Total number of printed pages: **Programme(UG)/6th Sem/UCSE603/2024**

MACHINE LEARNING

Full Marks: 100

Time : Three hours

The figures in the margin indicate full marks for the questions.

*Answer **any five** questions.*

Q1. Consider the following data of animals and construct a decision tree -

fins	aquatic	legs	tail	Type
No	No	Yes	No	Mammal
Yes	Yes	No	No	Fish
No	No	Yes	Yes	Bird
No	Yes	Yes	Yes	Mammal
No	No	Yes	No	Bird
No	Yes	Yes	Yes	Reptile
No	Yes	No	Yes	Fish
Yes	Yes	Yes	Yes	Reptile

Consider a case when an animal has the following properties -

fins= No, aquatic =No, legs =No, and tail = yes. Identify the type of the animal.

(20)

Q2.

a) Consider a case of non-linear regression with the degree of polynomial is 3 (n=3) returns an R^2 value of 83%. The R^2 changed to 98% when n =10. Out of these two cases which one will be better and why?

(5)

b) Consider the following data for K- means clustering (K=3). The initial three centroids are P2 (for Centroid 1, C1), P3 (for Centroid 2, C2), and P7 (for Centroid 3, C3), respectively. Use K-means clustering (K=3) to determine the final centroids of the three clusters. Identify the cluster for the point (X Value = 10, Y Value =5).

Point	X Value	Y Value
P1	0	0
P2	1	0
P3	5	3
P4	4	4
P5	10	1
P6	3	9
P7	8	8
P8	1	1

(15)

Q3. Consider the following truth table: A, B, and C are inputs, and D is output. Develop an ANN for realizing the truth table. Identify the number of coefficients and biases and their best corresponding values.

A	B	C	D
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

(20)

Q4. Consider the IPL run score dataset of the four stadiums as mentioned below

#Match	Guwahati Barsapara	Kolkata Eden Garden	Chinnaswamy stadium Bengaluru	wankhede stadium Mumbai
1	190	170	260	110
2	80	110	90	200
3	150	250	130	210
4	180	210	160	190
5	200	50	140	130
6	90	100	200	250
7	260	290	210	270
8	170	180	150	260

BCCI (Board of Control for Cricket in India (BCCI) is interested in checking whether all the pitches are equally batting friendly or not.

a) Write your NULL and Alternate Hypothesis.
b) Use an ANOVA study to check whether any of those pitches performed better in terms of runs scored with a significance level of $\alpha = 5\%$.

Critical values of F for the 0.05 significance level:

	1	2	3	4	5	6	7	8	9	10
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.39	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.97	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.10	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.97	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.56	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.33	3.47	3.07	2.84	2.69	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.38	2.32	2.28
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.26
25	4.24	3.39	2.99	2.76	2.60	2.49	2.41	2.34	2.28	2.24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18

(5+15)

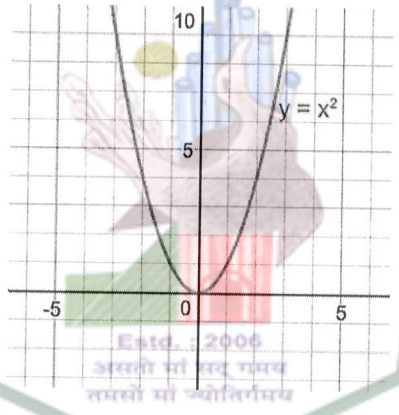
Q5. Consider the following data points with their class and develop an SVM Model. Compute the coefficient and bias of your model. Identify the support vectors of your model.

X Value	Y Value	Class
-5	-10	1(+)
-5	-15	1(+)
-4	9	-1 (-)
3	6	-1 (-)

(15+3+2)

Q6. Consider the error graph of an ANN. Assume you are at the point of y = 3.

a. Write the equation for updating y such that you can attain minimum error.
b. What will happen if you choose learning rate = 1?
c. Consider learning rate = 0.1, and trace up to 5 epochs to determine the y value.



(5+5+10)

Q7. Write short notes on

a. SIGMOID function
b. Degree of freedom
c. Rewards in Reinforcement Learning
d. t-test vs Z-test

(4*5)