Last Name	First Name	ID

Prob #	1	2	3	4	5
Points	10	20	20	25	25

Time: 80 Minutes

NOTES:

- a. Credit is only given to the correct numerical values.
- b. All numerical values must be calculated with three digits of accuracy after the decimal point.
- 1. Consider a convolutional neural network. Note: **Do NOT consider Biases**.

Input layer:

Input to this CNN are color images of size 64x52x3 with the batch size = 128 Note: Input image has different horizontal and vertical resolution.

Next layer is Conv2D layer:

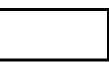
number of filters: 30, filter size: 11x11; stride: 3x3; padding: 2x2 What is the shape of the weight matrix for this layer? What is the shape of the output (tensor) of this layer?

Next layer is Flatten layer:

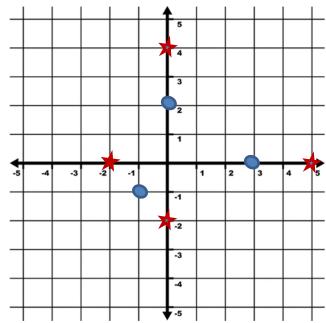
What is the shape of the output (tensor) for this layer?

Next layer is Dense layer:

number of nodes: 100 What is the shape of the weight matrix (tensor) for this layer? What is the shape of the output (tensor) for this layer?



2. Design a two-layer Perceptron neural network which will correctly classify the two classes (circles and stars) as shown below. Assume the activation (transfer) function for all the nodes are hard-limit with the output of **0** (star) and **1**(Circle).



Show the weight matrices and biases for both layers. Biases should be included in the weight matrices in the first column.

Weight matrix for the first layer =

Weight matrix for the second layer =

3. Consider the expression:

$$f(x,y) = \frac{xy}{\min(xy,y^2) - 6}$$

given the inputs: x = 5, y = 4

Draw the computational graph for this expression.

• You must show all the forward values and backward **partial derivatives for each node** in the graph and compute numerical values of the partial derivatives with respect to x and y

$\delta f(x,y)$		
$\delta x \equiv$		
$\delta f(x,y)$		
$\frac{\delta y}{\delta y}$		

4. Given the actual outputs and the desired output indexes of a neural network,

calculate and return the confusion matrix. y_hat is a matrix presenting the actual output. Each row of this matrix is the actual output of the neural network for an input sample. The number of rows in this matrix is equal to the number of samples in the input batch. Y is the desired output. Each number in in the y is the index of the desired class for the corresponding input sample.

Hint: This is similar to the function in Assignment 03.

import numpy as np

```
def calculate_confusion_matrix(y_hat, y):
""" Given the actual outputs and the desired output index of a neural
network, this function calculates the confusion matrix.
:param y_hat: Array of actual outputs [num_of samples,num_of classes]
:param y: Array of desired (target) outputs [num_of samples]. This
array includes the indexes of the desired (true) class.
:return confusion_matrix[number_of_classes,number_of_classes].
Confusion matrix should be shown as the number of times that an input
of class n is classified as class m."""
```

• Complete the code for the following method/function (this is the same function in assignment 04).

Assume that the object that you are working with, i.e. self, is the same as the object that you used in your assignment 04. For example, if you used a Keras model in your assignment 04 then assume your object (self) is a Keras model.

Hint: you may use layers[] and get_layer() and get_wights() property and methods

Import numpy as np				
<pre>def get_biases(self,biases,layer_number=None,layer_name=""):</pre>				
""" This function gets the biases for the layer specified by layer_number or by				
layer_name.				
:param layer_number: Layer number starting from layer 0.				
:param layer_name: Layer name (if both layer_number and layer_name are				
specified, layer number takes precedence).				
:return: biases for the given layer (If the given layer does not have bias then				
None should be returned) """				

5. Complete the following code using Tensorflow. The code should initialize the weights and biases and compute the output, loss, and gradients, and perform a single weight update of a neural network. Assume the dimension of each input sample is 400. There are 100 nodes with sigmoid activation in the first layer, and 10 output nodes with linear activation. Loss is MSE. Assume you have a batch of inputs called X that has dimensions [20, 400], and a batch of targets called y that has dimensions [20, 10]

<pre>import tensorflow as tf</pre>	
<pre>import numpy as np</pre>	
# Create random weights and biases	
W_1 = tf.Variable(np.random.randn(), trainable=True)
<pre>b_1 = tf.Variable(np.random.randn(</pre>), trainable=True)
W_2 = tf.Variable(np.random.randn(), trainable=True)
<pre>b_2 = tf.Variable(np.random.randn(</pre>), trainable=True)
<pre>with tf.GradientTape(persistent=True) as tape:</pre>	
# Calculate output	

output=

Calculate loss

Loss=

Calculate gradients

dW_1, db_1 =

dW_2, db_2 =

Calculate new values of weights and biases