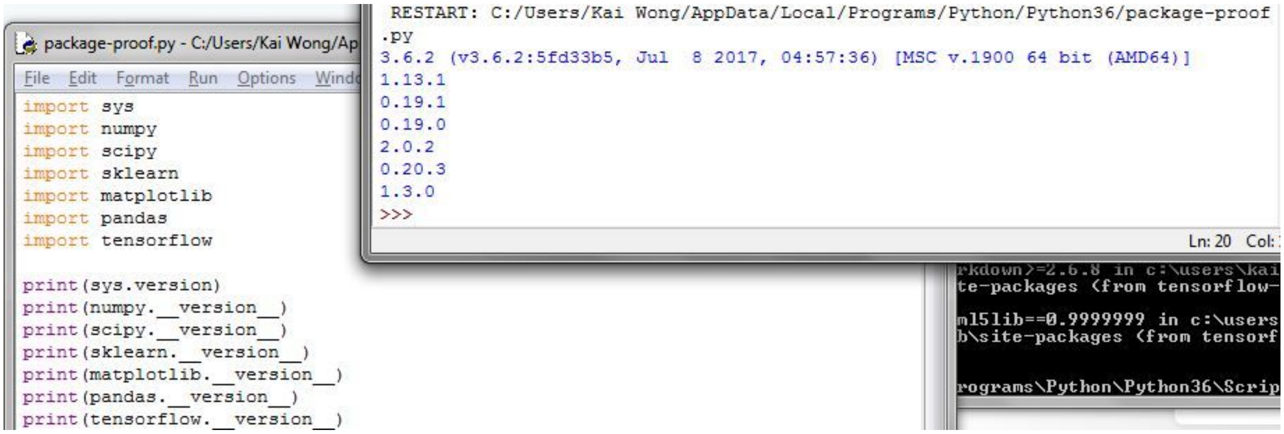
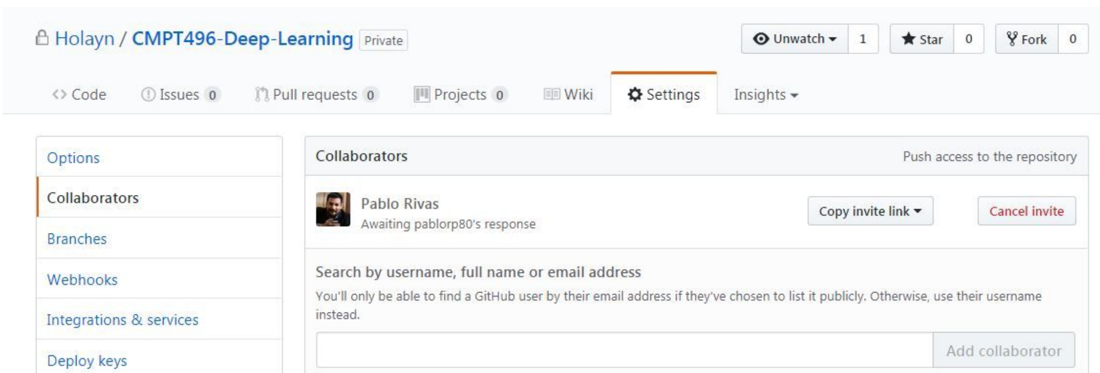
Kai Wong
CMPT469
9/11/2017

