```
# Imports math library
import numpy as np
# Imports plotting library
import matplotlib.pyplot as plt
```

```
# Define the Rectified Linear Unit (ReLU) function - Points: 10

def ReLU(preactivation):
    # TODO write code to implement the ReLU and compute the activation at the
    # hidden unit from the preactivation
    # Replace the given code with your code
    # To understand the input format, look at the code cell below this one
    return np.maximum(preactivation, 0)
```

```
# Make an array of inputs
z = np.arange(-5,5,0.1)
RelU_z = ReLU(z)

# Plot the ReLU function
fig, ax = plt.subplots()
ax.plot(z,RelU_z,'r-')
ax.set_xlim([-5,5]);ax.set_ylim([-5,5])
ax.set_xlabel('z'); ax.set_ylabel('ReLU[z]')
ax.set_aspect('equal')
plt.show()
```

```
# Define a shallow neural network with, one input, one output, and three hidden
units - Points: 12

def shallow_1_1_3(x, activation_fn, phi_0,phi_1,phi_2,phi_3, theta_10, theta_11,
theta_20, theta_21, theta_30, theta_31):
    # TODO Replace the code below to compute the three initial lines
    # from the theta parameters (i.e. implement equations given at bottom of figure
a-c below this Code Cell). These are the preactivations
    pre_1 = theta_10 + theta_11*x
    pre_2 = theta_20 + theta_21*x
    pre_3 = theta_30 + theta_31*x

# Pass these through the ReLU function to compute the activations as in figure
d-f given below. No need to replace anything here
    act_1 = activation_fn(pre_1)
```

```
act_2 = activation_fn(pre_2)
act_3 = activation_fn(pre_3)

# TODO Replace the code below to weight the activations using phi1, phi2 and
phi3 to create the equivalent of figure g-i
    w_act_1 = phi_1*act_1
    w_act_2 = phi_2*act_2
    w_act_3 = phi_3*act_3

# TODO Replace the code below to combining the weighted activations and add
phi_0 to create the output as in figure j
    y = phi_0 + w_act_1 + w_act_2 + w_act_3

# Return everything we have calculated
    return y, pre_1, pre_2, pre_3, act_1, act_2, act_3, w_act_1, w_act_2, w_act_3
```

```
# Plot the shallow neural network. We'll assume input in is range [0,1] and
output [-1,1]
# If the plot all flag is set to true, then we'll plot all the intermediate
def plot_neural(x, y, pre_1, pre_2, pre_3, act_1, act_2, act_3, w_act_1, w_act_2,
w_act_3, plot_all=False, x_data=None, y_data=None):
  # Plot intermediate plots if flag set
 if plot_all:
   fig, ax = plt.subplots(3,3)
    fig.set size inches(8.5, 8.5)
    fig.tight layout(pad=3.0)
    ax[0,0].plot(x,pre 1,'r-'); ax[0,0].set ylabel('Preactivation')
    ax[0,1].plot(x,pre_2,'b-'); ax[0,1].set_ylabel('Preactivation')
    ax[0,2].plot(x,pre_3,'g-'); ax[0,2].set_ylabel('Preactivation')
    ax[1,0].plot(x,act_1,'r-'); ax[1,0].set_ylabel('Activation')
    ax[1,1].plot(x,act_2,'b-'); ax[1,1].set_ylabel('Activation')
    ax[1,2].plot(x,act_3,'g-'); ax[1,2].set_ylabel('Activation')
    ax[2,0].plot(x,w_act_1,'r-'); ax[2,0].set_ylabel('Weighted Act')
    ax[2,1].plot(x,w_act_2,'b-'); ax[2,1].set_ylabel('Weighted Act')
    ax[2,2].plot(x,w_act_3,'g-'); ax[2,2].set_ylabel('Weighted Act')
    for plot y in range(3):
      for plot x in range(3):
```

```
ax[plot_y,plot_x].set_xlim([0,1]);ax[plot_x,plot_y].set_ylim([-1,1])
    ax[plot_y,plot_x].set_aspect(0.5)
    ax[2,plot_y].set_xlabel('Input, $x$');
    plt.show()

fig, ax = plt.subplots()
ax.plot(x,y)
ax.set_xlabel('Input, $x$'); ax.set_ylabel('Output, $y$')
ax.set_xlim([0,1]);ax.set_ylim([-1,1])
ax.set_aspect(0.5)
if x_data is not None:
    ax.plot(x_data, y_data, 'mo')
    for i in range(len(x_data)):
        ax.plot(x_data[i], y_data[i],)
plt.show()
```

```
# TODO - Points: 20 (2.5 for each part. Answer this question in the cell space
given below this cell)
# 1. Predict what effect changing phi_0 will have on the network.
# 2. Predict what effect multiplying phi_1, phi_2, phi_3 by 0.5 would
have. Check if you are correct
```

```
# 3. Predict what effect multiplying phi 1 by -1 will have. Check if you are
correct.
