

GRADIENT DESCENT:

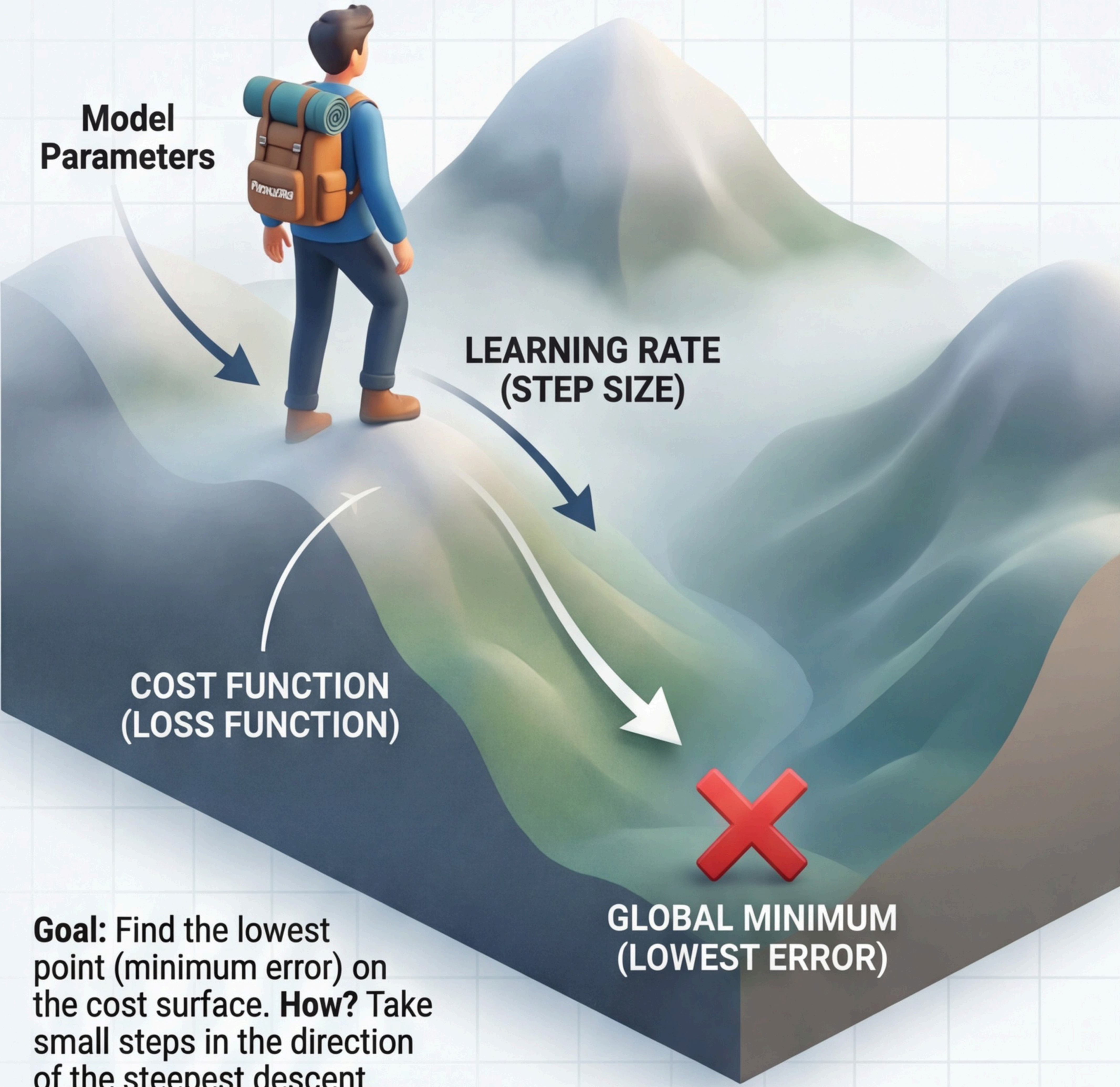
The Engine of Machine Learning Optimization



