ECE 763 Project 3 Report

April 29, 2022

1 Introduction

In this project, I will build up a four-layer fully connected neural network from scratch, including all the forward computation and backward propogation to update the network parameters.

We first import necessary packages,

```
[20]:  import numpy as np
       import h5py
       import matplotlib.pyplot as plt
       from sklearn import preprocessing
       from sklearn.model_selection import train_test_split
```

and define ReLu activation function, as well as its derivative.

```
[21]:  # ReLu
       def relu(x):
           x[x<0] = 0
           return x

       # ReLu prime
       def relu_p(x):
           x[x>0] = 1
           x[x<=0] = 0
           return x
```

Moreover, we consider normalization as data preprocessing such that the data has zero mean and unit variance.

```
[22]:  def data_preprossing(X_train):
           scaler = preprocessing.StandardScaler().fit(X_train)
           return scaler.transform(X_train)
```

Now we start explaining the model configuration with 2 hidden layers. Specifically, the network structure is given in the following figure:

Input layer $->$ 1st hidden layer with D_1 number of neurons $->$ Batch Normalization layer $->$ ReLu activation layer $->$ 2nd hidden layer with D_2 number of neurons $->$ Batch Normalization layer $->$ ReLu activation layer $->$ Output layer.

The weight of the model is initialized with Xavier method and updated according to a vanilla mini-batch SGD. Moreover, I make an option to allow the best model to be recorded.
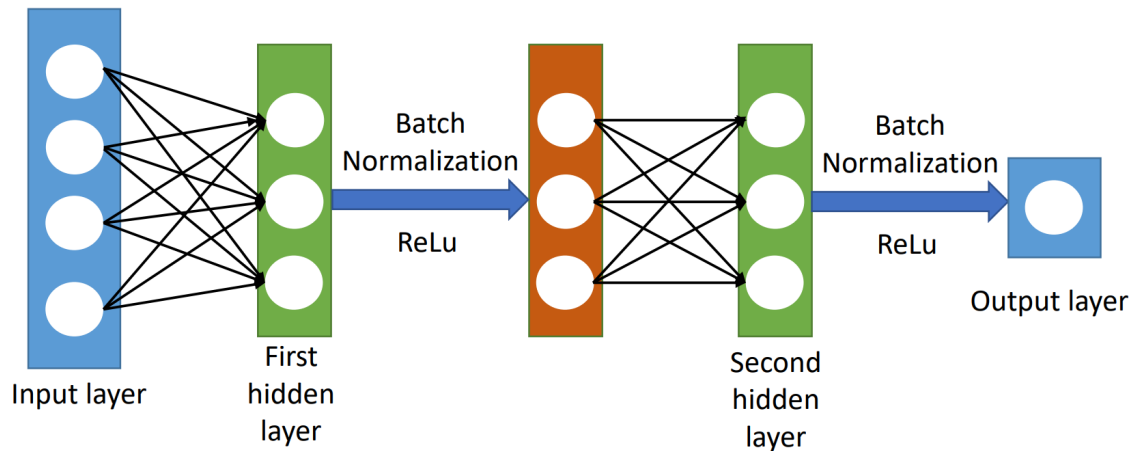

```
[23]: # The source code of batch normalization is referred to https://deepnotes.io/
      ↪batchnorm
      class Batchnorm():

          # X_dim is the dimension of each data point, par is the tuple of learnable
      ↪parameters [gamma, beta]
          def __init__(self,X_dim, par=None):
              self.params = par
              if self.params == None:
                  self.gamma = np.ones((1, int(X_dim)))
                  self.beta = np.zeros((1, int(X_dim)))
                  self.params = [self.gamma,self.beta]

          def forward(self,X):
              self.n_X = X.shape[0]
              self.X_shape = X.shape

              self.X_flat = X.ravel().reshape(self.n_X,-1)
              self.mu = np.mean(self.X_flat,axis=0)
              self.var = np.var(self.X_flat, axis=0)
              self.X_norm = (self.X_flat - self.mu)/np.sqrt(self.var + 1e-8)
              out = self.gamma * self.X_norm + self.beta

              return out.reshape(self.X_shape)

          def backward(self,dout):

              dout = dout.ravel().reshape(dout.shape[0],-1)
              X_mu = self.X_flat - self.mu
              var_inv = 1./np.sqrt(self.var + 1e-8)
```

```
        dbeta = np.sum(dout,axis=0)
        dgamma = (dout * self.X_norm).sum(axis=0)

        dX_norm = dout * self.gamma
        dvar = np.sum(dX_norm * X_mu,axis=0) * -0.5 * (self.var + 1e-8)**(-3/2)
        dmu = np.sum(dX_norm * -var_inv ,axis=0) + dvar * 1/self.n_X * np.
    →sum(-2.* X_mu, axis=0)
        dX = (dX_norm * var_inv) + (dmu / self.n_X) + (dvar * 2/self.n_X * X_mu)

        dX = dX.reshape(self.X_shape)
        return dX, dgamma, dbeta
```

[41]:
```
class four_layers_model():

    # initialize the number of neurons in each hidden layer
    def __init__(self, hidden1_size, hidden2_size):
        self.hidden1_size = hidden1_size
        self.hidden2_size = hidden2_size

    # function to create a list containing mini-batches
    def create_mini_batches(self, X, Y, batch_size):
        mini_batches = []
        data = np.hstack((X, Y))
        np.random.shuffle(data)
        n_minibatches = data.shape[0] // batch_size

        for i in range(n_minibatches):
            mini_batch = data[i * batch_size:(i + 1)*batch_size, :]
            X_mini = mini_batch[:, :-1]
            Y_mini = mini_batch[:, -1].reshape((-1, 1))
            mini_batches.append((X_mini, Y_mini))
        if data.shape[0] % batch_size != 0:
            mini_batch = data[n_minibatches * batch_size:data.shape[0], :]
            X_mini = mini_batch[:, :-1]
            Y_mini = mini_batch[:, -1].reshape((-1, 1))
            mini_batches.append((X_mini, Y_mini))
        return mini_batches



    def␣
    →train(self,X,Y,X_val,Y_val,epochs,h,batch_size,regulation,best_model=True,␣
    →verbose=True):

        self.input_dim =  len(X.T)
        self.output_dim = len(Y.T)
```

```python
        self.max_label = max(Y)
        self.min_label = min(Y)
        self.error_tolerance = regulation

        # store the result from each epoch
        self.accuracies = []
        self.losses = []
        self.val_accuracies = []
        self.val_losses = []

        # initialize the weight of each neuron with Xavier init
        np.random.seed(0)
        W1 = np.random.normal(loc=0.0, scale=(1/np.sqrt(len(X))),size=(self.
→input_dim, self.hidden1_size))
        b1 = np.zeros(shape=self.hidden1_size)
        W2 = np.random.normal(loc=0.0, scale=(1/np.sqrt(len(X))),size=(self.
→hidden1_size, self.hidden2_size))
        b2 = np.zeros(shape=self.hidden2_size)
        W3 = np.random.normal(loc=0.0, scale=(1/np.sqrt(len(X))),size=(self.
→hidden2_size, self.output_dim))
        b3 = np.zeros(shape=self.output_dim)
        batchnorm1 = Batchnorm(self.hidden1_size)
        batchnorm2 = Batchnorm(self.hidden2_size)


        for i in range(epochs):

            mini_batches = self.create_mini_batches(X, Y, batch_size)
            for mini_batch in mini_batches:
                N = len(mini_batch[0])
                # compute the loss
                z1 = np.dot(mini_batch[0], W1) + b1     # [N x hidden1_size]
                c1 = batchnorm1.forward(z1)             # [N x hidden1_size]
                y1 = relu(c1)                           # [N x hidden1_size]
                z2 = np.dot(y1, W2) + b2                # [N x hidden2_size]
                c2 = batchnorm2.forward(z2)             # [N x hidden2_size]
                y2 = relu(c2)                           # [N x hidden2_size]
                z3 = np.dot(y2, W3) + b3                # [N x output_dim]
                r = z3 - mini_batch[1]                  # [N x output_dim]
                L = np.dot(r.T, r)/2.0                  # [1 x 1]

                #
                # backward pass
                #
                # partial derivatives
                dLdr = r                                                  #␣
→[N x output_dim]
```

```
            drdz3 = 1                                                    #␣
↪[N x output_dim]
            dz3dW3 = y2                                                  #␣
↪[N x hidden2_size]
            dz3db3 = np.ones(shape = N)                                  #␣
↪[N, ]
            dz3dy2 = W3                                                  #␣
↪[hidden2_size x output_dim]
            dy2dc2 = relu_p(c2)                                          #␣
↪[N x hidden2_size]
            dc2dz2, dc2dgamma2, dc2dbeta2 = batchnorm2.backward(c2)      #␣
↪[N x hidden2_size], [hidden2_size, ], [hidden2_size, ]
            dz2dW2 = y1                                                  #␣
↪[N x hidden1_size]
            dz2db2 = np.ones(shape = N)                                  #␣
↪[N, ]
            dz2dy1 = W2                                                  #␣
↪[hidden1_size x hidden2_size]
            dy1dc1 = relu_p(c1)                                          #␣
↪[N x hidden1_size]
            dc1dz1, dc1dgamma1, dc1dbeta1 = batchnorm1.backward(c1)      #␣
↪[N x hidden1_size], [hidden1_size, ], [hidden1_size, ]
            dz1dW1 = mini_batch[0]                                       #␣
↪[N x input_dim]
            dz1db1 = np.ones(shape = N)                                  #␣
↪[N, ]


            # chain rule
            dLdW3 = np.dot(dLdr.T, dz3dW3).T
            dLdb3 = np.dot(dLdr.T, dz3db3)
            dLdW2 = np.dot((np.dot(dLdr,dz3dy2.T)*dy2dc2*dc2dz2).T, dz2dW2).
↪T                                          # dLdW2 =␣
↪dLdr*drdz3*dz3dy2*dy2dc2*dc2dz2*dz2dW2
            dLdb2 = np.dot((np.dot(dLdr,dz3dy2.T)*dy2dc2*dc2dz2).T, dz2db2)␣
↪                                           # dLdb2 =␣
↪dLdr*drdz3*dz3dy2*dy2dc2*dc2dz2*dz2db2
            dLdW1 = np.dot((np.dot((np.dot(dLdr,dz3dy2.T)*dy2dc2*dc2dz2),␣
↪dz2dy1.T)*dy1dc1*dc1dz1).T,dz1dW1).T        # dLdW1 =␣
↪dLdr*drdz3*dz3dy2*dy2dc2*dc2dz2*dz2dy1*dy1dc1*dc1dz1*dz1dW1
            dLdb1 = np.dot((np.dot((np.dot(dLdr,dz3dy2.T)*dy2dc2*dc2dz2),␣
↪dz2dy1.T)*dy1dc1*dc1dz1).T,dz1db1)          # dLdb1 =␣
↪dLdr*drdz3*dz3dy2*dy2dc2*dc2dz2*dz2dy1*dy1dc1*dc1dz1*dz1db1
            dLdgamma2 = np.dot(np.dot(np.dot(dLdr,dz3dy2.T).T,dy2dc2),␣
↪dc2dgamma2)                                 # dLdgamma2 =␣
↪dLdr*drdz3*dz3dy2*dy2dc2*dc2dgamma2
```

```python
            dLdbeta2 = np.dot(np.dot(np.dot(dLdr,dz3dy2.T).T,dy2dc2),
↪dc2dbeta2)                                      # dLdbeta2 =
↪dLdr*drdz3*dz3dy2*dy2dc2*dc2dbeta2
            dLdgamma1 = np.dot(np.dot(np.dot((np.dot(dLdr,dz3dy2.
↪T)*dy2dc2*dc2dz2), dz2dy1.T).T,dy1dc1), dc1dgamma1)  # dLdgamma2 =
↪dLdr*drdz3*dz3dy2*dy2dc2*dc2dz2*dz2dy1*dy1dc1*dc1dgamma1
            dLdbeta1 = np.dot(np.dot(np.dot((np.dot(dLdr,dz3dy2.
↪T)*dy2dc2*dc2dz2), dz2dy1.T).T,dy1dc1), dc1dbeta1)    # dLdbeta2 =
↪dLdr*drdz3*dz3dy2*dy2dc2*dc2dz2*dz2dy1*dy1dc1*dc1dbeta1


            # gradient update
            W3 = W3 - h*dLdW3
            b3 = b3 - h*dLdb3
            W2 = W2 - h*dLdW2
            b2 = b2 - h*dLdb2
            W1 = W1 - h*dLdW1
            b1 = b1 - h*dLdb1
            batchnorm2.gamma = batchnorm2.gamma - h*dLdgamma2
            batchnorm2.beta = batchnorm2.beta - h*dLdbeta2
            batchnorm1.gamma = batchnorm1.gamma - h*dLdgamma1
            batchnorm1.beta = batchnorm1.beta - h*dLdbeta1
            batchnorm2.params = [batchnorm2.gamma, batchnorm2.beta]
            batchnorm1.params = [batchnorm1.gamma, batchnorm1.beta]

        # compute the loss in training
        z1 = np.dot(X, W1) + b1                    # [N x hidden1_size]
        c1 = batchnorm1.forward(z1)                # [N x hidden1_size]
        y1 = relu(c1)                              # [N x hidden1_size]
        z2 = np.dot(y1, W2) + b2                   # [N x hidden2_size]
        c2 = batchnorm2.forward(z2)                # [N x hidden2_size]
        y2 = relu(c2)                              # [N x hidden2_size]
        z3 = np.dot(y2, W3) + b3                   # [N x output_dim]
        r = z3 - Y                                 # [N x output_dim]
        L = np.dot(r.T, r)/2.0                      # [1 x 1]


        # compute the loss in validation
        z1_val = np.dot(X_val, W1) + b1
        c1_val = batchnorm1.forward(z1_val)
        y1_val = relu(c1_val)
        z2_val = np.dot(y1_val, W2) + b2
        c2_val = batchnorm2.forward(z2_val)
        y2_val = relu(c2_val)
        z3_val = np.dot(y2_val, W3) + b3
        r_val = z3_val - Y_val
        L_val = np.dot(r_val.T, r_val)/2.0
```

```python
            # save the best model
            if best_model:
                if len(self.losses) == 0:

                    self.saved_par = [W1, b1, W2, b2, W3, b3]
                    self.batchnorm_par = [batchnorm1.params, batchnorm2.params]

                    # round the predictions to obtain class labels
                    predict = np.round(z3).clip(self.min_label, self.max_label)
                    predict_val = np.round(z3_val).clip(min(Y_val), max(Y_val))

                    # compute the error
                    error = float(np.count_nonzero(predict-Y)) / float(len(X))
                    error_val = float(np.count_nonzero(predict_val-Y_val)) /
→float(len(X_val))

                    # store the accuracy
                    self.accuracies.append(1.0 - error)
                    self.val_accuracies.append(1.0 - error_val)

                    # store the loss
                    self.losses.append(np.squeeze(L))
                    self.val_losses.append(np.squeeze(L_val))

                elif self.losses[-1] > np.squeeze(L):

                    self.saved_par = [W1, b1, W2, b2, W3, b3]
                    self.batchnorm_par = [batchnorm1.params, batchnorm2.params]

                    predict = np.round(z3).clip(self.min_label, self.max_label)
                    predict_val = np.round(z3_val).clip(min(Y_val), max(Y_val))

                    error = float(np.count_nonzero(predict-Y)) / float(len(X))
                    error_val = float(np.count_nonzero(predict_val-Y_val)) /
→float(len(X_val))

                    self.accuracies.append(1.0 - error)
                    self.val_accuracies.append(1.0 - error_val)

                    self.losses.append(np.squeeze(L))
                    self.val_losses.append(np.squeeze(L_val))

                else:
                    [W1,b1,W2,b2,W3,b3] = self.saved_par
                    [batchnorm1.params, batchnorm2.params] = self.batchnorm_par
```

```python
                self.accuracies.append(self.accuracies[-1])
                self.val_accuracies.append(self.val_accuracies[-1])
                self.losses.append(self.losses[-1])
                self.val_losses.append(self.val_losses[-1])
        else:
            self.saved_par = [W1, b1, W2, b2, W3, b3]
            self.batchnorm_par = [batchnorm1.params, batchnorm2.params]

            predict = np.round(z3).clip(self.min_label, self.max_label)
            predict_val = np.round(z3_val).clip(min(Y_val), max(Y_val))

            error = float(np.count_nonzero(predict-Y)) / float(len(X))
            error_val = float(np.count_nonzero(predict_val-Y_val)) /
→float(len(X_val))

            self.accuracies.append(1.0 - error)
            self.val_accuracies.append(1.0 - error_val)

            self.losses.append(np.squeeze(L))
            self.val_losses.append(np.squeeze(L_val))


        # print the training result of current epoch
        if verbose:
            print('Finishing epoch {}/{}: train loss: {}, train accuracy:
→{}, val accuracy: {}, lr: {}'.format(i+1,epochs,self.losses[-1],self.
→accuracies[-1],self.val_accuracies[-1],h))

            if self.losses[-1] < self.error_tolerance:
                print('Training ends early when the error satisfies the
→regularization.')
                break
            elif i == (epochs - 1):
                print('Training ends.')

    return self.accuracies, self.losses, self.val_accuracies, [self.
→saved_par,self.batchnorm_par]

  def prediction(self,X_test):

    [W1,b1,W2,b2,W3,b3] = self.saved_par

    batchnorm1 = Batchnorm(self.hidden1_size)
    batchnorm2 = Batchnorm(self.hidden2_size)
    [batchnorm1.gamma,batchnorm1.beta] = self.batchnorm_par[0]
    [batchnorm2.gamma,batchnorm2.beta] = self.batchnorm_par[1]
```

```
        # compute the loss in training
        z1 = np.dot(X_test, W1) + b1              # [N x neurons]
        c1 = batchnorm1.forward(z1)
        y1 = relu(c1)                             # [N x neurons]
        z2 = np.dot(y1, W2) + b2                  # [N x neurons]
        c2 = batchnorm2.forward(z2)
        y2 = relu(c2)                             # [N x neurons]
        z3 = np.dot(y2, W3) + b3                  # [N x output_dim]

        return np.round(z3).clip(self.min_label, self.max_label)
```

Parameter configuration. For this project, I use 50 neurons in each of the two hidden layers. I choose batch size to be 150 and step size 1e-5. I train the network for 200 epochs.

```
[25]: np.random.seed(0)
      epochs = 200 # total number of epochs
      neurons_1 = 50 # number of neurons in the first hidden layer
      neurons_2 = 50 # number of neurons in the second hidden layer
      batch_size = 150
      h = 1e-5 # step size for parameter update
```

Read the face dataset used from my project 1 and 2. Here I have an option to use normalization to preprocess the data before training. Then, I split the training dataset into two parts: training data and validation data with proportion 4:1.

```
[26]: # Read dataset
      with h5py.File('./dataset/train_dataset.h5', 'r') as file:
          data = np.array(file.get('data'))
          label = np.array(file.get('label'))
          X_train = data
          y_train = label

      with h5py.File('./dataset/test_dataset.h5', 'r') as file:
          data = np.array(file.get('data'))
          label = np.array(file.get('label'))
          X_test = data
          y_test = label

      is_data_prepossing = True
      # Data Preprossing
      if is_data_prepossing:
          X_train = data_preprossing(X_train)

      X_train, X_val, y_train, y_val = train_test_split(X_train, y_train, test_size=0.
       ↪25, random_state=66)
```

```
y_train = np.expand_dims(y_train, -1)
y_val = np.expand_dims(y_val, -1)
y_test = np.expand_dims(y_test, -1)

print(X_train.shape, y_train.shape, X_val.shape, y_val.shape, X_test.shape,␣
 →y_test.shape)
```

(1500, 1200) (1500, 1) (500, 1200) (500, 1) (200, 1200) (200, 1)

