Roll No. ..... Total Pages : 3

BT–8/M-24 **48426**

OPTIMIZATION METHOD IN ML

Paper–PC-CS-AIML–402A

Time Allowed : 3 Hours] [Maximum Marks : 75

Note : Attempt **five** questions in all, selecting at least **one** question from each Unit. All questions carry equal marks.

UNIT–I

1. (a) Explain the concept of Convexity-preserving Operations. Describe three such Operations and Illustrate each with a simple example. 7½

   (b) Briefly describe Second-order Cone Programming (SOCP) and Semi-definite Programming (SDP). Provide one example for each that highlights their formulation. 7½

2. (a) State the Karush-Kuhn-Tucker (KKT) conditions for a constrained optimization problem. Discuss the significance of these conditions in Convex optimization. 7½

   (b) Define the concept of duality in Convex Optimization. Explain the Primal and Dual problems and describe the Weak and Strong Duality Theorems. 7½

UNIT–II

3. (a) Explain the Basic Gradient Descent Algorithm. Define the Sub-gradient method and explain how it differs from the standard gradient descent method. Provide an example of a function where the Sub-gradient method is applicable. 8

   (b) Describe Nesterov's Accelerated gradient method. Explain how it improves the convergence rate over Standard gradient descent and provide the update equations. 7

4. (a) Discuss the Mirror Descent method. Explain the concept of Bregman divergence and how it is used in the update rule of Mirror descent. 7

   (b) Explain the Proximal Gradient method. Derive the proximal operator and provide an example of its use in optimizing a non-smooth objective function. Discuss Moreau-Yosida regularization. Define the Moreau envelope and proximal mapping, and explain their significance in optimization. 8

UNIT–III

5. (a) Explain the Augmented Lagrangian method for solving constrained optimization problems. Derive the Augmented Lagrangian function and discuss how it modifies the standard Lagrangian approach. 8

(b) Describe the Alternating Direction Method of Multipliers (ADMM). Explain the step-bystep procedure of ADMM and illustrate with a simple example how it decomposes the problem into smaller subproblems. 7

6. (a) Elaborate the Douglas-Rachford Splitting method. Derive the iteration steps involved and explain how this method can be applied to find a zero of the sum of two Monotone operators ? 7

(b) Discuss the concepts of Primal and Dual decomposition in the context of optimization. Explain how these decompositions help in solving large-scale optimization problems and provide an example where they are particularly useful. 8

UNIT–IV

7. (a) Elaborate the concept of dual averaging in Stochastic optimization. Derive the update rule for dual averaging and discuss its convergence properties. 7½

(b) Describe the Polyak-Juditsky averaging method. How does it differ from traditional Stochastic Gradient Descent (SGD), and what are its advantages? 7½

8. (a) Elaborate the Stochastic Variance Reduced Gradient (SVRG) method. Explain how SVRG improves the convergence rate of SGD and outline the key steps in the SVRG algorithm. 7½

(b) Discuss the application of Non-convex optimization methods in deep learning. 7½