracy = 0.810
 4, Train Loss = 0.381, Train Accuracy = 0.865, Test Loss = 0.594, Test Accuracy = 0.810
 5, Train Loss = 0.369, Train Accuracy = 0.871, Test Loss = 0.607, Test Accuracy = 0.807
 6, Train Loss = 0.319, Train Accuracy = 0.890, Test Loss = 0.581, Test Accuracy = 0.810
 7, Train Loss = 0.291, Train Accuracy = 0.893, Test Loss = 0.568, Test Accuracy = 0.812
 8, Train Loss = 0.289, Train Accuracy = 0.894, Test Loss = 0.598, Test

Accuracy = 0.814
 9, Train Loss = 0.271, Train Accuracy = 0.905, Test Loss = 0.596, Test Accuracy = 0.822
 10, Train Loss = 0.286, Train Accuracy = 0.897, Test Loss = 0.677, Test Accuracy = 0.813
 11, Train Loss = 0.244, Train Accuracy = 0.911, Test Loss = 0.648, Test Accuracy = 0.820
 12, Train Loss = 0.246, Train Accuracy = 0.910, Test Loss = 0.698, Test Accuracy = 0.808
 13, Train Loss = 0.213, Train Accuracy = 0.924, Test Loss = 0.698, Test Accuracy = 0.818
 14, Train Loss = 0.244, Train Accuracy = 0.918, Test Loss = 0.745, Test Accuracy = 0.816
 15, Train Loss = 0.212, Train Accuracy = 0.923, Test Loss = 0.725, Test Accuracy = 0.805
 16, Train Loss = 0.214, Train Accuracy = 0.926, Test Loss = 0.765, Test Accuracy = 0.815
 17, Train Loss = 0.223, Train Accuracy = 0.920, Test Loss = 0.822, Test Accuracy = 0.800
 18, Train Loss = 0.209, Train Accuracy = 0.925, Test Loss = 0.806, Test Accuracy = 0.806
 19, Train Loss = 0.169, Train Accuracy = 0.938, Test Loss = 0.810, Test Accuracy = 0.814
 20, Train Loss = 0.209, Train Accuracy = 0.923, Test Loss = 0.867, Test Accuracy = 0.799
 21, Train Loss = 0.191, Train Accuracy = 0.929, Test Loss = 0.864, Test Accuracy = 0.802
 22, Train Loss = 0.128, Train Accuracy = 0.956, Test Loss = 0.794, Test Accuracy = 0.819
 23, Train Loss = 0.144, Train Accuracy = 0.950, Test Loss = 0.879, Test Accuracy = 0.817
 24, Train Loss = 0.132, Train Accuracy = 0.952, Test Loss = 0.873, Test Accuracy = 0.813
 25, Train Loss = 0.137, Train Accuracy = 0.951, Test Loss = 0.913, Test Accuracy = 0.819
 26, Train Loss = 0.140, Train Accuracy = 0.950, Test Loss = 0.943, Test Accuracy = 0.811
 27, Train Loss = 0.122, Train Accuracy = 0.959, Test Loss = 0.952, Test Accuracy = 0.812
 28, Train Loss = 0.108, Train Accuracy = 0.965, Test Loss = 0.928, Test Accuracy = 0.808
 29, Train Loss = 0.122, Train Accuracy = 0.954, Test Loss = 1.024, Test Accuracy = 0.809
 30, Train Loss = 0.117, Train Accuracy = 0.957, Test Loss = 1.007, Test Accuracy = 0.809
 31, Train Loss = 0.123, Train Accuracy = 0.953, Test Loss = 1.052, Test Accuracy = 0.807
 32, Train Loss = 0.108, Train Accuracy = 0.959, Test Loss = 1.054, Test

Accuracy = 0.804
 33, Train Loss = 0.077, Train Accuracy = 0.974, Test Loss = 1.043, Test Accuracy = 0.816
 34, Train Loss = 0.080, Train Accuracy = 0.972, Test Loss = 1.085, Test Accuracy = 0.814
 35, Train Loss = 0.170, Train Accuracy = 0.943, Test Loss = 1.209, Test Accuracy = 0.793
 36, Train Loss = 0.079, Train Accuracy = 0.970, Test Loss = 1.074, Test Accuracy = 0.815
 37, Train Loss = 0.117, Train Accuracy = 0.961, Test Loss = 1.178, Test Accuracy = 0.799
 38, Train Loss = 0.072, Train Accuracy = 0.975, Test Loss = 1.155, Test Accuracy = 0.816
 39, Train Loss = 0.069, Train Accuracy = 0.977, Test Loss = 1.166, Test Accuracy = 0.816
 40, Train Loss = 0.098, Train Accuracy = 0.966, Test Loss = 1.308, Test Accuracy = 0.807
 41, Train Loss = 0.060, Train Accuracy = 0.979, Test Loss = 1.224, Test Accuracy = 0.812
 42, Train Loss = 0.080, Train Accuracy = 0.972, Test Loss = 1.306, Test Accuracy = 0.807
 43, Train Loss = 0.084, Train Accuracy = 0.974, Test Loss = 1.304, Test Accuracy = 0.811
 44, Train Loss = 0.101, Train Accuracy = 0.964, Test Loss = 1.311, Test Accuracy = 0.807
 45, Train Loss = 0.203, Train Accuracy = 0.947, Test Loss = 1.554, Test Accuracy = 0.807
 46, Train Loss = 0.112, Train Accuracy = 0.961, Test Loss = 1.399, Test Accuracy = 0.807
 47, Train Loss = 0.140, Train Accuracy = 0.957, Test Loss = 1.496, Test Accuracy = 0.803
 48, Train Loss = 0.068, Train Accuracy = 0.977, Test Loss = 1.414, Test Accuracy = 0.811
 49, Train Loss = 0.048, Train Accuracy = 0.983, Test Loss = 1.428, Test Accuracy = 0.815
 50, Train Loss = 0.038, Train Accuracy = 0.988, Test Loss = 1.450, Test Accuracy = 0.813
 51, Train Loss = 0.047, Train Accuracy = 0.983, Test Loss = 1.458, Test Accuracy = 0.817
 52, Train Loss = 0.048, Train Accuracy = 0.984, Test Loss = 1.555, Test Accuracy = 0.810
 53, Train Loss = 0.027, Train Accuracy = 0.992, Test Loss = 1.498, Test Accuracy = 0.813
 54, Train Loss = 0.030, Train Accuracy = 0.991, Test Loss = 1.537, Test Accuracy = 0.813
 55, Train Loss = 0.073, Train Accuracy = 0.973, Test Loss = 1.646, Test Accuracy = 0.798
 56, Train Loss = 0.036, Train Accuracy = 0.989, Test Loss = 1.568, Test

Accuracy = 0.810
 57, Train Loss = 0.032, Train Accuracy = 0.988, Test Loss = 1.579, Test Accuracy = 0.810
 58, Train Loss = 0.034, Train Accuracy = 0.988, Test Loss = 1.642, Test Accuracy = 0.808
 59, Train Loss = 0.241, Train Accuracy = 0.941, Test Loss = 1.935, Test Accuracy = 0.781
 60, Train Loss = 0.061, Train Accuracy = 0.981, Test Loss = 1.679, Test Accuracy = 0.804
 61, Train Loss = 0.028, Train Accuracy = 0.990, Test Loss = 1.641, Test Accuracy = 0.809
 62, Train Loss = 0.024, Train Accuracy = 0.993, Test Loss = 1.678, Test Accuracy = 0.812
 63, Train Loss = 0.021, Train Accuracy = 0.994, Test Loss = 1.694, Test Accuracy = 0.808
 64, Train Loss = 0.017, Train Accuracy = 0.994, Test Loss = 1.743, Test Accuracy = 0.812
 65, Train Loss = 0.013, Train Accuracy = 0.996, Test Loss = 1.737, Test Accuracy = 0.811
 66, Train Loss = 0.014, Train Accuracy = 0.997, Test Loss = 1.766, Test Accuracy = 0.810
 67, Train Loss = 0.010, Train Accuracy = 0.997, Test Loss = 1.745, Test Accuracy = 0.812
 68, Train Loss = 0.026, Train Accuracy = 0.991, Test Loss = 1.789, Test Accuracy = 0.814
 69, Train Loss = 0.052, Train Accuracy = 0.981, Test Loss = 1.840, Test Accuracy = 0.807
 70, Train Loss = 0.027, Train Accuracy = 0.991, Test Loss = 1.797, Test Accuracy = 0.809
 71, Train Loss = 0.017, Train Accuracy = 0.994, Test Loss = 1.801, Test Accuracy = 0.811
 72, Train Loss = 0.039, Train Accuracy = 0.985, Test Loss = 1.878, Test Accuracy = 0.805
 73, Train Loss = 0.045, Train Accuracy = 0.986, Test Loss = 1.855, Test Accuracy = 0.804
 74, Train Loss = 0.019, Train Accuracy = 0.994, Test Loss = 1.820, Test Accuracy = 0.806
 75, Train Loss = 0.010, Train Accuracy = 0.998, Test Loss = 1.807, Test Accuracy = 0.811
 76, Train Loss = 0.007, Train Accuracy = 0.999, Test Loss = 1.836, Test Accuracy = 0.810
 77, Train Loss = 0.006, Train Accuracy = 0.999, Test Loss = 1.853, Test Accuracy = 0.811
 78, Train Loss = 0.005, Train Accuracy = 0.999, Test Loss = 1.860, Test Accuracy = 0.811
 79, Train Loss = 0.014, Train Accuracy = 0.994, Test Loss = 1.892, Test Accuracy = 0.809
 80, Train Loss = 0.031, Train Accuracy = 0.990, Test Loss = 1.991, Test