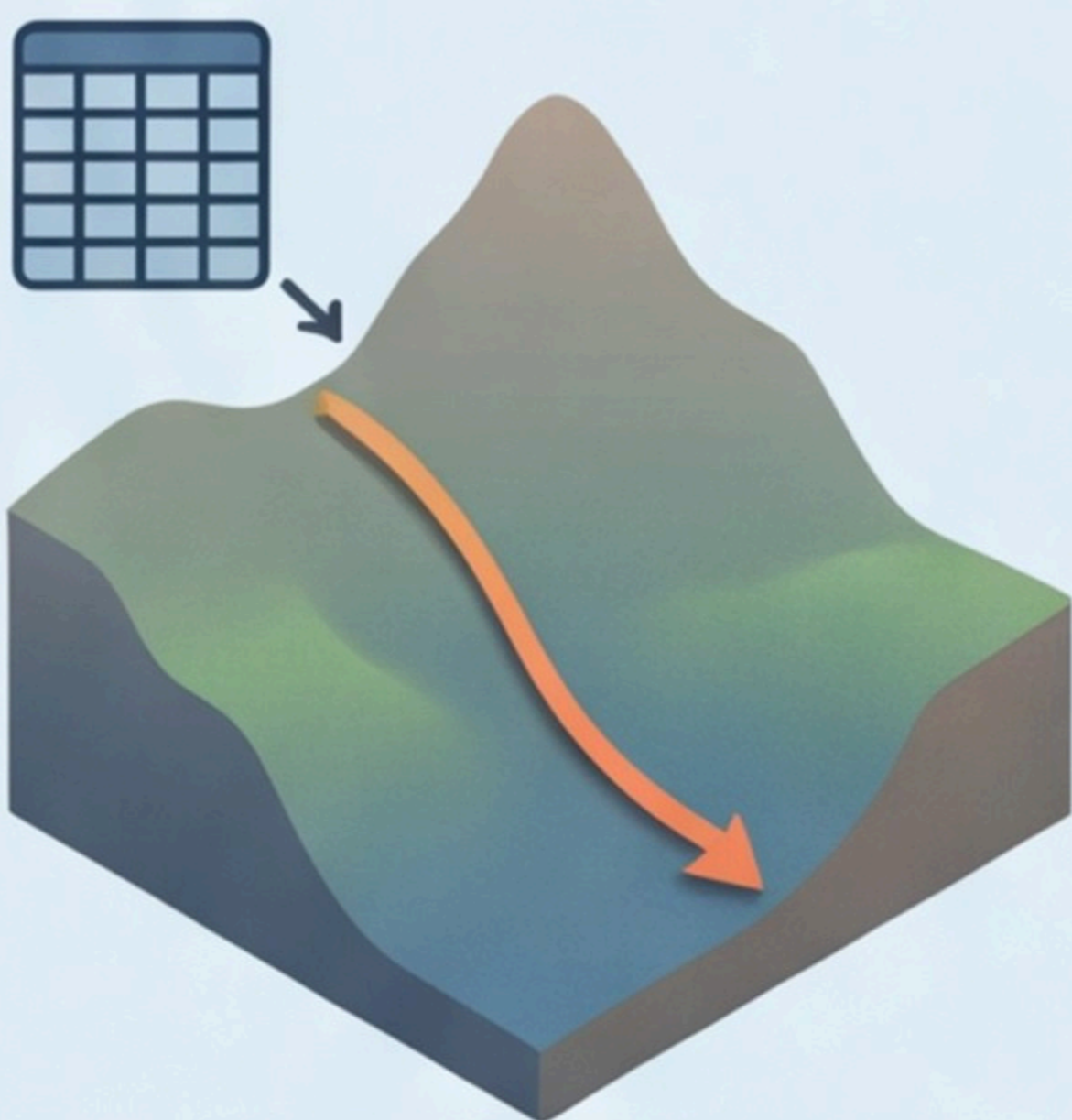
WHAT IS GRADIENT DESCENT?

An iterative optimization algorithm used to minimize a cost function by adjusting model parameters.