2 Step 1: First shot

I test the performance of my network and use data preprocessing. The best_model option is enabled and I choose learning rate to be 1e-5. I set regularization to be 0 for now.

```
[27]: model = four_layers_model(neurons_1, neurons_2)

      is_data_prepossing = True # set True if use data normalization
      regularization = 0.0

      # Read dataset
      with h5py.File('./dataset/train_dataset.h5', 'r') as file:
          data = np.array(file.get('data'))
          label = np.array(file.get('label'))
          X_train = data
          y_train = label

      with h5py.File('./dataset/test_dataset.h5', 'r') as file:
          data = np.array(file.get('data'))
          label = np.array(file.get('label'))
          X_test = data
          y_test = label

      # Data Preprossing
      if is_data_prepossing:
          X_train = data_preprossing(X_train)

      X_train, X_val, y_train, y_val = train_test_split(X_train, y_train, test_size=0.
       →25, random_state=66)

      y_train = np.expand_dims(y_train, -1)
      y_val = np.expand_dims(y_val, -1)
      y_test = np.expand_dims(y_test, -1)

      accuracies, losses, val_accuracies, par1 = model.
       →train(X_train,y_train,X_val,y_val,epochs,h,batch_size,regularization,best_model=True)
```

```
Finishing epoch 1/200: train loss: 216.57389976366974, train accuracy:
0.5173333333333333, val accuracy: 0.45999999999999996, lr: 1e-05
Finishing epoch 2/200: train loss: 194.27900596855577, train accuracy:
0.5473333333333333, val accuracy: 0.512, lr: 1e-05
Finishing epoch 3/200: train loss: 188.71878954987613, train accuracy:
0.5793333333333333, val accuracy: 0.5640000000000001, lr: 1e-05
Finishing epoch 4/200: train loss: 185.46207412110158, train accuracy:
0.5886666666666667, val accuracy: 0.5760000000000001, lr: 1e-05
Finishing epoch 5/200: train loss: 182.41419241439118, train accuracy:
0.5986666666666667, val accuracy: 0.5920000000000001, lr: 1e-05
Finishing epoch 6/200: train loss: 179.74246260492905, train accuracy:
0.6106666666666667, val accuracy: 0.594, lr: 1e-05
Finishing epoch 7/200: train loss: 177.39452907852703, train accuracy: 0.622,
val accuracy: 0.602, lr: 1e-05
Finishing epoch 8/200: train loss: 175.18870675840085, train accuracy:
0.6326666666666667, val accuracy: 0.614, lr: 1e-05
Finishing epoch 9/200: train loss: 173.4736208582077, train accuracy:
0.6406666666666667, val accuracy: 0.626, lr: 1e-05
Finishing epoch 10/200: train loss: 172.1447669315506, train accuracy:
0.6446666666666667, val accuracy: 0.628, lr: 1e-05
Finishing epoch 11/200: train loss: 171.02457591930641, train accuracy: 0.644,
val accuracy: 0.632, lr: 1e-05
Finishing epoch 12/200: train loss: 170.11559216103043, train accuracy:
0.6446666666666667, val accuracy: 0.626, lr: 1e-05
Finishing epoch 13/200: train loss: 169.21624215835897, train accuracy:
0.6466666666666667, val accuracy: 0.618, lr: 1e-05
Finishing epoch 14/200: train loss: 168.59777383820096, train accuracy:
0.6513333333333333, val accuracy: 0.622, lr: 1e-05
Finishing epoch 15/200: train loss: 168.09995224094257, train accuracy:
0.6493333333333333, val accuracy: 0.624, lr: 1e-05
Finishing epoch 16/200: train loss: 167.5892511529036, train accuracy:
0.6466666666666667, val accuracy: 0.628, lr: 1e-05
Finishing epoch 17/200: train loss: 167.30057730714708, train accuracy: 0.652,
val accuracy: 0.632, lr: 1e-05
Finishing epoch 18/200: train loss: 166.7747121574224, train accuracy:
0.6566666666666667, val accuracy: 0.634, lr: 1e-05
Finishing epoch 19/200: train loss: 166.43507176979227, train accuracy:
0.6593333333333333, val accuracy: 0.636, lr: 1e-05
Finishing epoch 20/200: train loss: 165.63962634269, train accuracy:
0.6599999999999999, val accuracy: 0.632, lr: 1e-05
Finishing epoch 21/200: train loss: 164.8406378970879, train accuracy:
0.6639999999999999, val accuracy: 0.638, lr: 1e-05
Finishing epoch 22/200: train loss: 163.9456624317781, train accuracy:
0.6639999999999999, val accuracy: 0.636, lr: 1e-05
Finishing epoch 23/200: train loss: 162.8688854528699, train accuracy:
0.6666666666666667, val accuracy: 0.632, lr: 1e-05
Finishing epoch 24/200: train loss: 161.90751775390498, train accuracy:
0.6699999999999999, val accuracy: 0.646, lr: 1e-05
```

Finishing epoch 25/200: train loss: 160.69243713530346, train accuracy: 0.6739999999999999, val accuracy: 0.646, lr: 1e-05
Finishing epoch 26/200: train loss: 159.29996617586545, train accuracy: 0.6766666666666667, val accuracy: 0.642, lr: 1e-05
Finishing epoch 27/200: train loss: 157.9775999837426, train accuracy: 0.6799999999999999, val accuracy: 0.648, lr: 1e-05
Finishing epoch 28/200: train loss: 156.69017171244838, train accuracy: 0.6846666666666666, val accuracy: 0.654, lr: 1e-05
Finishing epoch 29/200: train loss: 155.40073847972326, train accuracy: 0.6853333333333333, val accuracy: 0.656, lr: 1e-05
Finishing epoch 30/200: train loss: 154.08398630311575, train accuracy: 0.6893333333333334, val accuracy: 0.6639999999999999, lr: 1e-05
Finishing epoch 31/200: train loss: 152.77313943593055, train accuracy: 0.692, val accuracy: 0.6639999999999999, lr: 1e-05
Finishing epoch 32/200: train loss: 151.36205144228742, train accuracy: 0.696, val accuracy: 0.6679999999999999, lr: 1e-05
Finishing epoch 33/200: train loss: 150.08087259980596, train accuracy: 0.7, val accuracy: 0.6779999999999999, lr: 1e-05
Finishing epoch 34/200: train loss: 148.81299222717274, train accuracy: 0.7066666666666667, val accuracy: 0.6859999999999999, lr: 1e-05
Finishing epoch 35/200: train loss: 147.47416286464772, train accuracy: 0.712, val accuracy: 0.6839999999999999, lr: 1e-05
Finishing epoch 36/200: train loss: 146.19658078521965, train accuracy: 0.7153333333333334, val accuracy: 0.69, lr: 1e-05
Finishing epoch 37/200: train loss: 144.79779276601766, train accuracy: 0.7173333333333334, val accuracy: 0.696, lr: 1e-05
Finishing epoch 38/200: train loss: 143.50979267906598, train accuracy: 0.7233333333333334, val accuracy: 0.694, lr: 1e-05
Finishing epoch 39/200: train loss: 142.19149050973346, train accuracy: 0.7246666666666667, val accuracy: 0.698, lr: 1e-05
Finishing epoch 40/200: train loss: 140.9171431269292, train accuracy: 0.7273333333333334, val accuracy: 0.698, lr: 1e-05
Finishing epoch 41/200: train loss: 139.83282541489834, train accuracy: 0.73, val accuracy: 0.698, lr: 1e-05
Finishing epoch 42/200: train loss: 138.55953185706306, train accuracy: 0.734, val accuracy: 0.7, lr: 1e-05
Finishing epoch 43/200: train loss: 137.46323952283512, train accuracy: 0.738, val accuracy: 0.704, lr: 1e-05
Finishing epoch 44/200: train loss: 136.31086756667202, train accuracy: 0.7413333333333334, val accuracy: 0.704, lr: 1e-05
Finishing epoch 45/200: train loss: 135.0910060748654, train accuracy: 0.746, val accuracy: 0.71, lr: 1e-05
Finishing epoch 46/200: train loss: 133.85939513002467, train accuracy: 0.746, val accuracy: 0.71, lr: 1e-05
Finishing epoch 47/200: train loss: 132.76688552741214, train accuracy: 0.7526666666666667, val accuracy: 0.712, lr: 1e-05
Finishing epoch 48/200: train loss: 131.68845909333535, train accuracy: 0.7513333333333333, val accuracy: 0.716, lr: 1e-05

Finishing epoch 49/200: train loss: 130.8003533040271, train accuracy:
0.7573333333333333, val accuracy: 0.72, lr: 1e-05
Finishing epoch 50/200: train loss: 129.72631797337, train accuracy:
0.7573333333333333, val accuracy: 0.726, lr: 1e-05
Finishing epoch 51/200: train loss: 128.7905452419518, train accuracy:
0.7586666666666666, val accuracy: 0.734, lr: 1e-05
Finishing epoch 52/200: train loss: 127.87631421571413, train accuracy:
0.7606666666666666, val accuracy: 0.738, lr: 1e-05
Finishing epoch 53/200: train loss: 126.83250865735232, train accuracy:
0.7633333333333333, val accuracy: 0.734, lr: 1e-05
Finishing epoch 54/200: train loss: 125.83892267072099, train accuracy:
0.7646666666666666, val accuracy: 0.734, lr: 1e-05
Finishing epoch 55/200: train loss: 124.96111879538793, train accuracy:
0.7666666666666666, val accuracy: 0.736, lr: 1e-05
Finishing epoch 56/200: train loss: 124.20786911043501, train accuracy:
0.7693333333333333, val accuracy: 0.742, lr: 1e-05
Finishing epoch 57/200: train loss: 123.26432301317959, train accuracy:
0.7693333333333333, val accuracy: 0.74, lr: 1e-05
Finishing epoch 58/200: train loss: 122.54724270746233, train accuracy: 0.772,
val accuracy: 0.746, lr: 1e-05
Finishing epoch 59/200: train loss: 121.61444497672741, train accuracy: 0.774,
val accuracy: 0.742, lr: 1e-05
Finishing epoch 60/200: train loss: 120.91567804551595, train accuracy:
0.7746666666666666, val accuracy: 0.744, lr: 1e-05
Finishing epoch 61/200: train loss: 120.14627126890927, train accuracy:
0.7766666666666666, val accuracy: 0.744, lr: 1e-05
Finishing epoch 62/200: train loss: 119.53662016936924, train accuracy:
0.7826666666666666, val accuracy: 0.752, lr: 1e-05
Finishing epoch 63/200: train loss: 118.8698973936093, train accuracy: 0.788,
val accuracy: 0.754, lr: 1e-05
Finishing epoch 64/200: train loss: 118.06450031235104, train accuracy: 0.788,
val accuracy: 0.752, lr: 1e-05
Finishing epoch 65/200: train loss: 117.56232422702482, train accuracy:
0.7866666666666666, val accuracy: 0.758, lr: 1e-05
Finishing epoch 66/200: train loss: 116.83815188106804, train accuracy: 0.788,
val accuracy: 0.762, lr: 1e-05
Finishing epoch 67/200: train loss: 116.29970335894512, train accuracy: 0.79,
val accuracy: 0.762, lr: 1e-05
Finishing epoch 68/200: train loss: 115.87429460491332, train accuracy:
0.7913333333333333, val accuracy: 0.766, lr: 1e-05
Finishing epoch 69/200: train loss: 115.09819189681681, train accuracy: 0.794,
val accuracy: 0.764, lr: 1e-05
Finishing epoch 70/200: train loss: 114.49720669943211, train accuracy: 0.794,
val accuracy: 0.768, lr: 1e-05
Finishing epoch 71/200: train loss: 114.1882032118067, train accuracy:
0.7986666666666666, val accuracy: 0.772, lr: 1e-05
Finishing epoch 72/200: train loss: 113.48934259967727, train accuracy:
0.7966666666666666, val accuracy: 0.772, lr: 1e-05

Finishing epoch 73/200: train loss: 113.13829026258014, train accuracy:
0.8013333333333333, val accuracy: 0.774, lr: 1e-05
Finishing epoch 74/200: train loss: 112.73768006036045, train accuracy: 0.802,
val accuracy: 0.778, lr: 1e-05
Finishing epoch 75/200: train loss: 112.3831011868424, train accuracy: 0.802,
val accuracy: 0.782, lr: 1e-05
Finishing epoch 76/200: train loss: 111.8986787571666, train accuracy:
0.8033333333333333, val accuracy: 0.78, lr: 1e-05
Finishing epoch 77/200: train loss: 111.42798342686879, train accuracy:
0.8053333333333333, val accuracy: 0.782, lr: 1e-05
Finishing epoch 78/200: train loss: 111.11693915644352, train accuracy:
0.8066666666666666, val accuracy: 0.782, lr: 1e-05
Finishing epoch 79/200: train loss: 110.44927481295404, train accuracy:
0.8066666666666666, val accuracy: 0.77, lr: 1e-05
Finishing epoch 80/200: train loss: 110.08503884093324, train accuracy: 0.808,
val accuracy: 0.77, lr: 1e-05
Finishing epoch 81/200: train loss: 109.84692437668107, train accuracy:
0.8066666666666666, val accuracy: 0.78, lr: 1e-05
Finishing epoch 82/200: train loss: 109.62194227259309, train accuracy: 0.808,
val accuracy: 0.784, lr: 1e-05
Finishing epoch 83/200: train loss: 109.04922411350516, train accuracy: 0.808,
val accuracy: 0.772, lr: 1e-05
Finishing epoch 84/200: train loss: 108.82750110778244, train accuracy:
0.8086666666666666, val accuracy: 0.78, lr: 1e-05
Finishing epoch 85/200: train loss: 108.56940287377483, train accuracy:
0.8093333333333333, val accuracy: 0.78, lr: 1e-05
Finishing epoch 86/200: train loss: 108.24696701768258, train accuracy: 0.81,
val accuracy: 0.78, lr: 1e-05
Finishing epoch 87/200: train loss: 107.85260508785024, train accuracy:
0.8086666666666666, val accuracy: 0.776, lr: 1e-05
Finishing epoch 88/200: train loss: 107.85260508785024, train accuracy:
0.8086666666666666, val accuracy: 0.776, lr: 1e-05
Finishing epoch 89/200: train loss: 107.68852257226737, train accuracy:
0.8160000000000001, val accuracy: 0.786, lr: 1e-05
Finishing epoch 90/200: train loss: 107.34614683899366, train accuracy:
0.8160000000000001, val accuracy: 0.786, lr: 1e-05
Finishing epoch 91/200: train loss: 106.72797355843439, train accuracy:
0.8126666666666666, val accuracy: 0.774, lr: 1e-05
Finishing epoch 92/200: train loss: 106.4871856449396, train accuracy:
0.8153333333333334, val accuracy: 0.778, lr: 1e-05
Finishing epoch 93/200: train loss: 106.13343767779591, train accuracy:
0.8146666666666667, val accuracy: 0.772, lr: 1e-05
Finishing epoch 94/200: train loss: 105.87334227571884, train accuracy:
0.8166666666666667, val accuracy: 0.776, lr: 1e-05
Finishing epoch 95/200: train loss: 105.72049634658273, train accuracy:
0.8180000000000001, val accuracy: 0.78, lr: 1e-05
Finishing epoch 96/200: train loss: 105.5153372522241, train accuracy:
0.8186666666666667, val accuracy: 0.784, lr: 1e-05

```
Finishing epoch 97/200: train loss: 105.47880026158674, train accuracy:
0.8206666666666667, val accuracy: 0.788, lr: 1e-05
Finishing epoch 98/200: train loss: 104.91798794577443, train accuracy:
0.8213333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 99/200: train loss: 104.68517491151555, train accuracy:
0.8213333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 100/200: train loss: 104.53348975180953, train accuracy:
0.8233333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 101/200: train loss: 104.35891149457191, train accuracy:
0.8246666666666667, val accuracy: 0.782, lr: 1e-05
Finishing epoch 102/200: train loss: 104.3340853187251, train accuracy:
0.8273333333333334, val accuracy: 0.788, lr: 1e-05
Finishing epoch 103/200: train loss: 103.80889639542607, train accuracy:
0.8240000000000001, val accuracy: 0.782, lr: 1e-05
Finishing epoch 104/200: train loss: 103.73441745655202, train accuracy:
0.8286666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 105/200: train loss: 103.55072374271504, train accuracy:
0.8293333333333334, val accuracy: 0.786, lr: 1e-05
Finishing epoch 106/200: train loss: 103.55072374271504, train accuracy:
0.8293333333333334, val accuracy: 0.786, lr: 1e-05
Finishing epoch 107/200: train loss: 103.06103368276828, train accuracy:
0.8273333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 108/200: train loss: 103.05228642405899, train accuracy:
0.8306666666666667, val accuracy: 0.786, lr: 1e-05
Finishing epoch 109/200: train loss: 102.75400699770516, train accuracy:
0.8306666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 110/200: train loss: 102.67540273936774, train accuracy:
0.8306666666666667, val accuracy: 0.786, lr: 1e-05
Finishing epoch 111/200: train loss: 102.44366119097995, train accuracy:
0.8326666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 112/200: train loss: 102.44366119097995, train accuracy:
0.8326666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 113/200: train loss: 101.94320448974386, train accuracy:
0.8326666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 114/200: train loss: 101.94320448974386, train accuracy:
0.8326666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 115/200: train loss: 101.50481059186876, train accuracy:
0.8273333333333334, val accuracy: 0.784, lr: 1e-05
Finishing epoch 116/200: train loss: 101.50481059186876, train accuracy:
0.8273333333333334, val accuracy: 0.784, lr: 1e-05
Finishing epoch 117/200: train loss: 101.3431443082583, train accuracy: 0.834,
val accuracy: 0.786, lr: 1e-05
Finishing epoch 118/200: train loss: 101.0591162828261, train accuracy:
0.8313333333333334, val accuracy: 0.784, lr: 1e-05
Finishing epoch 119/200: train loss: 100.88459081183882, train accuracy:
0.8306666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 120/200: train loss: 100.5987349099272, train accuracy:
0.8260000000000001, val accuracy: 0.78, lr: 1e-05
```

