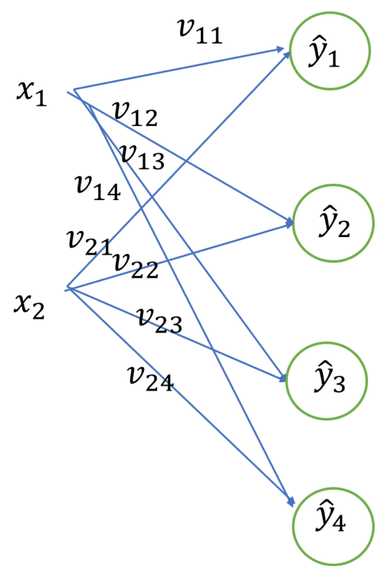
DSCI 6011: Deep Learning
Fall 2025

Assignment 1A

Q1. Consider an iris dataset with 3 samples for a multitask (classification and regression).

Observations		Target		
Sepal Length	Sepal width	Class	Petal Length	Petal Width
5.1	3.5	Setosa	2.0	3.0
3.4	3.2	Versicolor	0.9	1.2
5.6	2.8	Setosa	1.8	3.5

A one layer neural network is provided below to predict the class label, petal length and petal width. The layer consists of linear transformation with output z_k, which is subsequently transformed by the output function to $\hat{y}_i$. The outputs are $\hat{y}_1$ for setosa, $\hat{y}_2$ for versicolor, $\hat{y}_3$ for petal length, and $\hat{y}_4$ for petal width.



	Layer configuration
Weights,	$$v_{ij} = (r)\sqrt{\frac{2}{length\ of\ the\ layer\ input\ vector}}$$ where r is a random number between 0 and 1. Use RAND in your calculator to generate this random number
Bias	$c^T = [-0.1, 0.1, 0.0, -0.2]$
Activation/output function	Softmax for classification, $y_i = softmax(z_i)$ and identity for regression, $y_i = z_i$

Do the forward propagation of this network for the given training data by completing the following steps. Note that you must use matrix form unless an operation does not support a matrix.

a) Build the input matrix, X, from the training data.
b) Build the output label matrix, T, from the training data.
c) Convert the label matrix, T, to the actual output matrix, Y. (Hint: for classification, t, is class label and y is actual probability.)
d) Redraw the network and assign the edges random weights according to the formula in the table above.
e) Arrange the weights of the layer in a matrix, V
f) Compute the linear transformation of the output layer, $Z = XV + C$ and the layer output , $\hat{Y}$.
g) Calculate the total average loss, L. It is the sum of the average cross-entropy loss for classification, L_1, and the mean squared error for regression, L_2.
h) Find the predicted labels. Evaluate the performance of classifier and regressor with classification accuracy and R^2, respectively.

Q2. Convexity and derivatives

a. $f(a, b) = ax^3y + b^3zy^2 + b^2a$

Find the partial derivatives of this function w.r.t all its variables.

b. Is $f(x) = \max(x, 0)$ continuous at x=0? Is its derivative continuous? Show mathematically.

c. *Given,* $v_1 = x_1^2 + x_2$, $v_2 = x_1x_2 + 2x_2^3$, *and* $y = v_1v_2$,
Find $\frac{\partial y}{\partial v_1}, \frac{\partial y}{\partial v_2}, \frac{\partial v_1}{\partial x_1}$, *and* $\frac{\partial v_2}{\partial x_1}$. Use chain rule with these terms to calculate $\frac{\partial y}{\partial x_1}$.

d) *In question (c), replace* v_1 *and* v_2 *in the equation of* y *and then directly compute* $\partial y/\partial x_1$.

Q3. Given the network in Q1, find the following derivatives.

i. Derivatives of the loss,, $\frac{\partial L}{\partial \hat{y}_1}$ and $\frac{\partial L}{\partial \hat{y}_3}$

ii. $\frac{\partial z_1}{\partial v_{11}}$ and $\frac{\partial z_1}{\partial v_{14}}$

iii. $\frac{\partial z_1}{\partial x_1}$