Ref. No. Ex/IEE/PE/B/T/421B/2024

B.I.E.E. 4thYear. 2nd Semester Examination, 2023-24

Subject: Data Analysis for Instrumentation System.

Time: Three hours Full Marks : 100

Instructions: Figures in the margin indicate marks for all items of Questions.(in three Pages)

Module:I [Answer any one Question; Marks: 1x10 =10]

1.(a) What is Primary Data? [1]

(b) Write short notes on Philosophy of Primary Data Collection and Presentation for Pattern Recognition in Instrumentation System with special emphasis on specifying a type of sensor, calibration values, and a description of the digitising and recording equipment. [6]

(c) Define Clearly Data set for training (or design) and test (or validation set) and also data set for performance evaluation purposes. What is randomisation of data? [3]

2. Consider two events are not independent, suppose a Company plans to tests two new techniques for improving the extraction of Oil from Ground. First technique consists of setting off an explosion at bottom of well to fracture the strata and then testing seismically to determine the extent of the fracturing. The second technique consists of injecting hot brine in to the well to loosen the oil and then pumping to measure oil recovery. Let E be the event that the explosion successfully fractures the strata within a radius of 100 meters and let R be the event that oil can be recovered at a rate of 50 barrels per day after pumping in hot brine. In certain region P(E) is estimated to be 0.8. If E occurs probability of R occurring, P(R|E) is estimated to be 0.9. But if explosion is not a success the estimated probability of recovery is only $P(R|\bar{E}) = 0.3$. These three assumptions are sufficient to define the probabilities of any combination of outcomes given any set of constraints.

a)Calculate probabilities of four possible joint events: P(E and R) , P(E and $\bar{R}$), P($\bar{E}$ and R), P($\bar{E}$ and $\bar{R}$) [6]

b) Calculate probability of successful brine recovery: P(R) [2]

c) What is the probability that Explosion and Brine Test are both successful or both fail? [2]

Module :II [Answer any two Questions; Marks : 2x19 =38]

3. a)What is the Purpose of Fisher's Linear Discriminant Analysis. [3]

b) If the projected mean of class 1 is $\mathbf{m_{1p}}$ and the same for class 2 is $\mathbf{m_{2p}}$, and if $\mathbf{s^2_{1p}}$ & $\mathbf{s^2_{2p}}$ is the projected within class scatter of class 1 patterns class 2 patterns, then the Fisher's criterion is :

$$\mathbf{J(V)} = \frac{|\mathbf{m_{1p}} - \mathbf{m_{2p}}|^2}{(\mathbf{s^2_{1p}} + \mathbf{s^2_{2p}})}$$

Prove that the solution for V is: $V = W^{-1}(m_1 - m_2)$. [16]

4. Consider ten two dimensional feature vector points : $(1,2)^t$; $(3,5)^t$; $(4,3)^t$; $(5,6)^t$; $(7,5)^t$ of Class 1 and $(6,2)^t$; $(9,4)^t$; $(10,1)^t$; $(12,3)^t$; $(13,6)^t$ of Class 2.

a) Calculate normalised projection Vector V_N. [14]
b) Consider two feature vector points $P_1 = (4,4)^t$ and P_2 $(12,4)^t$ and confirm its Classification. [3]
c) Prove the projection of $(m_1 - m_2)$ onto $V_N = 4.875$ [2]

5. Consider ten two dimensional feature vector points : $(1,2)^t$; $(3,5)^t$; $(4,3)^t$; $(5,6)^t$; $(7,5)^t$; $(6,2)^t$; $(9,4)^t$; $(10,1)^t$; $(12,3)^t$; $(13,6)^t$.
a) Calculate Covariance Matrix. [13]
b) Calculate Eigen Vector and Eigen Values. [4]
c) Calculate PC1 and PC2 for the first feature vector point. [2]

Module :III [Marks : 1x10=10]

6. sklearn.preprocessing package provides several common utility functions and transformer classes to change raw feature vectors into representation that is more suitable for down stream estimators. Mention and describe briefly most important utility functions for data pre-processing jobs. [10]