# 4. Predict what effect setting theta 20 to -1.2 will have. Check if you are
correct.
# 5. Change the parameters so that there are only two "joints" (including outside
the range of the plot)
# There are actually three ways to do this. See if you can figure them all out
# 6. With the original parameters, the second line segment is flat (i.e. has
slope zero)
# How could you change theta 10 so that all of the segments have non-zero slopes
# 7. What do you predict would happen if you multiply theta 20 and theta21 by
0.5, and phi 2 by 2.0?
# Check if you are correct.
# 8. What do you predict would happen if you multiply theta 20 and theta21 by -
0.5, and phi 2 by -2.0?
# Check if you are correct.
theta 10 = 0.3; theta 11 = -1.0
theta 20 = -1.0; theta 21 = 2.0
theta_30 = -0.5 ; theta_31 = 0.65
phi 0 = -0.3; phi 1 = 2.0; phi 2 = -1.0; phi 3 = 7.0
# Define a range of input values
x = np.arange(0,1,0.01)
# We run the neural network for each of these input values
y, pre_1, pre_2, pre_3, act_1, act_2, act_3, w_act_1, w_act_2, w_act_3 = \
    shallow_1_1_3(x, ReLU, phi_0,phi_1,phi_2,phi_3, theta_10, theta_11, theta_20,
theta 21, theta 30, theta 31)
# And then plot it
plot_neural(x, y, pre_1, pre_2, pre_3, act_1, act_2, act_3, w_act_1, w_act_2,
w act 3, plot all=True)
```

```
# Least squares function - Points: 10
def least_squares_loss(y_train, y_predict):
    # TODO Replace the line below to compute the sum of squared
```

```
# differences between the real values of y and the predicted values from the
model f[x_i,phi] (figure given above)
  # you will need to use the function np.sum
  # Loop is not allowed
  loss = np.sum((y_train - y_predict)**2)
  return loss
```

```
# Now lets define some parameters, run the neural network, and compute the loss
theta_10 = 0.0 ; theta_11 = 0.5
theta 20 = -1.0; theta 21 = 2.0
theta 30 = -0.5; theta 31 = 0.65
phi 0 = -0.1; phi 1 = 2.0; phi 2 = -1.0; phi 3 = 7.0
# Define a range of input values
x = np.arange(0,1,0.01)
x train =
np.array([0.09291784,0.46809093,0.93089486,0.67612654,0.73441752,0.86847339,\
                   0.49873225,0.51083168,0.18343972,0.99380898,0.27840809,0.38028
817,\
                   0.12055708,0.56715537,0.92005746,0.77072270,0.85278176,0.05315
950,\
                   0.87168699,0.58858043])
y_train = np.array([-
0.15934537,0.18195445,0.451270150,0.13921448,0.09366691,0.30567674,\
                    0.372291170,0.40716968,-
0.08131792,0.41187806,0.36943738,0.3994327,\
                    0.019062570,0.35820410,0.452564960,-0.0183121,0.02957665,-
0.24354444, \
                    0.148038840,0.26824970])
# We run the neural network for each of these input values
y, pre 1, pre 2, pre 3, act 1, act 2, act 3, w act 1, w act 2, w act 3 = \setminus
    shallow_1_1_3(x, ReLU, phi_0,phi_1,phi_2,phi_3, theta_10, theta_11, theta_20,
theta 21, theta 30, theta 31)
# And then plot it
plot_neural(x, y, pre_1, pre_2, pre_3, act_1, act_2, act_3, w_act_1, w_act_2,
w_act_3, plot_all=True, x_data = x_train, y_data = y_train)
# Run the neural network on the training data
y_predict, *_ = shallow_1_1_3(x_train, ReLU, phi_0,phi_1,phi_2,phi_3, theta_10,
theta 11, theta 20, theta 21, theta 30, theta 31)
```

```
# Compute the least squares loss and print it out
loss = least_squares_loss(y_train,y_predict)
print('Your Loss = %3.3f, True value = 9.385'%(loss))

