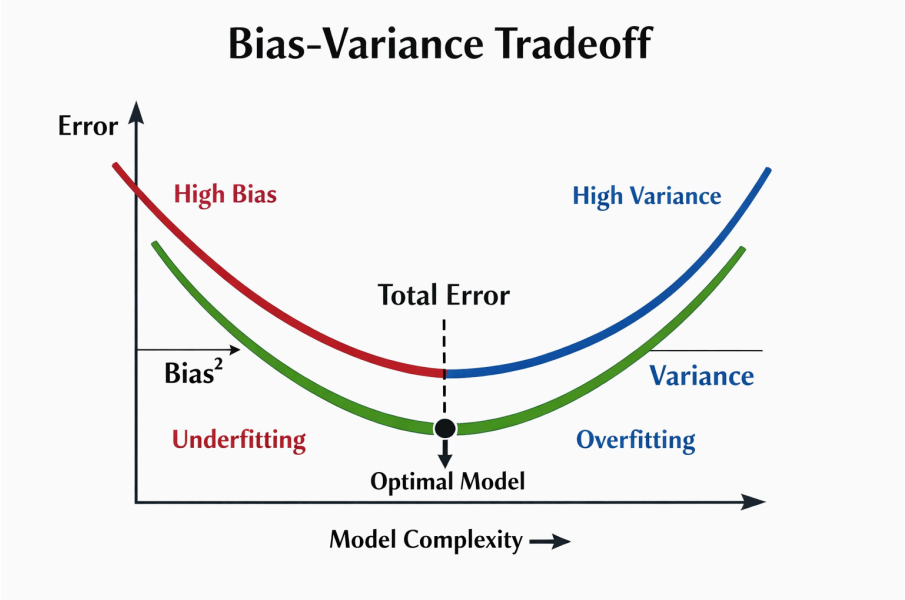
Machine Learning- Class 3

1.What is the Bias-Variance Tradeoff?

The bias–variance tradeoff describes the balance between two types of errors that affect model performance.

- **Bias** refers to error caused by overly simplistic assumptions in the model. High bias leads to underfitting.
- **Variance** refers to error caused by excessive sensitivity to training data. High variance leads to overfitting.

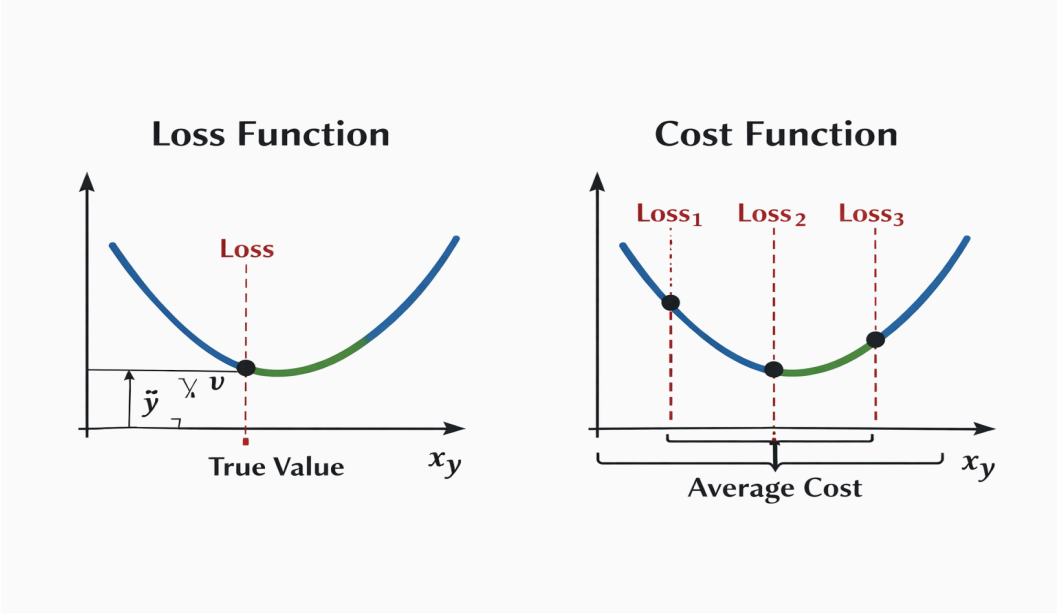
A good model finds the right balance between bias and variance to generalize well on unseen data.



