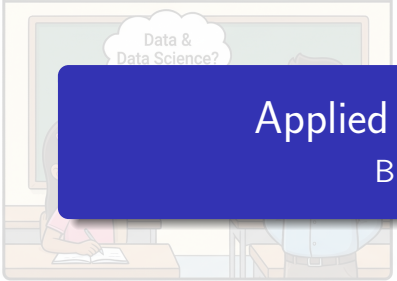
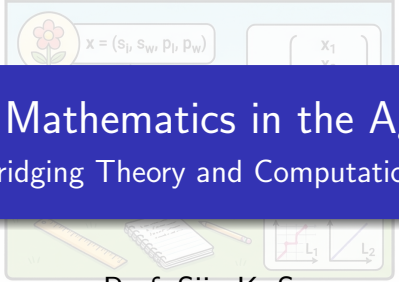
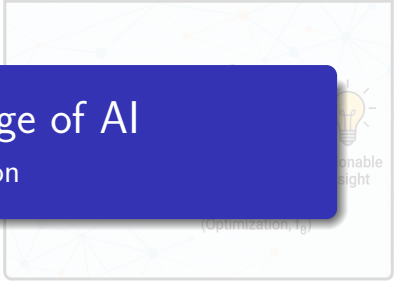


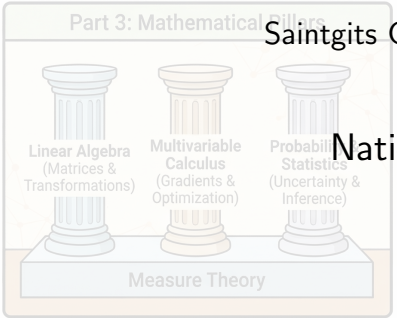
# Applied Mathematics in the Age of AI

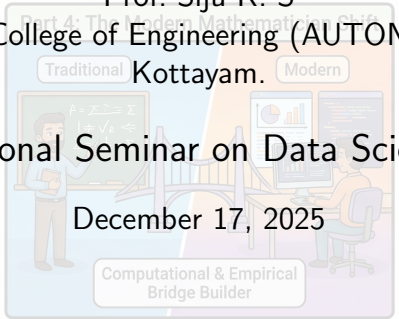
## Bridging Theory and Computation

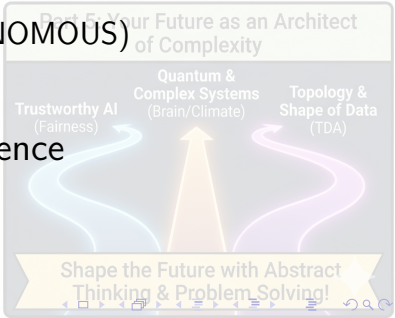
Prof. Siju K. S
Saintgits College of Engineering (AUTONOMOUS)
Kottayam.

