TOSHKENT SHAHRIDAGI INHA UNIVERSITETI
INHA UNIVERSITY IN TASHKENT

STUDENT NAME:			
STUDENT ID NUMBER:		TION NUMBER IN THIS COURSE:	

HW2 – FALL 2025

COURSE NAME:	*Big Data Analytics*
COURSE NUMBER:	SOC4170
EXAMINATION DATE:	TIME:
EXAMINATION DURATION:	Minutes
ADDITIONAL MATERIALS ALLOWED TO USE:	**Calculator**
SPECIAL INSTRUCTIONS:	*- Answers will only be evaluated if they are readable.* *- Show all your work. For some questions, you may get partial credit even if the end result is wrong due to a calculation mistake. If you make assumptions, state your assumptions clearly and precisely.*

Please do not open the examination paper until directed to do so.

READ INSTRUCTIONS FIRST:

Desks should be free from all unnecessary items (books, notes, technology, food, water, clothes)

Use of any electronic device (Phone, iPod, iPad, laptop) is not allowed during the examination.

Cheating, talking to fellow students, singing, turning back are not allowed

Write your Name (*capital letters*), ID number and Section number in each page of your examination paper.

Final answers must be written by **only blue or black, non-erasable pen**. Do not use highlighters or correction pen.

All answers should be written in the space provided for each question, unless specified the other way.

If you have a problem please raise your hand and wait quietly for a Proctor.

You are not allowed to leave the exam room until you submit the exam papers.

CALCULATORS ARE ALLOWED.

STUDENT NAME:

STUDENT ID NUMBER: CTION NUMBER IN THIS COURSE:

1.[16 points] (True/False questions) Write T if the following statement is True, write F otherwise inside the table below

a.	b.	c.	d.	e.	f.	g.	h.	i.	j.	k.	l.	m.	n.	o.	p.

a. In FP mining if minsup threshold gets bigger, it takes less time to find all the frequent itemsets.

b. The learning slowdown problems can be fixed by using the quadratic cost function.

c. Softmax is used for output layer neurons because it makes them independent of each other.

d. Sigmoid function is a step funcion.

e. When we sample data for mini-batch, we do random sampling without replacement

f. Unsupervised learning includes clustering

g. The Bootstrap sample is taken from the original by sampling with replacement

h. The first step of backpropagation is calculating "the error" of the output layer

i. The direction of a gradient vector is the opposite direction from a function's minimum point

j. When inputs are images, we can artificially increase the number of training inputs.

k. When using dropout, half of hidden layer neurons are disabled before start of every mini-batch

l. When deciding the learning rate, we need to start with big learning rate.

m. Boosting stands for Bootstrap Aggregation

n. Small changes of weights and biases of a perceptron causes small changes of output

o. Regularization decreases overfitting by decreasing bias values.

p. If a weight has a large magnitude, L1 regularization shrinks more weight.

STUDENT NAME:

STUDENT ID NUMBER: CTION NUMBER IN THIS COURSE:

2.[10 points] Confusion Matrix

We have 2 models, M1 and M2 which detect fraud credit card transactions.

Total number of test data transactions are 1000. And the test data contains 4 actual fraud transactions.

(a) [5 Points] M1 detected 5 fraud transactions, four of which are normal transactions.
Complete the confusion matrix and compute accuracy, precision, and recall.

Predicted class Actual class	Fraud = Yes	Fraud = No	Total
Fraud = Yes			
Fraud = No			
Total			

Accuracy: % (1 point)

Precision: % (1 point)

Recall: % (1 point)

=

(b) [5 Points] M2 detected 14 fraud transactions, 3 of which are fraud transactions.
Complete the confusion matrix and compute accuracy, precision, and recall.

Predicted class Actual class	Fraud = Yes	Fraud = No	Total
Fraud = Yes			
Fraud = No			
Total			

Accuracy: % (1 point)

Precision: % (1 point)

Recall: % (1 point)

STUDENT NAME:

STUDENT ID NUMBER: CTION NUMBER IN THIS COURSE:

3.[8 points] BP4 $\frac{\partial C}{\partial w^l_{jk}} = a^{l-1}_k \delta^l_j$

(a) [8 Points] Prove the BP4 formula

4.[16 points] Adaboost

Given a set of *d* class-labeled tuples, ($\mathbf{X_1}$, y_1), …, ($\mathbf{X_d}$, y_d)
Initially, all the weights of tuples are set the same (1/d)

Let d = 5

and here is the sequence of your random numbers for sampling D_1 **:** ($\mathbf{X_3}$ $\mathbf{X_1}$ $\mathbf{X_1}$ $\mathbf{X_2}$ $\mathbf{X_2}$)

From D_1**, your model** M_1 **is generated.** M_1 **predicted the class labels of** X_1 **and** X_2 **correctly but misclassified** X_3

(a) [4 Points] What is the $error\ (M_1)$? Refer to the following equation and show all your work.

$$error(M_i) = \sum_j^d w_j \times err(\mathbf{X_j})$$

$error$ is the sum of the weights of misclassified tuples.