THE ANALOGY: DESCENDING THE MOUNTAIN OF ERROR

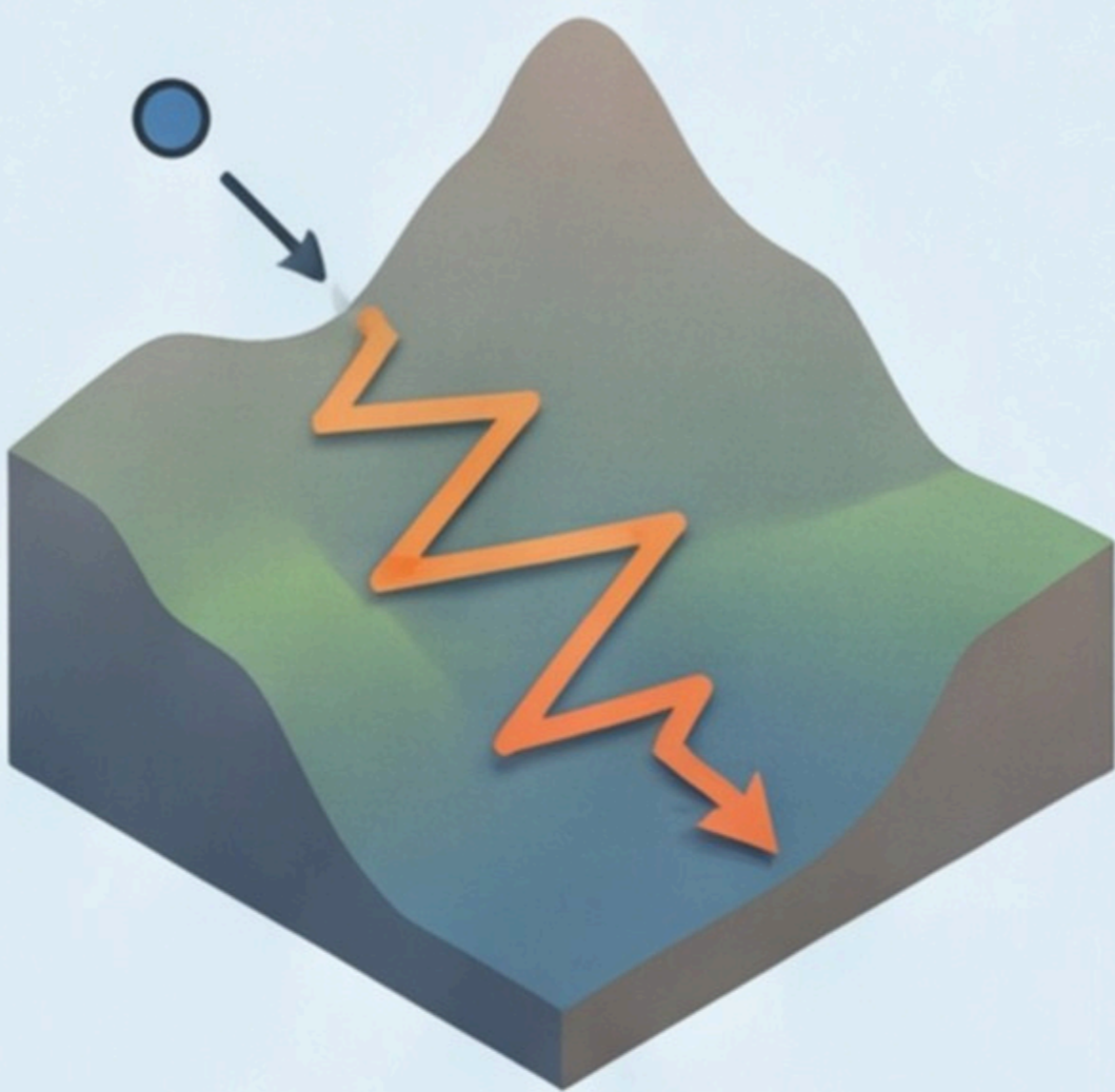


TYPES OF GRADIENT DESCENT



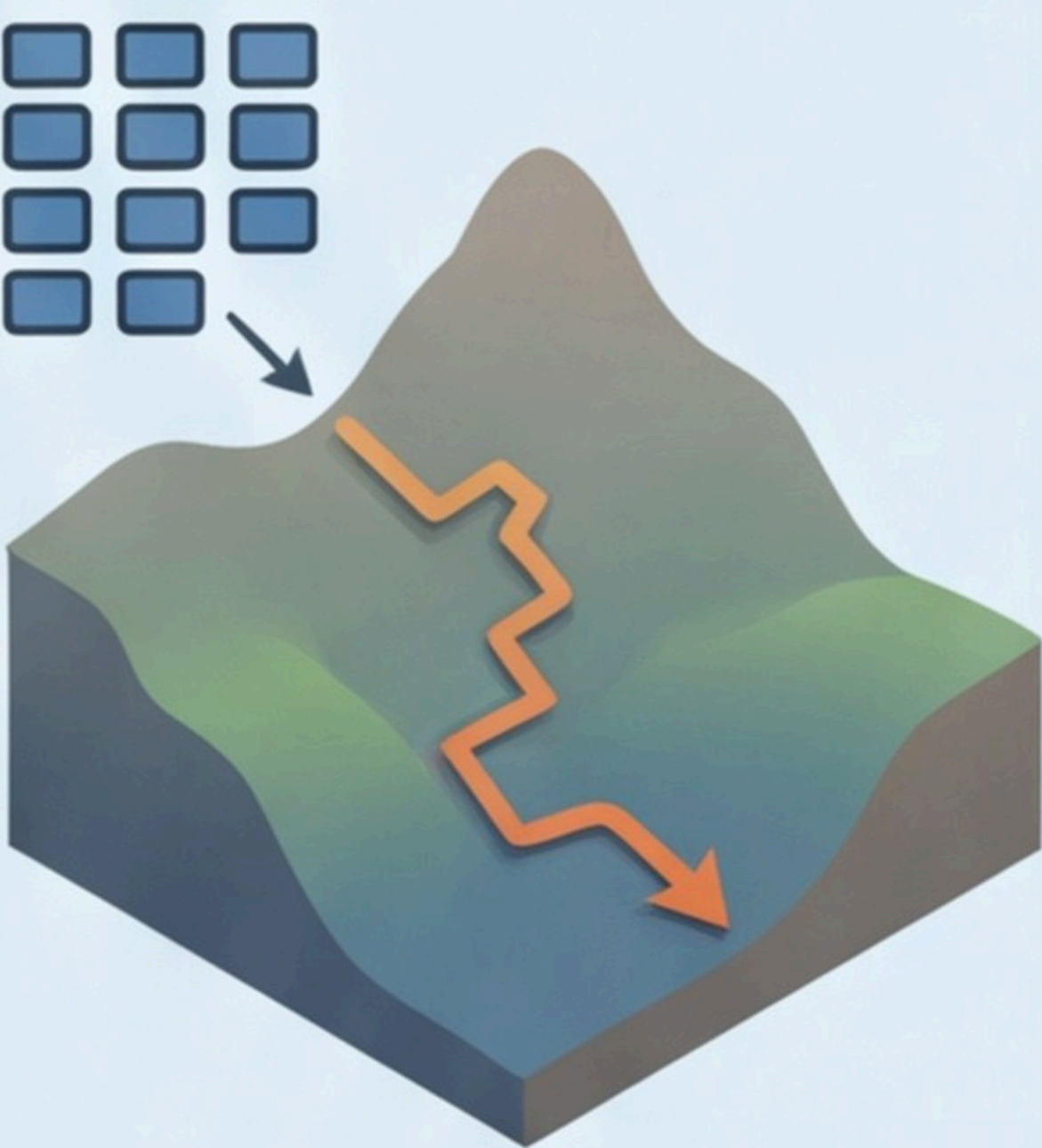
BATCH GD