Accuracy = 0.805
 81, Train Loss = 0.030, Train Accuracy = 0.990, Test Loss = 1.948, Test Accuracy = 0.812
 82, Train Loss = 0.041, Train Accuracy = 0.986, Test Loss = 1.994, Test Accuracy = 0.798
 83, Train Loss = 0.026, Train Accuracy = 0.991, Test Loss = 2.006, Test Accuracy = 0.805
 84, Train Loss = 0.008, Train Accuracy = 0.998, Test Loss = 1.937, Test Accuracy = 0.811
 85, Train Loss = 0.007, Train Accuracy = 0.999, Test Loss = 1.957, Test Accuracy = 0.809
 86, Train Loss = 0.010, Train Accuracy = 0.997, Test Loss = 1.969, Test Accuracy = 0.809
 87, Train Loss = 0.007, Train Accuracy = 0.998, Test Loss = 2.012, Test Accuracy = 0.808
 88, Train Loss = 0.008, Train Accuracy = 0.998, Test Loss = 2.013, Test Accuracy = 0.809
 89, Train Loss = 0.004, Train Accuracy = 0.999, Test Loss = 2.028, Test Accuracy = 0.810
 90, Train Loss = 0.004, Train Accuracy = 0.999, Test Loss = 2.039, Test Accuracy = 0.809
 91, Train Loss = 0.003, Train Accuracy = 0.999, Test Loss = 2.052, Test Accuracy = 0.809
 92, Train Loss = 0.004, Train Accuracy = 0.999, Test Loss = 2.073, Test Accuracy = 0.810
 93, Train Loss = 0.082, Train Accuracy = 0.975, Test Loss = 2.152, Test Accuracy = 0.803
 94, Train Loss = 0.074, Train Accuracy = 0.978, Test Loss = 2.183, Test Accuracy = 0.802
 95, Train Loss = 0.049, Train Accuracy = 0.985, Test Loss = 2.072, Test Accuracy = 0.808
 96, Train Loss = 0.035, Train Accuracy = 0.988, Test Loss = 2.009, Test Accuracy = 0.804
 97, Train Loss = 0.118, Train Accuracy = 0.965, Test Loss = 2.210, Test Accuracy = 0.794
 98, Train Loss = 0.041, Train Accuracy = 0.985, Test Loss = 2.071, Test Accuracy = 0.807
 99, Train Loss = 0.067, Train Accuracy = 0.979, Test Loss = 2.153, Test Accuracy = 0.805
 100, Train Loss = 0.042, Train Accuracy = 0.985, Test Loss = 2.103, Test Accuracy = 0.805

Hidden Size: 32

Train Loss = 2.303, Train Accuracy = 0.099, Test Loss = 2.303, Test Accuracy = 0.098

1, Train Loss = 0.574, Train Accuracy = 0.791, Test Loss = 0.621, Test Accuracy = 0.779

2, Train Loss = 0.483, Train Accuracy = 0.818, Test Loss = 0.605, Test

Accuracy = 0.771
 3, Train Loss = 0.376, Train Accuracy = 0.863, Test Loss = 0.568, Test Accuracy = 0.808
 4, Train Loss = 0.343, Train Accuracy = 0.879, Test Loss = 0.563, Test Accuracy = 0.816
 5, Train Loss = 0.343, Train Accuracy = 0.878, Test Loss = 0.597, Test Accuracy = 0.811
 6, Train Loss = 0.310, Train Accuracy = 0.886, Test Loss = 0.607, Test Accuracy = 0.822
 7, Train Loss = 0.273, Train Accuracy = 0.902, Test Loss = 0.584, Test Accuracy = 0.824
 8, Train Loss = 0.225, Train Accuracy = 0.917, Test Loss = 0.586, Test Accuracy = 0.830
 9, Train Loss = 0.210, Train Accuracy = 0.927, Test Loss = 0.613, Test Accuracy = 0.824
 10, Train Loss = 0.226, Train Accuracy = 0.917, Test Loss = 0.641, Test Accuracy = 0.816
 11, Train Loss = 0.197, Train Accuracy = 0.930, Test Loss = 0.668, Test Accuracy = 0.820
 12, Train Loss = 0.175, Train Accuracy = 0.938, Test Loss = 0.700, Test Accuracy = 0.823
 13, Train Loss = 0.216, Train Accuracy = 0.919, Test Loss = 0.731, Test Accuracy = 0.809
 14, Train Loss = 0.171, Train Accuracy = 0.936, Test Loss = 0.723, Test Accuracy = 0.818
 15, Train Loss = 0.151, Train Accuracy = 0.947, Test Loss = 0.709, Test Accuracy = 0.823
 16, Train Loss = 0.231, Train Accuracy = 0.921, Test Loss = 0.840, Test Accuracy = 0.804
 17, Train Loss = 0.125, Train Accuracy = 0.955, Test Loss = 0.756, Test Accuracy = 0.816
 18, Train Loss = 0.136, Train Accuracy = 0.952, Test Loss = 0.788, Test Accuracy = 0.819
 19, Train Loss = 0.133, Train Accuracy = 0.947, Test Loss = 0.807, Test Accuracy = 0.819
 20, Train Loss = 0.103, Train Accuracy = 0.962, Test Loss = 0.825, Test Accuracy = 0.828
 21, Train Loss = 0.076, Train Accuracy = 0.974, Test Loss = 0.836, Test Accuracy = 0.826
 22, Train Loss = 0.110, Train Accuracy = 0.960, Test Loss = 0.897, Test Accuracy = 0.822
 23, Train Loss = 0.090, Train Accuracy = 0.966, Test Loss = 0.950, Test Accuracy = 0.819
 24, Train Loss = 0.085, Train Accuracy = 0.970, Test Loss = 0.930, Test Accuracy = 0.828
 25, Train Loss = 0.101, Train Accuracy = 0.967, Test Loss = 0.943, Test Accuracy = 0.826
 26, Train Loss = 0.075, Train Accuracy = 0.974, Test Loss = 0.947, Test

