



B.E DEGREE EXAMINATIONS: NOV/DEC 2023

(Regulation 2018)

Seventh Semester

COMPUTER SCIENCE AND ENGINEERING

U18CST7002: Machine Learning Techniques

COURSE OUTCOMES

CO1: Differentiate the implementation of mathematical models to various machine learning methods.

CO2: Illustrate graphical models and multiple learners.

CO3: Develop projects using appropriate machine learning approaches for real life problems.

Time: Three Hours

Maximum Marks: 100

Answer all the Questions:-

PART A (10 x 2 = 20 Marks)

(Answer not more than 40 words)

- | | | |
|---|-----|-------------------|
| 1. State the difference between classification and clustering. | CO1 | [K ₂] |
| 2. Classify well posed and ill posed problems. | CO1 | [K ₂] |
| 3. Recall the steps in EM algorithm. | CO1 | [K ₁] |
| 4. Outline the working of running mean smoothing model. | CO1 | [K ₂] |
| 5. Classify weak learner and strong learner. | CO2 | [K ₂] |
| 6. Explain the cascading method used to combine multiple classifiers performance. | CO2 | [K ₁] |
| 7. How will you diagnosis graph use in graphical model? | CO2 | [K ₂] |
| 8. List the importance of learning factor in optimization technique. | CO3 | [K ₁] |
| 9. What is entropy in decision tree? | CO3 | [K ₁] |
| 10. Calculate Euclidian distance between the pair of points (5,10) and (11,2) | CO3 | [K ₃] |

Answer any FIVE Questions:-

PART B (5 x 16 = 80 Marks)

(Answer not more than 400 words)

11. a) Apply hierarchical clustering method to form clusters in the given distance matrix below: 10 CO1 [K₃]

	A	B	C	D	E	F
A	0	0.71	5.66	3.61	4.24	3.20
B	0.71	0	4.95	2.92	3.54	2.50
C	5.66	4.95	0	2.24	1.41	2.50
D	3.61	2.92	2.24	0	1.00	0.50
E	4.24	3.54	1.41	1.00	0	1.12
F	3.20	2.50	2.50	0.50	1.12	0

- b) Consider the given scenario. It rains on 40 percent of the days and when it rains, there is a 90 percent chance that the grass gets wet; maybe 10 percent of the time it does not rain long enough for us to really consider the grass wet enough. There 6 CO2 [K₃]

is a 20 percent probability that the grass gets wet without rain for example, when a sprinkler is used. Find the probability that the rain has caused the grass to get wet. Illustrate the above scenario in graphical model representation.

12. a) Generate a sample from multivariate normal density $N(\mu, \Sigma)$. Derive the joint bivariate density function for the above sample. 10 CO1 [K₃]
 b) Use the following data to construct a linear regression model for the auto insurance premium as a function of driving experience. 6 CO3 [K₃]

Driving Experience (in years)	5	2	12	9	15	6	25	16
Monthly auto insurance premium (\$)	64	87	50	71	44	56	42	60

13. a) Implement AdaBoost algorithm to combine base learners. Summarize the advantage of boosting over the bagging technique. 10 CO2 [K₂]
 b) In a two-class, two-action problem, if the loss function is $\lambda_{11} = \lambda_{22} = 0$, $\lambda_{12} = 10$, and $\lambda_{21} = 5$, write the optimal decision rule. How does the rule change if we add a third action of reject with $\lambda = 1$? 6 CO3 [K₃]
14. a) Explain the canonical cases for conditional independence events in graphical model. 10 CO2 [K₂]
 b) Illustrate how pairwise separation is used to solve multiple class problem. 6 CO1 [K₂]
15. a) Describe model selection approaches for balancing model complexity with appropriate terminologies and its applicability. 10 CO3 [K₂]
 b) Develop an algorithm for condensed nearest neighbor and write its augmented error function. 6 CO1 [K₃]
16. a) Apply k-means clustering in image compression. Write about the performance metrics. 10 CO3 [K₃]
 b) Summarize the discussion of error correcting output codes with example 6 CO2 [K₂]
