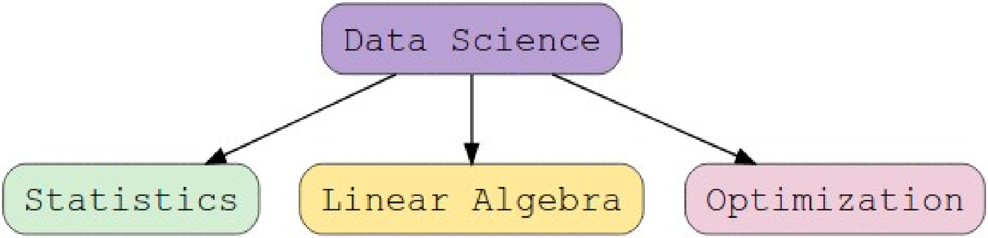
UNIT II	Optimization for Data Science :	06 - Hrs
	Optimization for Data Science, Unconstrained Multivariate Optimization Gradient (Steepest) Descent (OR) Learning Rule, Multivariate Optimization With Equality Constraints, Solving Data Analysis Problems.	

What is optimization for data science?

- Data science is a field used to analyze the vast volume of data through different techniques to make it understandable. To understand data science, we need to have a good understanding of these three concepts: statistics, linear algebra, and optimization.
- As a data scientist, you spend a lot of your time helping to make better decisions. You build predictive models to provide improved insights. You might be predicting whether an image is a cat or dog, store sales for the next month, or the likelihood if a part will fail.



Optimization

- **Optimization** is a procedure or technique which is used to find the most efficient resolution. Its value may be minimum or maximum, which depends upon the requirement.
- **For example,** if a firm is required to find a way to earn a maximum profit on their products, then the condition will be maximum, and if they want to find a way through which their firm gets minimum product cost on production, then the requirement will be minimum.

Usage of optimization

- Optimization works as a backbone for almost all techniques used in data science. It also helps the business leaders to predict their business plans, and it allows the government to find an appropriate resolution for human welfare by analyzing data they provide suitably optimization helps in numerous other sectors like economic, artificial intelligence, and many more.
- **So, let’s understand what the optimization problem is and its types.**

  1. Elements of the optimization problem are:

  2. The objective function

  3. Decision variables

  4. Constraints

The Objective Function

- **The objective function** is the first element of an optimization problem if the function f(y) is used to find the minimize or maximize value. As a general rule, we talk about minimization issues, this is essential since, in case you have a boost issue with f(y) , we can change it over to a minimization issue with - f(y). Thus, without loss of consensus, we can check.

Decision Variable and Continuous Variable

- **Decision variable** is the second element. The decision variable can observe the ideal variable to solve the optimization problem. Which can only be done by decision variable. $F(y)=3x+2$ in this y is a decision variable. There are three types of variables: continuous, integer, and mixed.
- **Continuous variable** should be ongoing, assuming if the y variable can take a limitless number of values. For this situation, y can handle an unlimited number of values between -3 to 5.

  $-f(y), y \in (-3,5)$

Types of optimization

- **Linear programming:** If the optimization element decision variable remains continuous and both of the remaining elements become linear, it will become a linear programming problem.
- **Non-linear programming:** If the optimization element decision variable remains continuous and either one of the remaining elements becomes non-linear, it will become a non-linear programming problem.
- **Linear integer programming:** If the object function and constraints are linear and the decision variable y is an integer, then it will be called a linear integer programming problem.
- **Non-linear programming:** If the object function and constraints are non-linear and the decision variable is an integer, then it will be called a non-linear integer programming problem.

Types of optimization

- **Binary integer programming:** If the objective function and constraints are linear and decision variable y can only have binary integer {0.1} options, then it will be called a binary integer programming problem.
- **Mixed-integer linear programming:** If the objective function f and the constraints are both linear, and decision variable y is a mixed variable that can be continuous or integer, it will be called a mixed-integer programming problem.
- **Mixed-integer non-linear programming:** If one of the objective function f or the constraints are non-linear, and the decision variable y is a mixed variable that can be continuous or integer, it will be called a mixed-integer non-linear programming problem.

Unconstrained Multivariable Optimization

- This is the model for the unconstrained general non-linear programming problem, where the objective function is **f X.** And the function involves n number of decision variables and we want to minimize the function **f of X.** Similarly, we can maximize the function as well.
- Now, this is the general form of unconstrained optimization problem. Now, our aim is to find out the values for **x 1, x 2, x n,** which satisfy the restrictions – means here only the range of the decision variables are given to us; there is no other constraint as such; and we want to minimize the function f here.