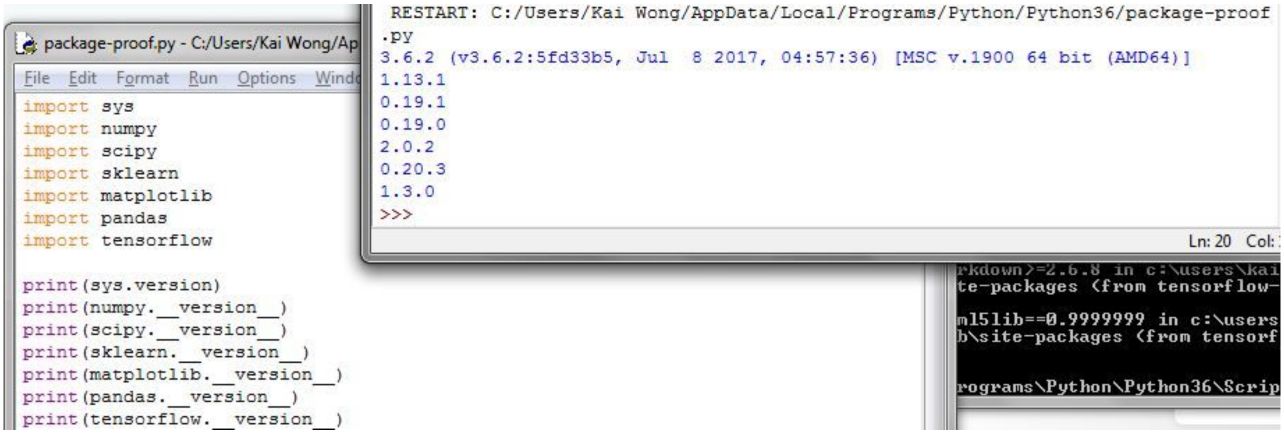
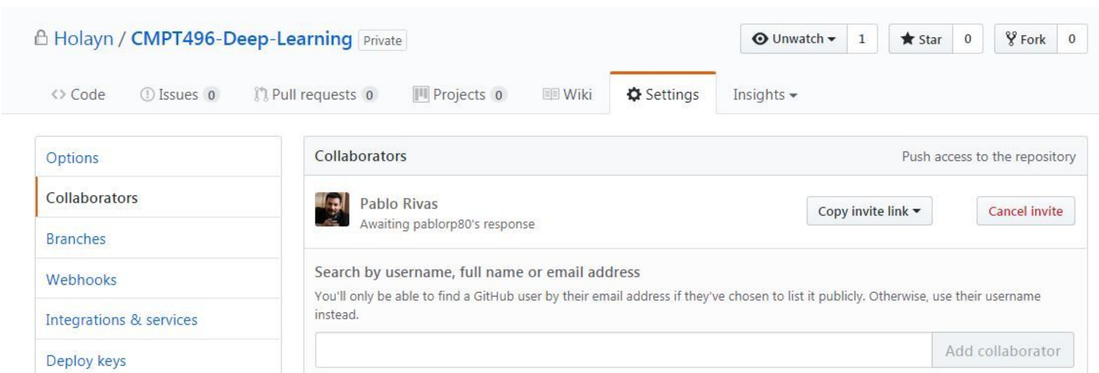
Kai Wong
CMPT496
9/11/2017

Homework 0

2.1 Proof:

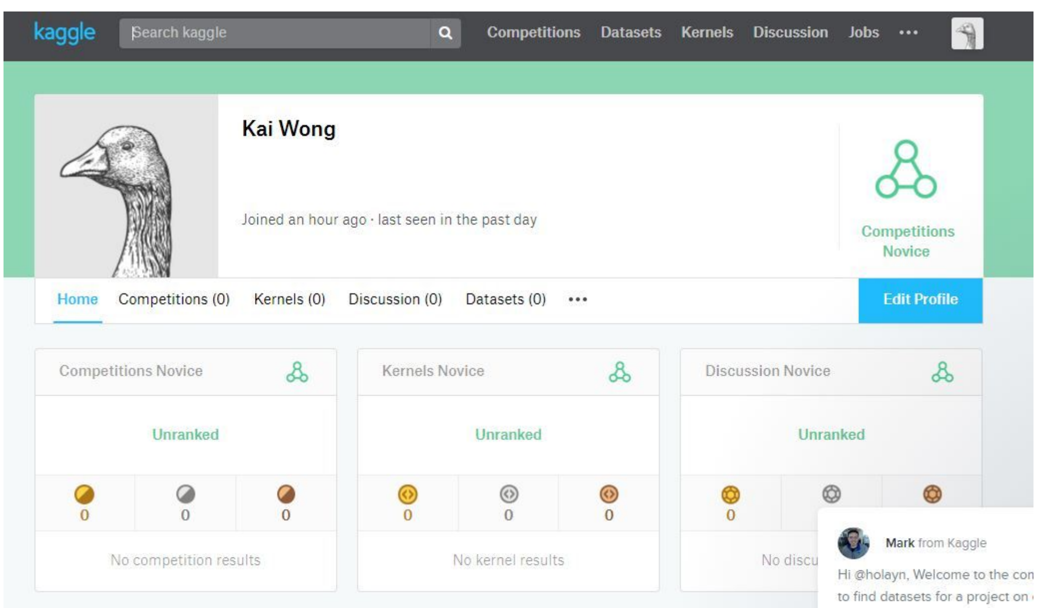


2.2 Proof:



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.
      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 linearly independent row vectors. This means that no row can be derived from a linear combination of any other rows. In A^TB shown above, there is one row vector that is not linearly independent: the third row. Subtracting the second row vector from the first row vector creates the third row vector. Therefore, A^TB has one non-linearly independent row vector. It has 3 row vectors, so its rank is 2, since it has two linearly independent row vectors.

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.