2. What is the difference between a loss function and a cost function?

- A **loss function** measures the error for a single data point by comparing the predicted output with the actual output.
- A **cost function** is the average (or total) of loss values over the entire dataset.

Key Difference:
Loss function evaluates individual predictions, while cost function evaluates overall model performance.

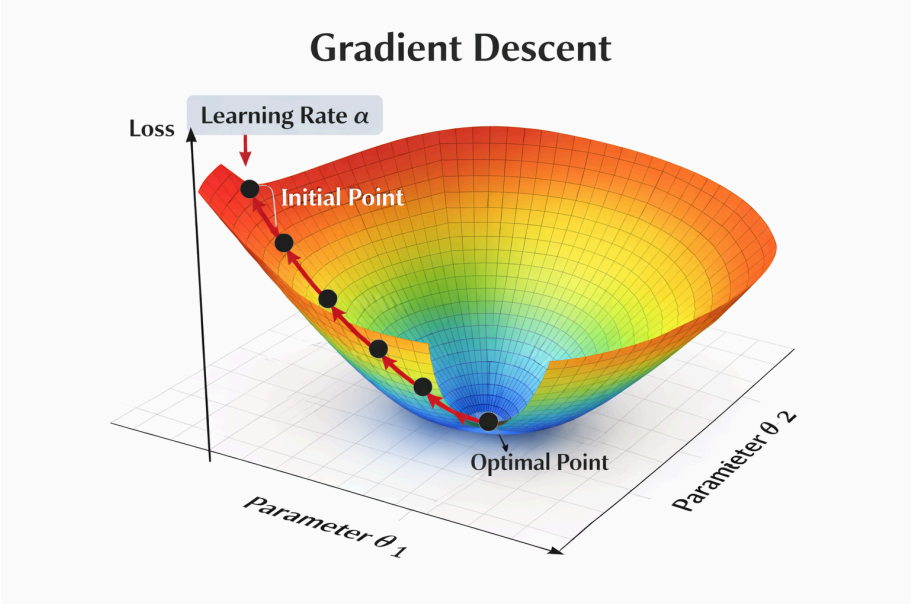


3. What is gradient descent?

Gradient descent is an optimization algorithm used to minimize the cost function of a machine learning model. It works by iteratively adjusting model parameters in the direction that reduces error the most.

Key Characteristics:

- Moves step by step toward minimum error
- Uses learning rate to control step size
- Widely used in regression and neural networks



Main Components

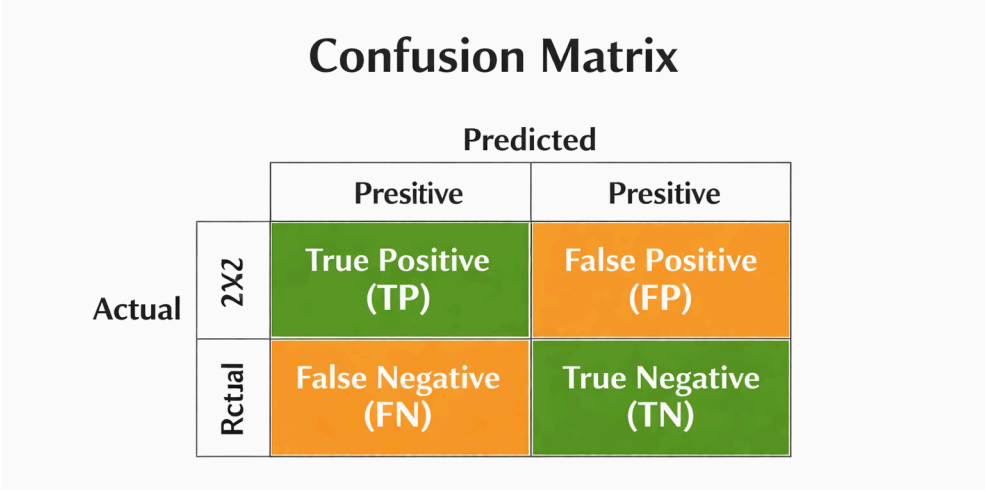
- **Loss Surface**: 3D surface showing how error changes with model parameters.
- **Parameters (θ_1, θ_2)**: Model weights updated during training.
- **Loss (Z-axis)**: Value of the error to be minimized.
- **Initial Point**: Starting values of parameters.
- **Gradient**: Direction of steepest increase in loss; descent moves opposite to it.
- **Learning Rate (α)**: Controls the step size of parameter updates.
- **Descent Path**: Iterative steps taken to reduce loss.
- **Optimal Point**: Point of minimum loss (best model).