National Seminar on Data Science

December 17, 2025



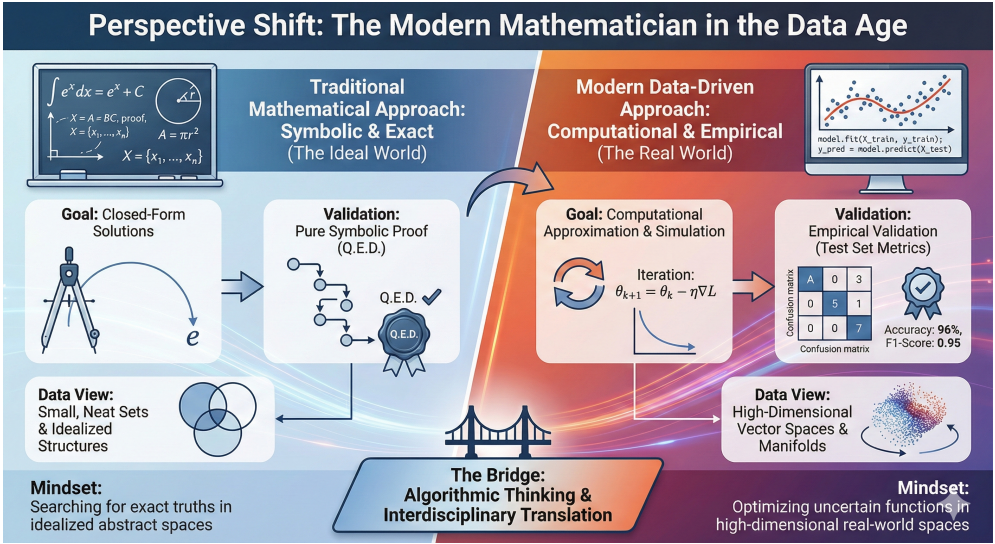



# Session Overview

1. Data Foundations & EDA
2. Mathematics of Machine Learning
3. Future Outlook

# The Modern Mathematician



# Mathematical Formalism of Data

**Set-Theoretic Definition:** Data can be considered as elements of a set:

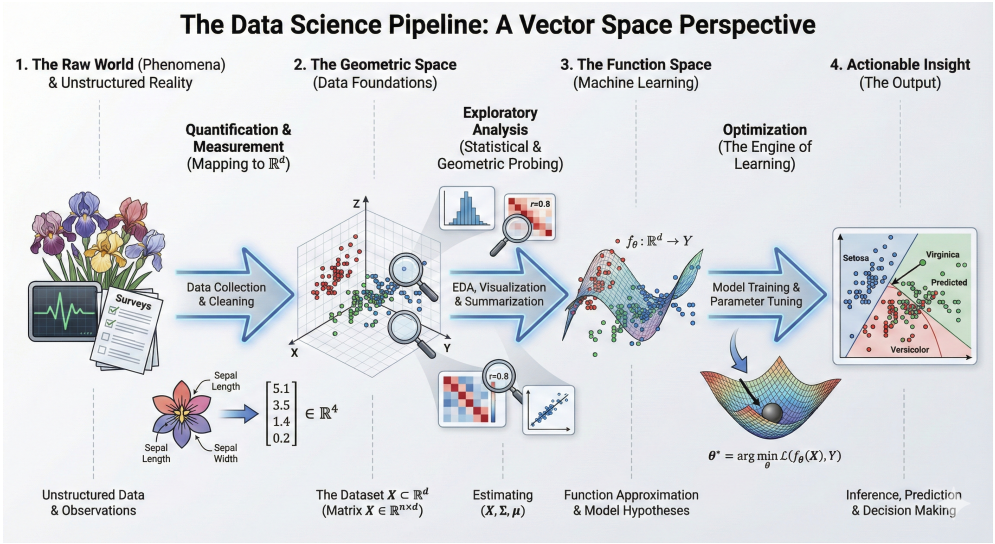
$$X = \{x_1, x_2, \ldots, x_n\} \subset \mathbb{R}^d$$

where each $x_i$ is a $d$-dimensional vector representing measurable features.

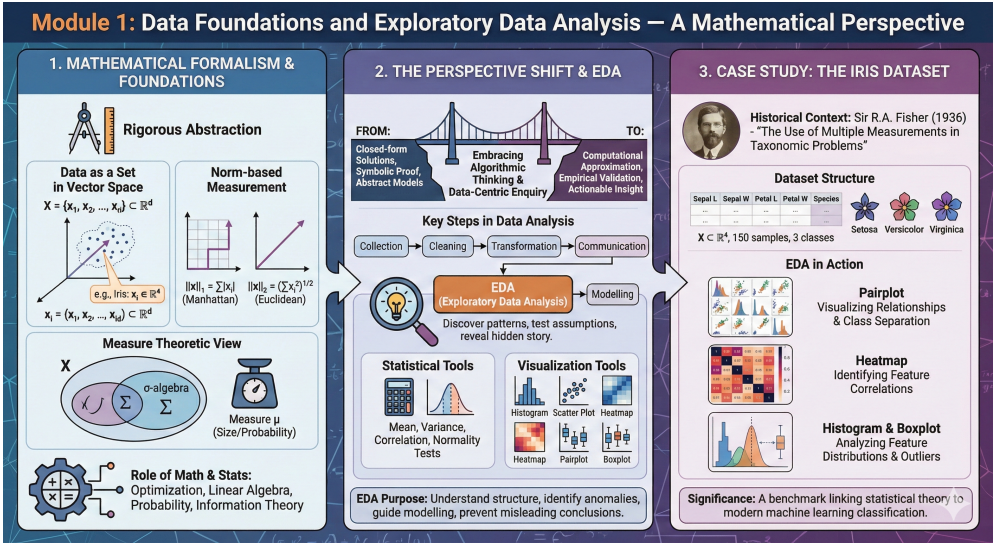
**Measure Theoretic View:** We view the data space as a triple $(X, \Sigma, \mu)$, where:

- $X$: The sample space.
- $\Sigma$: The $\sigma$-algebra of measurable subsets.
- $\mu$: The measure assigning probability/size.

# The Data Science Pipeline



# Exploratory Data Analysis (EDA)



# Metrics and Norms in Data Analysis

To analyze similarity (clustering) or error (regression), we require a metric space structure.