Uses entire dataset for each step. Slow but stable.



STOCHASTIC GD (SGD)

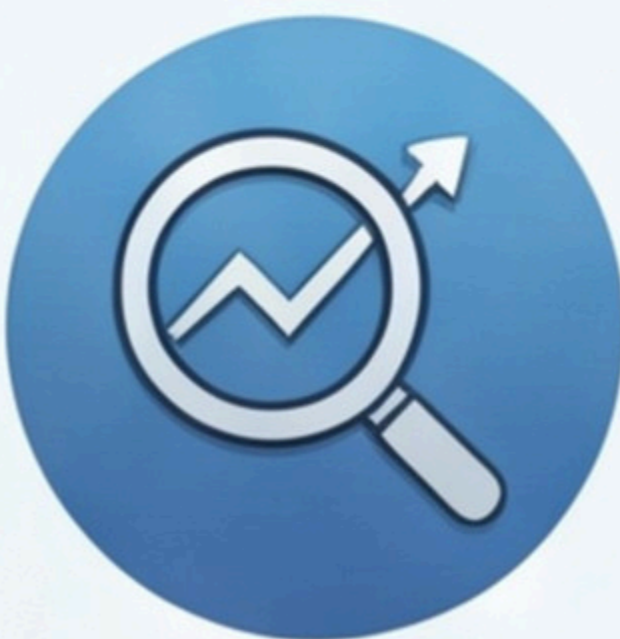
Uses one data point per step. Fast but noisy.