# TODO. Manipulate the parameters (by hand!) to make the function - Points: 8
# fit the data better and try to reduce the loss to as small a number
# as possible. The best that I could do was 0.181
# Tip... start by manipulating phi_0.
# It's not that easy, so don't spend too much time on this!
```

```
import plotly.graph objects as go
# Code to draw 2D function -- read it so you know what is going on, but you don't
have to change it
def draw 2D function(ax, x1 mesh, x2 mesh, y):
    pos = ax.contourf(x1_mesh, x2_mesh, y, levels=256 ,cmap = 'hot', vmin=-
10, vmax=10.0
    ax.set xlabel('x1');ax.set ylabel('x2')
    levels = np.arange(-10,10,1.0)
    ax.contour(x1 mesh, x2 mesh, y, levels, cmap='winter')
# Plot the shallow neural network. We'll assume input in is range [0,10],[0,10]
and output [-10,10]
def plot_neural_2_inputs(x1,x2, y, pre_1, pre_2, pre_3, act_1, act_2, act_3,
w_act_1, w_act_2, w_act_3):
  fig, ax = plt.subplots(3,3)
  fig.set size inches(8.5, 8.5)
  fig.tight_layout(pad=3.0)
  draw 2D function(ax[0,0], x1,x2,pre 1); ax[0,0].set title('Preactivation')
  draw_2D_function(ax[0,1], x1,x2,pre_2); ax[0,1].set_title('Preactivation')
  draw_2D_function(ax[0,2], x1,x2,pre_3); ax[0,2].set_title('Preactivation')
  draw_2D_function(ax[1,0], x1,x2,act_1); ax[1,0].set_title('Activation')
  draw_2D_function(ax[1,1], x1,x2,act_2); ax[1,1].set_title('Activation')
  draw_2D_function(ax[1,2], x1,x2,act_3); ax[1,2].set_title('Activation')
  draw_2D_function(ax[2,0], x1,x2,w_act_1); ax[2,0].set_title('Weighted Act')
  draw_2D_function(ax[2,1], x1,x2,w_act_2); ax[2,1].set_title('Weighted Act')
  draw_2D_function(ax[2,2], x1,x2,w_act_3); ax[2,2].set_title('Weighted Act')
  plt.show()
  fig, ax = plt.subplots()
```

```
draw_2D_function(ax,x1,x2,y)
ax.set_title('Network output, $y$')
ax.set_aspect(1.0)
plt.show()
```

```
# Define a shallow neural network with, two input, one output, and three hidden
units - Points: 12
def shallow 2 1 3(x1,x2, activation fn, phi 0,phi 1,phi 2,phi 3, theta 10,
theta 11,\
                  theta 12, theta 20, theta 21, theta 22, theta 30, theta 31,
theta_32):
 # TODO Replace the lines below to compute the three initial linear functions
 # (figure a-c) from the theta parameters. These are the preactivations
 pre 1 = theta 10 + theta 11*x1 + theta 12*x2
  pre 2 = theta 20 + theta 21*x1 + theta 22*x2
  pre_3 = theta_30 + theta_31*x1 + theta_32*x2
 # Pass these through the ReLU function to compute the activations as in figure
d-f
  act 1 = activation fn(pre 1)
 act_2 = activation_fn(pre_2)
  act_3 = activation_fn(pre_3)
 # TODO Replace the code below to weight the activations using phi1, phi2 and
phi3
 w act 1 = phi 1*act 1
 w_act_2 = phi_2*act_2
 w_act_3 = phi_3*act_3
 # TODO Replace the code below to combing the weighted activations and add
 # phi 0 to create the output as in j
 y = phi_0 + w_act_1 + w_act_2 + w_act_3
 # Return everything we have calculated
 return y, pre_1, pre_2, pre_3, act_1, act_2, act_3, w_act_1, w_act_2, w_act_3
