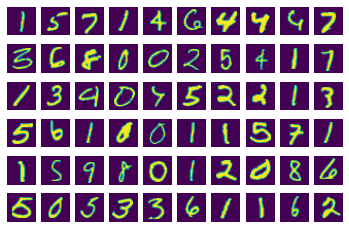
2.1

2.2



2.1 Changing loss function

a) **The negative log likelihood loss**

input\_size = 784

hidden\_sizes = [128, 64]

output\_size = 10

Sequential(

(0): Linear(in\_features=784, out\_features=128, bias=True)

(1): ReLU()

(2): Linear(in\_features=128, out\_features=64, bias=True)

(3): ReLU()

(4): Linear(in\_features=64, out\_features=10, bias=True)

(5): LogSoftmax(dim=1)

)

0s

print('Before backward pass: \n', model[0].weight.grad)

loss.backward()

print('After backward pass: \n', model[0].weight.grad)

Before backward pass:

None

After backward pass:

tensor([[ 1.2498e-03, 1.2498e-03, 1.2498e-03, ..., 1.2498e-03,

1.2498e-03, 1.2498e-03],

[-3.0802e-06, -3.0802e-06, -3.0802e-06, ..., -3.0802e-06,

-3.0802e-06, -3.0802e-06],

[ 2.0777e-03, 2.0777e-03, 2.0777e-03, ..., 2.0777e-03,

2.0777e-03, 2.0777e-03],

...,

[ 5.5131e-04, 5.5131e-04, 5.5131e-04, ..., 5.5131e-04,

5.5131e-04, 5.5131e-04],

[ 9.7681e-04, 9.7681e-04, 9.7681e-04, ..., 9.7681e-04,

9.7681e-04, 9.7681e-04],

[-3.3201e-03, -3.3201e-03, -3.3201e-03, ..., -3.3201e-03,

-3.3201e-03, -3.3201e-03]])

Epoch 0 - Training loss: 0.6456932500163629

Epoch 1 - Training loss: 0.28293575715027386

Epoch 2 - Training loss: 0.21947987223945575

Epoch 3 - Training loss: 0.1744409932502742

Epoch 4 - Training loss: 0.14630308691888794

Epoch 5 - Training loss: 0.12601863889138837

Epoch 6 - Training loss: 0.11006225234985764

Epoch 7 - Training loss: 0.09740535398202538

Epoch 8 - Training loss: 0.08776471037556653

Epoch 9 - Training loss: 0.08028588705767653

Epoch 10 - Training loss: 0.07204085356567794

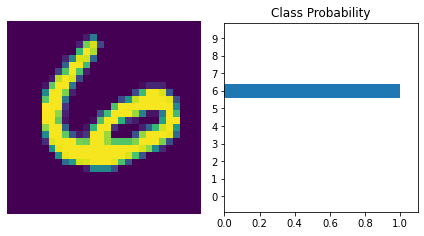
Epoch 11 - Training loss: 0.06585739059214875

Epoch 12 - Training loss: 0.06001676960082761

Epoch 13 - Training loss: 0.05645040457801167

Epoch 14 - Training loss: 0.05297171393893897

Training Time (in minutes) = 2.844489872455597



Number Of Images Tested = 10000

Model Accuracy = 0.9739

b. **Cross Entropy loss**

Sequential(

(0): Linear(in\_features=784, out\_features=128, bias=True)

(1): ReLU()

(2): Linear(in\_features=128, out\_features=64, bias=True)

(3): ReLU()

(4): Linear(in\_features=64, out\_features=10, bias=True)

(5): LogSoftmax(dim=1)

)

Before backward pass:

None

After backward pass:

tensor([[ 3.2894e-03, 3.2894e-03, 3.2894e-03, ..., 3.2894e-03,

3.2894e-03, 3.2894e-03],

[-2.6265e-03, -2.6265e-03, -2.6265e-03, ..., -2.6265e-03,

-2.6265e-03, -2.6265e-03],

[ 2.3543e-04, 2.3543e-04, 2.3543e-04, ..., 2.3543e-04,

2.3543e-04, 2.3543e-04],

...,

[ 1.4618e-04, 1.4618e-04, 1.4618e-04, ..., 1.4618e-04,

1.4618e-04, 1.4618e-04],

[-3.9080e-06, -3.9080e-06, -3.9080e-06, ..., -3.9080e-06,

-3.9080e-06, -3.9080e-06],

[-3.1418e-03, -3.1418e-03, -3.1418e-03, ..., -3.1418e-03,

-3.1418e-03, -3.1418e-03]])