4. What is a confusion matrix?

A confusion matrix is a performance evaluation tool used for classification models. It summarizes prediction results by comparing actual values with predicted values.

It Helps To:

- Measure accuracy, precision, recall, and F1-score
- Identify types of classification errors
- Evaluate model performance beyond accuracy



5. What are L1 (Lasso) and L2 (Ridge) Regularization?

Regularization techniques prevent overfitting by adding a penalty term to the cost function.

- **L1 (Lasso) Regularization:**
  Encourages sparsity by shrinking some coefficients to zero, effectively performing feature selection.

- **L2 (Ridge) Regularization:**
  Penalizes large coefficients but does not eliminate features.

Key Difference:
Lasso selects features, Ridge stabilizes coefficients.

6. What is the "Curse of Dimensionality"?

The curse of dimensionality refers to problems that arise when the number of features increases excessively. As dimensions grow, data becomes sparse, and models require more data to learn effectively.

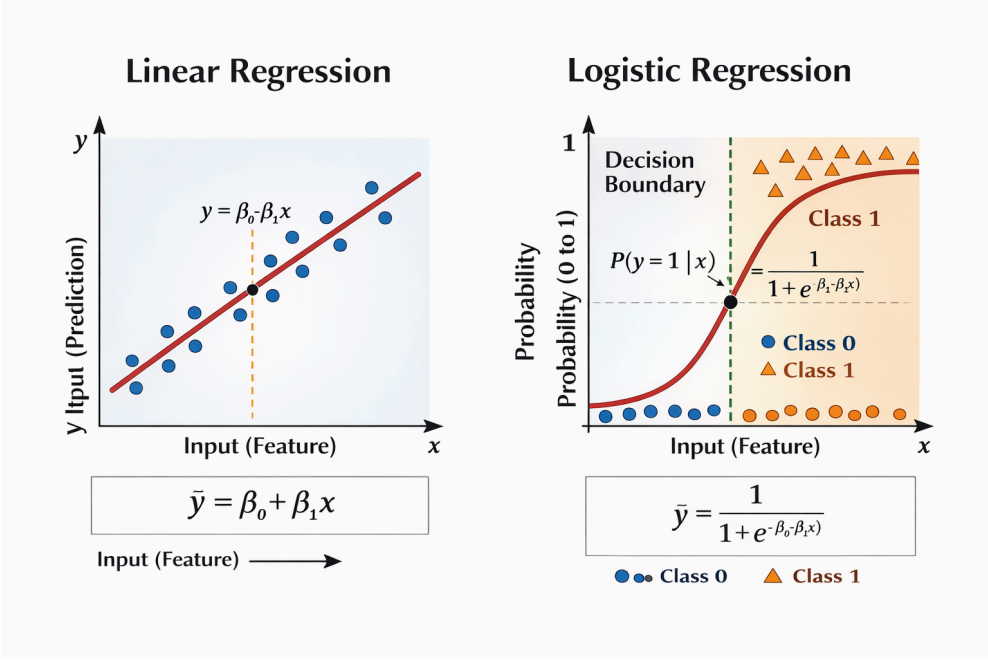
Effects:

- Increased computational cost
- Reduced model performance
- Difficulty in identifying meaningful patterns

7. How does Logistic Regression differ from Linear Regression?

- **Linear Regression** predicts continuous numerical values.
- **Logistic Regression** predicts probabilities and is used for classification tasks.

Key Difference:
Linear regression outputs real values, while logistic regression outputs probabilities between 0 and 1.



8. What is the difference between Parameters and Hyperparameters?

- **Parameters** are learned automatically by the model during training.
- **Hyperparameters** are set before training and control the learning process.

Examples:

- Parameters: weights, coefficients
- Hyperparameters: learning rate, number of layers, regularization strength

9. What is the difference between feature engineering and feature selection?

- **Feature Engineering** involves creating new features from existing data to improve model learning.

- **Feature Selection** involves choosing the most relevant features and removing irrelevant ones.

Key Difference:
Engineering creates features; selection reduces features.

10. What are the key assumptions of linear regression?

Linear regression relies on several assumptions for reliable results:

1. **Linearity:** Relationship between input and output is linear
2. **Independence:** Observations are independent of each other
3. **Homoscedasticity:** Constant variance of errors
4. **Normal Distribution of Errors:** Residuals follow a normal distribution
5. **No Multicollinearity:** Features are not highly correlated