Finishing epoch 121/200: train loss: 100.49965827202124, train accuracy: 0.83,
val accuracy: 0.782, lr: 1e-05
Finishing epoch 122/200: train loss: 100.25561749216551, train accuracy: 0.83,
val accuracy: 0.782, lr: 1e-05
Finishing epoch 123/200: train loss: 100.09612742134621, train accuracy:
0.8293333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 124/200: train loss: 99.91422924736656, train accuracy:
0.8280000000000001, val accuracy: 0.78, lr: 1e-05
Finishing epoch 125/200: train loss: 99.8873965973512, train accuracy: 0.834,
val accuracy: 0.784, lr: 1e-05
Finishing epoch 126/200: train loss: 99.8873965973512, train accuracy: 0.834,
val accuracy: 0.784, lr: 1e-05
Finishing epoch 127/200: train loss: 99.8873965973512, train accuracy: 0.834,
val accuracy: 0.784, lr: 1e-05
Finishing epoch 128/200: train loss: 99.71711923626893, train accuracy:
0.8373333333333334, val accuracy: 0.79, lr: 1e-05
Finishing epoch 129/200: train loss: 99.42815415878204, train accuracy:
0.8353333333333334, val accuracy: 0.786, lr: 1e-05
Finishing epoch 130/200: train loss: 99.10578228126622, train accuracy:
0.8266666666666667, val accuracy: 0.78, lr: 1e-05
Finishing epoch 131/200: train loss: 99.1026546066351, train accuracy:
0.8353333333333334, val accuracy: 0.786, lr: 1e-05
Finishing epoch 132/200: train loss: 99.1026546066351, train accuracy:
0.8353333333333334, val accuracy: 0.786, lr: 1e-05
Finishing epoch 133/200: train loss: 98.74518791263962, train accuracy:
0.8273333333333334, val accuracy: 0.786, lr: 1e-05
Finishing epoch 134/200: train loss: 98.74518791263962, train accuracy:
0.8273333333333334, val accuracy: 0.786, lr: 1e-05
Finishing epoch 135/200: train loss: 98.6782486713943, train accuracy:
0.8366666666666667, val accuracy: 0.788, lr: 1e-05
Finishing epoch 136/200: train loss: 98.42175893164838, train accuracy:
0.8280000000000001, val accuracy: 0.786, lr: 1e-05
Finishing epoch 137/200: train loss: 98.42175893164838, train accuracy:
0.8280000000000001, val accuracy: 0.786, lr: 1e-05
Finishing epoch 138/200: train loss: 98.42175893164838, train accuracy:
0.8280000000000001, val accuracy: 0.786, lr: 1e-05
Finishing epoch 139/200: train loss: 98.24335716234786, train accuracy:
0.8326666666666667, val accuracy: 0.79, lr: 1e-05
Finishing epoch 140/200: train loss: 98.2019908558091, train accuracy:
0.8366666666666667, val accuracy: 0.79, lr: 1e-05
Finishing epoch 141/200: train loss: 98.09812649768433, train accuracy:
0.8373333333333334, val accuracy: 0.79, lr: 1e-05
Finishing epoch 142/200: train loss: 98.09812649768433, train accuracy:
0.8373333333333334, val accuracy: 0.79, lr: 1e-05
Finishing epoch 143/200: train loss: 98.09184211192643, train accuracy: 0.84,
val accuracy: 0.79, lr: 1e-05
Finishing epoch 144/200: train loss: 97.73359181864477, train accuracy: 0.834,
val accuracy: 0.79, lr: 1e-05

```
Finishing epoch 145/200: train loss: 97.70536598995744, train accuracy:
0.8373333333333334, val accuracy: 0.792, lr: 1e-05
Finishing epoch 146/200: train loss: 97.38202713663263, train accuracy:
0.8326666666666667, val accuracy: 0.792, lr: 1e-05
Finishing epoch 147/200: train loss: 97.24039266488677, train accuracy: 0.83,
val accuracy: 0.792, lr: 1e-05
Finishing epoch 148/200: train loss: 97.24039266488677, train accuracy: 0.83,
val accuracy: 0.792, lr: 1e-05
Finishing epoch 149/200: train loss: 97.22171582834653, train accuracy: 0.836,
val accuracy: 0.794, lr: 1e-05
Finishing epoch 150/200: train loss: 96.94797669027217, train accuracy:
0.8306666666666667, val accuracy: 0.792, lr: 1e-05
Finishing epoch 151/200: train loss: 96.89426979062006, train accuracy: 0.832,
val accuracy: 0.796, lr: 1e-05
Finishing epoch 152/200: train loss: 96.65291533250135, train accuracy:
0.8280000000000001, val accuracy: 0.794, lr: 1e-05
Finishing epoch 153/200: train loss: 96.65291533250135, train accuracy:
0.8280000000000001, val accuracy: 0.794, lr: 1e-05
Finishing epoch 154/200: train loss: 96.53585155237516, train accuracy:
0.8313333333333334, val accuracy: 0.796, lr: 1e-05
Finishing epoch 155/200: train loss: 96.53585155237516, train accuracy:
0.8313333333333334, val accuracy: 0.796, lr: 1e-05
Finishing epoch 156/200: train loss: 96.53585155237516, train accuracy:
0.8313333333333334, val accuracy: 0.796, lr: 1e-05
Finishing epoch 157/200: train loss: 96.27160450894003, train accuracy: 0.83,
val accuracy: 0.798, lr: 1e-05
Finishing epoch 158/200: train loss: 96.27160450894003, train accuracy: 0.83,
val accuracy: 0.798, lr: 1e-05
Finishing epoch 159/200: train loss: 96.11391836085443, train accuracy:
0.8306666666666667, val accuracy: 0.802, lr: 1e-05
Finishing epoch 160/200: train loss: 96.05015251253292, train accuracy: 0.832,
val accuracy: 0.8, lr: 1e-05
Finishing epoch 161/200: train loss: 95.95086972841068, train accuracy:
0.8326666666666667, val accuracy: 0.8, lr: 1e-05
Finishing epoch 162/200: train loss: 95.84235509814826, train accuracy: 0.832,
val accuracy: 0.8, lr: 1e-05
Finishing epoch 163/200: train loss: 95.84235509814826, train accuracy: 0.832,
val accuracy: 0.8, lr: 1e-05
Finishing epoch 164/200: train loss: 95.84235509814826, train accuracy: 0.832,
val accuracy: 0.8, lr: 1e-05
Finishing epoch 165/200: train loss: 95.84235509814826, train accuracy: 0.832,
val accuracy: 0.8, lr: 1e-05
Finishing epoch 166/200: train loss: 95.67809391027248, train accuracy: 0.836,
val accuracy: 0.806, lr: 1e-05
Finishing epoch 167/200: train loss: 95.53723706825498, train accuracy:
0.8353333333333334, val accuracy: 0.804, lr: 1e-05
Finishing epoch 168/200: train loss: 95.41262093881592, train accuracy:
0.8333333333333334, val accuracy: 0.806, lr: 1e-05
```

```
Finishing epoch 169/200: train loss: 95.41262093881592, train accuracy:
0.8333333333333334, val accuracy: 0.806, lr: 1e-05
Finishing epoch 170/200: train loss: 95.30968051827301, train accuracy: 0.836,
val accuracy: 0.806, lr: 1e-05
Finishing epoch 171/200: train loss: 95.20073720621714, train accuracy:
0.8366666666666667, val accuracy: 0.806, lr: 1e-05
Finishing epoch 172/200: train loss: 95.1315157933384, train accuracy:
0.8353333333333334, val accuracy: 0.808, lr: 1e-05
Finishing epoch 173/200: train loss: 94.97254203482845, train accuracy:
0.8353333333333334, val accuracy: 0.806, lr: 1e-05
Finishing epoch 174/200: train loss: 94.97254203482845, train accuracy:
0.8353333333333334, val accuracy: 0.806, lr: 1e-05
Finishing epoch 175/200: train loss: 94.97254203482845, train accuracy:
0.8353333333333334, val accuracy: 0.806, lr: 1e-05
Finishing epoch 176/200: train loss: 94.9129040897725, train accuracy:
0.8386666666666667, val accuracy: 0.81, lr: 1e-05
Finishing epoch 177/200: train loss: 94.9129040897725, train accuracy:
0.8386666666666667, val accuracy: 0.81, lr: 1e-05
Finishing epoch 178/200: train loss: 94.80203025556705, train accuracy:
0.8386666666666667, val accuracy: 0.812, lr: 1e-05
Finishing epoch 179/200: train loss: 94.7274839653279, train accuracy:
0.8386666666666667, val accuracy: 0.8140000000000001, lr: 1e-05
Finishing epoch 180/200: train loss: 94.7274839653279, train accuracy:
0.8386666666666667, val accuracy: 0.8140000000000001, lr: 1e-05
Finishing epoch 181/200: train loss: 94.45782710438446, train accuracy: 0.834,
val accuracy: 0.808, lr: 1e-05
Finishing epoch 182/200: train loss: 94.37827785027645, train accuracy: 0.834,
val accuracy: 0.81, lr: 1e-05
Finishing epoch 183/200: train loss: 94.37827785027645, train accuracy: 0.834,
val accuracy: 0.81, lr: 1e-05
Finishing epoch 184/200: train loss: 94.2479933911314, train accuracy: 0.834,
val accuracy: 0.804, lr: 1e-05
Finishing epoch 185/200: train loss: 94.18272863474363, train accuracy: 0.834,
val accuracy: 0.804, lr: 1e-05
Finishing epoch 186/200: train loss: 94.0977674721228, train accuracy:
0.8333333333333334, val accuracy: 0.806, lr: 1e-05
Finishing epoch 187/200: train loss: 94.02537091766874, train accuracy: 0.834,
val accuracy: 0.812, lr: 1e-05
Finishing epoch 188/200: train loss: 94.02537091766874, train accuracy: 0.834,
val accuracy: 0.812, lr: 1e-05
Finishing epoch 189/200: train loss: 94.02537091766874, train accuracy: 0.834,
val accuracy: 0.812, lr: 1e-05
Finishing epoch 190/200: train loss: 94.02537091766874, train accuracy: 0.834,
val accuracy: 0.812, lr: 1e-05
Finishing epoch 191/200: train loss: 94.02537091766874, train accuracy: 0.834,
val accuracy: 0.812, lr: 1e-05
Finishing epoch 192/200: train loss: 93.81640808737919, train accuracy:
0.8346666666666667, val accuracy: 0.812, lr: 1e-05
```

```
Finishing epoch 193/200: train loss: 93.81640808737919, train accuracy:
0.8346666666666667, val accuracy: 0.812, lr: 1e-05
Finishing epoch 194/200: train loss: 93.71862341489694, train accuracy:
0.8333333333333334, val accuracy: 0.804, lr: 1e-05
Finishing epoch 195/200: train loss: 93.63226369421808, train accuracy:
0.8346666666666667, val accuracy: 0.812, lr: 1e-05
Finishing epoch 196/200: train loss: 93.63226369421808, train accuracy:
0.8346666666666667, val accuracy: 0.812, lr: 1e-05
Finishing epoch 197/200: train loss: 93.63226369421808, train accuracy:
0.8346666666666667, val accuracy: 0.812, lr: 1e-05
Finishing epoch 198/200: train loss: 93.63226369421808, train accuracy:
0.8346666666666667, val accuracy: 0.812, lr: 1e-05
Finishing epoch 199/200: train loss: 93.55985979046261, train accuracy: 0.838,
val accuracy: 0.8200000000000001, lr: 1e-05
Finishing epoch 200/200: train loss: 93.41593340055, train accuracy: 0.838, val
accuracy: 0.8140000000000001, lr: 1e-05
Training ends.
```

```python
[28]: #Plot the loss and accuracy
      fig = plt.figure(figsize=(15,5))

      ax = fig.add_subplot(121)
      plt.semilogy(losses)
      plt.title('Loss')
      plt.xlabel('Epochs')
      plt.ylabel('Sum of Squares Error')

      ax = fig.add_subplot(122)
      plt.plot(accuracies,'.')
      plt.title('Accuracy')
      plt.xlabel('Epochs')
      plt.ylabel('Proportion of Correct Classifications')
      #plt.ylim([0,1])
      plt.grid()

      plt.show()
```
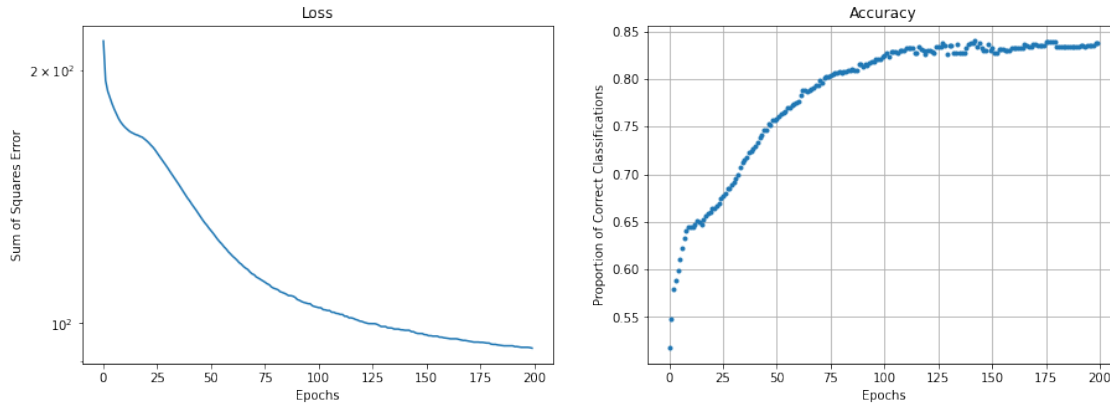
3 Step 2: Check model correctness

I crank up regularization to 1e2 to see if my model is setup properly.

```
[29]: model = four_layers_model(neurons_1, neurons_2)

      is_data_prepossing = True # set True if use data normalization
      regularization = 1e2

      accuracies, losses, val_accuracies, par2 = model.
      ↪train(X_train,y_train,X_val,y_val,epochs,h,batch_size,regularization,best_model=True)
```

Finishing epoch 1/200: train loss: 216.57389976366974, train accuracy:
0.5173333333333333, val accuracy: 0.45999999999999996, lr: 1e-05
Finishing epoch 2/200: train loss: 194.27900596855577, train accuracy:
0.5473333333333333, val accuracy: 0.512, lr: 1e-05
Finishing epoch 3/200: train loss: 188.71878954987613, train accuracy:
0.5793333333333333, val accuracy: 0.5640000000000001, lr: 1e-05
Finishing epoch 4/200: train loss: 185.46207412110158, train accuracy:
0.5886666666666667, val accuracy: 0.5760000000000001, lr: 1e-05
Finishing epoch 5/200: train loss: 182.41419241439118, train accuracy:
0.5986666666666667, val accuracy: 0.5920000000000001, lr: 1e-05
Finishing epoch 6/200: train loss: 179.74246260492905, train accuracy:
0.6106666666666667, val accuracy: 0.594, lr: 1e-05
Finishing epoch 7/200: train loss: 177.39452907852703, train accuracy: 0.622,
val accuracy: 0.602, lr: 1e-05
Finishing epoch 8/200: train loss: 175.18870675840085, train accuracy:
0.6326666666666667, val accuracy: 0.614, lr: 1e-05
Finishing epoch 9/200: train loss: 173.4736208582077, train accuracy:
0.6406666666666667, val accuracy: 0.626, lr: 1e-05
Finishing epoch 10/200: train loss: 172.1447669315506, train accuracy:
0.6446666666666667, val accuracy: 0.628, lr: 1e-05

Finishing epoch 11/200: train loss: 171.02457591930641, train accuracy: 0.644,
val accuracy: 0.632, lr: 1e-05
Finishing epoch 12/200: train loss: 170.11559216103043, train accuracy:
0.6446666666666667, val accuracy: 0.626, lr: 1e-05
Finishing epoch 13/200: train loss: 169.21624215835897, train accuracy:
0.6466666666666667, val accuracy: 0.618, lr: 1e-05
Finishing epoch 14/200: train loss: 168.59777383820096, train accuracy:
0.6513333333333333, val accuracy: 0.622, lr: 1e-05
Finishing epoch 15/200: train loss: 168.09995224094257, train accuracy:
0.6493333333333333, val accuracy: 0.624, lr: 1e-05
Finishing epoch 16/200: train loss: 167.5892511529036, train accuracy:
0.6466666666666667, val accuracy: 0.628, lr: 1e-05
Finishing epoch 17/200: train loss: 167.30057730714708, train accuracy: 0.652,
val accuracy: 0.632, lr: 1e-05
Finishing epoch 18/200: train loss: 166.7747121574224, train accuracy:
0.6566666666666667, val accuracy: 0.634, lr: 1e-05
Finishing epoch 19/200: train loss: 166.43507176979227, train accuracy:
0.6593333333333333, val accuracy: 0.636, lr: 1e-05
Finishing epoch 20/200: train loss: 165.63962634269, train accuracy:
0.6599999999999999, val accuracy: 0.632, lr: 1e-05
Finishing epoch 21/200: train loss: 164.8406378970879, train accuracy:
0.6639999999999999, val accuracy: 0.638, lr: 1e-05
Finishing epoch 22/200: train loss: 163.9456624317781, train accuracy:
0.6639999999999999, val accuracy: 0.636, lr: 1e-05
Finishing epoch 23/200: train loss: 162.8688854528699, train accuracy:
0.6666666666666667, val accuracy: 0.632, lr: 1e-05
Finishing epoch 24/200: train loss: 161.90751775390498, train accuracy:
0.6699999999999999, val accuracy: 0.646, lr: 1e-05
Finishing epoch 25/200: train loss: 160.69243713530346, train accuracy:
0.6739999999999999, val accuracy: 0.646, lr: 1e-05
Finishing epoch 26/200: train loss: 159.29996617586545, train accuracy:
0.6766666666666667, val accuracy: 0.642, lr: 1e-05
Finishing epoch 27/200: train loss: 157.9775999837426, train accuracy:
0.6799999999999999, val accuracy: 0.648, lr: 1e-05
Finishing epoch 28/200: train loss: 156.69017171244838, train accuracy:
0.6846666666666666, val accuracy: 0.654, lr: 1e-05
Finishing epoch 29/200: train loss: 155.40073847972326, train accuracy:
0.6853333333333333, val accuracy: 0.656, lr: 1e-05
Finishing epoch 30/200: train loss: 154.08398630311575, train accuracy:
0.6893333333333334, val accuracy: 0.6639999999999999, lr: 1e-05
Finishing epoch 31/200: train loss: 152.77313943593055, train accuracy: 0.692,
val accuracy: 0.6639999999999999, lr: 1e-05
Finishing epoch 32/200: train loss: 151.36205144228742, train accuracy: 0.696,
val accuracy: 0.6679999999999999, lr: 1e-05
Finishing epoch 33/200: train loss: 150.08087259980596, train accuracy: 0.7, val
accuracy: 0.6779999999999999, lr: 1e-05
Finishing epoch 34/200: train loss: 148.81299222717274, train accuracy:
0.7066666666666667, val accuracy: 0.6859999999999999, lr: 1e-05

