Roll No.												

PRESIDENCY UNIVERSITY

BENGALURU

End - Term Examinations – May/June 2026

Date: 14 – 05- 2026 **Time:** 01:00pm – 04:00pm

School: SOCSE	**Program:** B.Tech Information Science & Technology	
Course Code:IST2501	**Course Name:** Optimization methods in Machine Learning	
Semester: VI	**Max Marks**: 100	**Weightage**: 50%

CO - Levels	**CO1**	**CO2**	**CO3**	**CO4**
Marks	**26**	**24**	**26**	**24**

Instructions:

(i) Read all questions carefully and answer accordingly.

(ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 2marks. **10Q x 2M=20M**

1.	Explain the term "Empirical Risk" and state its mathematical expression.	**2 Marks**	**L2**	**CO1**
2.	Illustrate the primary goal of Structural Risk Minimization (SRM) over ERM.	**2 Marks**	**L2**	**CO1**
3.	Differentiate between standard PCA and Sparse PCA regarding model interpretability.	**2 Marks**	**L2**	**CO2**
4.	Outline any four stages of the 8-step Machine Learning Paradigm.	**2 Marks**	**L2**	**CO1**
5.	Describe the role of Cross-Entropy as a loss function in classification.	**2 Marks**	**L2**	**CO2**
6.	Explain about convex set and give one example.	**2 Marks**	**L2**	**CO3**
7.	Explain the Second Order Cone Programming (SOCP).	**2 Marks**	**L2**	**CO3**
8.	Show the constraints in optimization problems?	**2 Marks**	**L2**	**CO3**
9.	Explain the Gradient Descent and write its update rule.	**2 Marks**	**L2**	**CO4**
10.	Classify the learning rate? What happens if it is too large?	**2 Marks**	**L2**	**CO4**

Part B

Answer the Questions. **Total Marks 80M**

11.	**a.**	Apply the Machine Learning Paradigm to describe the 8-step workflow from data representation to model evaluation.	**10 Marks**	**L3**	**CO1**
	b.	Apply Empirical Risk Minimization (ERM) and Structural Risk Minimization (SRM) to compare and contrast them with suitable examples.	**10 Marks**	**L3**	**CO1**
		Or			
12.	**a.**	Apply the concept of Curse of Dimensionality to interpret how Principal Component Analysis (PCA) reduces dimensionality.	**10 Marks**	**L3**	**CO1**
	b.	Apply the concept of VC-Dimension to interpret its significance in providing learning guarantees for a model's performance.	**10 Marks**	**L3**	**CO1**

13.	**a.**	Apply Low Rank Matrix Factorization to explain its significance and utility in building recommendation engines.	**10 Marks**	**L3**	**CO2**
	b.	Apply Support Vector Machines (SVM) to interpret the role of the kernel trick in handling non-linear data separation.	**10 Marks**	**L3**	**CO2**
		Or			
14.	**a.**	Apply Logistic Regression to explain how the sigmoid function maps input features to probabilities.	**10 Marks**	**L3**	**CO2**
	b.	Apply Low Rank Matrix Factorization to explain its importance and applications in modern machine learning models.	**10 Marks**	**L3**	**CO2**

15.	**a.**	Solve the Numerical Problem – KKT Minimize $\boldsymbol{f(x) = x^2}$ Subject to $\boldsymbol{x \geq 1}$ Use Karush-Kuhn-Tucker (KKT) conditions to find the optimal solution.	**10 Marks**	**L3**	**CO3**
	b.	Solve whether the function below is convex using the second derivative test. $f(x) = 3x^2 + 2x + 1$ a) Find the first derivative b) Find the second derivative c) Verify convexity	**10 Marks**	**L3**	**CO3**
		Or			
16.	**a.**	Apply Semi-Definite Programming to explain its concept and applications in kernel learning and Support Vector Machines (SVM).	**10 Marks**	**L3**	**CO3**
	b.	Apply Second Order Cone Optimization to explain its concept and advantages compared to Linear Optimization.	**10 Marks**	**L3**	**CO3**

17.	**a.**	Solve the function: $f(x, y) = x^2 + 2y^2$ Given: Initial point: $(x_0, y_0) = (4, 7)$ Learning rate: $\eta = 0.4$ Solve 3 iterations of Gradient Descent .	**10 Marks**	**L3**	**CO4**
	b.	**Solve the Newton Method** Minimize: $f(x) = x^2 - 7x + 14$ Initial value: $x_0 = 7$ a) Find first and second derivatives b) Solve 4 iteration of Newton's method c) Find updated value of x	**10 Marks**	**L3**	**CO4**
		Or			
18.	**a.**	Apply the Proximal Gradient Method to derive its update rule and solve L1-regularized optimization problems.	**10 Marks**	**L3**	**CO4**
	b.	Apply Coordinate Descent to explain its working with a suitable example and determine when it is preferred over Gradient Descent.	**10 Marks**	**L3**	**CO4**