Module :IV [Answer any three Questions; Marks : 3x14=42]

7. Prove that for d-dimensional decision boundary of Multivariate Normal classes having equal Covariance Matrices is hyperplane : $\mathbf{b}^T.\mathbf{x}+c = 0$ [14]

8. (a) Construct a Neural Net having two input nodes (excluding Bias), two hidden nodes (excluding Bias), and one output node. Consider x_0 values for both input and hidden layer is unity. [3]

(b) Compute all gradient back flow (upto right side of hidden layer) as per back propagation Algorithm and Sequential MSE with Steepest (Gradient) Descent Algorithm and indicate in the Net drawing. [11]

9.(a) Consider Turbine supervisory Instrumentation and control system having Polygonal Decision Boundary [ABCDEFA anti clockwise] bounded by the following equations.

Line AB: y =0.7769x+96; Line BC: x =520; Line CD: y =530;

Line DE: y =0.7647x+247.06; Line EF: y =400; Line FA: x =0

Operating points inside the polygon are in Recommended zone and outside points of the polygon are Not- Recommended zone.

Coordinate of lowest point/vertex A is (0, 96). Coordinates other points/vertex are B(520,500) ; C(520,530); D(370,530); E(200,400); F(0,400); Construct Neural Net with Hidden Layer for classification of Recommended and Not-Recommended zone of safe

operation of the Turbine. Where x indicate the Middle wall temperature of Main control valve casing and y indicate the Main Steam Temperature. [11]

(b).Also check the operating points (200,260) and (300,490) and classify (Recommended or Not-Recommended). [3]

10.(a) What is the Advantage of Adaptive Decision Boundary Algorithm over Simple Perceptron Algorithm. [2]

(b).Write a Program in Python for calculating final weights for Non linear Decision boundary with adaptive algorithm with the following feature vector points in Allowed and Not-allowed operating zone of Turbine supervisory Instrumentation and control system (Allowable maximum main Steam Pressure ahead of Turbine when opening of MS Stop Valve.)

Allowed zone :

$(2 ,0.8 ,4, 1.6,0.64)^t$; $(5 ,1.1 ,25, 5.5,1.21)^t$; $(8 ,1.8 ,64, 14.4,3.24)^t$; $(10 ,2.5 ,100, 25,6.25)^t$;

$(13,4,169,52,16)^t$; $(15,5 ,225, 75,25)^t$; $(20 ,8.5 ,400, 170,72.25)^t$; $(30 ,20.5 ,900, 615,420.25)^t$

Not-Allowed zone :

$(2 ,1.2 ,4, 2.4,1.44)^t$; $(5 ,1.5 ,25, 7.5,2.25)^t$; $(8 ,2.3 ,64, 18.4,5.29)^t$; $(10 ,3.2 ,100, 32,10.24)^t$;

$(13,4.8,169,62.4,23.04)^t$;$(15,5.8,225,87,33.64)^t$;$(20,9.2,400,184,84.64)^t$;$(30,21.8,900, 654,475.24)^t$

[12]

11.a) Write a program in Python for calculating final weights for linear Decision boundary with simple Perceptron algorithm with the following feature vector points in Recommended and Not-Recommended operating zone of Turbine supervisory Instrumentation and control system (Recommended maximum reheat temperature before Turbine is loaded).

Recommended zone :

$(0,190)^t$; $(40,210)^t$; $(100,245)^t$; $(140,280)^t$; $(200,325)^t$; $(240,350)^t$; $(300,390)^t$; $(400,460)^t$;

Not -Recommended zone :

$(0,210)^t$; $(40,230)^t$; $(100,275)^t$; $(140,300)^t$; $(200,340)^t$; $(240,370)^t$; $(300,410)^t$; $(400,480)^t$

[9]

b) Construct Neural Net with Hidden layer for Realisation of Exclusive-OR Logic for two dimension feature vectors. [5]

12. Consider Novel Feature Vector for ECG Classification Technique using Deep Learning:

a) Draw scheme of Methodology that performs the Classification. [3]

b) Construct with block diagram the Scheme of the utilised CNN. [11]