**Norm-based Measurements:**

The $L_1$ Norm (Manhattan - useful for sparse data):

$$\|x\|_1 = \sum_{i=1}^{d} |x_i|$$

The $L_2$ Norm (Euclidean - standard geometric distance):

$$\|x\|_2 = \left( \sum_{i=1}^{d} x_i^2 \right)^{1/2}$$

# Machine Learning: An Optimization Problem

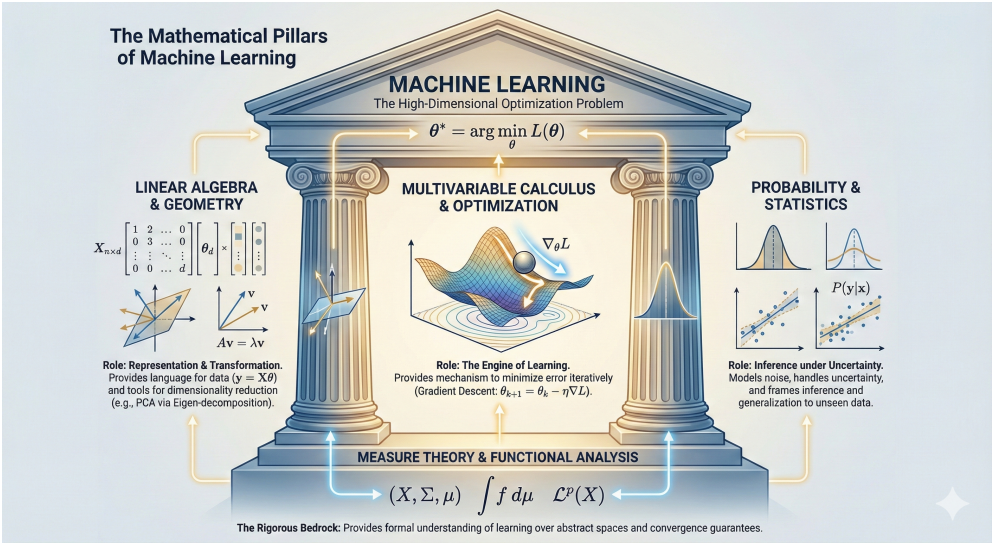
Given inputs $X$ and outputs $Y$, we seek a function $f_\theta : \mathbb{R}^d \to Y$.
The goal is to find parameters $\theta^*$ that minimize a Loss Function $L$:

$$\theta^* = \arg\min_\theta L(f_\theta(X), Y)$$

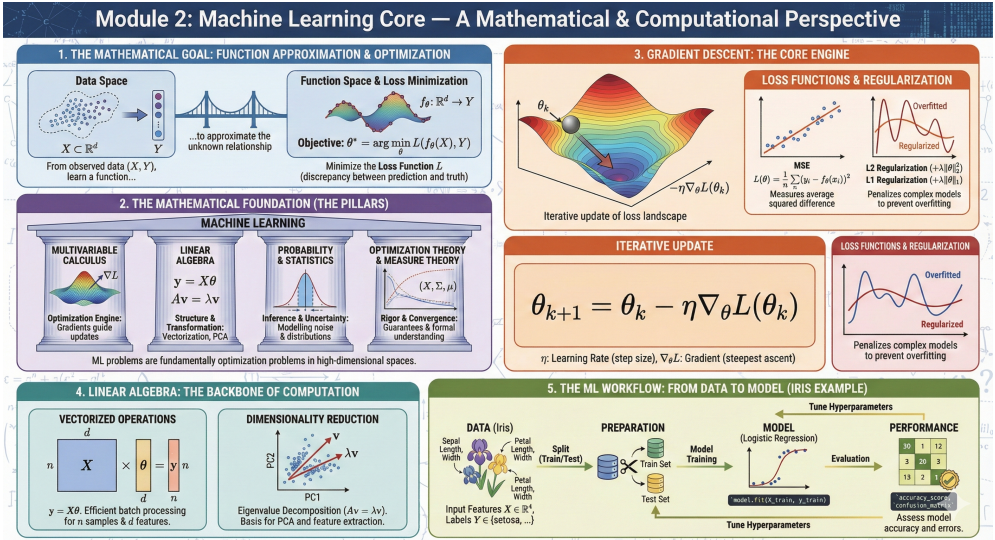
**Key Insight**

ML is fundamentally an optimization problem in a high-dimensional vector space, constrained by computational resources and statistical uncertainty.

# The Mathematical Pillars of Machine Learning



# Machine Learning Core



# The Engine: Gradient Descent

To solve $\min_\theta L(\theta)$, we use iterative updates based on Multivariable Calculus.

**Update Rule:**

$$\theta_{k+1} = \theta_k - \eta \nabla_\theta L(\theta_k)$$

Where:

- $\eta$: Learning rate (step size).
- $\nabla_\theta L$: The gradient vector (direction of steepest ascent).