Accuracy = 0.825
 27, Train Loss = 0.061, Train Accuracy = 0.978, Test Loss = 0.990, Test Accuracy = 0.825
 28, Train Loss = 0.045, Train Accuracy = 0.986, Test Loss = 0.961, Test Accuracy = 0.827
 29, Train Loss = 0.083, Train Accuracy = 0.970, Test Loss = 1.065, Test Accuracy = 0.817
 30, Train Loss = 0.081, Train Accuracy = 0.972, Test Loss = 1.080, Test Accuracy = 0.820
 31, Train Loss = 0.050, Train Accuracy = 0.981, Test Loss = 1.053, Test Accuracy = 0.829
 32, Train Loss = 0.097, Train Accuracy = 0.972, Test Loss = 1.151, Test Accuracy = 0.823
 33, Train Loss = 0.048, Train Accuracy = 0.984, Test Loss = 1.129, Test Accuracy = 0.823
 34, Train Loss = 0.032, Train Accuracy = 0.990, Test Loss = 1.128, Test Accuracy = 0.828
 35, Train Loss = 0.029, Train Accuracy = 0.992, Test Loss = 1.135, Test Accuracy = 0.827
 36, Train Loss = 0.025, Train Accuracy = 0.993, Test Loss = 1.145, Test Accuracy = 0.828
 37, Train Loss = 0.103, Train Accuracy = 0.968, Test Loss = 1.286, Test Accuracy = 0.812
 38, Train Loss = 0.024, Train Accuracy = 0.992, Test Loss = 1.168, Test Accuracy = 0.832
 39, Train Loss = 0.026, Train Accuracy = 0.993, Test Loss = 1.210, Test Accuracy = 0.828
 40, Train Loss = 0.019, Train Accuracy = 0.994, Test Loss = 1.187, Test Accuracy = 0.833
 41, Train Loss = 0.046, Train Accuracy = 0.988, Test Loss = 1.278, Test Accuracy = 0.826
 42, Train Loss = 0.023, Train Accuracy = 0.993, Test Loss = 1.223, Test Accuracy = 0.828
 43, Train Loss = 0.015, Train Accuracy = 0.996, Test Loss = 1.228, Test Accuracy = 0.828
 44, Train Loss = 0.017, Train Accuracy = 0.996, Test Loss = 1.228, Test Accuracy = 0.828
 45, Train Loss = 0.025, Train Accuracy = 0.993, Test Loss = 1.283, Test Accuracy = 0.828
 46, Train Loss = 0.009, Train Accuracy = 0.998, Test Loss = 1.250, Test Accuracy = 0.832
 47, Train Loss = 0.009, Train Accuracy = 0.998, Test Loss = 1.255, Test Accuracy = 0.831
 48, Train Loss = 0.017, Train Accuracy = 0.994, Test Loss = 1.279, Test Accuracy = 0.827
 49, Train Loss = 0.005, Train Accuracy = 0.999, Test Loss = 1.274, Test Accuracy = 0.831
 50, Train Loss = 0.005, Train Accuracy = 0.999, Test Loss = 1.291, Test

Accuracy = 0.831
 51, Train Loss = 0.004, Train Accuracy = 1.000, Test Loss = 1.304, Test Accuracy = 0.832
 52, Train Loss = 0.003, Train Accuracy = 1.000, Test Loss = 1.307, Test Accuracy = 0.832
 53, Train Loss = 0.004, Train Accuracy = 0.999, Test Loss = 1.327, Test Accuracy = 0.830
 54, Train Loss = 0.003, Train Accuracy = 1.000, Test Loss = 1.330, Test Accuracy = 0.831
 55, Train Loss = 0.003, Train Accuracy = 1.000, Test Loss = 1.340, Test Accuracy = 0.833
 56, Train Loss = 0.005, Train Accuracy = 0.999, Test Loss = 1.345, Test Accuracy = 0.830
 57, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.353, Test Accuracy = 0.833
 58, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.359, Test Accuracy = 0.833
 59, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.360, Test Accuracy = 0.833
 60, Train Loss = 0.003, Train Accuracy = 1.000, Test Loss = 1.377, Test Accuracy = 0.833
 61, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.375, Test Accuracy = 0.832
 62, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.383, Test Accuracy = 0.832
 63, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.390, Test Accuracy = 0.831
 64, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.388, Test Accuracy = 0.832
 65, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.397, Test Accuracy = 0.833
 66, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.401, Test Accuracy = 0.831
 67, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.410, Test Accuracy = 0.832
 68, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.414, Test Accuracy = 0.832
 69, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.416, Test Accuracy = 0.831
 70, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.421, Test Accuracy = 0.832
 71, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.424, Test Accuracy = 0.832
 72, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.430, Test Accuracy = 0.832
 73, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.434, Test Accuracy = 0.832
 74, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.436, Test

Accuracy = 0.832
 75, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.439, Test Accuracy = 0.831
 76, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.446, Test Accuracy = 0.832
 77, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.450, Test Accuracy = 0.832
 78, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.451, Test Accuracy = 0.832
 79, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.456, Test Accuracy = 0.831
 80, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.460, Test Accuracy = 0.832
 81, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.465, Test Accuracy = 0.831
 82, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.466, Test Accuracy = 0.832
 83, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.468, Test Accuracy = 0.831
 84, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.472, Test Accuracy = 0.831
 85, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.473, Test Accuracy = 0.831
 86, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.479, Test Accuracy = 0.832
 87, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.483, Test Accuracy = 0.831
 88, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.488, Test Accuracy = 0.831
 89, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.490, Test Accuracy = 0.832
 90, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.490, Test Accuracy = 0.832
 91, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.492, Test Accuracy = 0.831
 92, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.496, Test Accuracy = 0.831
 93, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.499, Test Accuracy = 0.831
 94, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.503, Test Accuracy = 0.832
 95, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.504, Test Accuracy = 0.832
 96, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.508, Test Accuracy = 0.832
 97, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.511, Test Accuracy = 0.830
 98, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.513, Test

Accuracy = 0.832

99, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.515, Test Accuracy = 0.832

100, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.518, Test Accuracy = 0.831

Hidden Size: 128

Train Loss = 2.303, Train Accuracy = 0.077, Test Loss = 2.303, Test Accuracy = 0.083

1, Train Loss = 0.583, Train Accuracy = 0.788, Test Loss = 0.646, Test Accuracy = 0.775

2, Train Loss = 0.415, Train Accuracy = 0.854, Test Loss = 0.558, Test Accuracy = 0.812

3, Train Loss = 0.358, Train Accuracy = 0.873, Test Loss = 0.530, Test Accuracy = 0.826

4, Train Loss = 0.361, Train Accuracy = 0.868, Test Loss = 0.610, Test Accuracy = 0.811

5, Train Loss = 0.435, Train Accuracy = 0.840, Test Loss = 0.736, Test Accuracy = 0.775

6, Train Loss = 0.321, Train Accuracy = 0.882, Test Loss = 0.641, Test Accuracy = 0.812

7, Train Loss = 0.252, Train Accuracy = 0.913, Test Loss = 0.597, Test Accuracy = 0.822

8, Train Loss = 0.258, Train Accuracy = 0.908, Test Loss = 0.672, Test Accuracy = 0.816

9, Train Loss = 0.204, Train Accuracy = 0.927, Test Loss = 0.642, Test Accuracy = 0.828

10, Train Loss = 0.179, Train Accuracy = 0.935, Test Loss = 0.644, Test Accuracy = 0.831

11, Train Loss = 0.166, Train Accuracy = 0.941, Test Loss = 0.682, Test Accuracy = 0.833

12, Train Loss = 0.162, Train Accuracy = 0.942, Test Loss = 0.712, Test Accuracy = 0.830

13, Train Loss = 0.170, Train Accuracy = 0.941, Test Loss = 0.812, Test Accuracy = 0.819

14, Train Loss = 0.149, Train Accuracy = 0.943, Test Loss = 0.750, Test Accuracy = 0.827

15, Train Loss = 0.116, Train Accuracy = 0.956, Test Loss = 0.766, Test Accuracy = 0.831

16, Train Loss = 0.091, Train Accuracy = 0.968, Test Loss = 0.753, Test Accuracy = 0.830

17, Train Loss = 0.079, Train Accuracy = 0.973, Test Loss = 0.787, Test Accuracy = 0.833

18, Train Loss = 0.098, Train Accuracy = 0.966, Test Loss = 0.838, Test Accuracy = 0.828

19, Train Loss = 0.072, Train Accuracy = 0.977, Test Loss = 0.815, Test Accuracy = 0.832