Epoch 0 - Training loss: 0.6297319704519787

Epoch 1 - Training loss: 0.2848597545105257

Epoch 2 - Training loss: 0.22586333360086117

Epoch 3 - Training loss: 0.1793513515316792

Epoch 4 - Training loss: 0.15040577223488708

Epoch 5 - Training loss: 0.1275358759124936

Epoch 6 - Training loss: 0.11116498099294489

Epoch 7 - Training loss: 0.09788053657418724

Epoch 8 - Training loss: 0.08912354648442887

Epoch 9 - Training loss: 0.07904308137180073

Epoch 10 - Training loss: 0.07188303624021648

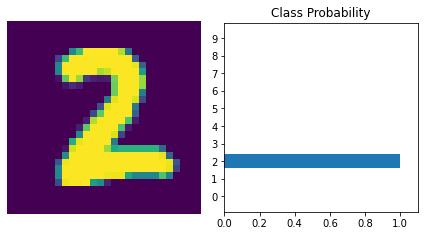
Epoch 11 - Training loss: 0.06568696206288614

Epoch 12 - Training loss: 0.06145310521459402

Epoch 13 - Training loss: 0.0559332811368951

Epoch 14 - Training loss: 0.05127944200828886

Training Time (in minutes) = 3.2410942673683167



Number Of Images Tested = 10000

Model Accuracy = 0.9771

ii. **Change in learning rate**

1. LR= 0.1

Sequential(

(0): Linear(in\_features=784, out\_features=128, bias=True)

(1): ReLU()

(2): Linear(in\_features=128, out\_features=64, bias=True)

(3): ReLU()

(4): Linear(in\_features=64, out\_features=10, bias=True)

(5): LogSoftmax(dim=1)

)

Before backward pass:

None

After backward pass:

tensor([[ 0.0023, 0.0023, 0.0023, ..., 0.0023, 0.0023, 0.0023],

[-0.0003, -0.0003, -0.0003, ..., -0.0003, -0.0003, -0.0003],

[-0.0002, -0.0002, -0.0002, ..., -0.0002, -0.0002, -0.0002],

...,

[-0.0056, -0.0056, -0.0056, ..., -0.0056, -0.0056, -0.0056],

[ 0.0001, 0.0001, 0.0001, ..., 0.0001, 0.0001, 0.0001],

[ 0.0006, 0.0006, 0.0006, ..., 0.0006, 0.0006, 0.0006]])

Epoch 0 - Training loss: 0.41664912060760995

Epoch 1 - Training loss: 0.17303483619261334

Epoch 2 - Training loss: 0.12815656289974572

Epoch 3 - Training loss: 0.1018889129493457

Epoch 4 - Training loss: 0.0869689962624618

Epoch 5 - Training loss: 0.07363336860561676

Epoch 6 - Training loss: 0.06482651721495492

Epoch 7 - Training loss: 0.05877277136557916

Epoch 8 - Training loss: 0.05020272468586814

Epoch 9 - Training loss: 0.047716485159193626

Epoch 10 - Training loss: 0.04172822655347893

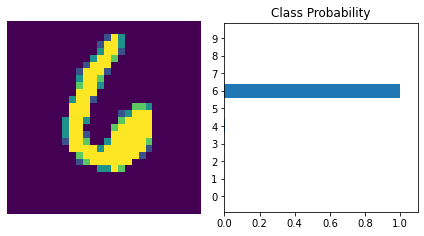
Epoch 11 - Training loss: 0.03751989115750157

Epoch 12 - Training loss: 0.03324498401022455

Epoch 13 - Training loss: 0.029393297238416893

Epoch 14 - Training loss: 0.02993799018490412

Training Time (in minutes) = 3.2686506191889446



Number Of Images Tested = 10000

Model Accuracy = 0.9749

1. R =0.45

Sequential(

(0): Linear(in\_features=784, out\_features=128, bias=True)

(1): ReLU()

(2): Linear(in\_features=128, out\_features=64, bias=True)

(3): ReLU()

(4): Linear(in\_features=64, out\_features=10, bias=True)

(5): LogSoftmax(dim=1)

)

Before backward pass:

None

After backward pass:

tensor([[ 0.0000, 0.0000, 0.0000, ..., 0.0000, 0.0000, 0.0000],

[ 0.0004, 0.0004, 0.0004, ..., 0.0004, 0.0004, 0.0004],

[-0.0005, -0.0005, -0.0005, ..., -0.0005, -0.0005, -0.0005],

...,

[-0.0009, -0.0009, -0.0009, ..., -0.0009, -0.0009, -0.0009],

[-0.0022, -0.0022, -0.0022, ..., -0.0022, -0.0022, -0.0022],

[-0.0065, -0.0065, -0.0065, ..., -0.0065, -0.0065, -0.0065]])

Epoch 0 - Training loss: 2.4741359102700566

Epoch 1 - Training loss: 2.317003215553918

Epoch 2 - Training loss: 2.31789896076422

Epoch 3 - Training loss: 2.317745109356797

Epoch 4 - Training loss: 2.317470037352556

Epoch 5 - Training loss: 2.3166150139338932

Epoch 6 - Training loss: 2.3159745156383718

Epoch 7 - Training loss: 2.3166307320218604

Epoch 8 - Training loss: 2.3175774789822383

Epoch 9 - Training loss: 2.316831034383794

Epoch 10 - Training loss: 2.3177149727908786

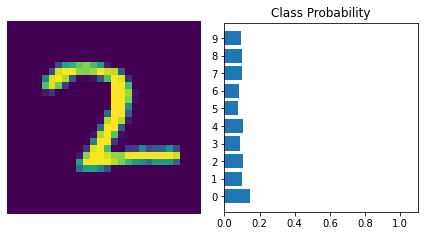
Epoch 11 - Training loss: 2.3181167523235655

Epoch 12 - Training loss: 2.3182983525526293

Epoch 13 - Training loss: 2.318401521202852

Epoch 14 - Training loss: 2.317895726100214

Training Time (in minutes) = 3.4732287605603536



Number Of Images Tested = 10000

Model Accuracy = 0.098

**iii. Change in Number of hidden layers**

1. Hidden\_sizes = [256, 128]

Sequential(

(0): Linear(in\_features=784, out\_features=256, bias=True)

(1): ReLU()

(2): Linear(in\_features=256, out\_features=128, bias=True)

(3): ReLU()

(4): Linear(in\_features=128, out\_features=10, bias=True)

(5): LogSoftmax(dim=1)

)

Before backward pass:

None

After backward pass:

tensor([[ 0.0000e+00, 0.0000e+00, 0.0000e+00, ..., 0.0000e+00,

0.0000e+00, 0.0000e+00],

[-1.5278e-03, -1.5278e-03, -1.5278e-03, ..., -1.5278e-03,

-1.5278e-03, -1.5278e-03],

[ 0.0000e+00, 0.0000e+00, 0.0000e+00, ..., 0.0000e+00,

0.0000e+00, 0.0000e+00],

...,

[-5.6990e-05, -5.6990e-05, -5.6990e-05, ..., -5.6990e-05,

-5.6990e-05, -5.6990e-05],

[-1.2500e-03, -1.2500e-03, -1.2500e-03, ..., -1.2500e-03,

-1.2500e-03, -1.2500e-03],

[-9.3862e-04, -9.3862e-04, -9.3862e-04, ..., -9.3862e-04,

-9.3862e-04, -9.3862e-04]])

Epoch 0 - Training loss: 0.6098155913068287

Epoch 1 - Training loss: 0.2754661773702801

Epoch 2 - Training loss: 0.21043347916814056

Epoch 3 - Training loss: 0.16618681483383754

Epoch 4 - Training loss: 0.13627917203567683

Epoch 5 - Training loss: 0.11587918775358688

Epoch 6 - Training loss: 0.10123397137270744

Epoch 7 - Training loss: 0.08905687926412582

Epoch 8 - Training loss: 0.07887759042391454

Epoch 9 - Training loss: 0.07171642368506473

Epoch 10 - Training loss: 0.06447171453865114

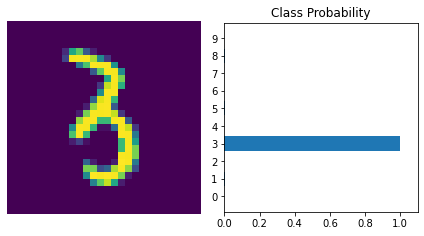
Epoch 11 - Training loss: 0.05793842594466929