Finishing epoch 35/200: train loss: 147.47416286464772, train accuracy: 0.712, val accuracy: 0.6839999999999999, lr: 1e-05
Finishing epoch 36/200: train loss: 146.19658078521965, train accuracy: 0.7153333333333334, val accuracy: 0.69, lr: 1e-05
Finishing epoch 37/200: train loss: 144.79779276601766, train accuracy: 0.7173333333333334, val accuracy: 0.696, lr: 1e-05
Finishing epoch 38/200: train loss: 143.50979267906598, train accuracy: 0.7233333333333334, val accuracy: 0.694, lr: 1e-05
Finishing epoch 39/200: train loss: 142.19149050973346, train accuracy: 0.7246666666666667, val accuracy: 0.698, lr: 1e-05
Finishing epoch 40/200: train loss: 140.9171431269292, train accuracy: 0.7273333333333334, val accuracy: 0.698, lr: 1e-05
Finishing epoch 41/200: train loss: 139.83282541489834, train accuracy: 0.73, val accuracy: 0.698, lr: 1e-05
Finishing epoch 42/200: train loss: 138.55953185706306, train accuracy: 0.734, val accuracy: 0.7, lr: 1e-05
Finishing epoch 43/200: train loss: 137.46323952283512, train accuracy: 0.738, val accuracy: 0.704, lr: 1e-05
Finishing epoch 44/200: train loss: 136.31086756667202, train accuracy: 0.7413333333333334, val accuracy: 0.704, lr: 1e-05
Finishing epoch 45/200: train loss: 135.0910060748654, train accuracy: 0.746, val accuracy: 0.71, lr: 1e-05
Finishing epoch 46/200: train loss: 133.85939513002467, train accuracy: 0.746, val accuracy: 0.71, lr: 1e-05
Finishing epoch 47/200: train loss: 132.76688552741214, train accuracy: 0.7526666666666667, val accuracy: 0.712, lr: 1e-05
Finishing epoch 48/200: train loss: 131.68845909333535, train accuracy: 0.7513333333333333, val accuracy: 0.716, lr: 1e-05
Finishing epoch 49/200: train loss: 130.8003533040271, train accuracy: 0.7573333333333333, val accuracy: 0.72, lr: 1e-05
Finishing epoch 50/200: train loss: 129.72631797337, train accuracy: 0.7573333333333333, val accuracy: 0.726, lr: 1e-05
Finishing epoch 51/200: train loss: 128.7905452419518, train accuracy: 0.7586666666666666, val accuracy: 0.734, lr: 1e-05
Finishing epoch 52/200: train loss: 127.87631421571413, train accuracy: 0.7606666666666666, val accuracy: 0.738, lr: 1e-05
Finishing epoch 53/200: train loss: 126.83250865735232, train accuracy: 0.7633333333333333, val accuracy: 0.734, lr: 1e-05
Finishing epoch 54/200: train loss: 125.83892267072099, train accuracy: 0.7646666666666666, val accuracy: 0.734, lr: 1e-05
Finishing epoch 55/200: train loss: 124.96111879538793, train accuracy: 0.7666666666666666, val accuracy: 0.736, lr: 1e-05
Finishing epoch 56/200: train loss: 124.20786911043501, train accuracy: 0.7693333333333333, val accuracy: 0.742, lr: 1e-05
Finishing epoch 57/200: train loss: 123.26432301317959, train accuracy: 0.7693333333333333, val accuracy: 0.74, lr: 1e-05
Finishing epoch 58/200: train loss: 122.54724270746233, train accuracy: 0.772, val accuracy: 0.746, lr: 1e-05

```
Finishing epoch 59/200: train loss: 121.61444497672741, train accuracy: 0.774,
val accuracy: 0.742, lr: 1e-05
Finishing epoch 60/200: train loss: 120.91567804551595, train accuracy:
0.7746666666666666, val accuracy: 0.744, lr: 1e-05
Finishing epoch 61/200: train loss: 120.14627126890927, train accuracy:
0.7766666666666666, val accuracy: 0.744, lr: 1e-05
Finishing epoch 62/200: train loss: 119.53662016936924, train accuracy:
0.7826666666666666, val accuracy: 0.752, lr: 1e-05
Finishing epoch 63/200: train loss: 118.8698973936093, train accuracy: 0.788,
val accuracy: 0.754, lr: 1e-05
Finishing epoch 64/200: train loss: 118.06450031235104, train accuracy: 0.788,
val accuracy: 0.752, lr: 1e-05
Finishing epoch 65/200: train loss: 117.56232422702482, train accuracy:
0.7866666666666666, val accuracy: 0.758, lr: 1e-05
Finishing epoch 66/200: train loss: 116.83815188106804, train accuracy: 0.788,
val accuracy: 0.762, lr: 1e-05
Finishing epoch 67/200: train loss: 116.29970335894512, train accuracy: 0.79,
val accuracy: 0.762, lr: 1e-05
Finishing epoch 68/200: train loss: 115.87429460491332, train accuracy:
0.7913333333333333, val accuracy: 0.766, lr: 1e-05
Finishing epoch 69/200: train loss: 115.09819189681681, train accuracy: 0.794,
val accuracy: 0.764, lr: 1e-05
Finishing epoch 70/200: train loss: 114.49720669943211, train accuracy: 0.794,
val accuracy: 0.768, lr: 1e-05
Finishing epoch 71/200: train loss: 114.1882032118067, train accuracy:
0.7986666666666666, val accuracy: 0.772, lr: 1e-05
Finishing epoch 72/200: train loss: 113.48934259967727, train accuracy:
0.7966666666666666, val accuracy: 0.772, lr: 1e-05
Finishing epoch 73/200: train loss: 113.13829026258014, train accuracy:
0.8013333333333333, val accuracy: 0.774, lr: 1e-05
Finishing epoch 74/200: train loss: 112.73768006036045, train accuracy: 0.802,
val accuracy: 0.778, lr: 1e-05
Finishing epoch 75/200: train loss: 112.3831011868424, train accuracy: 0.802,
val accuracy: 0.782, lr: 1e-05
Finishing epoch 76/200: train loss: 111.8986787571666, train accuracy:
0.8033333333333333, val accuracy: 0.78, lr: 1e-05
Finishing epoch 77/200: train loss: 111.42798342686879, train accuracy:
0.8053333333333333, val accuracy: 0.782, lr: 1e-05
Finishing epoch 78/200: train loss: 111.11693915644352, train accuracy:
0.8066666666666666, val accuracy: 0.782, lr: 1e-05
Finishing epoch 79/200: train loss: 110.44927481295404, train accuracy:
0.8066666666666666, val accuracy: 0.77, lr: 1e-05
Finishing epoch 80/200: train loss: 110.08503884093324, train accuracy: 0.808,
val accuracy: 0.77, lr: 1e-05
Finishing epoch 81/200: train loss: 109.84692437668107, train accuracy:
0.8066666666666666, val accuracy: 0.78, lr: 1e-05
Finishing epoch 82/200: train loss: 109.62194227259309, train accuracy: 0.808,
val accuracy: 0.784, lr: 1e-05
```

```
Finishing epoch 83/200: train loss: 109.04922411350516, train accuracy: 0.808,
val accuracy: 0.772, lr: 1e-05
Finishing epoch 84/200: train loss: 108.82750110778244, train accuracy:
0.8086666666666666, val accuracy: 0.78, lr: 1e-05
Finishing epoch 85/200: train loss: 108.56940287377483, train accuracy:
0.8093333333333333, val accuracy: 0.78, lr: 1e-05
Finishing epoch 86/200: train loss: 108.24696701768258, train accuracy: 0.81,
val accuracy: 0.78, lr: 1e-05
Finishing epoch 87/200: train loss: 107.85260508785024, train accuracy:
0.8086666666666666, val accuracy: 0.776, lr: 1e-05
Finishing epoch 88/200: train loss: 107.85260508785024, train accuracy:
0.8086666666666666, val accuracy: 0.776, lr: 1e-05
Finishing epoch 89/200: train loss: 107.68852257226737, train accuracy:
0.8160000000000001, val accuracy: 0.786, lr: 1e-05
Finishing epoch 90/200: train loss: 107.34614683899366, train accuracy:
0.8160000000000001, val accuracy: 0.786, lr: 1e-05
Finishing epoch 91/200: train loss: 106.72797355843439, train accuracy:
0.8126666666666666, val accuracy: 0.774, lr: 1e-05
Finishing epoch 92/200: train loss: 106.4871856449396, train accuracy:
0.8153333333333334, val accuracy: 0.778, lr: 1e-05
Finishing epoch 93/200: train loss: 106.13343767779591, train accuracy:
0.8146666666666667, val accuracy: 0.772, lr: 1e-05
Finishing epoch 94/200: train loss: 105.87334227571884, train accuracy:
0.8166666666666667, val accuracy: 0.776, lr: 1e-05
Finishing epoch 95/200: train loss: 105.72049634658273, train accuracy:
0.8180000000000001, val accuracy: 0.78, lr: 1e-05
Finishing epoch 96/200: train loss: 105.5153372522241, train accuracy:
0.8186666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 97/200: train loss: 105.47880026158674, train accuracy:
0.8206666666666667, val accuracy: 0.788, lr: 1e-05
Finishing epoch 98/200: train loss: 104.91798794577443, train accuracy:
0.8213333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 99/200: train loss: 104.68517491151555, train accuracy:
0.8213333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 100/200: train loss: 104.53348975180953, train accuracy:
0.8233333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 101/200: train loss: 104.35891149457191, train accuracy:
0.8246666666666667, val accuracy: 0.782, lr: 1e-05
Finishing epoch 102/200: train loss: 104.3340853187251, train accuracy:
0.8273333333333334, val accuracy: 0.788, lr: 1e-05
Finishing epoch 103/200: train loss: 103.80889639542607, train accuracy:
0.8240000000000001, val accuracy: 0.782, lr: 1e-05
Finishing epoch 104/200: train loss: 103.73441745655202, train accuracy:
0.8286666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 105/200: train loss: 103.55072374271504, train accuracy:
0.8293333333333334, val accuracy: 0.786, lr: 1e-05
Finishing epoch 106/200: train loss: 103.55072374271504, train accuracy:
0.8293333333333334, val accuracy: 0.786, lr: 1e-05
```

```
Finishing epoch 107/200: train loss: 103.06103368276828, train accuracy:
0.8273333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 108/200: train loss: 103.05228642405899, train accuracy:
0.8306666666666667, val accuracy: 0.786, lr: 1e-05
Finishing epoch 109/200: train loss: 102.75400699770516, train accuracy:
0.8306666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 110/200: train loss: 102.67540273936774, train accuracy:
0.8306666666666667, val accuracy: 0.786, lr: 1e-05
Finishing epoch 111/200: train loss: 102.44366119097995, train accuracy:
0.8326666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 112/200: train loss: 102.44366119097995, train accuracy:
0.8326666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 113/200: train loss: 101.94320448974386, train accuracy:
0.8326666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 114/200: train loss: 101.94320448974386, train accuracy:
0.8326666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 115/200: train loss: 101.50481059186876, train accuracy:
0.8273333333333334, val accuracy: 0.784, lr: 1e-05
Finishing epoch 116/200: train loss: 101.50481059186876, train accuracy:
0.8273333333333334, val accuracy: 0.784, lr: 1e-05
Finishing epoch 117/200: train loss: 101.3431443082583, train accuracy: 0.834,
val accuracy: 0.786, lr: 1e-05
Finishing epoch 118/200: train loss: 101.0591162828261, train accuracy:
0.8313333333333334, val accuracy: 0.784, lr: 1e-05
Finishing epoch 119/200: train loss: 100.88459081183882, train accuracy:
0.8306666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 120/200: train loss: 100.5987349099272, train accuracy:
0.8260000000000001, val accuracy: 0.78, lr: 1e-05
Finishing epoch 121/200: train loss: 100.49965827202124, train accuracy: 0.83,
val accuracy: 0.782, lr: 1e-05
Finishing epoch 122/200: train loss: 100.25561749216551, train accuracy: 0.83,
val accuracy: 0.782, lr: 1e-05
Finishing epoch 123/200: train loss: 100.09612742134621, train accuracy:
0.8293333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 124/200: train loss: 99.91422924736656, train accuracy:
0.8280000000000001, val accuracy: 0.78, lr: 1e-05
Training ends early when the error satisfies the regularization.
```

```python
[30]: #Plot the loss and accuracy
fig = plt.figure(figsize=(15,5))

ax = fig.add_subplot(121)
plt.semilogy(losses)
plt.title('Loss')
plt.xlabel('Epochs')
plt.ylabel('Sum of Squares Error')
```

```
ax = fig.add_subplot(122)
plt.plot(accuracies,'.')
plt.title('Accuracy')
plt.xlabel('Epochs')
plt.ylabel('Proportion of Correct Classifications')
#plt.ylim([0,1])
plt.grid()

plt.show()
```
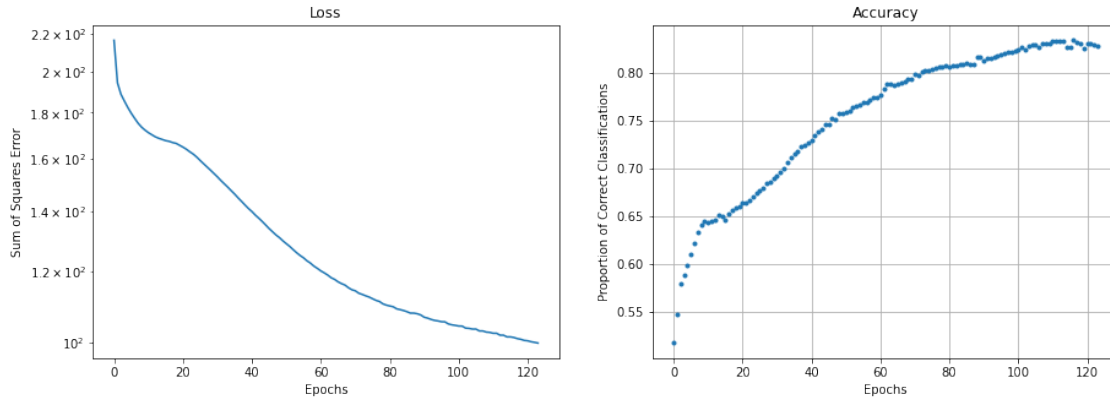

We can see from above figure that the training process ends early when the error is below 1e2. This shows that my code works properly.

4 Step 3: Test data normalization

I disabled the data preprocessing option and see how it affects the final training result.

```
[31]: model = four_layers_model(neurons_1, neurons_2)

is_data_prepossing = False # set True if use data normalization
regularization = 0.0

# Read dataset
with h5py.File('./dataset/train_dataset.h5', 'r') as file:
    data = np.array(file.get('data'))
    label = np.array(file.get('label'))
    X_train = data
    y_train = label

with h5py.File('./dataset/test_dataset.h5', 'r') as file:
    data = np.array(file.get('data'))
    label = np.array(file.get('label'))
    X_test = data
```

```python
    y_test = label

# Data Preprossing
if is_data_prepossing:
    X_train = data_preprossing(X_train)


X_train, X_val, y_train, y_val = train_test_split(X_train, y_train, test_size=0.
 ↪25, random_state=66)


y_train = np.expand_dims(y_train, -1)
y_val = np.expand_dims(y_val, -1)
y_test = np.expand_dims(y_test, -1)


accuracies, losses, val_accuracies, par3 = model.
 ↪train(X_train,y_train,X_val,y_val,epochs,h,batch_size,regularization,best_model=True)
```

Finishing epoch 1/200: train loss: 214.10103167914815, train accuracy:
0.5226666666666666, val accuracy: 0.44799999999999995, lr: 1e-05
Finishing epoch 2/200: train loss: 190.99249775220562, train accuracy:
0.5553333333333333, val accuracy: 0.518, lr: 1e-05
Finishing epoch 3/200: train loss: 184.75835789133868, train accuracy:
0.5873333333333333, val accuracy: 0.562, lr: 1e-05
Finishing epoch 4/200: train loss: 180.97852829890655, train accuracy:
0.5973333333333333, val accuracy: 0.5820000000000001, lr: 1e-05
Finishing epoch 5/200: train loss: 177.53236630378018, train accuracy:
0.6133333333333333, val accuracy: 0.602, lr: 1e-05
Finishing epoch 6/200: train loss: 174.31016794643017, train accuracy:
0.6253333333333333, val accuracy: 0.612, lr: 1e-05
Finishing epoch 7/200: train loss: 171.4337756545761, train accuracy: 0.64, val
accuracy: 0.62, lr: 1e-05
Finishing epoch 8/200: train loss: 168.54713106828328, train accuracy:
0.6526666666666667, val accuracy: 0.636, lr: 1e-05
Finishing epoch 9/200: train loss: 165.73433979463937, train accuracy:
0.6613333333333333, val accuracy: 0.64, lr: 1e-05
Finishing epoch 10/200: train loss: 163.27560818607714, train accuracy:
0.6693333333333333, val accuracy: 0.65, lr: 1e-05
Finishing epoch 11/200: train loss: 160.87435987404461, train accuracy:
0.6779999999999999, val accuracy: 0.6639999999999999, lr: 1e-05
Finishing epoch 12/200: train loss: 158.72527366849982, train accuracy:
0.6846666666666666, val accuracy: 0.6739999999999999, lr: 1e-05
Finishing epoch 13/200: train loss: 156.74681984491852, train accuracy:
0.6906666666666667, val accuracy: 0.6759999999999999, lr: 1e-05
Finishing epoch 14/200: train loss: 155.02034981343655, train accuracy:
0.6966666666666667, val accuracy: 0.6779999999999999, lr: 1e-05
Finishing epoch 15/200: train loss: 153.44544344907223, train accuracy:
0.6973333333333334, val accuracy: 0.6759999999999999, lr: 1e-05
Finishing epoch 16/200: train loss: 151.99794001380906, train accuracy:

0.7053333333333334, val accuracy: 0.6759999999999999, lr: 1e-05
Finishing epoch 17/200: train loss: 150.67335224189907, train accuracy:
0.7066666666666667, val accuracy: 0.6799999999999999, lr: 1e-05
Finishing epoch 18/200: train loss: 149.51777165303125, train accuracy:
0.7073333333333334, val accuracy: 0.688, lr: 1e-05
Finishing epoch 19/200: train loss: 148.40021135145372, train accuracy:
0.7153333333333334, val accuracy: 0.69, lr: 1e-05
Finishing epoch 20/200: train loss: 147.5519522735263, train accuracy: 0.714,
val accuracy: 0.69, lr: 1e-05
Finishing epoch 21/200: train loss: 146.7620510979054, train accuracy:
0.7153333333333334, val accuracy: 0.694, lr: 1e-05
Finishing epoch 22/200: train loss: 146.107738725788, train accuracy: 0.718, val
accuracy: 0.6839999999999999, lr: 1e-05
Finishing epoch 23/200: train loss: 145.5021462104682, train accuracy:
0.7226666666666667, val accuracy: 0.6799999999999999, lr: 1e-05
Finishing epoch 24/200: train loss: 145.09708142209877, train accuracy:
0.7246666666666667, val accuracy: 0.6859999999999999, lr: 1e-05
Finishing epoch 25/200: train loss: 144.465857592785, train accuracy:
0.7246666666666667, val accuracy: 0.6859999999999999, lr: 1e-05
Finishing epoch 26/200: train loss: 143.65835933095687, train accuracy:
0.7266666666666667, val accuracy: 0.6839999999999999, lr: 1e-05
Finishing epoch 27/200: train loss: 142.87420776406992, train accuracy:
0.7293333333333334, val accuracy: 0.6859999999999999, lr: 1e-05
Finishing epoch 28/200: train loss: 141.7796880622666, train accuracy:
0.7306666666666667, val accuracy: 0.688, lr: 1e-05
Finishing epoch 29/200: train loss: 140.59022860939638, train accuracy: 0.732,
val accuracy: 0.6859999999999999, lr: 1e-05
Finishing epoch 30/200: train loss: 139.33033755142353, train accuracy:
0.7333333333333334, val accuracy: 0.688, lr: 1e-05
Finishing epoch 31/200: train loss: 138.1320141597759, train accuracy: 0.734,
val accuracy: 0.696, lr: 1e-05
Finishing epoch 32/200: train loss: 137.0187135979782, train accuracy:
0.7353333333333334, val accuracy: 0.698, lr: 1e-05
Finishing epoch 33/200: train loss: 135.80424657706732, train accuracy:
0.7366666666666667, val accuracy: 0.696, lr: 1e-05
Finishing epoch 34/200: train loss: 134.80042130048378, train accuracy:
0.7426666666666666, val accuracy: 0.698, lr: 1e-05
Finishing epoch 35/200: train loss: 133.8592297102832, train accuracy: 0.744,
val accuracy: 0.7, lr: 1e-05
Finishing epoch 36/200: train loss: 132.92322459665291, train accuracy:
0.7473333333333334, val accuracy: 0.7, lr: 1e-05
Finishing epoch 37/200: train loss: 131.92697222627447, train accuracy:
0.7473333333333334, val accuracy: 0.702, lr: 1e-05
Finishing epoch 38/200: train loss: 131.00881325641132, train accuracy:
0.7466666666666666, val accuracy: 0.704, lr: 1e-05
Finishing epoch 39/200: train loss: 130.16762091937025, train accuracy:
0.7453333333333334, val accuracy: 0.71, lr: 1e-05
Finishing epoch 40/200: train loss: 129.29194193772753, train accuracy:

0.7493333333333334, val accuracy: 0.712, lr: 1e-05
Finishing epoch 41/200: train loss: 128.65521652958824, train accuracy:
0.7633333333333333, val accuracy: 0.724, lr: 1e-05
Finishing epoch 42/200: train loss: 127.85544158970654, train accuracy:
0.7626666666666666, val accuracy: 0.712, lr: 1e-05
Finishing epoch 43/200: train loss: 127.19491392071771, train accuracy:
0.7673333333333333, val accuracy: 0.72, lr: 1e-05
Finishing epoch 44/200: train loss: 126.51142065681115, train accuracy:
0.7706666666666666, val accuracy: 0.72, lr: 1e-05
Finishing epoch 45/200: train loss: 125.78748782169332, train accuracy:
0.7713333333333333, val accuracy: 0.72, lr: 1e-05
Finishing epoch 46/200: train loss: 125.10044943888103, train accuracy: 0.772,
val accuracy: 0.718, lr: 1e-05
Finishing epoch 47/200: train loss: 124.55082156915829, train accuracy: 0.774,
val accuracy: 0.72, lr: 1e-05
Finishing epoch 48/200: train loss: 123.94457954898732, train accuracy: 0.774,
val accuracy: 0.718, lr: 1e-05
Finishing epoch 49/200: train loss: 123.54319804661682, train accuracy: 0.778,
val accuracy: 0.722, lr: 1e-05
Finishing epoch 50/200: train loss: 122.97489522451679, train accuracy:
0.7786666666666666, val accuracy: 0.722, lr: 1e-05
Finishing epoch 51/200: train loss: 122.59661325037348, train accuracy:
0.7806666666666666, val accuracy: 0.726, lr: 1e-05
Finishing epoch 52/200: train loss: 122.17316459969848, train accuracy: 0.782,
val accuracy: 0.726, lr: 1e-05
Finishing epoch 53/200: train loss: 121.57929412521318, train accuracy:
0.7866666666666666, val accuracy: 0.728, lr: 1e-05
Finishing epoch 54/200: train loss: 121.12164054528455, train accuracy:
0.7866666666666666, val accuracy: 0.726, lr: 1e-05
Finishing epoch 55/200: train loss: 120.77512831342185, train accuracy: 0.788,
val accuracy: 0.73, lr: 1e-05
Finishing epoch 56/200: train loss: 120.43297908979216, train accuracy:
0.7933333333333333, val accuracy: 0.734, lr: 1e-05
Finishing epoch 57/200: train loss: 119.99756573122636, train accuracy:
0.7946666666666666, val accuracy: 0.736, lr: 1e-05
Finishing epoch 58/200: train loss: 119.66032198044806, train accuracy:
0.7953333333333333, val accuracy: 0.742, lr: 1e-05
Finishing epoch 59/200: train loss: 119.19222371523678, train accuracy: 0.796,
val accuracy: 0.74, lr: 1e-05
Finishing epoch 60/200: train loss: 118.85397735863656, train accuracy:
0.7966666666666666, val accuracy: 0.744, lr: 1e-05
Finishing epoch 61/200: train loss: 118.54656530334645, train accuracy:
0.7966666666666666, val accuracy: 0.742, lr: 1e-05
Finishing epoch 62/200: train loss: 118.32928115388576, train accuracy: 0.798,
val accuracy: 0.744, lr: 1e-05
Finishing epoch 63/200: train loss: 118.07071794617653, train accuracy:
0.7973333333333333, val accuracy: 0.744, lr: 1e-05
Finishing epoch 64/200: train loss: 117.64054793192324, train accuracy: 0.798,

val accuracy: 0.742, lr: 1e-05
Finishing epoch 65/200: train loss: 117.4448246149062, train accuracy: 0.7993333333333333, val accuracy: 0.742, lr: 1e-05
Finishing epoch 66/200: train loss: 117.03279958935028, train accuracy: 0.7993333333333333, val accuracy: 0.742, lr: 1e-05
Finishing epoch 67/200: train loss: 116.82163890150909, train accuracy: 0.8, val accuracy: 0.742, lr: 1e-05
Finishing epoch 68/200: train loss: 116.68426650394079, train accuracy: 0.8, val accuracy: 0.748, lr: 1e-05
Finishing epoch 69/200: train loss: 116.29915176537607, train accuracy: 0.798, val accuracy: 0.744, lr: 1e-05
Finishing epoch 70/200: train loss: 115.95226245987925, train accuracy: 0.8013333333333333, val accuracy: 0.744, lr: 1e-05
Finishing epoch 71/200: train loss: 115.8609368540078, train accuracy: 0.7986666666666666, val accuracy: 0.748, lr: 1e-05
Finishing epoch 72/200: train loss: 115.48643797660748, train accuracy: 0.8006666666666666, val accuracy: 0.746, lr: 1e-05
Finishing epoch 73/200: train loss: 115.37551875204629, train accuracy: 0.8, val accuracy: 0.75, lr: 1e-05
Finishing epoch 74/200: train loss: 115.20688715218948, train accuracy: 0.8013333333333333, val accuracy: 0.75, lr: 1e-05
Finishing epoch 75/200: train loss: 115.0206808492526, train accuracy: 0.802, val accuracy: 0.752, lr: 1e-05
Finishing epoch 76/200: train loss: 114.77116849915355, train accuracy: 0.8026666666666666, val accuracy: 0.754, lr: 1e-05
Finishing epoch 77/200: train loss: 114.49029046199371, train accuracy: 0.8033333333333333, val accuracy: 0.752, lr: 1e-05
Finishing epoch 78/200: train loss: 114.30270224808042, train accuracy: 0.804, val accuracy: 0.754, lr: 1e-05
Finishing epoch 79/200: train loss: 113.86857896112332, train accuracy: 0.804, val accuracy: 0.75, lr: 1e-05
Finishing epoch 80/200: train loss: 113.66780745036216, train accuracy: 0.8053333333333333, val accuracy: 0.75, lr: 1e-05
Finishing epoch 81/200: train loss: 113.57845425268349, train accuracy: 0.8026666666666666, val accuracy: 0.75, lr: 1e-05
Finishing epoch 82/200: train loss: 113.47732896368296, train accuracy: 0.8046666666666666, val accuracy: 0.752, lr: 1e-05
Finishing epoch 83/200: train loss: 113.05003669883172, train accuracy: 0.8033333333333333, val accuracy: 0.748, lr: 1e-05
Finishing epoch 84/200: train loss: 112.93432187448458, train accuracy: 0.8046666666666666, val accuracy: 0.748, lr: 1e-05
Finishing epoch 85/200: train loss: 112.77849717884561, train accuracy: 0.8053333333333333, val accuracy: 0.75, lr: 1e-05
Finishing epoch 86/200: train loss: 112.60735309963968, train accuracy: 0.8053333333333333, val accuracy: 0.75, lr: 1e-05
Finishing epoch 87/200: train loss: 112.29099743479681, train accuracy: 0.8053333333333333, val accuracy: 0.752, lr: 1e-05
Finishing epoch 88/200: train loss: 112.29099743479681, train accuracy:

0.8053333333333333, val accuracy: 0.752, lr: 1e-05
Finishing epoch 89/200: train loss: 112.1656853814159, train accuracy:
0.8066666666666666, val accuracy: 0.756, lr: 1e-05
Finishing epoch 90/200: train loss: 111.95717089415163, train accuracy:
0.8086666666666666, val accuracy: 0.756, lr: 1e-05
Finishing epoch 91/200: train loss: 111.50985168125598, train accuracy:
0.8066666666666666, val accuracy: 0.756, lr: 1e-05
Finishing epoch 92/200: train loss: 111.36753425382767, train accuracy:
0.8066666666666666, val accuracy: 0.758, lr: 1e-05
Finishing epoch 93/200: train loss: 111.0923396214348, train accuracy:
0.8053333333333333, val accuracy: 0.756, lr: 1e-05
Finishing epoch 94/200: train loss: 110.94501011911706, train accuracy: 0.806,
val accuracy: 0.76, lr: 1e-05
Finishing epoch 95/200: train loss: 110.88409496740587, train accuracy: 0.808,
val accuracy: 0.762, lr: 1e-05
Finishing epoch 96/200: train loss: 110.7834588816203, train accuracy:
0.8106666666666666, val accuracy: 0.766, lr: 1e-05
Finishing epoch 97/200: train loss: 110.7834588816203, train accuracy:
0.8106666666666666, val accuracy: 0.766, lr: 1e-05
Finishing epoch 98/200: train loss: 110.39920530264055, train accuracy: 0.81,
val accuracy: 0.766, lr: 1e-05
Finishing epoch 99/200: train loss: 110.23130531573976, train accuracy:
0.8106666666666666, val accuracy: 0.768, lr: 1e-05
Finishing epoch 100/200: train loss: 110.14254882384463, train accuracy:
0.8126666666666666, val accuracy: 0.766, lr: 1e-05
Finishing epoch 101/200: train loss: 109.96286117659292, train accuracy:
0.8146666666666667, val accuracy: 0.768, lr: 1e-05
Finishing epoch 102/200: train loss: 109.96286117659292, train accuracy:
0.8146666666666667, val accuracy: 0.768, lr: 1e-05
Finishing epoch 103/200: train loss: 109.64417746797471, train accuracy:
0.8140000000000001, val accuracy: 0.77, lr: 1e-05
Finishing epoch 104/200: train loss: 109.59438285674682, train accuracy:
0.8160000000000001, val accuracy: 0.774, lr: 1e-05
Finishing epoch 105/200: train loss: 109.39782389576544, train accuracy:
0.8160000000000001, val accuracy: 0.774, lr: 1e-05
Finishing epoch 106/200: train loss: 109.39782389576544, train accuracy:
0.8160000000000001, val accuracy: 0.774, lr: 1e-05
Finishing epoch 107/200: train loss: 108.96712704986886, train accuracy:
0.8126666666666666, val accuracy: 0.772, lr: 1e-05
Finishing epoch 108/200: train loss: 108.96712704986886, train accuracy:
0.8126666666666666, val accuracy: 0.772, lr: 1e-05
Finishing epoch 109/200: train loss: 108.7665318332553, train accuracy:
0.8133333333333334, val accuracy: 0.772, lr: 1e-05
Finishing epoch 110/200: train loss: 108.69429706005806, train accuracy:
0.8133333333333334, val accuracy: 0.774, lr: 1e-05
Finishing epoch 111/200: train loss: 108.56211093594783, train accuracy:
0.8146666666666667, val accuracy: 0.774, lr: 1e-05
Finishing epoch 112/200: train loss: 108.56211093594783, train accuracy:

0.8146666666666667, val accuracy: 0.774, lr: 1e-05
Finishing epoch 113/200: train loss: 108.16207104579, train accuracy:
0.8173333333333334, val accuracy: 0.774, lr: 1e-05
Finishing epoch 114/200: train loss: 108.16207104579, train accuracy:
0.8173333333333334, val accuracy: 0.774, lr: 1e-05
Finishing epoch 115/200: train loss: 107.80224423880898, train accuracy:
0.8160000000000001, val accuracy: 0.774, lr: 1e-05
Finishing epoch 116/200: train loss: 107.80224423880898, train accuracy:
0.8160000000000001, val accuracy: 0.774, lr: 1e-05
Finishing epoch 117/200: train loss: 107.66947451305649, train accuracy:
0.8180000000000001, val accuracy: 0.774, lr: 1e-05
Finishing epoch 118/200: train loss: 107.49298292491292, train accuracy:
0.8173333333333334, val accuracy: 0.774, lr: 1e-05
Finishing epoch 119/200: train loss: 107.3748046883614, train accuracy:
0.8186666666666667, val accuracy: 0.776, lr: 1e-05
Finishing epoch 120/200: train loss: 107.1914694612628, train accuracy: 0.812,
val accuracy: 0.772, lr: 1e-05
Finishing epoch 121/200: train loss: 107.1086923997058, train accuracy:
0.8173333333333334, val accuracy: 0.776, lr: 1e-05
Finishing epoch 122/200: train loss: 106.91113835493326, train accuracy:
0.8140000000000001, val accuracy: 0.774, lr: 1e-05
Finishing epoch 123/200: train loss: 106.8269750223808, train accuracy:
0.8153333333333334, val accuracy: 0.776, lr: 1e-05
Finishing epoch 124/200: train loss: 106.6979838675463, train accuracy:
0.8140000000000001, val accuracy: 0.774, lr: 1e-05
Finishing epoch 125/200: train loss: 106.67418805439532, train accuracy:
0.8166666666666667, val accuracy: 0.776, lr: 1e-05
Finishing epoch 126/200: train loss: 106.67418805439532, train accuracy:
0.8166666666666667, val accuracy: 0.776, lr: 1e-05
Finishing epoch 127/200: train loss: 106.67418805439532, train accuracy:
0.8166666666666667, val accuracy: 0.776, lr: 1e-05
Finishing epoch 128/200: train loss: 106.4872939111444, train accuracy:
0.8200000000000001, val accuracy: 0.778, lr: 1e-05
Finishing epoch 129/200: train loss: 106.29561088961412, train accuracy:
0.8186666666666667, val accuracy: 0.776, lr: 1e-05
Finishing epoch 130/200: train loss: 106.0840044225944, train accuracy:
0.8146666666666667, val accuracy: 0.776, lr: 1e-05
Finishing epoch 131/200: train loss: 106.05097782357197, train accuracy:
0.8186666666666667, val accuracy: 0.776, lr: 1e-05
Finishing epoch 132/200: train loss: 106.05097782357197, train accuracy:
0.8186666666666667, val accuracy: 0.776, lr: 1e-05
Finishing epoch 133/200: train loss: 105.76894554298983, train accuracy:
0.8146666666666667, val accuracy: 0.776, lr: 1e-05
Finishing epoch 134/200: train loss: 105.7556735256097, train accuracy:
0.8206666666666667, val accuracy: 0.78, lr: 1e-05
Finishing epoch 135/200: train loss: 105.6879541564422, train accuracy:
0.8200000000000001, val accuracy: 0.78, lr: 1e-05
Finishing epoch 136/200: train loss: 105.5258449879875, train accuracy:

0.8173333333333334, val accuracy: 0.776, lr: 1e-05
Finishing epoch 137/200: train loss: 105.5258449879875, train accuracy:
0.8173333333333334, val accuracy: 0.776, lr: 1e-05
Finishing epoch 138/200: train loss: 105.5258449879875, train accuracy:
0.8173333333333334, val accuracy: 0.776, lr: 1e-05
Finishing epoch 139/200: train loss: 105.33160679263301, train accuracy:
0.8193333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 140/200: train loss: 105.30542242654272, train accuracy:
0.8213333333333334, val accuracy: 0.782, lr: 1e-05
Finishing epoch 141/200: train loss: 105.24617371662649, train accuracy:
0.8220000000000001, val accuracy: 0.782, lr: 1e-05
Finishing epoch 142/200: train loss: 105.24617371662649, train accuracy:
0.8220000000000001, val accuracy: 0.782, lr: 1e-05
Finishing epoch 143/200: train loss: 105.16651204819044, train accuracy:
0.8220000000000001, val accuracy: 0.786, lr: 1e-05
Finishing epoch 144/200: train loss: 104.9469422432038, train accuracy:
0.8213333333333334, val accuracy: 0.782, lr: 1e-05
Finishing epoch 145/200: train loss: 104.92881009743252, train accuracy:
0.8226666666666667, val accuracy: 0.782, lr: 1e-05
Finishing epoch 146/200: train loss: 104.66910454449902, train accuracy:
0.8186666666666667, val accuracy: 0.78, lr: 1e-05
Finishing epoch 147/200: train loss: 104.62650931727043, train accuracy:
0.8193333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 148/200: train loss: 104.62650931727043, train accuracy:
0.8193333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 149/200: train loss: 104.57667044445486, train accuracy:
0.8226666666666667, val accuracy: 0.782, lr: 1e-05
Finishing epoch 150/200: train loss: 104.37124074225548, train accuracy:
0.8213333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 151/200: train loss: 104.36995001134812, train accuracy:
0.8226666666666667, val accuracy: 0.778, lr: 1e-05
Finishing epoch 152/200: train loss: 104.16203713276151, train accuracy:
0.8193333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 153/200: train loss: 104.16203713276151, train accuracy:
0.8193333333333334, val accuracy: 0.78, lr: 1e-05
Finishing epoch 154/200: train loss: 104.05919398621543, train accuracy:
0.8220000000000001, val accuracy: 0.78, lr: 1e-05
Finishing epoch 155/200: train loss: 104.05919398621543, train accuracy:
0.8220000000000001, val accuracy: 0.78, lr: 1e-05
Finishing epoch 156/200: train loss: 104.05919398621543, train accuracy:
0.8220000000000001, val accuracy: 0.78, lr: 1e-05
Finishing epoch 157/200: train loss: 103.86548563428951, train accuracy:
0.8220000000000001, val accuracy: 0.782, lr: 1e-05
Finishing epoch 158/200: train loss: 103.86548563428951, train accuracy:
0.8220000000000001, val accuracy: 0.782, lr: 1e-05
Finishing epoch 159/200: train loss: 103.73555007531058, train accuracy:
0.8226666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 160/200: train loss: 103.69159517282895, train accuracy:

0.8220000000000001, val accuracy: 0.784, lr: 1e-05
Finishing epoch 161/200: train loss: 103.64381436412283, train accuracy:
0.8226666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 162/200: train loss: 103.54468041405775, train accuracy:
0.8226666666666667, val accuracy: 0.786, lr: 1e-05
Finishing epoch 163/200: train loss: 103.54468041405775, train accuracy:
0.8226666666666667, val accuracy: 0.786, lr: 1e-05
Finishing epoch 164/200: train loss: 103.54468041405775, train accuracy:
0.8226666666666667, val accuracy: 0.786, lr: 1e-05
Finishing epoch 165/200: train loss: 103.52493524637731, train accuracy:
0.8266666666666667, val accuracy: 0.784, lr: 1e-05
Finishing epoch 166/200: train loss: 103.389598407746, train accuracy:
0.8266666666666667, val accuracy: 0.786, lr: 1e-05
Finishing epoch 167/200: train loss: 103.2745324848489, train accuracy:
0.8260000000000001, val accuracy: 0.786, lr: 1e-05
Finishing epoch 168/200: train loss: 103.16804050406958, train accuracy:
0.8260000000000001, val accuracy: 0.786, lr: 1e-05
Finishing epoch 169/200: train loss: 103.16804050406958, train accuracy:
0.8260000000000001, val accuracy: 0.786, lr: 1e-05
Finishing epoch 170/200: train loss: 103.08210604263344, train accuracy:
0.8253333333333334, val accuracy: 0.788, lr: 1e-05
Finishing epoch 171/200: train loss: 103.03210485101741, train accuracy:
0.8266666666666667, val accuracy: 0.788, lr: 1e-05
Finishing epoch 172/200: train loss: 102.95926901360338, train accuracy:
0.8260000000000001, val accuracy: 0.788, lr: 1e-05
Finishing epoch 173/200: train loss: 102.84705427386893, train accuracy:
0.8266666666666667, val accuracy: 0.788, lr: 1e-05
Finishing epoch 174/200: train loss: 102.84705427386893, train accuracy:
0.8266666666666667, val accuracy: 0.788, lr: 1e-05
Finishing epoch 175/200: train loss: 102.79295228947703, train accuracy:
0.8273333333333334, val accuracy: 0.79, lr: 1e-05
Finishing epoch 176/200: train loss: 102.76264923127059, train accuracy:
0.8286666666666667, val accuracy: 0.79, lr: 1e-05
Finishing epoch 177/200: train loss: 102.76264923127059, train accuracy:
0.8286666666666667, val accuracy: 0.79, lr: 1e-05
Finishing epoch 178/200: train loss: 102.66574784332818, train accuracy:
0.8293333333333334, val accuracy: 0.792, lr: 1e-05
Finishing epoch 179/200: train loss: 102.55003503404592, train accuracy:
0.8273333333333334, val accuracy: 0.792, lr: 1e-05
Finishing epoch 180/200: train loss: 102.55003503404592, train accuracy:
0.8273333333333334, val accuracy: 0.792, lr: 1e-05
Finishing epoch 181/200: train loss: 102.34561652760003, train accuracy:
0.8253333333333334, val accuracy: 0.79, lr: 1e-05
Finishing epoch 182/200: train loss: 102.29185806821343, train accuracy:
0.8253333333333334, val accuracy: 0.788, lr: 1e-05
Finishing epoch 183/200: train loss: 102.29185806821343, train accuracy:
0.8253333333333334, val accuracy: 0.788, lr: 1e-05
Finishing epoch 184/200: train loss: 102.17850552604924, train accuracy:

```
0.8253333333333334, val accuracy: 0.79, lr: 1e-05
Finishing epoch 185/200: train loss: 102.12255914200045, train accuracy:
0.8246666666666667, val accuracy: 0.79, lr: 1e-05
Finishing epoch 186/200: train loss: 102.08163063359734, train accuracy:
0.8253333333333334, val accuracy: 0.79, lr: 1e-05
Finishing epoch 187/200: train loss: 102.0619751566992, train accuracy:
0.8246666666666667, val accuracy: 0.79, lr: 1e-05
Finishing epoch 188/200: train loss: 102.0619751566992, train accuracy:
0.8246666666666667, val accuracy: 0.79, lr: 1e-05
Finishing epoch 189/200: train loss: 102.0619751566992, train accuracy:
0.8246666666666667, val accuracy: 0.79, lr: 1e-05
Finishing epoch 190/200: train loss: 102.0619751566992, train accuracy:
0.8246666666666667, val accuracy: 0.79, lr: 1e-05
Finishing epoch 191/200: train loss: 102.0619751566992, train accuracy:
0.8246666666666667, val accuracy: 0.79, lr: 1e-05
Finishing epoch 192/200: train loss: 101.84573719019713, train accuracy:
0.8246666666666667, val accuracy: 0.792, lr: 1e-05
Finishing epoch 193/200: train loss: 101.84573719019713, train accuracy:
0.8246666666666667, val accuracy: 0.792, lr: 1e-05
Finishing epoch 194/200: train loss: 101.79257206835639, train accuracy:
0.8246666666666667, val accuracy: 0.794, lr: 1e-05
Finishing epoch 195/200: train loss: 101.71434921596187, train accuracy:
0.8253333333333334, val accuracy: 0.792, lr: 1e-05
Finishing epoch 196/200: train loss: 101.71434921596187, train accuracy:
0.8253333333333334, val accuracy: 0.792, lr: 1e-05
Finishing epoch 197/200: train loss: 101.71434921596187, train accuracy:
0.8253333333333334, val accuracy: 0.792, lr: 1e-05
Finishing epoch 198/200: train loss: 101.71434921596187, train accuracy:
0.8253333333333334, val accuracy: 0.792, lr: 1e-05
Finishing epoch 199/200: train loss: 101.62537511094605, train accuracy:
0.8306666666666667, val accuracy: 0.798, lr: 1e-05
Finishing epoch 200/200: train loss: 101.51241303207607, train accuracy:
0.8266666666666667, val accuracy: 0.798, lr: 1e-05
Training ends.
```

```python
[32]: #Plot the loss and accuracy
fig = plt.figure(figsize=(15,5))

ax = fig.add_subplot(121)
plt.semilogy(losses)
plt.title('Loss')
plt.xlabel('Epochs')
plt.ylabel('Sum of Squares Error')

ax = fig.add_subplot(122)
plt.plot(accuracies,'.')
plt.title('Accuracy')
```

```
plt.xlabel('Epochs')
plt.ylabel('Proportion of Correct Classifications')
#plt.ylim([0,1])
plt.grid()

plt.show()
```
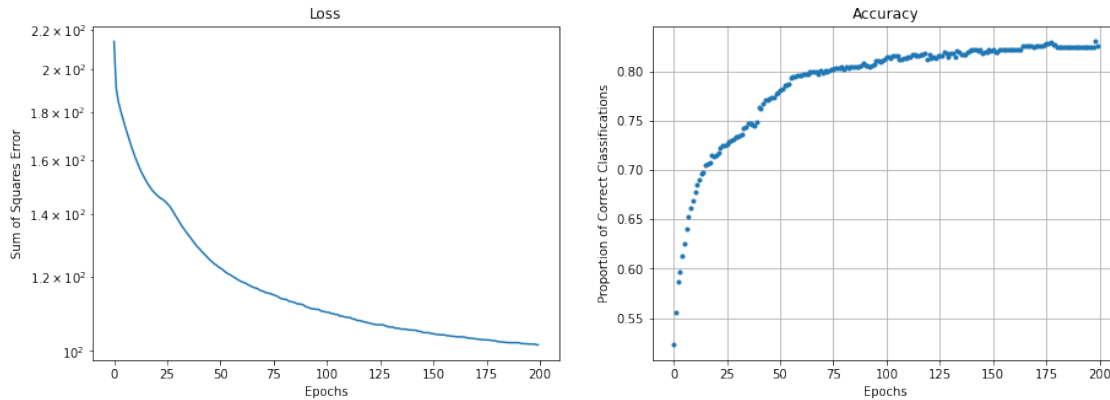