20, Train Loss = 0.075, Train Accuracy = 0.975, Test Loss = 0.866, Test

Accuracy = 0.829
 21, Train Loss = 0.054, Train Accuracy = 0.983, Test Loss = 0.848, Test Accuracy = 0.834
 22, Train Loss = 0.040, Train Accuracy = 0.988, Test Loss = 0.885, Test Accuracy = 0.836
 23, Train Loss = 0.048, Train Accuracy = 0.985, Test Loss = 0.909, Test Accuracy = 0.830
 24, Train Loss = 0.034, Train Accuracy = 0.990, Test Loss = 0.894, Test Accuracy = 0.837
 25, Train Loss = 0.060, Train Accuracy = 0.980, Test Loss = 0.980, Test Accuracy = 0.830
 26, Train Loss = 0.037, Train Accuracy = 0.988, Test Loss = 0.979, Test Accuracy = 0.831
 27, Train Loss = 0.032, Train Accuracy = 0.991, Test Loss = 0.980, Test Accuracy = 0.836
 28, Train Loss = 0.084, Train Accuracy = 0.975, Test Loss = 1.066, Test Accuracy = 0.825
 29, Train Loss = 0.018, Train Accuracy = 0.994, Test Loss = 0.978, Test Accuracy = 0.838
 30, Train Loss = 0.018, Train Accuracy = 0.995, Test Loss = 1.024, Test Accuracy = 0.839
 31, Train Loss = 0.014, Train Accuracy = 0.996, Test Loss = 1.023, Test Accuracy = 0.841
 32, Train Loss = 0.114, Train Accuracy = 0.975, Test Loss = 1.229, Test Accuracy = 0.815
 33, Train Loss = 0.009, Train Accuracy = 0.999, Test Loss = 1.045, Test Accuracy = 0.842
 34, Train Loss = 0.013, Train Accuracy = 0.997, Test Loss = 1.063, Test Accuracy = 0.838
 35, Train Loss = 0.020, Train Accuracy = 0.994, Test Loss = 1.084, Test Accuracy = 0.835
 36, Train Loss = 0.004, Train Accuracy = 1.000, Test Loss = 1.062, Test Accuracy = 0.842
 37, Train Loss = 0.005, Train Accuracy = 0.999, Test Loss = 1.079, Test Accuracy = 0.842
 38, Train Loss = 0.004, Train Accuracy = 1.000, Test Loss = 1.087, Test Accuracy = 0.842
 39, Train Loss = 0.003, Train Accuracy = 1.000, Test Loss = 1.088, Test Accuracy = 0.843
 40, Train Loss = 0.003, Train Accuracy = 1.000, Test Loss = 1.096, Test Accuracy = 0.842
 41, Train Loss = 0.003, Train Accuracy = 1.000, Test Loss = 1.103, Test Accuracy = 0.841
 42, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.113, Test Accuracy = 0.842
 43, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.115, Test Accuracy = 0.842
 44, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.123, Test

Accuracy = 0.841
 45, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.131, Test Accuracy = 0.843
 46, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.136, Test Accuracy = 0.841
 47, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.141, Test Accuracy = 0.841
 48, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.147, Test Accuracy = 0.842
 49, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.150, Test Accuracy = 0.842
 50, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.157, Test Accuracy = 0.842
 51, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.159, Test Accuracy = 0.841
 52, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.161, Test Accuracy = 0.842
 53, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.168, Test Accuracy = 0.842
 54, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.170, Test Accuracy = 0.842
 55, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.172, Test Accuracy = 0.842
 56, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.177, Test Accuracy = 0.842
 57, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.180, Test Accuracy = 0.842
 58, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.183, Test Accuracy = 0.842
 59, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.188, Test Accuracy = 0.842
 60, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.191, Test Accuracy = 0.842
 61, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.196, Test Accuracy = 0.842
 62, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.197, Test Accuracy = 0.842
 63, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.200, Test Accuracy = 0.841
 64, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.202, Test Accuracy = 0.842
 65, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.203, Test Accuracy = 0.842
 66, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.208, Test Accuracy = 0.842
 67, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.212, Test Accuracy = 0.842
 68, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.212, Test

Accuracy = 0.842
 69, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.216, Test Accuracy = 0.841
 70, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.220, Test Accuracy = 0.841
 71, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.221, Test Accuracy = 0.842
 72, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.225, Test Accuracy = 0.842
 73, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.231, Test Accuracy = 0.842
 74, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.229, Test Accuracy = 0.842
 75, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.231, Test Accuracy = 0.842
 76, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.230, Test Accuracy = 0.842
 77, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.235, Test Accuracy = 0.842
 78, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.237, Test Accuracy = 0.843
 79, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.243, Test Accuracy = 0.843
 80, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.242, Test Accuracy = 0.842
 81, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.245, Test Accuracy = 0.842
 82, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.249, Test Accuracy = 0.842
 83, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.249, Test Accuracy = 0.842
 84, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.250, Test Accuracy = 0.841
 85, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.252, Test Accuracy = 0.842
 86, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.255, Test Accuracy = 0.843
 87, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.257, Test Accuracy = 0.841
 88, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.258, Test Accuracy = 0.842
 89, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.262, Test Accuracy = 0.842
 90, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.263, Test Accuracy = 0.842
 91, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.264, Test Accuracy = 0.842
 92, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.268, Test

Accuracy = 0.842
 93, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.269, Test Accuracy = 0.843
 94, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.270, Test Accuracy = 0.843
 95, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.274, Test Accuracy = 0.842
 96, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.275, Test Accuracy = 0.842
 97, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.275, Test Accuracy = 0.842
 98, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.279, Test Accuracy = 0.842
 99, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.278, Test Accuracy = 0.842
 100, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.280, Test Accuracy = 0.842

Hidden Size: 512

Train Loss = 2.303, Train Accuracy = 0.100, Test Loss = 2.303, Test Accuracy = 0.101

1, Train Loss = 0.543, Train Accuracy = 0.799, Test Loss = 0.600, Test Accuracy = 0.780
 2, Train Loss = 0.413, Train Accuracy = 0.853, Test Loss = 0.538, Test Accuracy = 0.808
 3, Train Loss = 0.351, Train Accuracy = 0.869, Test Loss = 0.513, Test Accuracy = 0.822
 4, Train Loss = 0.318, Train Accuracy = 0.882, Test Loss = 0.526, Test Accuracy = 0.818
 5, Train Loss = 0.343, Train Accuracy = 0.877, Test Loss = 0.609, Test Accuracy = 0.805
 6, Train Loss = 0.258, Train Accuracy = 0.906, Test Loss = 0.555, Test Accuracy = 0.829
 7, Train Loss = 0.225, Train Accuracy = 0.920, Test Loss = 0.562, Test Accuracy = 0.831
 8, Train Loss = 0.319, Train Accuracy = 0.885, Test Loss = 0.714, Test Accuracy = 0.795
 9, Train Loss = 0.192, Train Accuracy = 0.933, Test Loss = 0.616, Test Accuracy = 0.832
 10, Train Loss = 0.165, Train Accuracy = 0.942, Test Loss = 0.612, Test Accuracy = 0.832
 11, Train Loss = 0.153, Train Accuracy = 0.943, Test Loss = 0.643, Test Accuracy = 0.826
 12, Train Loss = 0.138, Train Accuracy = 0.948, Test Loss = 0.659, Test Accuracy = 0.836
 13, Train Loss = 0.114, Train Accuracy = 0.962, Test Loss = 0.666, Test Accuracy = 0.833
 14, Train Loss = 0.146, Train Accuracy = 0.945, Test Loss = 0.754, Test

Accuracy = 0.826
 15, Train Loss = 0.148, Train Accuracy = 0.948, Test Loss = 0.805, Test Accuracy = 0.816
 16, Train Loss = 0.073, Train Accuracy = 0.976, Test Loss = 0.696, Test Accuracy = 0.840
 17, Train Loss = 0.076, Train Accuracy = 0.975, Test Loss = 0.760, Test Accuracy = 0.830
 18, Train Loss = 0.094, Train Accuracy = 0.965, Test Loss = 0.817, Test Accuracy = 0.827
 19, Train Loss = 0.051, Train Accuracy = 0.984, Test Loss = 0.762, Test Accuracy = 0.837
 20, Train Loss = 0.052, Train Accuracy = 0.982, Test Loss = 0.796, Test Accuracy = 0.833
 21, Train Loss = 0.033, Train Accuracy = 0.990, Test Loss = 0.792, Test Accuracy = 0.842
 22, Train Loss = 0.063, Train Accuracy = 0.981, Test Loss = 0.889, Test Accuracy = 0.835
 23, Train Loss = 0.035, Train Accuracy = 0.990, Test Loss = 0.847, Test Accuracy = 0.838
 24, Train Loss = 0.046, Train Accuracy = 0.991, Test Loss = 0.866, Test Accuracy = 0.837
 25, Train Loss = 0.029, Train Accuracy = 0.990, Test Loss = 0.889, Test Accuracy = 0.841
 26, Train Loss = 0.029, Train Accuracy = 0.992, Test Loss = 0.914, Test Accuracy = 0.837
 27, Train Loss = 0.037, Train Accuracy = 0.992, Test Loss = 0.965, Test Accuracy = 0.841
 28, Train Loss = 0.015, Train Accuracy = 0.996, Test Loss = 0.923, Test Accuracy = 0.842
 29, Train Loss = 0.010, Train Accuracy = 0.999, Test Loss = 0.946, Test Accuracy = 0.845
 30, Train Loss = 0.010, Train Accuracy = 0.998, Test Loss = 0.970, Test Accuracy = 0.843
 31, Train Loss = 0.008, Train Accuracy = 0.999, Test Loss = 0.964, Test Accuracy = 0.841
 32, Train Loss = 0.005, Train Accuracy = 0.999, Test Loss = 0.964, Test Accuracy = 0.842
 33, Train Loss = 0.004, Train Accuracy = 1.000, Test Loss = 0.989, Test Accuracy = 0.841
 34, Train Loss = 0.003, Train Accuracy = 1.000, Test Loss = 0.982, Test Accuracy = 0.845
 35, Train Loss = 0.004, Train Accuracy = 1.000, Test Loss = 0.996, Test Accuracy = 0.841
 36, Train Loss = 0.003, Train Accuracy = 1.000, Test Loss = 1.002, Test Accuracy = 0.844
 37, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.004, Test Accuracy = 0.844
 38, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.014, Test