**Regularization ($L_2$ - Ridge):** Prevents overfitting by penalizing large norms:

$$L(\theta) = \frac{1}{n} \sum (y_i - f_\theta(x_i))^2 + \lambda \|\theta\|_2^2$$

# Linear Algebra: The Computational Backbone

**Vectorization:**

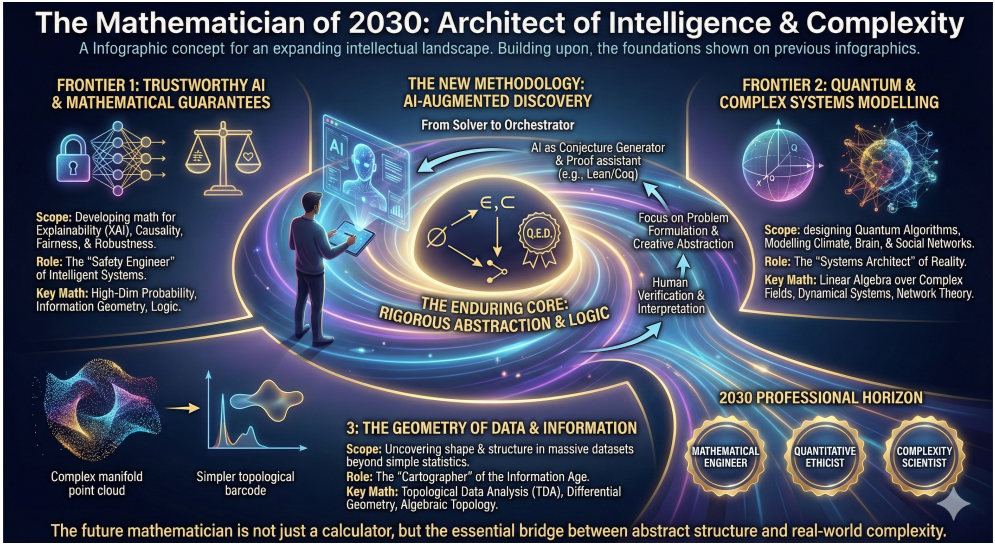
$$y = X\theta$$

$$X \in \mathbb{R}^{n \times d}, \quad \theta \in \mathbb{R}^{d}$$

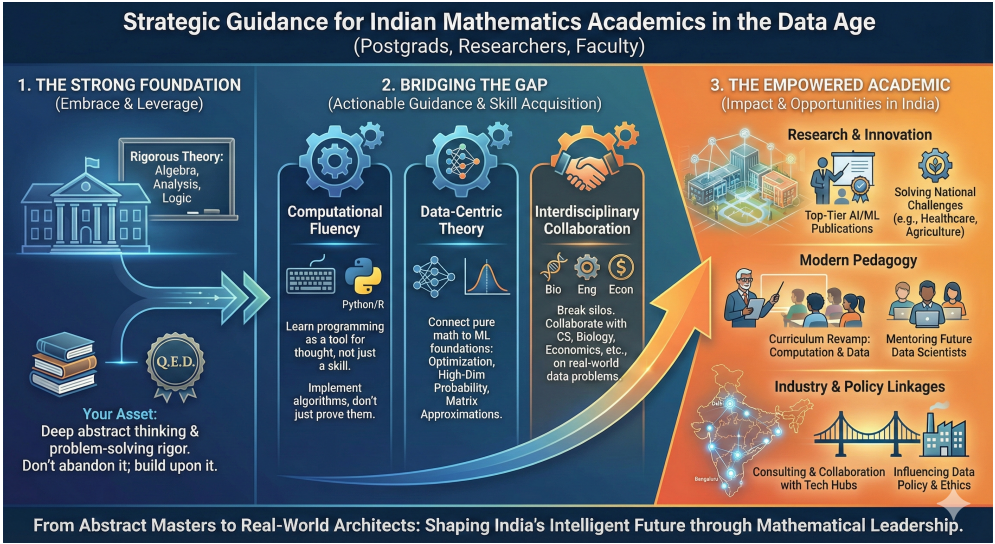
**Dimensionality Reduction:**
Eigenvalue Decomposition helps us find principal components:

$$Av = \lambda v$$

# The Future: Mathematics in 2030



# Strategic Guidance for Indian Mathematics Academics



# Summary & Discussion

**To contribute effectively, we must:**

1. Move from closed-form solutions to computational approximations.
2. Accept empirical validation alongside symbolic proof.
3. Embrace interdisciplinary collaboration.

## Thank You

# Clear Vision