As we can see from above two figures, disabling data normalization reduces the train accuracy from 83.8% to 82.6%.

5 Step 4: Test overfitting

Now we select a small portion of the training dataset (first 30 data points) and see how it affects the validation accuracy.

```
[33]: model = four_layers_model(neurons_1, neurons_2)

is_data_prepossing = True # set True if use data normalization
regularization = 0.0

# Read dataset
with h5py.File('./dataset/train_dataset.h5', 'r') as file:
    data = np.array(file.get('data'))
    label = np.array(file.get('label'))
    X_train = data
    y_train = label

with h5py.File('./dataset/test_dataset.h5', 'r') as file:
    data = np.array(file.get('data'))
    label = np.array(file.get('label'))
    X_test = data
    y_test = label
```

```python
# Data Preprossing
if is_data_prepossing:
    X_train = data_preprossing(X_train)

X_train, X_val, y_train, y_val = train_test_split(X_train, y_train, test_size=0.
 →25, random_state=66)

y_train = np.expand_dims(y_train, -1)
y_val = np.expand_dims(y_val, -1)
y_test = np.expand_dims(y_test, -1)

h = 1e-4 # step size for parameter update

accuracies, losses, val_accuracies, par4 = model.train(X_train[:30],y_train[:
 →30],X_val,y_val,epochs,h,batch_size,regularization,best_model=False)
```

Finishing epoch 1/200: train loss: 6.606461322562772, train accuracy:
0.6333333333333333, val accuracy: 0.492, lr: 0.0001
Finishing epoch 2/200: train loss: 6.378347927864325, train accuracy: 0.6, val
accuracy: 0.522, lr: 0.0001
Finishing epoch 3/200: train loss: 6.245600158424419, train accuracy:
0.5333333333333333, val accuracy: 0.518, lr: 0.0001
Finishing epoch 4/200: train loss: 6.197518049125836, train accuracy:
0.5666666666666667, val accuracy: 0.52, lr: 0.0001
Finishing epoch 5/200: train loss: 6.164639110959747, train accuracy: 0.5, val
accuracy: 0.522, lr: 0.0001
Finishing epoch 6/200: train loss: 6.162948835636403, train accuracy:
0.5333333333333333, val accuracy: 0.522, lr: 0.0001
Finishing epoch 7/200: train loss: 6.1380779155731675, train accuracy: 0.5, val
accuracy: 0.526, lr: 0.0001
Finishing epoch 8/200: train loss: 6.171984553777369, train accuracy: 0.5, val
accuracy: 0.52, lr: 0.0001
Finishing epoch 9/200: train loss: 6.2031588444661, train accuracy: 0.5, val
accuracy: 0.518, lr: 0.0001
Finishing epoch 10/200: train loss: 6.243138481035977, train accuracy:
0.5333333333333333, val accuracy: 0.518, lr: 0.0001
Finishing epoch 11/200: train loss: 6.274168522986495, train accuracy:
0.5333333333333333, val accuracy: 0.518, lr: 0.0001
Finishing epoch 12/200: train loss: 6.306828538965645, train accuracy:
0.5666666666666667, val accuracy: 0.512, lr: 0.0001
Finishing epoch 13/200: train loss: 6.308173224854158, train accuracy:
0.5666666666666667, val accuracy: 0.498, lr: 0.0001
Finishing epoch 14/200: train loss: 6.282187142029784, train accuracy: 0.6, val
accuracy: 0.498, lr: 0.0001
Finishing epoch 15/200: train loss: 6.210571482174226, train accuracy:
0.5666666666666667, val accuracy: 0.5, lr: 0.0001
Finishing epoch 16/200: train loss: 6.143439685466978, train accuracy:

0.5666666666666667, val accuracy: 0.504, lr: 0.0001
Finishing epoch 17/200: train loss: 6.070289207915778, train accuracy:
0.5666666666666667, val accuracy: 0.502, lr: 0.0001
Finishing epoch 18/200: train loss: 5.983749934060652, train accuracy:
0.5666666666666667, val accuracy: 0.502, lr: 0.0001
Finishing epoch 19/200: train loss: 5.92200020698418, train accuracy:
0.5666666666666667, val accuracy: 0.506, lr: 0.0001
Finishing epoch 20/200: train loss: 5.936899799174089, train accuracy:
0.5333333333333333, val accuracy: 0.514, lr: 0.0001
Finishing epoch 21/200: train loss: 5.955983413140467, train accuracy:
0.5333333333333333, val accuracy: 0.514, lr: 0.0001
Finishing epoch 22/200: train loss: 5.954115421741761, train accuracy:
0.5333333333333333, val accuracy: 0.514, lr: 0.0001
Finishing epoch 23/200: train loss: 5.969449062866748, train accuracy:
0.5333333333333333, val accuracy: 0.512, lr: 0.0001
Finishing epoch 24/200: train loss: 5.918211882590086, train accuracy:
0.5333333333333333, val accuracy: 0.512, lr: 0.0001
Finishing epoch 25/200: train loss: 5.867443371215426, train accuracy:
0.5333333333333333, val accuracy: 0.512, lr: 0.0001
Finishing epoch 26/200: train loss: 5.815157354754151, train accuracy:
0.5333333333333333, val accuracy: 0.514, lr: 0.0001
Finishing epoch 27/200: train loss: 5.7690712863343006, train accuracy:
0.5333333333333333, val accuracy: 0.512, lr: 0.0001
Finishing epoch 28/200: train loss: 5.67360842444209, train accuracy:
0.5333333333333333, val accuracy: 0.51, lr: 0.0001
Finishing epoch 29/200: train loss: 5.552978639756834, train accuracy:
0.5333333333333333, val accuracy: 0.514, lr: 0.0001
Finishing epoch 30/200: train loss: 5.430775528824727, train accuracy:
0.5333333333333333, val accuracy: 0.514, lr: 0.0001
Finishing epoch 31/200: train loss: 5.315897937720328, train accuracy:
0.5333333333333333, val accuracy: 0.516, lr: 0.0001
Finishing epoch 32/200: train loss: 5.227159602241304, train accuracy:
0.5333333333333333, val accuracy: 0.526, lr: 0.0001
Finishing epoch 33/200: train loss: 5.158945666089228, train accuracy:
0.5333333333333333, val accuracy: 0.524, lr: 0.0001
Finishing epoch 34/200: train loss: 5.082625710506176, train accuracy:
0.5666666666666667, val accuracy: 0.526, lr: 0.0001
Finishing epoch 35/200: train loss: 4.989507982925159, train accuracy:
0.5666666666666667, val accuracy: 0.528, lr: 0.0001
Finishing epoch 36/200: train loss: 4.877413886526865, train accuracy:
0.5666666666666667, val accuracy: 0.536, lr: 0.0001
Finishing epoch 37/200: train loss: 4.7796474917593095, train accuracy:
0.5666666666666667, val accuracy: 0.536, lr: 0.0001
Finishing epoch 38/200: train loss: 4.7043234578406885, train accuracy: 0.6, val
accuracy: 0.546, lr: 0.0001
Finishing epoch 39/200: train loss: 4.666404946438474, train accuracy: 0.6, val
accuracy: 0.536, lr: 0.0001
Finishing epoch 40/200: train loss: 4.631020106297905, train accuracy: 0.6, val

accuracy: 0.536, lr: 0.0001
Finishing epoch 41/200: train loss: 4.590933392433474, train accuracy: 0.6, val
accuracy: 0.542, lr: 0.0001
Finishing epoch 42/200: train loss: 4.552279575539099, train accuracy:
0.6333333333333333, val accuracy: 0.534, lr: 0.0001
Finishing epoch 43/200: train loss: 4.526835788872965, train accuracy:
0.6333333333333333, val accuracy: 0.538, lr: 0.0001
Finishing epoch 44/200: train loss: 4.482130237991514, train accuracy:
0.6333333333333333, val accuracy: 0.536, lr: 0.0001
Finishing epoch 45/200: train loss: 4.436558746122367, train accuracy:
0.6333333333333333, val accuracy: 0.534, lr: 0.0001
Finishing epoch 46/200: train loss: 4.389000700795929, train accuracy:
0.6333333333333333, val accuracy: 0.536, lr: 0.0001
Finishing epoch 47/200: train loss: 4.350434663519199, train accuracy:
0.6333333333333333, val accuracy: 0.534, lr: 0.0001
Finishing epoch 48/200: train loss: 4.310116286103817, train accuracy:
0.6333333333333333, val accuracy: 0.536, lr: 0.0001
Finishing epoch 49/200: train loss: 4.278519964720982, train accuracy:
0.6333333333333333, val accuracy: 0.536, lr: 0.0001
Finishing epoch 50/200: train loss: 4.244351880758566, train accuracy:
0.6666666666666667, val accuracy: 0.536, lr: 0.0001
Finishing epoch 51/200: train loss: 4.209601161599997, train accuracy:
0.6666666666666667, val accuracy: 0.538, lr: 0.0001
Finishing epoch 52/200: train loss: 4.1767186853530855, train accuracy:
0.6666666666666667, val accuracy: 0.532, lr: 0.0001
Finishing epoch 53/200: train loss: 4.133300716572657, train accuracy:
0.6666666666666667, val accuracy: 0.526, lr: 0.0001
Finishing epoch 54/200: train loss: 4.100511245423583, train accuracy:
0.6666666666666667, val accuracy: 0.534, lr: 0.0001
Finishing epoch 55/200: train loss: 4.06482379419846, train accuracy:
0.6666666666666667, val accuracy: 0.53, lr: 0.0001
Finishing epoch 56/200: train loss: 4.036550128048578, train accuracy: 0.7, val
accuracy: 0.528, lr: 0.0001
Finishing epoch 57/200: train loss: 4.005324147027791, train accuracy: 0.7, val
accuracy: 0.53, lr: 0.0001
Finishing epoch 58/200: train loss: 3.975648581392282, train accuracy: 0.7, val
accuracy: 0.532, lr: 0.0001
Finishing epoch 59/200: train loss: 3.9475450415228224, train accuracy: 0.7, val
accuracy: 0.53, lr: 0.0001
Finishing epoch 60/200: train loss: 3.9220166579840505, train accuracy: 0.7, val
accuracy: 0.53, lr: 0.0001
Finishing epoch 61/200: train loss: 3.891965375521223, train accuracy: 0.7, val
accuracy: 0.53, lr: 0.0001
Finishing epoch 62/200: train loss: 3.863721018783405, train accuracy: 0.7, val
accuracy: 0.532, lr: 0.0001
Finishing epoch 63/200: train loss: 3.837879016139267, train accuracy: 0.7, val
accuracy: 0.532, lr: 0.0001
Finishing epoch 64/200: train loss: 3.8153260664011843, train accuracy: 0.7, val

accuracy: 0.534, lr: 0.0001
Finishing epoch 65/200: train loss: 3.7955359747695994, train accuracy: 0.7, val accuracy: 0.53, lr: 0.0001
Finishing epoch 66/200: train loss: 3.772189516276037, train accuracy: 0.7, val accuracy: 0.534, lr: 0.0001
Finishing epoch 67/200: train loss: 3.7356272604146477, train accuracy: 0.7333333333333334, val accuracy: 0.534, lr: 0.0001
Finishing epoch 68/200: train loss: 3.704269575736231, train accuracy: 0.7333333333333334, val accuracy: 0.536, lr: 0.0001
Finishing epoch 69/200: train loss: 3.669922434734377, train accuracy: 0.7333333333333334, val accuracy: 0.538, lr: 0.0001
Finishing epoch 70/200: train loss: 3.6378207903130395, train accuracy: 0.7333333333333334, val accuracy: 0.54, lr: 0.0001
Finishing epoch 71/200: train loss: 3.6101667128219788, train accuracy: 0.7333333333333334, val accuracy: 0.538, lr: 0.0001
Finishing epoch 72/200: train loss: 3.588887645968681, train accuracy: 0.7333333333333334, val accuracy: 0.54, lr: 0.0001
Finishing epoch 73/200: train loss: 3.5710219688792955, train accuracy: 0.7333333333333334, val accuracy: 0.54, lr: 0.0001
Finishing epoch 74/200: train loss: 3.552265395028984, train accuracy: 0.7333333333333334, val accuracy: 0.54, lr: 0.0001
Finishing epoch 75/200: train loss: 3.5392933423803594, train accuracy: 0.7333333333333334, val accuracy: 0.54, lr: 0.0001
Finishing epoch 76/200: train loss: 3.5263353975495493, train accuracy: 0.7333333333333334, val accuracy: 0.54, lr: 0.0001
Finishing epoch 77/200: train loss: 3.51747274870765, train accuracy: 0.7333333333333334, val accuracy: 0.538, lr: 0.0001
Finishing epoch 78/200: train loss: 3.5117639210307656, train accuracy: 0.7333333333333334, val accuracy: 0.54, lr: 0.0001
Finishing epoch 79/200: train loss: 3.505260441777528, train accuracy: 0.7333333333333334, val accuracy: 0.54, lr: 0.0001
Finishing epoch 80/200: train loss: 3.4990059213772837, train accuracy: 0.7, val accuracy: 0.542, lr: 0.0001
Finishing epoch 81/200: train loss: 3.4934372488951584, train accuracy: 0.7, val accuracy: 0.538, lr: 0.0001
Finishing epoch 82/200: train loss: 3.4875412670305237, train accuracy: 0.7, val accuracy: 0.538, lr: 0.0001
Finishing epoch 83/200: train loss: 3.483245086871309, train accuracy: 0.7, val accuracy: 0.538, lr: 0.0001
Finishing epoch 84/200: train loss: 3.4784139824595113, train accuracy: 0.7, val accuracy: 0.536, lr: 0.0001
Finishing epoch 85/200: train loss: 3.474155533988056, train accuracy: 0.7, val accuracy: 0.538, lr: 0.0001
Finishing epoch 86/200: train loss: 3.4681827607184985, train accuracy: 0.7, val accuracy: 0.538, lr: 0.0001
Finishing epoch 87/200: train loss: 3.4663730446751164, train accuracy: 0.6666666666666667, val accuracy: 0.536, lr: 0.0001
Finishing epoch 88/200: train loss: 3.463249012659525, train accuracy:

0.6666666666666667, val accuracy: 0.538, lr: 0.0001
Finishing epoch 89/200: train loss: 3.45596162629524265, train accuracy: 0.7, val accuracy: 0.54, lr: 0.0001
Finishing epoch 90/200: train loss: 3.4581720591074308, train accuracy: 0.7, val accuracy: 0.536, lr: 0.0001
Finishing epoch 91/200: train loss: 3.455793999857028, train accuracy: 0.7, val accuracy: 0.536, lr: 0.0001
Finishing epoch 92/200: train loss: 3.4489608062433974, train accuracy: 0.7, val accuracy: 0.538, lr: 0.0001
Finishing epoch 93/200: train loss: 3.440104407689661, train accuracy: 0.7333333333333334, val accuracy: 0.54, lr: 0.0001
Finishing epoch 94/200: train loss: 3.43390756943902, train accuracy: 0.7333333333333334, val accuracy: 0.542, lr: 0.0001
Finishing epoch 95/200: train loss: 3.4260013871389434, train accuracy: 0.7333333333333334, val accuracy: 0.544, lr: 0.0001
Finishing epoch 96/200: train loss: 3.419941808452329, train accuracy: 0.7, val accuracy: 0.546, lr: 0.0001
Finishing epoch 97/200: train loss: 3.4128055342053196, train accuracy: 0.7, val accuracy: 0.546, lr: 0.0001
Finishing epoch 98/200: train loss: 3.4065366269128736, train accuracy: 0.7, val accuracy: 0.548, lr: 0.0001
Finishing epoch 99/200: train loss: 3.400141112803362, train accuracy: 0.7, val accuracy: 0.548, lr: 0.0001
Finishing epoch 100/200: train loss: 3.3959038055177, train accuracy: 0.7, val accuracy: 0.548, lr: 0.0001
Finishing epoch 101/200: train loss: 3.3906573932030932, train accuracy: 0.7, val accuracy: 0.546, lr: 0.0001
Finishing epoch 102/200: train loss: 3.3834789562521133, train accuracy: 0.7, val accuracy: 0.544, lr: 0.0001
Finishing epoch 103/200: train loss: 3.385716110405143, train accuracy: 0.7, val accuracy: 0.544, lr: 0.0001
Finishing epoch 104/200: train loss: 3.386843446316744, train accuracy: 0.7, val accuracy: 0.544, lr: 0.0001
Finishing epoch 105/200: train loss: 3.3906800240351203, train accuracy: 0.7, val accuracy: 0.544, lr: 0.0001
Finishing epoch 106/200: train loss: 3.395831178423827, train accuracy: 0.7, val accuracy: 0.542, lr: 0.0001
Finishing epoch 107/200: train loss: 3.4004613435919793, train accuracy: 0.7, val accuracy: 0.544, lr: 0.0001
Finishing epoch 108/200: train loss: 3.404993054185523, train accuracy: 0.7, val accuracy: 0.544, lr: 0.0001
Finishing epoch 109/200: train loss: 3.407876602454216, train accuracy: 0.7, val accuracy: 0.546, lr: 0.0001
Finishing epoch 110/200: train loss: 3.4139323818764122, train accuracy: 0.7, val accuracy: 0.55, lr: 0.0001
Finishing epoch 111/200: train loss: 3.4164813186554817, train accuracy: 0.7, val accuracy: 0.55, lr: 0.0001
Finishing epoch 112/200: train loss: 3.4233019372239, train accuracy: 0.7, val

accuracy: 0.55, lr: 0.0001
Finishing epoch 113/200: train loss: 3.425733459950118, train accuracy: 0.7, val accuracy: 0.55, lr: 0.0001
Finishing epoch 114/200: train loss: 3.430145745931786, train accuracy: 0.7, val accuracy: 0.552, lr: 0.0001
Finishing epoch 115/200: train loss: 3.4325611734245394, train accuracy: 0.6666666666666667, val accuracy: 0.55, lr: 0.0001
Finishing epoch 116/200: train loss: 3.438897316165728, train accuracy: 0.6666666666666667, val accuracy: 0.55, lr: 0.0001
Finishing epoch 117/200: train loss: 3.444380616801472, train accuracy: 0.6666666666666667, val accuracy: 0.554, lr: 0.0001
Finishing epoch 118/200: train loss: 3.447707535394359, train accuracy: 0.6666666666666667, val accuracy: 0.554, lr: 0.0001
Finishing epoch 119/200: train loss: 3.450811245909908, train accuracy: 0.7, val accuracy: 0.554, lr: 0.0001
Finishing epoch 120/200: train loss: 3.4512815630068525, train accuracy: 0.7, val accuracy: 0.556, lr: 0.0001
Finishing epoch 121/200: train loss: 3.4505201043481057, train accuracy: 0.7, val accuracy: 0.556, lr: 0.0001
Finishing epoch 122/200: train loss: 3.4525110266310963, train accuracy: 0.7, val accuracy: 0.56, lr: 0.0001
Finishing epoch 123/200: train loss: 3.454365280546846, train accuracy: 0.7, val accuracy: 0.562, lr: 0.0001
Finishing epoch 124/200: train loss: 3.454746658215849, train accuracy: 0.7, val accuracy: 0.562, lr: 0.0001
Finishing epoch 125/200: train loss: 3.456945042445501, train accuracy: 0.7, val accuracy: 0.562, lr: 0.0001
Finishing epoch 126/200: train loss: 3.4584377532176824, train accuracy: 0.7, val accuracy: 0.56, lr: 0.0001
Finishing epoch 127/200: train loss: 3.4635572889846293, train accuracy: 0.7, val accuracy: 0.562, lr: 0.0001
Finishing epoch 128/200: train loss: 3.466766683594689, train accuracy: 0.7, val accuracy: 0.562, lr: 0.0001
Finishing epoch 129/200: train loss: 3.47209973589899, train accuracy: 0.7, val accuracy: 0.56, lr: 0.0001
Finishing epoch 130/200: train loss: 3.4789784075119594, train accuracy: 0.7333333333333334, val accuracy: 0.56, lr: 0.0001
Finishing epoch 131/200: train loss: 3.488749169068901, train accuracy: 0.7333333333333334, val accuracy: 0.56, lr: 0.0001
Finishing epoch 132/200: train loss: 3.497709260201506, train accuracy: 0.7333333333333334, val accuracy: 0.56, lr: 0.0001
Finishing epoch 133/200: train loss: 3.5067691696242465, train accuracy: 0.7666666666666666, val accuracy: 0.558, lr: 0.0001
Finishing epoch 134/200: train loss: 3.515558242827262, train accuracy: 0.8, val accuracy: 0.56, lr: 0.0001
Finishing epoch 135/200: train loss: 3.525396380372228, train accuracy: 0.8, val accuracy: 0.5640000000000001, lr: 0.0001
Finishing epoch 136/200: train loss: 3.5342571381918417, train accuracy:

0.7666666666666666, val accuracy: 0.56, lr: 0.0001
Finishing epoch 137/200: train loss: 3.546815626767504, train accuracy:
0.7666666666666666, val accuracy: 0.5660000000000001, lr: 0.0001
Finishing epoch 138/200: train loss: 3.557387102346045, train accuracy:
0.7666666666666666, val accuracy: 0.5680000000000001, lr: 0.0001
Finishing epoch 139/200: train loss: 3.5667952728436076, train accuracy:
0.7666666666666666, val accuracy: 0.5700000000000001, lr: 0.0001
Finishing epoch 140/200: train loss: 3.5755965423889062, train accuracy: 0.8,
val accuracy: 0.5720000000000001, lr: 0.0001
Finishing epoch 141/200: train loss: 3.5823146418989533, train accuracy: 0.8,
val accuracy: 0.5720000000000001, lr: 0.0001
Finishing epoch 142/200: train loss: 3.5930901051494892, train accuracy: 0.8,
val accuracy: 0.5780000000000001, lr: 0.0001
Finishing epoch 143/200: train loss: 3.6031249757361934, train accuracy: 0.8,
val accuracy: 0.5800000000000001, lr: 0.0001
Finishing epoch 144/200: train loss: 3.615467659161194, train accuracy: 0.8, val
accuracy: 0.5860000000000001, lr: 0.0001
Finishing epoch 145/200: train loss: 3.630723870475281, train accuracy: 0.8, val
accuracy: 0.5900000000000001, lr: 0.0001
Finishing epoch 146/200: train loss: 3.644279167203788, train accuracy: 0.8, val
accuracy: 0.5900000000000001, lr: 0.0001
Finishing epoch 147/200: train loss: 3.658988894092836, train accuracy: 0.8, val
accuracy: 0.5880000000000001, lr: 0.0001
Finishing epoch 148/200: train loss: 3.67990442169709, train accuracy: 0.8, val
accuracy: 0.5880000000000001, lr: 0.0001
Finishing epoch 149/200: train loss: 3.6957484911277696, train accuracy: 0.8,
val accuracy: 0.5880000000000001, lr: 0.0001
Finishing epoch 150/200: train loss: 3.717550350198022, train accuracy: 0.8, val
accuracy: 0.5880000000000001, lr: 0.0001
Finishing epoch 151/200: train loss: 3.7322755539648926, train accuracy: 0.8,
val accuracy: 0.5920000000000001, lr: 0.0001
Finishing epoch 152/200: train loss: 3.7508717636275883, train accuracy: 0.8,
val accuracy: 0.5920000000000001, lr: 0.0001
Finishing epoch 153/200: train loss: 3.7638347705717337, train accuracy: 0.8,
val accuracy: 0.5900000000000001, lr: 0.0001
Finishing epoch 154/200: train loss: 3.773924511264913, train accuracy: 0.8, val
accuracy: 0.5900000000000001, lr: 0.0001
Finishing epoch 155/200: train loss: 3.786593433535135, train accuracy: 0.8, val
accuracy: 0.5920000000000001, lr: 0.0001
Finishing epoch 156/200: train loss: 3.797341127895872, train accuracy: 0.8, val
accuracy: 0.5900000000000001, lr: 0.0001
Finishing epoch 157/200: train loss: 3.8107448086701377, train accuracy: 0.8,
val accuracy: 0.5880000000000001, lr: 0.0001
Finishing epoch 158/200: train loss: 3.820595540374927, train accuracy: 0.8, val
accuracy: 0.5920000000000001, lr: 0.0001
Finishing epoch 159/200: train loss: 3.83081048700745, train accuracy: 0.8, val
accuracy: 0.598, lr: 0.0001
Finishing epoch 160/200: train loss: 3.8423422932613716, train accuracy: 0.8,

val accuracy: 0.602, lr: 0.0001
Finishing epoch 161/200: train loss: 3.853152446127102, train accuracy: 0.8, val accuracy: 0.606, lr: 0.0001
Finishing epoch 162/200: train loss: 3.8645982903625473, train accuracy: 0.8, val accuracy: 0.608, lr: 0.0001
Finishing epoch 163/200: train loss: 3.8760612397289957, train accuracy: 0.8, val accuracy: 0.608, lr: 0.0001
Finishing epoch 164/200: train loss: 3.8861429554149325, train accuracy: 0.8, val accuracy: 0.608, lr: 0.0001
Finishing epoch 165/200: train loss: 3.895462418777474, train accuracy: 0.8, val accuracy: 0.61, lr: 0.0001
Finishing epoch 166/200: train loss: 3.9062996442064866, train accuracy: 0.8, val accuracy: 0.61, lr: 0.0001
Finishing epoch 167/200: train loss: 3.9195211219123314, train accuracy: 0.8, val accuracy: 0.616, lr: 0.0001
Finishing epoch 168/200: train loss: 3.9306455260981896, train accuracy: 0.8, val accuracy: 0.612, lr: 0.0001
Finishing epoch 169/200: train loss: 3.942246867540024, train accuracy: 0.8, val accuracy: 0.612, lr: 0.0001
Finishing epoch 170/200: train loss: 3.9529663540385664, train accuracy: 0.8, val accuracy: 0.61, lr: 0.0001
Finishing epoch 171/200: train loss: 3.959407722896625, train accuracy: 0.8, val accuracy: 0.612, lr: 0.0001
Finishing epoch 172/200: train loss: 3.975186577982079, train accuracy: 0.8, val accuracy: 0.61, lr: 0.0001
Finishing epoch 173/200: train loss: 3.989085286851263, train accuracy: 0.8, val accuracy: 0.612, lr: 0.0001
Finishing epoch 174/200: train loss: 4.008065003851739, train accuracy: 0.8, val accuracy: 0.614, lr: 0.0001
Finishing epoch 175/200: train loss: 4.018763571190638, train accuracy: 0.8, val accuracy: 0.618, lr: 0.0001
Finishing epoch 176/200: train loss: 4.028357658170219, train accuracy: 0.8, val accuracy: 0.618, lr: 0.0001
Finishing epoch 177/200: train loss: 4.034911605902526, train accuracy: 0.8, val accuracy: 0.622, lr: 0.0001
Finishing epoch 178/200: train loss: 4.04254281789934, train accuracy: 0.8, val accuracy: 0.624, lr: 0.0001
Finishing epoch 179/200: train loss: 4.053859562432221, train accuracy: 0.8, val accuracy: 0.622, lr: 0.0001
Finishing epoch 180/200: train loss: 4.065490051168895, train accuracy: 0.8, val accuracy: 0.624, lr: 0.0001
Finishing epoch 181/200: train loss: 4.074108704430035, train accuracy: 0.8, val accuracy: 0.622, lr: 0.0001
Finishing epoch 182/200: train loss: 4.085035381667734, train accuracy: 0.8, val accuracy: 0.624, lr: 0.0001
Finishing epoch 183/200: train loss: 4.099852403565226, train accuracy: 0.8, val accuracy: 0.626, lr: 0.0001
Finishing epoch 184/200: train loss: 4.107787602262039, train accuracy: 0.8, val

```
accuracy: 0.63, lr: 0.0001
Finishing epoch 185/200: train loss: 4.118556838582208, train accuracy: 0.8, val
accuracy: 0.63, lr: 0.0001
Finishing epoch 186/200: train loss: 4.1250271045010205, train accuracy: 0.8,
val accuracy: 0.63, lr: 0.0001
Finishing epoch 187/200: train loss: 4.134900839783985, train accuracy: 0.8, val
accuracy: 0.634, lr: 0.0001
Finishing epoch 188/200: train loss: 4.143279260933232, train accuracy: 0.8, val
accuracy: 0.636, lr: 0.0001
Finishing epoch 189/200: train loss: 4.151962959674552, train accuracy: 0.8, val
accuracy: 0.638, lr: 0.0001
Finishing epoch 190/200: train loss: 4.162010361159248, train accuracy: 0.8, val
accuracy: 0.64, lr: 0.0001
Finishing epoch 191/200: train loss: 4.1737166907932215, train accuracy: 0.8,
val accuracy: 0.638, lr: 0.0001
Finishing epoch 192/200: train loss: 4.178982600724535, train accuracy: 0.8, val
accuracy: 0.636, lr: 0.0001
Finishing epoch 193/200: train loss: 4.18505997013202, train accuracy: 0.8, val
accuracy: 0.636, lr: 0.0001
Finishing epoch 194/200: train loss: 4.189805324401196, train accuracy:
0.8333333333333334, val accuracy: 0.64, lr: 0.0001
Finishing epoch 195/200: train loss: 4.193573004497507, train accuracy:
0.8333333333333334, val accuracy: 0.638, lr: 0.0001
Finishing epoch 196/200: train loss: 4.196524141047311, train accuracy:
0.8333333333333334, val accuracy: 0.638, lr: 0.0001
Finishing epoch 197/200: train loss: 4.200849669413172, train accuracy:
0.8333333333333334, val accuracy: 0.638, lr: 0.0001
Finishing epoch 198/200: train loss: 4.204905026776153, train accuracy:
0.8333333333333334, val accuracy: 0.634, lr: 0.0001
Finishing epoch 199/200: train loss: 4.206324283029954, train accuracy:
0.8333333333333334, val accuracy: 0.632, lr: 0.0001
Finishing epoch 200/200: train loss: 4.207022423917373, train accuracy:
0.8333333333333334, val accuracy: 0.634, lr: 0.0001
Training ends.
```

[34]:
```python
#Plot the loss and accuracy
fig = plt.figure(figsize=(15,5))

ax = fig.add_subplot(121)
plt.semilogy(losses)
plt.title('Loss')
plt.xlabel('Epochs')
plt.ylabel('Sum of Squares Error')

ax = fig.add_subplot(122)
plt.plot(accuracies,'.')
plt.title('Accuracy')
```

```
plt.xlabel('Epochs')
plt.ylabel('Proportion of Correct Classifications')
#plt.ylim([0,1])
plt.grid()

plt.show()
```
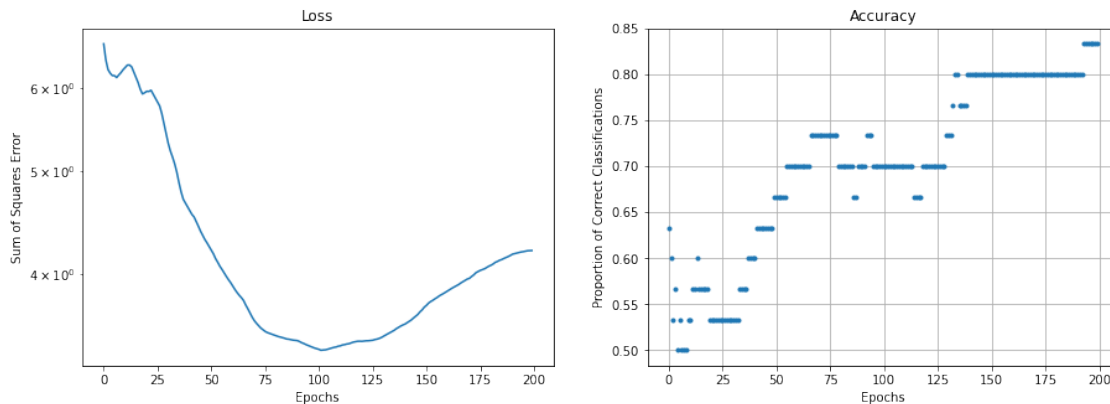

As we can see in the above figures, when epochs go beyong 100, the loss rises again. Although we still get about 83% train accuracy, the validation accuracy decreases from 81% to 63.4%. This shows that the current network is overfitted to the current train dataset.

6   Step 5: Hyperparameter tuning

In the previous steps, I use the step size 1e-5 by default and disable regularization. Now, I will train the network for 20 epochs so as to tune the step size and regularization to search for the suitable parameters.

```
[35]: is_data_prepossing = True # set True if use data normalization

      # Read dataset
      with h5py.File('./dataset/train_dataset.h5', 'r') as file:
          data = np.array(file.get('data'))
          label = np.array(file.get('label'))
          X_train = data
          y_train = label

      with h5py.File('./dataset/test_dataset.h5', 'r') as file:
          data = np.array(file.get('data'))
          label = np.array(file.get('label'))
          X_test = data
          y_test = label
```

```python
# Data Preprossing
if is_data_prepossing:
    X_train = data_preprossing(X_train)

X_train, X_val, y_train, y_val = train_test_split(X_train, y_train, test_size=0.
 ↪25, random_state=66)


y_train = np.expand_dims(y_train, -1)
y_val = np.expand_dims(y_val, -1)
y_test = np.expand_dims(y_test, -1)




max_count = 100
model_save = []
epochs = 20
neurons_1 = 50
neurons_2 = 50
batch_size = 150

reg = 10**np.random.uniform(-5,0,max_count)
lr = 10**np.random.uniform(-3,-6,max_count)

for count in range(max_count):
    model = four_layers_model(neurons_1, neurons_2)
    accuracies, losses, val_accuracies, params = model.
 ↪train(X_train,y_train,X_val,y_val,epochs,lr[count],batch_size,reg[count],best_model=True,␣
 ↪verbose=False)
    model_save.append(params)
    print('Round {}/{}: val_acc: {}, lr: {}, reg: {}'.
 ↪format((count+1),max_count,val_accuracies[-1],lr[count],reg[count]))
```

```
Round 1/100: val_acc: 0.612, lr: 0.00035335650584983975, reg:
0.0003001959544457955
Round 2/100: val_acc: 0.628, lr: 9.738534787767991e-06, reg:
7.384454987543527e-05
Round 3/100: val_acc: 0.604, lr: 3.388166935136662e-06, reg:
0.0006802243292630383
Round 4/100: val_acc: 0.772, lr: 0.00012233313919268827, reg:
0.01227918196709563
Round 5/100: val_acc: 0.622, lr: 6.5836468664911736e-06, reg:
2.1856525894297923e-05
Round 6/100: val_acc: 0.754, lr: 0.00014317888802358792, reg:
0.003596260401529178
Round 7/100: val_acc: 0.784, lr: 0.0002936028242368535, reg:
```