Accuracy = 0.844
 39, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.020, Test Accuracy = 0.844
 40, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.026, Test Accuracy = 0.844
 41, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.030, Test Accuracy = 0.844
 42, Train Loss = 0.002, Train Accuracy = 1.000, Test Loss = 1.037, Test Accuracy = 0.844
 43, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.039, Test Accuracy = 0.844
 44, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.045, Test Accuracy = 0.844
 45, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.051, Test Accuracy = 0.844
 46, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.052, Test Accuracy = 0.844
 47, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.057, Test Accuracy = 0.843
 48, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.062, Test Accuracy = 0.845
 49, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.066, Test Accuracy = 0.844
 50, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.068, Test Accuracy = 0.844
 51, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.071, Test Accuracy = 0.844
 52, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.075, Test Accuracy = 0.843
 53, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.076, Test Accuracy = 0.845
 54, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.080, Test Accuracy = 0.844
 55, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.084, Test Accuracy = 0.844
 56, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.086, Test Accuracy = 0.844
 57, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.086, Test Accuracy = 0.844
 58, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.091, Test Accuracy = 0.844
 59, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.094, Test Accuracy = 0.844
 60, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.097, Test Accuracy = 0.844
 61, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.100, Test Accuracy = 0.844
 62, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.101, Test

Accuracy = 0.843
 63, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.104, Test Accuracy = 0.844
 64, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.106, Test Accuracy = 0.843
 65, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.111, Test Accuracy = 0.844
 66, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.112, Test Accuracy = 0.844
 67, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.112, Test Accuracy = 0.844
 68, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.113, Test Accuracy = 0.843
 69, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.116, Test Accuracy = 0.843
 70, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.120, Test Accuracy = 0.844
 71, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.123, Test Accuracy = 0.843
 72, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.124, Test Accuracy = 0.844
 73, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.125, Test Accuracy = 0.843
 74, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.127, Test Accuracy = 0.845
 75, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.129, Test Accuracy = 0.844
 76, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.133, Test Accuracy = 0.844
 77, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.134, Test Accuracy = 0.845
 78, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.136, Test Accuracy = 0.843
 79, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.138, Test Accuracy = 0.843
 80, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.140, Test Accuracy = 0.843
 81, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.143, Test Accuracy = 0.843
 82, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.145, Test Accuracy = 0.843
 83, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.146, Test Accuracy = 0.844
 84, Train Loss = 0.001, Train Accuracy = 1.000, Test Loss = 1.146, Test Accuracy = 0.842
 85, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.147, Test Accuracy = 0.844
 86, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.151, Test

```

Accuracy = 0.844
    87, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.150, Test
Accuracy = 0.844
    88, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.151, Test
Accuracy = 0.843
    89, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.153, Test
Accuracy = 0.844
    90, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.155, Test
Accuracy = 0.844
    91, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.157, Test
Accuracy = 0.843
    92, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.159, Test
Accuracy = 0.844
    93, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.159, Test
Accuracy = 0.844
    94, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.162, Test
Accuracy = 0.844
    95, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.163, Test
Accuracy = 0.844
    96, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.164, Test
Accuracy = 0.844
    97, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.165, Test
Accuracy = 0.843
    98, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.166, Test
Accuracy = 0.843
    99, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.167, Test
Accuracy = 0.843
    100, Train Loss = 0.000, Train Accuracy = 1.000, Test Loss = 1.170, Test
Accuracy = 0.843

```

Save the information for the various passes of data: `train_loss`, `train_accuracy`, `test_loss`, `test_accuracy`

```

[12]: if not os.path.exists(MLP_Q1_RESULTS_FILE):
        torch.save(hidden_size_logs, MLP_Q1_RESULTS_FILE)

```

```

[13]: if os.path.exists(MLP_Q1_RESULTS_FILE):
        hidden_size_logs = torch.load(MLP_Q1_RESULTS_FILE)
        # print(hidden_size_logs)

```

1.1.1 Deliverables

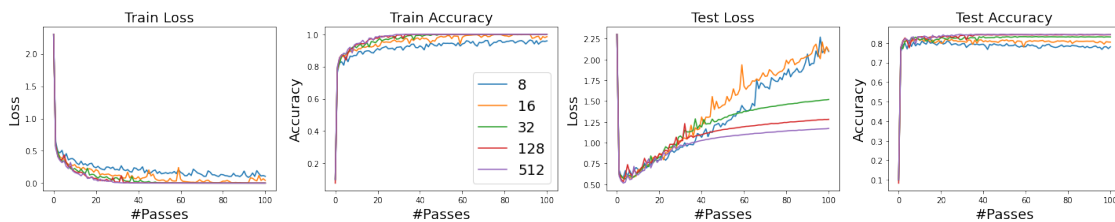
1.1 Make 4 plots, one each for the train loss, train accuracy, test loss and test accuracy over the course of training (i.e., the metric on the y-axis and number of effective passes on the x-axis). Plot all 4 lines, one for each value of `h` on the same plot.


```
[14]: f, ax = plt.subplots(1, 4, figsize=(20, 4))

ax[0].set_title('Train Loss', fontsize=18)
ax[1].set_title('Train Accuracy', fontsize=18)
ax[2].set_title('Test Loss', fontsize=18)
ax[3].set_title('Test Accuracy', fontsize=18)

for i in range(4):
    ax[i].set_xlabel('#Passes', fontsize=18)
    ax[i].set_ylabel('Loss' if i%2==0 else 'Accuracy', fontsize=18)
    for h, log in hidden_size_logs.items():
        ax[i].plot(list(map(lambda x: x[i], log)), label=h)

ax[1].legend(fontsize=18)
plt.tight_layout()
```



1.2 When the training accuracy is 100%, the model is said to interpolate the training data. What is the smallest width at which we observe perfect interpolation of the training data?

32 is the smallest width at which we see interpolation

1.3 As we vary the width of the network, at which training epoch do we observe perfect interpolation of the data? That is, make a plot with h on the x-axis and number of passes over the data required for interpolation on the y axis.

```
[15]: interpolation_epoch_value = []

for hidden_size in hidden_sizes:
    log = hidden_size_logs[hidden_size]
    train_accuracy = list(map(lambda x: x[1].item(), log))
    try:
        interpolation_epoch_value.append(train_accuracy.index(1))
    except ValueError as e:
        interpolation_epoch_value.append(-1)

f = plt.figure()
ax = f.gca()
```

```

ax.set_title('Interpolation Epoch vs Width', fontsize=18)
ax.set_xlabel('Width', fontsize=18)
ax.set_ylabel('Interpolation Epoch', fontsize=18)

interpolation_epoch_value = np.asarray(interpolation_epoch_value)
hidden_sizes = np.asarray(hidden_sizes)

# Now let's extract only the part of the data we're interested in...
x_filt = hidden_sizes[interpolation_epoch_value > -1]
y_filt = (interpolation_epoch_value[interpolation_epoch_value > -1]).astype(int)
ax.plot(x_filt, y_filt, 'x-')

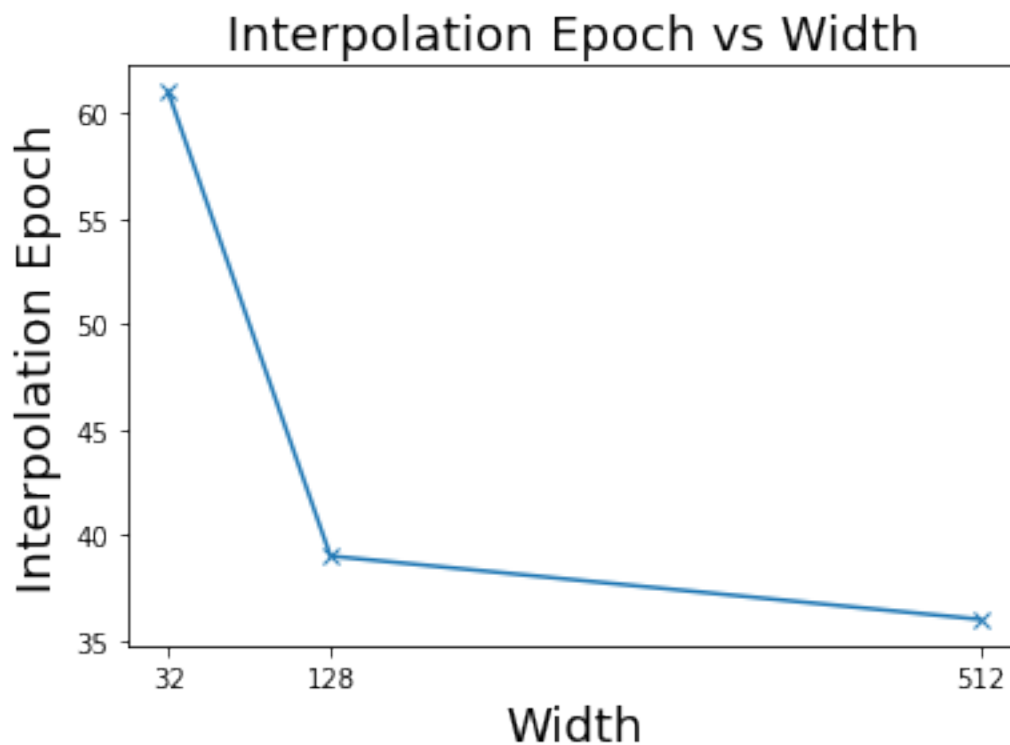
ax.set_xticks(x_filt)