STUDENT NAME:

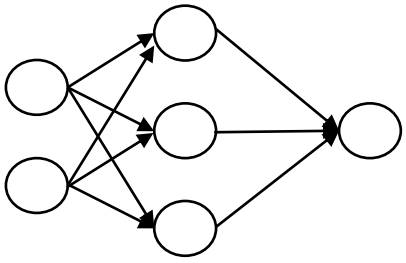
STUDENT ID NUMBER: CTION NUMBER IN THIS COURSE:

(b) [5 Points] What are the weights of 5 tuples $(x_1 - x_5)$ before normalization? Show all your work

(c) [7 Points] Write the sum of old weights and the sum of new weights. And write the weights of 5 tuples $(x_1 - x_5)$ after normalization? Show all your work

5.[50 points] Feedforward and back propagation

Below is a neural network:



All the matrices and vectors are as follows:

$$w^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \end{bmatrix} \quad w^3 = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix} \quad b^2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad b^3 = [0]$$

We are doing the mini-batch of size 4. And four inputs are

$$x_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad x_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad x_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad x_4 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

STUDENT NAME:

STUDENT ID NUMBER: CTION NUMBER IN THIS COURSE:

Labels are

$y_1 = [0]$ $y_2 = [0]$ $y_3 = [1]$ $y_4 = [1]$

And the learning rate is $\eta = 0.6$

We are processing 4 inputs at the same time by having all 4 of them as a matrix.

Round off all the number to the nearest 1000th. For example, 6.6543 is 6.654 and 3.3355 is 3.336.

The cost function C is a quadratic cost function.

And the back propagation equations are as follows:

Summary: the equations of backpropagation

$$\delta^L = \nabla_a C \odot \sigma'(z^L) \quad \text{(BP1)}$$

$$\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l) \quad \text{(BP2)}$$

$$\frac{\partial C}{\partial b^l_j} = \delta^l_j \quad \text{(BP3)}$$

$$\frac{\partial C}{\partial w^l_{jk}} = a^{l-1}_k \delta^l_j \quad \text{(BP4)}$$

z	-0.262	-0.362	-0.462	-0.962	-1.133	0.745	1.233	1.425	1.518
a	0.435	0.410	0.387	0.276	0.244	0.678	0.774	0.806	0.820

z	0	0.5	1	1.5	2	-0.5	-1	-1.5	-2
a	0.5	0.622	0.731	0.818	0.881	0.378	0.269	0.182	0.119

(a) [4 Points] **What is Z^2 matrix?** First represent it with matrix format and write the result. Show all your work.

STUDENT NAME:

STUDENT ID NUMBER: CTION NUMBER IN THIS COURSE:

(b) [4 Points] What is a^2 matrix? Show all your work.

(c) [4 Points] What is Z^3 matrix? First represent it with matrix format and write the result. Show all your work.

(d) [4 Points] What is a^3 matrix? Show all your work.

STUDENT NAME:

STUDENT ID NUMBER: CTION NUMBER IN THIS COURSE:

(e) [3 Points] What is ${}^{1}\delta^{3}$ value? ${}^{1}\delta^{3}$ is the layer 3 delta for input 1. Show all your work.

(f) [3 Points] What is ${}^{2}\delta^{3}$ value? ${}^{2}\delta^{3}$ is the layer 3 delta for input 2. Show all your work.

(g) [3 Points] What is ${}^{3}\delta^{3}$ value? ${}^{3}\delta^{3}$ is the layer 3 delta for input 3. Show all your work.

(h) [3 Points] What is ${}^{4}\delta^{3}$ value? ${}^{4}\delta^{3}$ is the layer 3 delta for input 4. Show all your work.

STUDENT NAME:

STUDENT ID NUMBER: CTION NUMBER IN THIS COURSE:

(i) [4 Points] What is average $\frac{\partial C}{\partial b^3}$ vector? Show all your work.

(j) [12 Points] What is average $\frac{\partial C}{\partial w^3}$ matrix? Show all your work.

(k) [3 Points] What is new b^3 vector? Show all your work.

STUDENT NAME:

STUDENT ID NUMBER: CTION NUMBER IN THIS COURSE:

(l) [3 Points] What is new w^3 matrix? Show all your work.