Epoch 12 - Training loss: 0.0528094763451788

Epoch 13 - Training loss: 0.048904169123492706

Epoch 14 - Training loss: 0.04487180012116419

Training Time (in minutes) = 3.6403919061024985



Number Of Images Tested = 10000

Model Accuracy = 0.977

Classwise Accuracy Score: [0.9966 0.9975 0.9953 0.9961 0.9948 0.996 0.9948 0.9943 0.995 0.9936]

b. **HIDDEN\_SIZES**= [512,256]

Sequential(

(0): Linear(in\_features=784, out\_features=512, bias=True)

(1): ReLU()

(2): Linear(in\_features=512, out\_features=256, bias=True)

(3): ReLU()

(4): Linear(in\_features=256, out\_features=10, bias=True)

(5): LogSoftmax(dim=1)

)

Before backward pass:

None

After backward pass:

tensor([[ 1.5961e-03, 1.5961e-03, 1.5961e-03, ..., 1.5961e-03,

1.5961e-03, 1.5961e-03],

[ 1.0158e-03, 1.0158e-03, 1.0158e-03, ..., 1.0158e-03,

1.0158e-03, 1.0158e-03],

[-7.1420e-04, -7.1420e-04, -7.1420e-04, ..., -7.1420e-04,

-7.1420e-04, -7.1420e-04],

...,

[-8.5954e-04, -8.5954e-04, -8.5954e-04, ..., -8.5954e-04,

-8.5954e-04, -8.5954e-04],

[ 1.9038e-05, 1.9038e-05, 1.9038e-05, ..., 1.9038e-05,

1.9038e-05, 1.9038e-05],

[-1.8606e-03, -1.8606e-03, -1.8606e-03, ..., -1.8606e-03,

-1.8606e-03, -1.8606e-03]])

Epoch 0 - Training loss: 0.5930933729926152

Epoch 1 - Training loss: 0.2669196735535349

Epoch 2 - Training loss: 0.20110233723005252

Epoch 3 - Training loss: 0.16129433003458768

Epoch 4 - Training loss: 0.13310488877988763

Epoch 5 - Training loss: 0.1123019320084088

Epoch 6 - Training loss: 0.09689417471890764

Epoch 7 - Training loss: 0.08529497721770615

Epoch 8 - Training loss: 0.07515561789460877

Epoch 9 - Training loss: 0.06623732568924504

Epoch 10 - Training loss: 0.05923827055708638

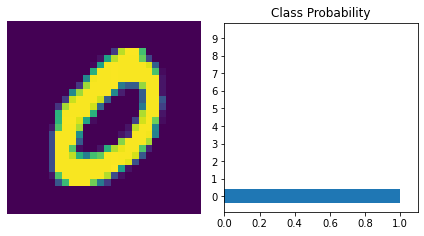
Epoch 11 - Training loss: 0.053927106782297556

Epoch 12 - Training loss: 0.04769227051708315

Epoch 13 - Training loss: 0.04351733108562455

Epoch 14 - Training loss: 0.03788999579092269

Training Time (in minutes) = 4.592447829246521



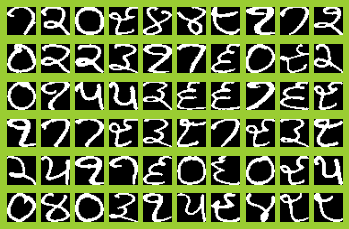
Number Of Images Tested = 10000

Model Accuracy = 0.9779

**2.3**

torch.Size([64, 1, 32, 32])

torch.Size([64])



Sequential(

(0): Linear(in\_features=1024, out\_features=128, bias=True)

(1): ReLU()

(2): Linear(in\_features=128, out\_features=64, bias=True)

(3): ReLU()

(4): Linear(in\_features=64, out\_features=10, bias=True)

(5): LogSoftmax(dim=1)

)

Before backward pass:

None

After backward pass:

tensor([[-5.3787e-04, -7.6704e-04, 1.2724e-05, ..., -2.8915e-03,

1.4057e-03, 3.8941e-03],

[ 2.1727e-03, 2.6997e-03, 3.0847e-03, ..., 6.8069e-03,

5.9685e-03, 5.0272e-03],

[ 1.3529e-03, 1.3529e-03, 1.3529e-03, ..., 1.0969e-03,

1.8658e-03, 1.5373e-03],

...,

[ 1.1308e-03, 1.1308e-03, 1.8279e-03, ..., -6.0356e-03,

-7.9852e-03, -6.1157e-03],

[-1.8019e-03, -1.8019e-03, -2.6233e-03, ..., -1.7803e-03,

-1.8019e-03, -1.8019e-03],

[-1.1698e-03, -1.1698e-03, -8.6464e-04, ..., -5.9030e-03,

-6.1789e-03, -5.0939e-03]])