```

```

[15]: [<matplotlib.axis.XTick at 0x7ff5af524250>,
      <matplotlib.axis.XTick at 0x7ff5af524220>,
      <matplotlib.axis.XTick at 0x7ff5af591e80>]

```



1.2 2. Divergent Learning Rate, Accuracy and Width

2.1 Find the divergent learning rate $*$ for width $h \in [4, 8, 16, 32, 128, 512, 2048]$. Make a plot for the divergent learning rate versus the hidden width.

```
[39]: hidden_sizes = [4, 8, 16, 32, 128, 512, 2048]
      learning_rates = []
```

```
[40]: learning_rate = 0.01
      learning_rates.append(learning_rate)

      hidden_size = 4
      ws = initialize_ws(hidden_size, n_class)
      bs = initialize_bs(hidden_size, n_class)

      _ = compute_logs(ws, bs, reg_param, verbose=True)

      ws, bs = sgd_one_pass(ws, bs, X_train, y_train, reg_param, learning_rate,
      ↪ verbose=True)
      _ = compute_logs(ws, bs, reg_param, verbose=True)
```

Train Loss = 2.303, Train Accuracy = 0.124, Test Loss = 2.303, Test Accuracy = 0.117

1.4621044326453378
1.9946212077286125
2.188290160071258
2.2673074699550444
2.2903342637293
2.2646656132070837
2.1035636986285033
1.8945605392492433
1.7045190959660383
1.6371252121000215
1.5706503565792471
1.5610587990509959
1.4428363601237224
1.2894367756875142
1.3289315551323706
1.2783734737657833
1.3040195461800386
1.3098727536752979
1.229342388631518
1.1851768984894018
1.1007731837999803
1.0931233074103104
1.0023980322681443
0.9756025283052179
0.998777254755458
0.9446234296111989
0.8794612793677193
1.0072491304920221
0.9866456339087731

```

1.0100517850363144
1.1039090411079666
1.0158672807237736
0.9025406167029669
0.9644502151316204
0.8613271149045321
0.8821630735740116
1.102523546364203
0.9035240123509657
1.0451258132625525
0.9088392632896937
0.8865732764537279
0.8471266561226818
0.7876959062020937
0.9362172896733943
0.7989035890233582
0.8858505090635214
0.8413130093114279
0.7123075687525289
0.8785647097653942
1.0224708740747879
0.8239897561756573
0.8294860352660162
0.8895013289287907
1.0477829802569196
0.8863447828697938
0.9264401683836575
0.7813873079443566
0.8109260823714468
0.7653973500546871
0.8598354142978076
Train Loss = 0.814, Train Accuracy = 0.705, Test Loss = 0.920, Test Accuracy =
0.686

```

```

[41]: learning_rate = 0.01
      learning_rates.append(learning_rate)

      hidden_size = 8
      ws = initialize_ws(hidden_size, n_class)
      bs = initialize_bs(hidden_size, n_class)

      _ = compute_logs(ws, bs, reg_param, verbose=True)

      ws, bs = sgd_one_pass(ws, bs, X_train, y_train, reg_param, learning_rate,
      ↪ verbose=True)
      _ = compute_logs(ws, bs, reg_param, verbose=True)

```

```

Train Loss = 2.303, Train Accuracy = 0.114, Test Loss = 2.303, Test Accuracy =

```

0.121
1.4594962983238338
1.994896669061386
2.19375505919179
2.263830598568168
2.270548670023029
2.129709586081141
1.9362766437383325
1.903442909672524
1.6834396941810519
1.4901799456384752
1.318910067048513
1.2380481216851393
1.19663744118872
1.2546736612064526
1.1046192528545489
0.9473943424715875
1.0774741985711862
0.978620198861561
0.9417639093785614
0.8803898129023999
0.9062765063081857
0.816659021690723
0.8540408363701852
0.8591681221519685
0.799925091072663
0.808321502965813
0.8141060130245283
0.6791167051764908
0.8025217660919476
0.7160209528662224
0.7470291940656685
0.7507430452863255
0.72662505616599
0.5518993270339667
0.6008608999766666
0.5779985573999937
0.7315067201025783
0.7512895114694139
0.7173518127124143
0.8456169550107429
0.8356060093964157
0.9454311718636675
0.9651192072284233
0.7477398996106921
0.5784488049034594
0.5619160998259286
0.7108016128398765

```

0.8171880802933742
0.848595086537765
0.9010985971096387
0.7205512121800163
0.9125870099660213
0.5391823927963894
0.5971671876685821
0.7371846485847406
0.7144617963109703
0.7902113080711434
0.5975278106577719
0.6250487178504821
0.523434914830702
Train Loss = 0.688, Train Accuracy = 0.754, Test Loss = 0.753, Test Accuracy =
0.739

```

```

[42]: learning_rate = 0.01
      learning_rates.append(learning_rate)

      hidden_size = 16
      ws = initialize_ws(hidden_size, n_class)
      bs = initialize_bs(hidden_size, n_class)

      _ = compute_logs(ws, bs, reg_param, verbose=True)

      ws, bs = sgd_one_pass(ws, bs, X_train, y_train, reg_param, learning_rate,
      ↪ verbose=True)
      _ = compute_logs(ws, bs, reg_param, verbose=True)

```

```

Train Loss = 2.303, Train Accuracy = 0.107, Test Loss = 2.303, Test Accuracy =
0.110
1.4586210769947818
1.995226417739537
2.1929056559375852
2.263722282457678
2.2899173843611567
2.2724305518860746
2.1058800048456305
1.875686551614814
1.5995967016659327
1.447633076281185
1.3343460819577724
1.1310162733890587
1.2435034250490513
1.1816860213231943
1.0875003096375828
1.0147646086683426
0.951172670716814

```

1.020679499985649
0.9815905852181919
0.8769080502232983
0.7114055748857689
0.705164166256284
0.6929650883418526
0.7869911148293648
0.7541833844337456
0.6953013365620873
0.8595848770148532
0.5987970263768012
0.6079001977077422
0.7571482515070442
0.6959483761535751
0.7415350028143547
0.762470332987473
0.7851127456058425
0.7642842542620728
0.8927166743641451
0.718240612196927
0.6616673362920175
0.7178821166474975
0.7970430961316594
0.7645088196181975
0.8064610026972202
0.7595335472893429
0.7026881578212822
0.7170782660264974
0.6927019856966086
0.6541970982914934
0.6362451882132276
0.7674625282332893
0.7791293344155633
0.778812697892731
0.8188360520453116
0.6955369290241553
0.6173783903483024
0.7085253230644285
0.6571796966911333
0.6283879282952105
0.7550578358694675
0.6929348935628198
0.6263666856222146
Train Loss = 0.629, Train Accuracy = 0.763, Test Loss = 0.708, Test Accuracy = 0.751

```
[43]: learning_rate = 0.01
learning_rates.append(learning_rate)

hidden_size = 32
ws = initialize_ws(hidden_size, n_class)
bs = initialize_bs(hidden_size, n_class)

_ = compute_logs(ws, bs, reg_param, verbose=True)

ws, bs = sgd_one_pass(ws, bs, X_train, y_train, reg_param, learning_rate,
    verbose=True)
_ = compute_logs(ws, bs, reg_param, verbose=True)
```