Homework 0

2.1 Proof:

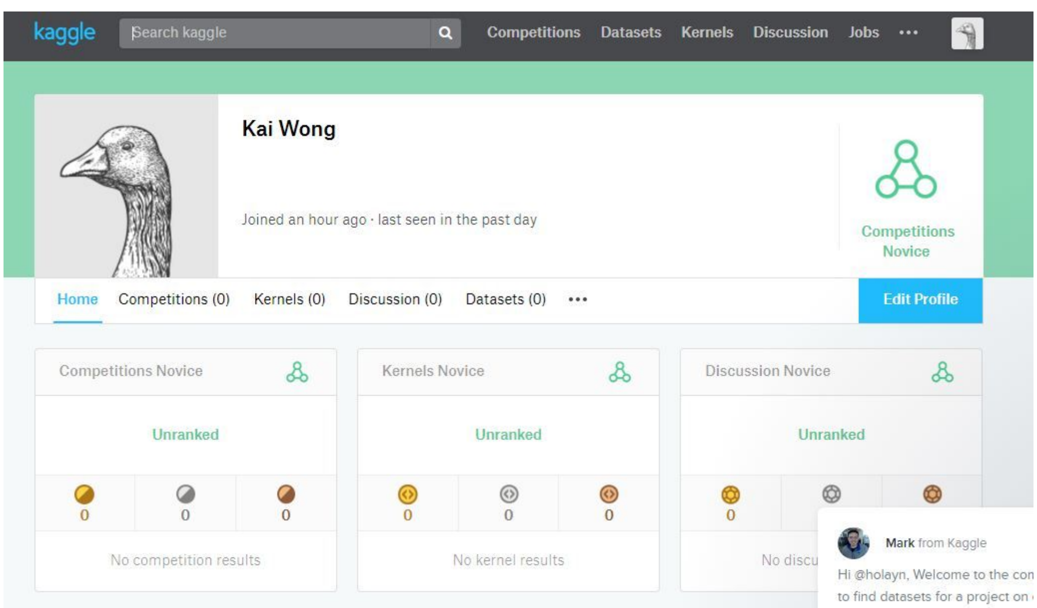


2.2 Proof:



Link to repo: https://github.com/Holayn/CMPT496-Deep-Learning

2.3 Proof:



Kaggle username: Holayn @ https://www.kaggle.com/holayn

4 Derivations and Solutions to Problems:

1. In order to find the value of x that maximizes g(x), we first take the derivative of g(x) like so:
   $g(x) = -3x^2 + 24x - 30$
   $g'(x) = -6x + 24$
   Next, we determine what values of x make g'(x) = 0. This signifies that the slope is equal to zero.
   $0 = -6x + 24$
   $6x = 24$
   $x = 4$
   We can see that at x = 4, the slope of the graph is 0. When the slope of a graph is 0, we know that it is neither increasing nor decreasing; thus it must be flat, so we can test if it is a maximum or minimum. We can test the value of the slope of the graph at x=3 and x=5. If the value of the slope at x=3 is positive, and the slope at x=5 is negative, we can conclude that x=4 maximizes g(x), since the graph has positive slope, becomes flat, then has negative slope, indicating a maximum.
   $g'(3) = -6(3) + 24 = 6$
   $g'(5) = -6(5) + 24 = -6$
   The slope of x=3 is positive, and the slope of x=5 is negative. Thus, x=4 maximizes g(x).
2. In order to take the partial derivatives of f(x) with respect to x_0 and x_1, we first take the derivative of f(x) with respect to x_0:
   $f(x) = 3x_0^3 - 2x_0x_1^2 + 4x_1 - 8$
   $f_{x0} = 9x_0^2 - 2x_1^2$
   Now we take the derivative of f_{x0} with respect to x_1:
   $f_{x0x1} = -4x_1$
3. Considering the matrix A and B where…
   A = [[1 4 -3]
   [2 -1 3]]
   B = [[-2 0 5]
   [0 -1 4]]
   a. You cannot multiply the two matrices. In order to perform matrix multiplication, matrix A's number of columns has to equal matrix B's number of rows. Matrix A has 3 columns. Matrix B has 2 rows. 3 != 2, so one cannot multiply these matrices together.
   b. Taking the transpose of A results in A = [[1,2]
   [4,-1]
   [-3,3]]
   To perform matrix multiplication, multiply the members of a column in one matrix to the members of a row in the other matrix, and then add the results to get the new value in the matrix. Repeat for all rows and columns. See below:

A^TB = [[((1*-2)+(2*0)),((1*0)+(2*-1)),((1*5)+(2*4))]
[((4*-2)+(-1*0)),((4*0)+(-1*-1)),((4*5)+(-1*4))]
[((-3*-2)+(3*0)),((-3*0)+(3*-1)),((-3*5)+(3*4))]]
= [[-2 -2 13]
[-8 1 16]
[6 -3 -3]]

In order to get the matrix's rank, we must find the number of non-zero rows (the row's elements are not all zero) when the matrix has been converted to upper triangular form (all elements under the main diagonal of the matrix is zero) using elementary row operations. Here is the process of converting the matrix to upper triangular form using elementary row operations, through a process called row reduction:

[[-2 -2 13]
[-8 1 16]
[6 -3 -3]]

Multiply the first row by -4 and add it to the second row, and multiply the first row by 3 and add it to the third row:

[[-2 -2 13]
[0 9 -36]
[0 -9 36]]

Now add the second row to the third row:

[[-2 -2 13]
[0 9 -36]
[0 0 0]]

Now the matrix is in upper triangular form. We can see that there are two non-zero rows out of three rows (one of the rows has all zeroes), so the rank of the matrix is 2.

```
import numpy as np
import tensorflow as tf

a = np.array([[1,2],[4,-1],[-3,3]])
print("Array A transposed: ")
print(a)
b = np.array([[-2, 0, 5], [0, -1, 4]])
print("Array B: ")
print(b)

a = tf.constant(a)
b = tf.constant(b)

with tf.Session() as sess:
    t = tf.matmul(a,b).eval()
    print("Multiplication of matrices: ")
    print(t)
    print("Rank: (using numpy)")
    print(np.linalg.matrix_rank(t))
    print("Rank: (using tensorflow)")
    print(tf.rank(t).eval())

    ## Note on TensorFlow.org's api doc for tf.rank...
    # The rank of a tensor is not the same as the rank of a matrix.
    # The rank of a tensor is the number of indices required to uniquely select each element of the tensor.
    # Rank is also known as "order", "degree", or "ndims."
    # https://www.tensorflow.org/api_docs/python/tf/rank
```

```
Array A transposed:
[[ 1  2]
 [ 4 -1]
 [-3  3]]
Array B:
[[-2  0  5]
 [ 0 -1  4]]
Multiplication of matrices:
[[-2 -2 13]
 [-8  1 16]
 [ 6 -3 -3]]
Rank: (using numpy)
2
Rank: (using tensorflow)
2

Process finished with exit code 0
```

4. If we are to find the expected value of X of a normal distribution with a mean or expected value of 2 and a variance of 3, we can see that the expected value of X of this normal distribution is 2, as the expected value, or the mean, of the normal distribution is 2.