0.00022758326707296786
Round 8/100: val_acc: 0.778, lr: 9.450088038393e-05, reg: 0.14358065755405142
Round 9/100: val_acc: 0.69, lr: 0.000827813638330122, reg: 0.0001413342901721463
Round 10/100: val_acc: 0.752, lr: 0.0001819458408405684, reg:
0.004369829781345267
Round 11/100: val_acc: 0.63, lr: 5.006279286895316e-06, reg:
0.049446220272742765
Round 12/100: val_acc: 0.554, lr: 0.0005168887750422141, reg:
0.0016873909173678867
Round 13/100: val_acc: 0.6859999999999999, lr: 1.775349566316394e-05, reg:
0.0005446606990948928
Round 14/100: val_acc: 0.632, lr: 5.311456444894522e-06, reg:
0.0006836135446120462
Round 15/100: val_acc: 0.602, lr: 3.070776022677189e-06, reg:
0.02254658339039537
Round 16/100: val_acc: 0.6719999999999999, lr: 0.00040514445914211595, reg:
0.009399548819938713
Round 17/100: val_acc: 0.652, lr: 0.0009280203513533005, reg:
0.06167609406792775
Round 18/100: val_acc: 0.6799999999999999, lr: 1.75060201102952233e-05, reg:
1.3918997204354817e-05
Round 19/100: val_acc: 0.608, lr: 3.769193297864706e-06, reg:
0.019593427489537855
Round 20/100: val_acc: 0.758, lr: 3.6783104715077385e-05, reg:
0.00128886364910346
Round 21/100: val_acc: 0.598, lr: 2.725985710028663e-06, reg:
0.05509364041881942
Round 22/100: val_acc: 0.706, lr: 0.0008409810110544093, reg:
0.00011599994147191901
Round 23/100: val_acc: 0.516, lr: 1.1101116577727213e-06, reg:
0.16636585736507764
Round 24/100: val_acc: 0.758, lr: 0.00014217798759247165, reg:
0.07577153535690868
Round 25/100: val_acc: 0.782, lr: 8.661163615875086e-05, reg: 0.2484889168597558
Round 26/100: val_acc: 0.652, lr: 0.0009930740601137356, reg:
0.0001416239652282429
Round 27/100: val_acc: 0.6579999999999999, lr: 1.4789196970880551e-05, reg:
0.0012078057969845907
Round 28/100: val_acc: 0.77, lr: 0.00012068255612237557, reg:
3.8829077481152696e-05
Round 29/100: val_acc: 0.766, lr: 0.00023624240939548751, reg:
0.0047479689008633535
Round 30/100: val_acc: 0.774, lr: 6.293003549670438e-05, reg:
0.0005534832562727593
Round 31/100: val_acc: 0.6659999999999999, lr: 1.6108584728491262e-05, reg:
0.004080578343540587
Round 32/100: val_acc: 0.6639999999999999, lr: 1.5816366533057045e-05, reg:
0.014870504794216144

```
Round 33/100: val_acc: 0.628, lr: 1.0544987779001688e-05, reg:
0.003116835606540032
Round 34/100: val_acc: 0.52, lr: 0.0005004460589473129, reg:
0.0013835724316663236
Round 35/100: val_acc: 0.714, lr: 0.0003656442489969515, reg:
0.006056451834862236
Round 36/100: val_acc: 0.696, lr: 1.9985831329604876e-05, reg:
8.880530060926964e-05
Round 37/100: val_acc: 0.644, lr: 0.0005728034546295083, reg:
0.01065881969950628
Round 38/100: val_acc: 0.774, lr: 5.9328941184352167e-05, reg:
0.0001366325906899988
Round 39/100: val_acc: 0.624, lr: 4.819282209637043e-06, reg:
0.14841142563710788
Round 40/100: val_acc: 0.62, lr: 6.672670518859799e-06, reg: 0.4350500165671914
Round 41/100: val_acc: 0.768, lr: 0.00017954830232016204, reg:
0.3066721622178784
Round 42/100: val_acc: 0.714, lr: 2.5185138458909746e-05, reg:
0.006273604709279105
Round 43/100: val_acc: 0.6699999999999999, lr: 1.6326676230545213e-05, reg:
0.13951466385421496
Round 44/100: val_acc: 0.628, lr: 1.0165071444980757e-05, reg:
0.016752857102487455
Round 45/100: val_acc: 0.76, lr: 0.00020750225806864787, reg:
1.3584965063606107e-05
Round 46/100: val_acc: 0.5840000000000001, lr: 2.3282560531062628e-06, reg:
0.2296084645672551
Round 47/100: val_acc: 0.508, lr: 1.0155104777767432e-06, reg:
1.226376139984077e-05
Round 48/100: val_acc: 0.478, lr: 0.000680217918945404, reg: 0.13867365142146582
Round 49/100: val_acc: 0.782, lr: 7.08921194660884e-05, reg: 0.10812240159139032
Round 50/100: val_acc: 0.606, lr: 3.608248305003565e-06, reg:
1.603515856889375e-05
Round 51/100: val_acc: 0.482, lr: 0.0007240747514143083, reg:
4.734071713962365e-05
Round 52/100: val_acc: 0.78, lr: 0.00018985305480054285, reg:
1.4281338451760208e-05
Round 53/100: val_acc: 0.52, lr: 0.0005276501512881268, reg:
0.00014088165691183828
Round 54/100: val_acc: 0.758, lr: 5.858455149092164e-05, reg:
0.00933021588776681
Round 55/100: val_acc: 0.512, lr: 0.0004926176874445555, reg: 0.0365351758309317
Round 56/100: val_acc: 0.516, lr: 1.126922189505302e-06, reg:
5.602993789525594e-05
Round 57/100: val_acc: 0.6759999999999999, lr: 0.00044409840274337774, reg:
0.06483803882810849
Round 58/100: val_acc: 0.642, lr: 1.332832829690164e-05, reg:
0.039349097109995554
```

```
Round 59/100: val_acc: 0.776, lr: 0.00022541663120125493, reg: 0.051077333118692
Round 60/100: val_acc: 0.758, lr: 4.5924152497084855e-05, reg:
1.9119244209913804e-05
Round 61/100: val_acc: 0.5780000000000001, lr: 2.1424185986341756e-06, reg:
0.0002456216814973249
Round 62/100: val_acc: 0.74, lr: 0.00030759133315376463, reg:
0.0033937501770057257
Round 63/100: val_acc: 0.628, lr: 7.48889338726841e-06, reg:
0.00029231788111368565
Round 64/100: val_acc: 0.688, lr: 0.0005384931113807589, reg:
0.00047834188207819664
Round 65/100: val_acc: 0.626, lr: 7.3292015841264915e-06, reg:
0.6726955980730862
Round 66/100: val_acc: 0.626, lr: 7.197603172519336e-06, reg:
8.497698928041516e-05
Round 67/100: val_acc: 0.742, lr: 0.0002344631826473017, reg:
0.0001510593302684319
Round 68/100: val_acc: 0.602, lr: 0.0009504915278794442, reg:
0.14436139116913022
Round 69/100: val_acc: 0.63, lr: 1.079093369096099e-05, reg: 0.8136961603931996
Round 70/100: val_acc: 0.776, lr: 0.00021676005385397532, reg:
0.0013834814093421565
Round 71/100: val_acc: 0.6799999999999999, lr: 0.0003909901038989237, reg:
0.09604044053915876
Round 72/100: val_acc: 0.782, lr: 8.583197453520482e-05, reg:
0.11968474992536456
Round 73/100: val_acc: 0.764, lr: 0.0003358768721745252, reg:
0.13070803855576663
Round 74/100: val_acc: 0.762, lr: 5.648057610565199e-05, reg:
0.04168849349333729
Round 75/100: val_acc: 0.5700000000000001, lr: 0.000494260068154903, reg:
0.031049206128356783
Round 76/100: val_acc: 0.7, lr: 2.1183277773631462e-05, reg:
0.042861520797262104
Round 77/100: val_acc: 0.762, lr: 3.9396996002950126e-05, reg:
0.0001490910472534869
Round 78/100: val_acc: 0.528, lr: 0.0005411429468348775, reg:
0.00019989555885275884
Round 79/100: val_acc: 0.754, lr: 0.000176290135853912, reg: 0.06371277935742746
Round 80/100: val_acc: 0.784, lr: 0.00013186158431527132, reg:
0.0018687631068362717
Round 81/100: val_acc: 0.614, lr: 3.991808428989363e-06, reg:
0.27077630159524146
Round 82/100: val_acc: 0.8, lr: 0.00012500323838883322, reg:
0.010035557536518327
Round 83/100: val_acc: 0.786, lr: 9.727752564399343e-05, reg:
0.013169355900996028
Round 84/100: val_acc: 0.602, lr: 3.170826663407878e-06, reg:
```

5.959928834128086e-05
Round 85/100: val_acc: 0.54, lr: 0.0006049659779912351, reg: 0.1839559260174279
Round 86/100: val_acc: 0.6, lr: 2.9877446548609863e-06, reg: 0.01260775594092302
Round 87/100: val_acc: 0.556, lr: 1.4903816251270174e-06, reg:
0.0262420980535794
Round 88/100: val_acc: 0.618, lr: 4.178295664535119e-06, reg: 0.9772784585297325
Round 89/100: val_acc: 0.52, lr: 1.1390949185993834e-06, reg:
0.00010791863317515481
Round 90/100: val_acc: 0.644, lr: 1.3222737672493466e-05, reg:
0.0008045783949257835
Round 91/100: val_acc: 0.6779999999999999, lr: 0.0005668119332369673, reg:
0.35235307010820405
Round 92/100: val_acc: 0.784, lr: 0.00017248987513456007, reg:
0.007621170403484562
Round 93/100: val_acc: 0.5660000000000001, lr: 1.5987022788720815e-06, reg:
0.30555837012285714
Round 94/100: val_acc: 0.5760000000000001, lr: 1.9257895153093322e-06, reg:
1.1912585939599736e-05
Round 95/100: val_acc: 0.744, lr: 3.3411371813868935e-05, reg:
7.780478570926214e-05
Round 96/100: val_acc: 0.778, lr: 6.806147433305496e-05, reg:
0.002057167026756189
Round 97/100: val_acc: 0.624, lr: 4.327388844970964e-06, reg:
0.00027304204564616803
Round 98/100: val_acc: 0.606, lr: 3.684530896633849e-06, reg:
1.3766498372010842e-05
Round 99/100: val_acc: 0.758, lr: 4.023819055214163e-05, reg:
0.47061223565645877
Round 100/100: val_acc: 0.612, lr: 3.856585966977233e-06, reg:
9.273694096085599e-05

Round 82 has the highest validation accuracy. Let's try the parameters: lr=0.0001, reg=1e-2

```
[42]: model = four_layers_model(neurons_1, neurons_2)

      is_data_prepossing = True # set True if use data normalization
      h = 1e-4
      regularization = 2e-5
      epochs = 200

      # Read dataset
      with h5py.File('./dataset/train_dataset.h5', 'r') as file:
          data = np.array(file.get('data'))
          label = np.array(file.get('label'))
          X_train = data
          y_train = label

      with h5py.File('./dataset/test_dataset.h5', 'r') as file:
```

```python
    data = np.array(file.get('data'))
    label = np.array(file.get('label'))
    X_test = data
    y_test = label

# Data Preprossing
if is_data_prepossing:
    X_train = data_preprossing(X_train)

X_train, X_val, y_train, y_val = train_test_split(X_train, y_train, test_size=0.
 ↪25, random_state=66)

y_train = np.expand_dims(y_train, -1)
y_val = np.expand_dims(y_val, -1)
y_test = np.expand_dims(y_test, -1)

accuracies, losses, val_accuracies, par6 = model.
 ↪train(X_train,y_train,X_val,y_val,epochs,h,batch_size,regularization,best_model=True)
```

```
Finishing epoch 1/200: train loss: 171.5503349666422, train accuracy: 0.64, val
accuracy: 0.636, lr: 0.0001
Finishing epoch 2/200: train loss: 162.43645199549937, train accuracy:
0.6733333333333333, val accuracy: 0.646, lr: 0.0001
Finishing epoch 3/200: train loss: 153.3265041419426, train accuracy:
0.6926666666666667, val accuracy: 0.648, lr: 0.0001
Finishing epoch 4/200: train loss: 146.4081676097307, train accuracy:
0.7453333333333334, val accuracy: 0.732, lr: 0.0001
Finishing epoch 5/200: train loss: 131.2127421580758, train accuracy:
0.7586666666666666, val accuracy: 0.738, lr: 0.0001
Finishing epoch 6/200: train loss: 122.97786220764019, train accuracy:
0.7846666666666666, val accuracy: 0.776, lr: 0.0001
Finishing epoch 7/200: train loss: 120.87585245567026, train accuracy:
0.8093333333333333, val accuracy: 0.792, lr: 0.0001
Finishing epoch 8/200: train loss: 112.27422954176288, train accuracy:
0.7906666666666666, val accuracy: 0.732, lr: 0.0001
Finishing epoch 9/200: train loss: 109.74559476303966, train accuracy:
0.8180000000000001, val accuracy: 0.798, lr: 0.0001
Finishing epoch 10/200: train loss: 105.58075398674114, train accuracy:
0.8146666666666667, val accuracy: 0.762, lr: 0.0001
Finishing epoch 11/200: train loss: 105.58075398674114, train accuracy:
0.8146666666666667, val accuracy: 0.762, lr: 0.0001
Finishing epoch 12/200: train loss: 105.58075398674114, train accuracy:
0.8146666666666667, val accuracy: 0.762, lr: 0.0001
Finishing epoch 13/200: train loss: 104.74234528978286, train accuracy: 0.796,
val accuracy: 0.756, lr: 0.0001
Finishing epoch 14/200: train loss: 102.20144588894155, train accuracy:
0.8346666666666667, val accuracy: 0.794, lr: 0.0001
```

```
Finishing epoch 15/200: train loss: 101.98424015405682, train accuracy: 0.84,
val accuracy: 0.81, lr: 0.0001
Finishing epoch 16/200: train loss: 101.98424015405682, train accuracy: 0.84,
val accuracy: 0.81, lr: 0.0001
Finishing epoch 17/200: train loss: 101.98424015405682, train accuracy: 0.84,
val accuracy: 0.81, lr: 0.0001
Finishing epoch 18/200: train loss: 101.83550778133494, train accuracy:
0.8253333333333334, val accuracy: 0.786, lr: 0.0001
Finishing epoch 19/200: train loss: 101.83550778133494, train accuracy:
0.8253333333333334, val accuracy: 0.786, lr: 0.0001
Finishing epoch 20/200: train loss: 101.83550778133494, train accuracy:
0.8253333333333334, val accuracy: 0.786, lr: 0.0001
Finishing epoch 21/200: train loss: 101.83550778133494, train accuracy:
0.8253333333333334, val accuracy: 0.786, lr: 0.0001
Finishing epoch 22/200: train loss: 101.83550778133494, train accuracy:
0.8253333333333334, val accuracy: 0.786, lr: 0.0001
Finishing epoch 23/200: train loss: 101.83550778133494, train accuracy:
0.8253333333333334, val accuracy: 0.786, lr: 0.0001
Finishing epoch 24/200: train loss: 101.83550778133494, train accuracy:
0.8253333333333334, val accuracy: 0.786, lr: 0.0001
Finishing epoch 25/200: train loss: 101.83550778133494, train accuracy:
0.8253333333333334, val accuracy: 0.786, lr: 0.0001
Finishing epoch 26/200: train loss: 101.83550778133494, train accuracy:
0.8253333333333334, val accuracy: 0.786, lr: 0.0001
Finishing epoch 27/200: train loss: 101.83550778133494, train accuracy:
0.8253333333333334, val accuracy: 0.786, lr: 0.0001
Finishing epoch 28/200: train loss: 101.83550778133494, train accuracy:
0.8253333333333334, val accuracy: 0.786, lr: 0.0001
Finishing epoch 29/200: train loss: 101.83550778133494, train accuracy:
0.8253333333333334, val accuracy: 0.786, lr: 0.0001
Finishing epoch 30/200: train loss: 100.26822411840767, train accuracy:
0.8240000000000001, val accuracy: 0.79, lr: 0.0001
Finishing epoch 31/200: train loss: 99.89631462500301, train accuracy:
0.8146666666666667, val accuracy: 0.778, lr: 0.0001
Finishing epoch 32/200: train loss: 99.89631462500301, train accuracy:
0.8146666666666667, val accuracy: 0.778, lr: 0.0001
Finishing epoch 33/200: train loss: 99.89631462500301, train accuracy:
0.8146666666666667, val accuracy: 0.778, lr: 0.0001
Finishing epoch 34/200: train loss: 99.89631462500301, train accuracy:
0.8146666666666667, val accuracy: 0.778, lr: 0.0001
Finishing epoch 35/200: train loss: 99.89631462500301, train accuracy:
0.8146666666666667, val accuracy: 0.778, lr: 0.0001
Finishing epoch 36/200: train loss: 99.89631462500301, train accuracy:
0.8146666666666667, val accuracy: 0.778, lr: 0.0001
Finishing epoch 37/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 38/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
```

```
Finishing epoch 39/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 40/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 41/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 42/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 43/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 44/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 45/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 46/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 47/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 48/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 49/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 50/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 51/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 52/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 53/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 54/200: train loss: 99.60191517665311, train accuracy:
0.8306666666666667, val accuracy: 0.8, lr: 0.0001
Finishing epoch 55/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 56/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 57/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 58/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 59/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 60/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 61/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 62/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
```

```
Finishing epoch 63/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 64/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 65/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 66/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 67/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 68/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 69/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 70/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 71/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 72/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 73/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 74/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 75/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 76/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 77/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 78/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 79/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 80/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 81/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 82/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 83/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 84/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 85/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 86/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
```

```
Finishing epoch 87/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 88/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 89/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 90/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 91/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 92/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 93/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 94/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 95/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 96/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 97/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 98/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 99/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 100/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 101/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 102/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 103/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 104/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 105/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 106/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 107/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 108/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 109/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 110/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
```

```
Finishing epoch 111/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 112/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 113/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 114/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 115/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 116/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 117/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 118/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 119/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 120/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 121/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 122/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 123/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 124/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 125/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 126/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 127/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 128/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 129/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 130/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 131/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 132/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 133/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 134/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
```

Finishing epoch 135/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 136/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 137/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 138/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 139/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 140/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 141/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 142/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 143/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 144/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 145/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 146/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 147/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 148/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 149/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 150/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 151/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 152/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 153/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 154/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 155/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 156/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 157/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 158/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001

```
Finishing epoch 159/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 160/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 161/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 162/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 163/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 164/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 165/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 166/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 167/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 168/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 169/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 170/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 171/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 172/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 173/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 174/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 175/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 176/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 177/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 178/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 179/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 180/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 181/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 182/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
```

```
Finishing epoch 183/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 184/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 185/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 186/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 187/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 188/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 189/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 190/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 191/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 192/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 193/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 194/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 195/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 196/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 197/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 198/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 199/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Finishing epoch 200/200: train loss: 99.4300600191751, train accuracy:
0.8306666666666667, val accuracy: 0.796, lr: 0.0001
Training ends.
```

[43]:
```python
#Plot the loss and accuracy
fig = plt.figure(figsize=(15,5))

ax = fig.add_subplot(121)
plt.semilogy(losses)
plt.title('Loss')
plt.xlabel('Epochs')
plt.ylabel('Sum of Squares Error')
```
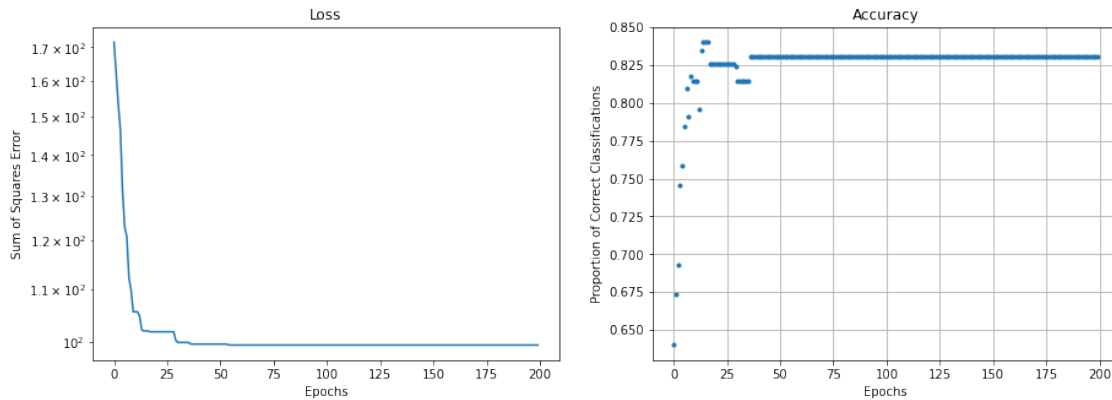
```
ax = fig.add_subplot(122)
plt.plot(accuracies,'.')
plt.title('Accuracy')
plt.xlabel('Epochs')
plt.ylabel('Proportion of Correct Classifications')
#plt.ylim([0,1])
plt.grid()

plt.show()
```

As shown above, even though the validation accuracy of the parameter set (lr:1e-4, reg:1e-2) is worse than the initial parameter set (lr:1e-5, reg:0) used in step 1 part, which is 83% vesus 83.8%, we can see the loss decreases much faster than the initial choice. The reason is that we have a larger learning rate. This means with smaller training epochs, we can quickly get the trained network with good performance. We further use test dataset to observe the accuracy can reach 80%.

```
[44]: predict = model.prediction(X_test)
      error = float(np.count_nonzero(predict-y_test)) / float(len(X_test))
      accuracy = 1.0 - error
      print('The test accuracy: {}'.format(accuracy))
```

The test accuracy: 0.8