Train Loss = 2.303, Train Accuracy = 0.114, Test Loss = 2.303, Test Accuracy = 0.106

1.4604296819786085
1.9937657026718718
2.191738612974964
2.2644980935102224
2.2895176527289216
2.268097149501768
1.9772957633286015
1.7253355874293008
1.5172367079427904
1.417063096087028
1.3024711580989299
1.276871615517528
1.1530726089111902
0.9282455268589835
0.88789431617351
0.8767393161920373
0.7657869790510068
0.8415954182012605
0.7434935835940699
0.8007385331369903
0.7815702963515885
0.7044436682587145
0.9074608660344218
0.7428019483871825
0.7170108900729025
0.6346618663913212
0.5822804778476975
0.7509245387612579
0.6856550335613025
0.761151877735095
0.7144731935828339
0.635200774445079


```

0.5843186330780359
0.6548712174480074
0.6599587211427264
0.6314480233341946
0.6959560543330002
0.6195002837087029
0.6992840293436821
0.7155905467144729
0.7469620303700519
0.8447939909637416
0.8276317714134559
0.7244280582365361
0.7030355697453063
0.5787154100989578
0.6507923037288098
0.6620100353940508
0.6806110581107527
0.6357654293898546
0.603507903675567
0.533819422461352
0.51225755915546
0.6529627837646746
0.6106445899269338
0.5519294145119399
0.5649292596466875
0.5695627984753262
0.45494305112374406
0.5299588903476419
Train Loss = 0.605, Train Accuracy = 0.782, Test Loss = 0.702, Test Accuracy = 0.769

```

```

[44]: learning_rate = 0.01
      learning_rates.append(learning_rate)

      hidden_size = 128
      ws = initialize_ws(hidden_size, n_class)
      bs = initialize_bs(hidden_size, n_class)

      _ = compute_logs(ws, bs, reg_param, verbose=True)

      ws, bs = sgd_one_pass(ws, bs, X_train, y_train, reg_param, learning_rate,
                             verbose=True)
      _ = compute_logs(ws, bs, reg_param, verbose=True)

```

```

Train Loss = 2.303, Train Accuracy = 0.078, Test Loss = 2.303, Test Accuracy = 0.076
1.4615504592571429
1.993241957408438

```

2.190435011942983
2.255443778915452
2.2705396121949706
2.0943142108643342
1.8783420196076466
1.660605759589135
1.476659261073228
1.3658297285344978
1.2922629823781664
1.1778248325790857
1.105667316084556
0.9671811102048473
0.8938494739315878
0.9126096957072981
0.763845309934718
0.8938048575197636
0.884066756582517
0.7949466909522835
0.7320578506948233
0.8341715149008695
0.7620635708638497
0.6729750230768966
0.6679829576279736
0.5775086873355579
0.6418161574479437
0.7701622527206765
0.8905306161587581
0.8510803330232642
0.753822768723063
0.7323974540171412
0.6052681837446096
0.5748896634061669
0.5410780550253043
0.5648009314544795
0.6367358209995782
0.663195898420038
0.5181604805123099
0.5913963662011166
0.65191183209491
0.7778110062106462
0.6129380100725589
0.6707763308776411
0.5895340868383555
0.6527199301089447
0.4628614797517581
0.5577811932761737
0.5869805070695948
0.64239434645619

```

0.6594149681827
0.6286934946865494
0.6594818671537055
0.6294672840021909
0.5595705346993586
0.6877459236954419
0.6380172257201256
0.650261022407181
0.6219016863575447
0.6275857844694599
Train Loss = 0.515, Train Accuracy = 0.817, Test Loss = 0.624, Test Accuracy =
0.789

```

```

[45]: learning_rate = 0.01
learning_rates.append(learning_rate)

hidden_size = 512
ws = initialize_ws(hidden_size, n_class)
bs = initialize_bs(hidden_size, n_class)

_ = compute_logs(ws, bs, reg_param, verbose=True)

ws, bs = sgd_one_pass(ws, bs, X_train, y_train, reg_param, learning_rate,
    verbose=True)
_ = compute_logs(ws, bs, reg_param, verbose=True)

```

```

Train Loss = 2.303, Train Accuracy = 0.126, Test Loss = 2.303, Test Accuracy =
0.126
1.4597177920234474
1.9940502134518898
2.190872995862269
2.2585039596859384
2.2481070603276105
2.083544451197122
1.842461363755616
1.5807559849977881
1.5253958142286648
1.4443802453906855
1.2885429029203634
1.1424751591067674
1.0085218783668486
0.8590315342198536
0.7867350124903643
0.7553873750760918
0.7960617987758482
0.6999666190350569
0.8159961693810619
0.6760270732423118

```

0.7381686205686209
0.8338934980382535
0.7726005496909174
0.7737335917326233
0.8133573589896871
0.7540432487659275
0.6499986343472077
0.5741427273421201
0.6076657660052898
0.5063845632424301
0.6403783771069719
0.6160596527765757
0.7716337698911755
0.6116327248176433
0.6036100506602174
0.5945839970813396
0.6520596895284186
0.5566870498443773
0.491207179156509
0.6340719264201714
0.5213239625437053
0.6371497632453987
0.6881954979101522
0.5805454977857449
0.5538171840936916
0.5782052051292756
0.7328340628331363
0.7876827809270361
0.8241044240483807
0.623832811313575
0.6648442290002811
0.6390042080984265
0.5836529765878936
0.6930553653200157
0.7104057293437039
0.5234683925138255
0.551695162031812
0.5253046209964247
0.5413307277072674
0.568943974799053
Train Loss = 0.524, Train Accuracy = 0.817, Test Loss = 0.625, Test Accuracy = 0.796

```
[46]: learning_rate = 0.01
      learning_rates.append(learning_rate)

      hidden_size = 2048
```

```

ws = initialize_ws(hidden_size, n_class)
bs = initialize_bs(hidden_size, n_class)

_ = compute_logs(ws, bs, reg_param, verbose=True)

ws, bs = sgd_one_pass(ws, bs, X_train, y_train, reg_param, learning_rate,
↳ verbose=True)
_ = compute_logs(ws, bs, reg_param, verbose=True)

```

Train Loss = 2.303, Train Accuracy = 0.109, Test Loss = 2.303, Test Accuracy = 0.104

1.4576087718447308
1.9955924440802584
2.18941118636607
2.2518131439116575
2.104809765034614
1.8284063341332026
1.6012037905094947
1.4505866679854038
1.2702324828123666
1.093306054408796
1.0280606633569631
1.0477610348092168
1.0478427626107387
0.9047132104560268
0.777297268106388
0.762789229678752
0.8344420961062277
0.8793097110318278
0.7492191442031185
0.541014008323092
0.7649644099954688
0.8278679053239989
0.7244995259122148
0.7744482149184827
0.7516112766193254
0.6605136349838191
0.5452872280841438
0.5592623373339352
0.622515098381035
0.6505740423704861
0.6235394742942802
0.6820165318358403
0.5999144811476509
0.6157584757337607
0.5291017899188006
0.4938435750233573

