Date: July 8, 2025

From G. Pang , **File is BRCA1-Simulation4**

-------------------------------------------------------------------------------------------------------------------

**Dataset description:**

Training dataset: 2000 inputs, Ref vector is 1x1000, same for variant vector;

one binary output (0 or 1)

**2000 training samples: 1000 label 1, 1000 label 0**

**Testing dataset: 2000 testing samples: 1000 label 1, 1000 label 0**

**================================================================================**

**Variant vector of label 1 is obtained by randomly choosing 8% (80 locations) in Ref vector, and multiply that element by plus or minus [1.2, 1.5]. One round would generate 10 cases.**

**Variant vector of label 0 is obtained by randomly choosing 2% (20 locations) in Ref vector, and multiply that element by plus or minus [1.2, 1.5]. One round would generate 5 cases.**

**Variant vector of label 0 is also obtained by Ref vector, and multiply each element by plus or minus 1%. One round would generate 5 cases.**

**In one round,10 label\_1 training samples, plus 10 label\_0 samples are obtained.**

**-----------------------------------------------------------------------------------------------------------------------------**

**ANN model is a typical 3 hidden layer structure (512;128;32 hidden nodes) with dropout 0.3.**

(2000, 2000)

(2000, 1)

(2000, 1)

\*\* CHECK Average BCE Loss for multiple samples: [0.0191]

Accuracy Score: 0.9945

error0to1 = 10 ; label is 0

error1to0 = 1 ; label is 1

Testing total error = 11 percentError = 0.55

2000 test cases: 1000 label 0; 1000 label 1

**-----------------------------------------------------------------------------------------------------------------------------------**

test\_x.shape (2000, 2000)

test\_y.shape (2000, 1)

(2000, 1)

Compare between test\_y and y\_pred\_binary

[0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 2, 0, 1, 1, 1, 1, 1, 1]

Sum of errors = 11 Percent Error 0.55

**==================================================================================**

train\_x.shape (2000, 2000)

train\_y.shape (2000, 1)

(2000, 1)

Compare between train\_y and y\_pred\_binary

[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

len(error) 20

Sum of errors = 0 Percent Error 0.0

**-----------------------------------------------------------------------------------------------------------------------------------**

**# Difference vector**

**#**

**# Getting the difference vector of the training cases for training**

**# input is now 1 x 1000; binary output**

**#**

**ref\_vector = train\_x[:,0:1000]**

**var\_vector = train\_x[:,1000:2000]**

**diff\_vector = np.abs(ref\_vector - var\_vector)**

**print(ref\_vector.shape,var\_vector.shape,diff\_vector.shape )**

(2000, 1000) (2000, 1000) (2000, 1000)

63/63 - 0s - loss: 2.6855e-04 - accuracy: 1.0000

0.0002685519284568727 1.0