```

```
# Now lets define some parameters and run the neural network
theta_10 = -4.0; theta_11 = 0.9; theta_12 = 0.0
```

```
def plot_neural_2_inputs_2_outputs(x1,x2, y1, y2, pre_1, pre_2, pre_3, act_1,
act_2, act_3, w_act_11, w_act_12, w_act_13, w_act_21, w_act_22, w_act_23):
  fig, ax = plt.subplots(4,3)
  fig.set size inches(8.5, 8.5)
 fig.tight_layout(pad=3.0)
  draw_2D_function(ax[0,0], x1,x2,pre_1); ax[0,0].set_title('Preactivation')
  draw 2D function(ax[0,1], x1,x2,pre 2); ax[0,1].set title('Preactivation')
  draw_2D_function(ax[0,2], x1,x2,pre_3); ax[0,2].set_title('Preactivation')
  draw_2D_function(ax[1,0], x1,x2,act_1); ax[1,0].set_title('Activation')
  draw_2D_function(ax[1,1], x1,x2,act_2); ax[1,1].set_title('Activation')
  draw_2D_function(ax[1,2], x1,x2,act_3); ax[1,2].set_title('Activation')
  draw 2D function(ax[2,0], x1,x2,w act 11); ax[2,0].set title('Weighted Act 1')
  draw_2D_function(ax[2,1], x1,x2,w_act_12); ax[2,1].set_title('Weighted Act 1')
  draw_2D_function(ax[2,2], x1,x2,w_act_13); ax[2,2].set_title('Weighted Act 1')
  draw_2D_function(ax[3,0], x1,x2,w_act_21); ax[3,0].set_title('Weighted Act 2')
  draw_2D_function(ax[3,1], x1,x2,w_act_22); ax[3,1].set_title('Weighted Act 2')
  draw_2D_function(ax[3,2], x1,x2,w_act_23); ax[3,2].set_title('Weighted Act 2')
  plt.show()
  fig, ax = plt.subplots()
  draw 2D function(ax,x1,x2,y1)
  ax.set_title('Network output, $y_1$')
  ax.set aspect(1.0)
```

```
plt.show()

fig, ax = plt.subplots()
draw_2D_function(ax,x1,x2,y2)
ax.set_title('Network output, $y_2$')
ax.set_aspect(1.0)
plt.show()
```

```
# Define a shallow neural network with, two inputs, two outputs, and three hidden
units - Points: 14
def shallow_2_2_3(x1,x2, activation_fn, phi_10,phi_11,phi_12,phi_13,
phi_20,phi_21,phi_22,phi_23, theta_10, theta_11,\
                  theta_12, theta_20, theta_21, theta_22, theta_30, theta_31,
theta 32):
 # TODO -- write this function -- replace the dummy code below
  pre 1 = theta 10 + theta 11*x1 + theta 12*x2
  pre_2 = theta_20 + theta_21*x1 + theta_22*x2
 pre 3 = theta 30 + theta 31*x1 + theta 32*x2
  act_1 = activation_fn(pre_1)
  act_2 = activation_fn(pre_2)
  act 3 = activation fn(pre 3)
 w_act_11 = phi_11*act_1
 w act 12 = phi 12*act 2
 w_act_13 = phi_13*act_3
 w_act_21 = phi_21*act_1
 w act 22 = phi 22*act 2
 w_act_23 = phi_23*act_3
 y1 = phi 10 + w act 11 + w act 12 + w act 13
 y2 = phi_20 + w_act_21 + w_act_22 + w_act_23
 # Return everything we have calculated
 return y1,y2, pre_1, pre_2, pre_3, act_1, act_2, act_3, w_act_11, w_act_12,
w_act_13, w_act_21, w_act_22, w_act_23
```

```
# Now lets define some parameters and run the neural network
theta_10 = -4.0; theta_11 = 0.9; theta_12 = 0.0
theta 20 = 5.0; theta 21 = -0.9; theta 22 = -0.5
```

```
# Plot the shallow neural network. We'll assume input is in range [0,1] and
output [-1,1]
# If the plot all flag is set to true, then we'll plot all the intermediate
stages as in activity 2 (also given in image above).
