Birla Institute of Technology & Science, Pilani
Work Integrated Learning Programmes Division
First Semester 2024-2025

Comprehensive Examination
(EC-3 Regular)

Course No. : **SS ZG568**
Course Title : **Applied Machine Learning**
Nature of Exam : Open Book
Weightage : 40%
Duration : 2.5 Hours
Date of Exam : 30/11/2024 (AN)

No. of Pages = 1
No. of Questions = 4

Note to Students:
1. Please follow all the *Instructions to Candidates* given on the cover page of the answer book.
2. All parts of a question should be answered consecutively. Each answer should start from a fresh page.
3. Assumptions made if any, should be stated clearly at the beginning of your answer.

Q.1 Draw a NAND logic perceptron using the following activation and verify its output. Prediction (y`) = 1 if $Wx+b > 0$ and 0 if $Wx+b \leq 0$. [10]

Q.2 Consider the two-dimensional patterns (2, 1), (3, 5), (4, 3), (5, 6), (6, 7), (7, 8). Compute the 1st principal component using PCA Algorithm. [10]

Q.3 Explain what effect will the following operations have on the bias and variance of your model. Fill in one of 'increases', 'decreases' or 'no change' in each of the cells: [10]

	Bias	Variance
Regularizing the weights in a linear/logistic regression model		
Pruning a decision tree (to a certain depth)		
Increase the number of hidden units in a neural network		
Use dropout to train a deep learning model		
Removing all non-support vectors in SVM		

Q.4 [10]

(a) Between SVM and logistic regression, which algorithm is most likely to work better in the presence of outliers? Why?

(b) If you observe that the cost function decreases rapidly before increasing or stagnating at a specific high value, what could you infer?

(c) How many binary classifiers would you need to implement one-vs-one for four classes? How does it work?

(d) In classification problems like logistic regression, classification accuracy alone is not considered a good measure. Why?

(e) Suppose we train a classier to learn f : X ->Y, where X is the feature vector X =< X,1X2, X3 >. Which classier contains sucient information to allow calculating P(X1X2X3 Y)? Explain.

*****