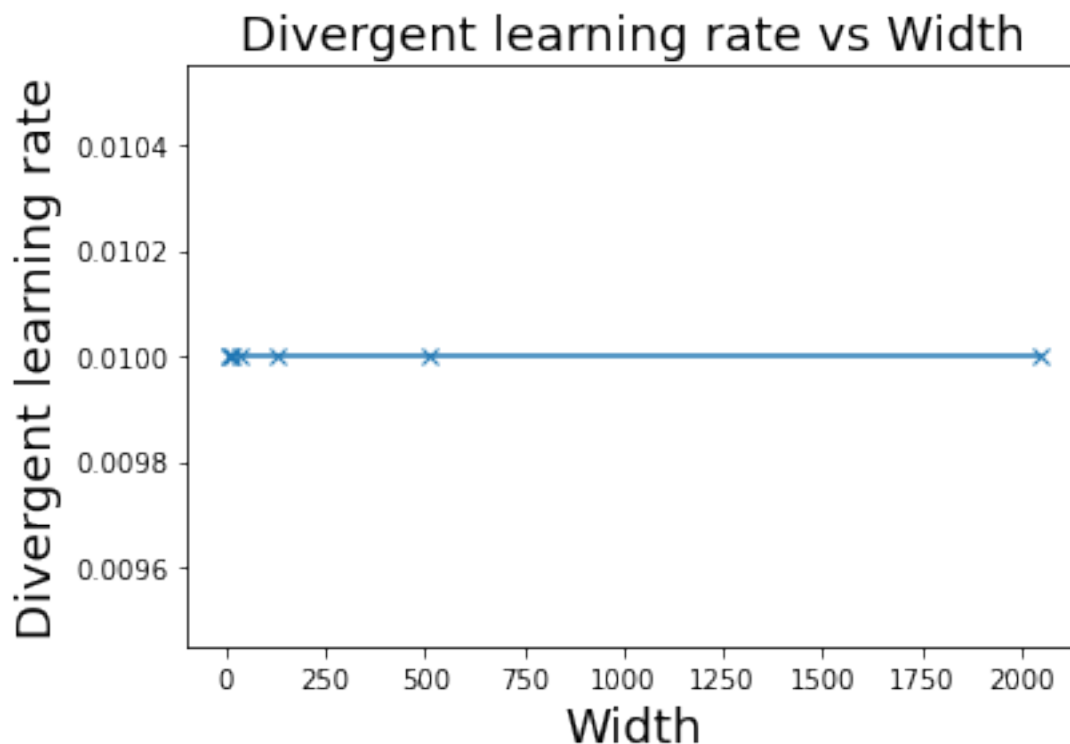
```
0.5395718730974007
0.6090631946586
0.633781475399923
0.6080900613764938
0.6475392315504407
0.5737953565020454
0.5946843518509863
0.6877970313363403
0.595222965779592
0.6182343199745135
0.4442231472680893
0.5400556946961178
0.5387641898334208
0.5111999651315541
0.4307242054141669
0.5716935370853947
0.582144315521034
0.5699694170609325
0.43723611974849735
0.6819603531711842
0.6284796309661069
0.6812728482712171
0.6002593525268825
0.7042676526305351
Train Loss = 0.644, Train Accuracy = 0.785, Test Loss = 0.748, Test Accuracy =
0.771
```

```
[51]: f = plt.figure()
      ax = f.gca()

      ax.set_title('Divergent learning rate vs Width', fontsize=18)
      ax.set_ylabel('Divergent learning rate', fontsize=18)
      ax.set_xlabel('Width', fontsize=18)

      ax.plot(hidden_sizes, learning_rates, 'x-')
      # ax.set_xticks(hidden_sizes)
```

```
[51]: [<matplotlib.lines.Line2D at 0x7ff5af8f8850>]
```



2.2 For a given width h , run SGD for 1 pass over the data with learning rate $[*h, *h/2, *h/4, *h/8]$, where $*h$ is from Part 2.1.

Measure the test accuracy for each of these learning rates. Let A_h denote the best accuracy obtained here. Repeat this procedure for each of the widths considered in Part 2.1.

```
[62]: divisors = [1, 2, 4, 8]
num_passes = 1

if not os.path.exists(MLP_Q2_RESULTS_FILE):
    hidden_size_lr_logs = {}
    for i, hidden_size in enumerate(hidden_sizes):
        hidden_size_lr_logs[hidden_size] = {}
        learning_rate = learning_rates[i]

        for divisor in divisors:
            lr = learning_rate / divisor
            print(f'Hidden Size: {hidden_size}, Learning Rate: {lr}')
            hidden_size_lr_logs[hidden_size][lr] = []

        ws = initialize_ws(hidden_size, n_class)
        bs = initialize_bs(hidden_size, n_class)
```

```

        # hidden_size_lr_logs[hidden_size][lr].append(compute_logs(ws, bs,
        ↪ reg_param, verbose=False))
        compute_logs(ws, bs, reg_param, verbose=False)

        for j in range(num_passes):
            ws, bs = sgd_one_pass(ws, bs, X_train, y_train, reg_param, lr,
            ↪ verbose=False)
            hidden_size_lr_logs[hidden_size][lr].append(compute_logs(ws,
            ↪ bs, reg_param, verbose=False))

```

```

Hidden Size: 4, Learning Rate: 0.01
Hidden Size: 4, Learning Rate: 0.005
Hidden Size: 4, Learning Rate: 0.0025
Hidden Size: 4, Learning Rate: 0.00125
Hidden Size: 8, Learning Rate: 0.01
Hidden Size: 8, Learning Rate: 0.005
Hidden Size: 8, Learning Rate: 0.0025
Hidden Size: 8, Learning Rate: 0.00125
Hidden Size: 16, Learning Rate: 0.01
Hidden Size: 16, Learning Rate: 0.005
Hidden Size: 16, Learning Rate: 0.0025
Hidden Size: 16, Learning Rate: 0.00125
Hidden Size: 32, Learning Rate: 0.01
Hidden Size: 32, Learning Rate: 0.005
Hidden Size: 32, Learning Rate: 0.0025
Hidden Size: 32, Learning Rate: 0.00125
Hidden Size: 128, Learning Rate: 0.01
Hidden Size: 128, Learning Rate: 0.005
Hidden Size: 128, Learning Rate: 0.0025
Hidden Size: 128, Learning Rate: 0.00125
Hidden Size: 512, Learning Rate: 0.01
Hidden Size: 512, Learning Rate: 0.005
Hidden Size: 512, Learning Rate: 0.0025
Hidden Size: 512, Learning Rate: 0.00125
Hidden Size: 2048, Learning Rate: 0.01
Hidden Size: 2048, Learning Rate: 0.005
Hidden Size: 2048, Learning Rate: 0.0025
Hidden Size: 2048, Learning Rate: 0.00125

```

```

[75]: # Find all the test accuries for each lr
hidden_size_lr_logs_test_acc = {}
for h, lr_d in hidden_size_lr_logs.items():
    hidden_size_lr_logs_test_acc[h] = {}
    for lr, vals in lr_d.items():
        hidden_size_lr_logs_test_acc[h][lr] = vals[0][3].item()

```

Test accuracy for each hidden size and learning rate value


```
[76]: pp(hidden_size_lr_logs_test_acc)
```

```
{4: {0.00125: 0.28220000863075256,
      0.0025: 0.4514999985694885,
      0.005: 0.703000009059906,
      0.01: 0.6636000275611877},
 8: {0.00125: 0.3702999949455261,
      0.0025: 0.6700999736785889,
      0.005: 0.7125999927520752,
      0.01: 0.7164999842643738},
16: {0.00125: 0.35910001397132874,
      0.0025: 0.6927000284194946,
      0.005: 0.7829999923706055,
      0.01: 0.7782999873161316},
32: {0.00125: 0.35749998688697815,
      0.0025: 0.7128000259399414,
      0.005: 0.7821999788284302,
      0.01: 0.7835999727249146},
128: {0.00125: 0.41190001368522644,
       0.0025: 0.7190999984741211,
       0.005: 0.7860000133514404,
       0.01: 0.7630000114440918},
512: {0.00125: 0.4494999945163727,
       0.0025: 0.7433000206947327,
       0.005: 0.8004999756813049,
       0.01: 0.7759000062942505},
2048: {0.00125: 0.5026999711990356,
        0.0025: 0.7515000104904175,
        0.005: 0.7896000146865845,
        0.01: 0.730400025844574}}
```

2.3 Make a plot of the best test accuracy A_h at the end of one pass over the data as the width h is varied.

```
[81]: # find the lr with the max test accuracy
hidden_size_lr_logs_test_acc_max = {}
for h, lr_d in hidden_size_lr_logs_test_acc.items():
    max_key = max(lr_d, key=lr_d.get)
    hidden_size_lr_logs_test_acc_max[h] = lr_d[max_key]
```

Max accuracy for each hidden size

```
[82]: pp(hidden_size_lr_logs_test_acc_max)
```

```
{4: 0.703000009059906,
 8: 0.7164999842643738,
16: 0.7829999923706055,
32: 0.7835999727249146,
```

```
128: 0.7860000133514404,  
512: 0.8004999756813049,  
2048: 0.7896000146865845}
```

```
[86]: f = plt.figure()  
      ax = f.gca()  
  
      ax.set_title('Divergent learning rate vs Width', fontsize=18)  
      ax.set_ylabel('Test Accuracy', fontsize=18)  
      ax.set_xlabel('Width', fontsize=18)  
  
      ax.plot(hidden_sizes, list(hidden_size_lr_logs_test_acc_max.values()), 'x-')
```

```
[86]: [<matplotlib.lines.Line2D at 0x7ff5af752b80>]
```