MINI-BATCH GD

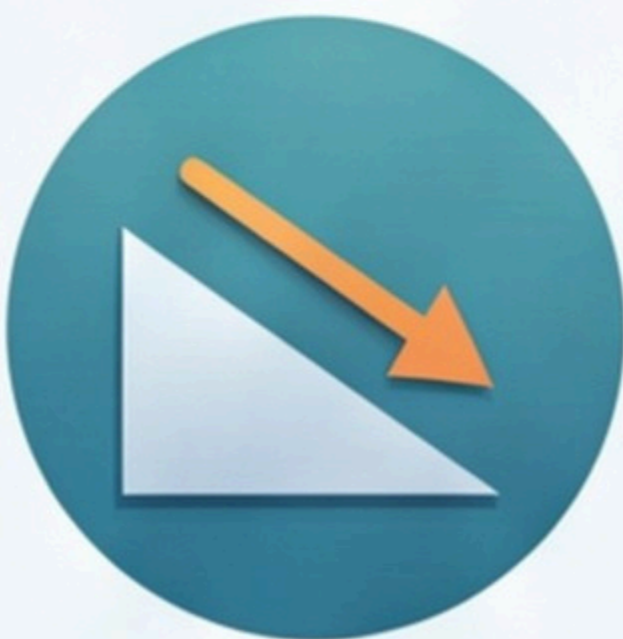
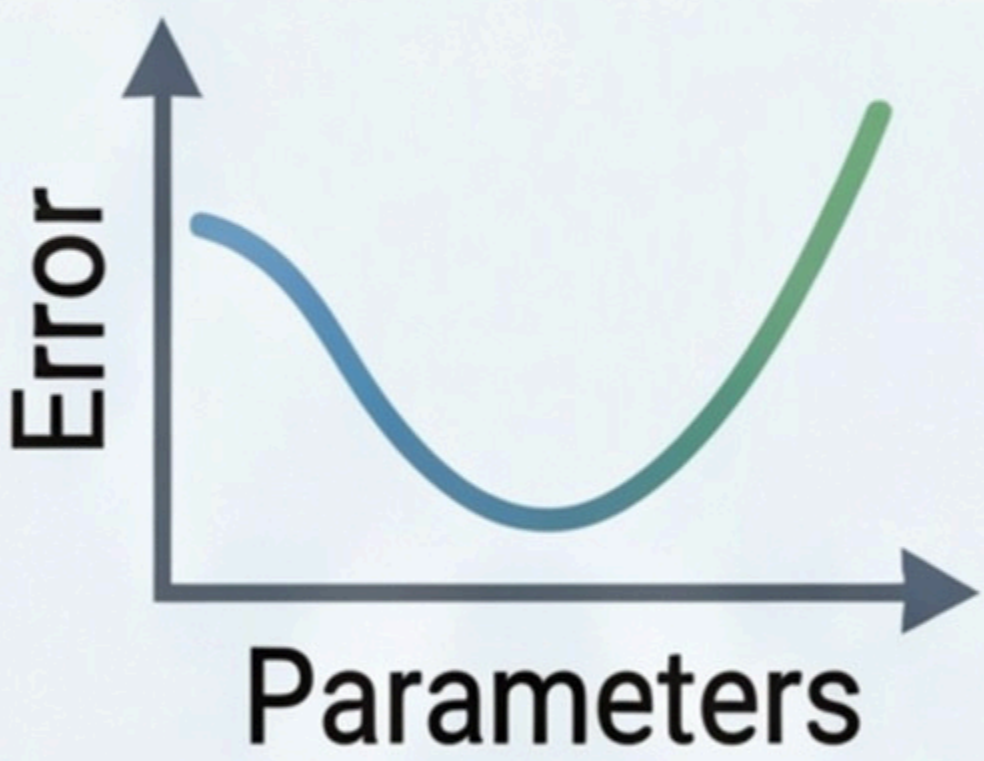
Uses a small subset of data. Balance of speed and stability.

KEY COMPONENTS & PROCESS



1. COST FUNCTION

Measures how wrong the model's predictions are. We want to minimize this.



2. GRADIENT

Calculate the slope (derivative) at the current position. Points uphill.



3. UPDATE PARAMETERS

Move parameters in the opposite direction of the gradient, multiplied by the learning rate.

$$\theta_{\text{new}} = \theta_{\text{old}} - \alpha * \nabla J(\theta)$$

KEY TAKEAWAYS

- ✓ Crucial for training Neural Networks & many ML models.
- ✓ Learning rate is a critical hyperparameter.
- ✓ Goal is to find optimal weights & biases.