Epoch 0 - Training loss: 2.1719422936439514

Epoch 1 - Training loss: 1.2746108956634998

Epoch 2 - Training loss: 0.3687690431252122

Epoch 3 - Training loss: 0.184416807256639

Epoch 4 - Training loss: 0.10426743910647929

Epoch 5 - Training loss: 0.08063909923657775

Epoch 6 - Training loss: 0.06083584774751216

Epoch 7 - Training loss: 0.038311883283313364

Epoch 8 - Training loss: 0.026431030710227787

Epoch 9 - Training loss: 0.01938108803005889

Epoch 10 - Training loss: 0.016824848571559414

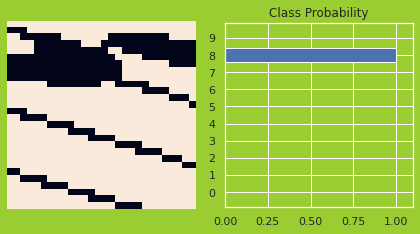
Epoch 11 - Training loss: 0.011113360989838839

Epoch 12 - Training loss: 0.009889105829643086

Epoch 13 - Training loss: 0.008325264323502779

Epoch 14 - Training loss: 0.007474763260688633

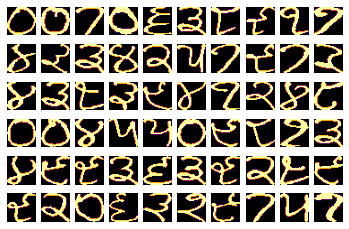
Training Time (in minutes) = 0.1390815774599711



Number Of Images Tested = 178

Model Accuracy = 0.9550561797752809

**Question 4.**



Linear Kernel: 0.9719101123595506

Polynomial Kernel: 0.9662921348314607

Radial Basis Kernel: 0.25280898876404495

Sigmoid Kernel: 0.9550561797752809

features = Linear kernel

Prediction of Gurumukhi Alphabet = 1

features = RBF kernel

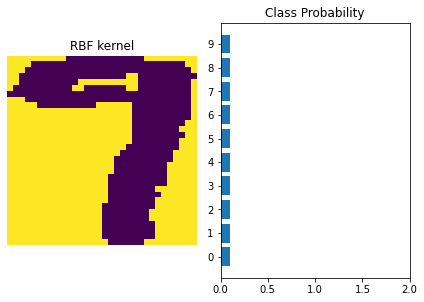
Prediction of Gurumukhi Alphabet = 2

features = Polynomial kernel

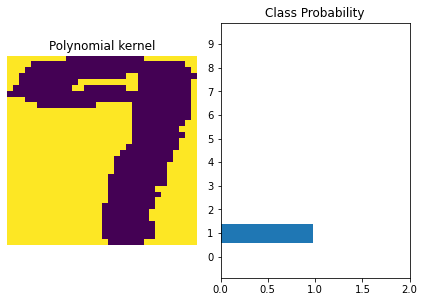
Prediction of Gurumukhi Alphabet = 1

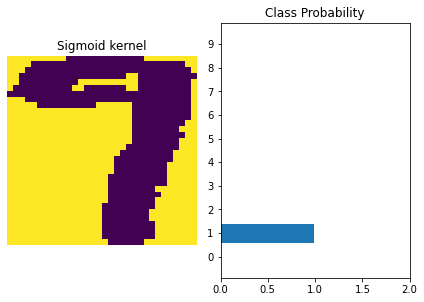
features = Sigmoid kernel

Prediction of Gurumukhi Alphabet = 1









Please Note

(Kindly note that the dataset used in the program are in Zip mode, you are requested to use zip format to load the program in 2.3 and 4)

Resources used:

1. <http://neuralnetworksanddeeplearning.com/chap2.html>
2. <http://neuralnetworksanddeeplearning.com/chap2.html>
3. <http://neuralnetworksanddeeplearning.com/chap2.html>
4. <https://towardsdatascience.com/mse-is-cross-entropy-at-heart-maximum-likelihood-estimation-explained-181a29450a0b>
5. <http://neuralnetworksanddeeplearning.com/chap2.html>
6. <https://vitalflux.com/mean-squared-error-vs-cross-entropy-loss-function/>
7. Youtube.com
8. NPTEL