def plot_neural(x, y, pre_1, pre_2, pre_3, act_1, act_2, act_3, w_act_1, w_act_2,
w_act_3, plot_all=False, x_data=None, y_data=None):
 # Plot intermediate plots if flag set
 if plot all:
   fig, ax = plt.subplots(3,3)
    fig.set size inches(8.5, 8.5)
    fig.tight layout(pad=3.0)
    ax[0,0].plot(x,pre_1,'r-'); ax[0,0].set_ylabel('Preactivation')
    ax[0,1].plot(x,pre_2,'b-'); ax[0,1].set_ylabel('Preactivation')
    ax[0,2].plot(x,pre_3,'g-'); ax[0,2].set_ylabel('Preactivation')
    ax[1,0].plot(x,act_1,'r-'); ax[1,0].set_ylabel('Activation')
    ax[1,1].plot(x,act_2,'b-'); ax[1,1].set_ylabel('Activation')
    ax[1,2].plot(x,act_3,'g-'); ax[1,2].set_ylabel('Activation')
    ax[2,0].plot(x,w_act_1,'r-'); ax[2,0].set_ylabel('Weighted Act')
    ax[2,1].plot(x,w act 2,'b-'); ax[2,1].set ylabel('Weighted Act')
    ax[2,2].plot(x,w_act_3,'g-'); ax[2,2].set_ylabel('Weighted Act')
    for plot y in range(3):
```

```
for plot_x in range(3):
    ax[plot_y,plot_x].set_xlim([0,1]);ax[plot_x,plot_y].set_ylim([-1,1])
    ax[plot_y,plot_x].set_aspect(0.5)
    ax[2,plot_y].set_xlabel('Input, $x$');
    plt.show()

fig, ax = plt.subplots()
    ax.plot(x,y)
    ax.set_xlabel('Input, $x$'); ax.set_ylabel('Output, $y$')
    ax.set_xlim([0,1]);ax.set_ylim([-1,1])
    ax.set_aspect(0.5)
    if x_data is not None:
        ax.plot(x_data, y_data, 'mo')
        for i in range(len(x_data)):
            ax.plot(x_data[i], y_data[i],)
    plt.show()
```

```
# Define the sigmoid function - Points: 10
def sigmoid(preactivation):
    # TODO write code to implement the sigmoid function and compute the activation
at the
    # hidden unit from the preactivation. Use the np.exp() function.
    return 1/(1+np.exp(-10*preactivation))
```

```
# Make an array of inputs
z = np.arange(-1,1,0.01)
sig_z = sigmoid(z)

# Plot the sigmoid function
fig, ax = plt.subplots()
ax.plot(z,sig_z,'r-')
ax.set_xlim([-1,1]);ax.set_ylim([0,1])
ax.set_xlabel('z'); ax.set_ylabel('sig[z]')
plt.show()
```

```
# Define the heaviside function - Points: 10
def heaviside(preactivation):
    # TODO write code to implement the heaviside function and compute the
activation at the
    # hidden unit from the preactivation.
    return np.where(preactivation>=0,1,0)
```

```
# Make an array of inputs
z = np.arange(-1,1,0.01)
heav_z = heaviside(z)
```

```
# Plot the heaviside function
fig, ax = plt.subplots()
ax.plot(z,heav_z,'r-')
ax.set_xlim([-1,1]);ax.set_ylim([-2,2])
ax.set_xlabel('z'); ax.set_ylabel('heaviside[z]')
plt.show()
```

```
# Define the linear activation function
def lin(preactivation):
    a = 0.5
    b = -.4
    # Compute linear function
    activation = a+b * preactivation
    # Return
    return activation
```

```
# TODO - Points: 2
# 1. The linear activation function above just returns the input: (0+1*z) = z
```

```
# Before running the code Make a prediction about what the ten panels of the
drawing will look like
# Now run the code below to see if you were right. What family of functions can
this represent?
# Answer: Since the z remains unchanged after activation, the activation function
essentially has no effect on the model
# 2. What happens if you change the parameters (a,b) to different values?
# Try a=0.5, b=-0.4 Don't forget to run the cell again to update the function
# Answer: When the values of a and b are change the preactivation and activation
lines are essentially completely different lines
theta_10 = 0.3; theta_11 = -1.0
theta_20 = -1.0 ; theta_21 = 2.0
theta_30 = -0.5 ; theta_31 = 0.65
phi_0 = 0.3; phi_1 = 0.5; phi_2 = -1.0; phi_3 = 0.9
# Define a range of input values
x = np.arange(0,1,0.01)
# We run the neural network for each of these input values
y, pre_1, pre_2, pre_3, act_1, act_2, act_3, w_act_1, w_act_2, w_act_3 = \
    shallow_1_1_3(x, lin, phi_0,phi_1,phi_2,phi_3, theta_10, theta_11, theta_20,
theta_21, theta_30, theta_31)
# And then plot it
plot_neural(x, y, pre_1, pre_2, pre_3, act_1, act_2, act_3, w_act_1, w_act_2,
w_act_3, plot all=True)
```