**What is Linear Regression?**

Linear Regression is a supervised machine learning algorithm used to predict continuous values (real numbers) based on the linear relationship between input features (independent variables) and the output (dependent variable).

It fits a straight line (or hyperplane) through the data, defined by:

Where:

* y: Predicted output (continuous)
* x1​,x2​,...,xn: Input features
* β0: Intercept
* β1​,β2​,...,βn​: Coefficients
* ϵ: Error term (residuals)

**Types:**

* **Simple Linear Regression**: One input variable.
* **Multiple Linear Regression**: Multiple input variables.

“Is the loss function in linear regression used to find the best-fitting line by minimizing the error between predicted and actual values?”

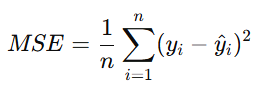
✅ Answer: Yes.

In more detail:

* The loss function — most commonly Mean Squared Error (MSE) — is what the linear regression model tries to minimize during training. This function quantifies how far off the model’s predictions (ŷ) are from the actual values (y).

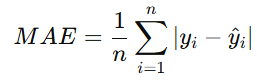
**Common Loss Functions for Linear Regression:**

1. **Mean Squared Error (MSE):**



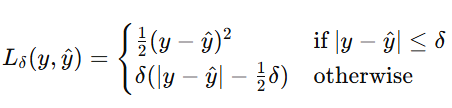
* Penalizes large errors
* Sensitive to outliers
* Most common loss function for linear regression

1. **Mean Absolute Error (MAE)**

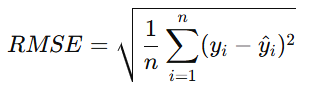


* Less sensitive to outliers than MSE
* Linear penalty on errors

**3. Huber Loss**

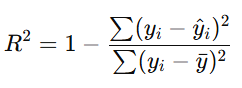


* Combines MSE and MAE
* Robust to outliers, smooth to optimize

**4. Root Mean Squared Error (RMSE)**

* Same unit as target variable
* Easy to interpret

**5. R-Squared (R²) – Evaluation Metric**

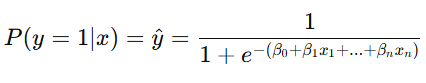


* Measures proportion of variance explained by the model
* R2=1R^2 = 1R2=1: Perfect prediction
* R2=0R^2 = 0R2=0: Predicts only mean

**What is Logistic Regression?**

Logistic Regression is a **supervised classification algorithm** used to **predict probabilities** for binary or multi-class outcomes. It models the **probability** that an instance belongs to a class using the **sigmoid (logistic) function**.

The output is between 0 and 1:



Where:

* y^​: Probability of class 1
* Output can be thresholded (e.g., 0.5) to classify

**Types:**

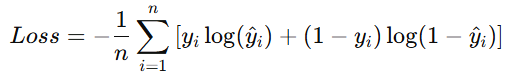
* **Binary Logistic Regression**: Two classes (0 or 1)
* **Multinomial Logistic Regression**: More than two classes
* **Ordinal Logistic Regression**: Ordered categories

**Goal**:

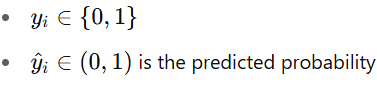
To maximize the likelihood (or equivalently, minimize the log loss) that the model correctly predicts the class.

**Common Loss Function for Logistic Regression:**

1. **Binary Cross-Entropy (Log Loss)**



Where:

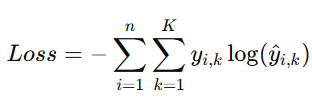


**Explanation**:

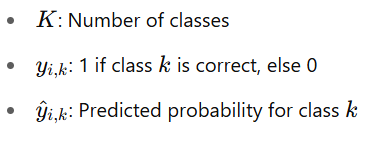
* Strongly penalizes confident but wrong predictions
* Best suited for **binary classification**

1. **Categorical Cross-Entropy (Multiclass)**

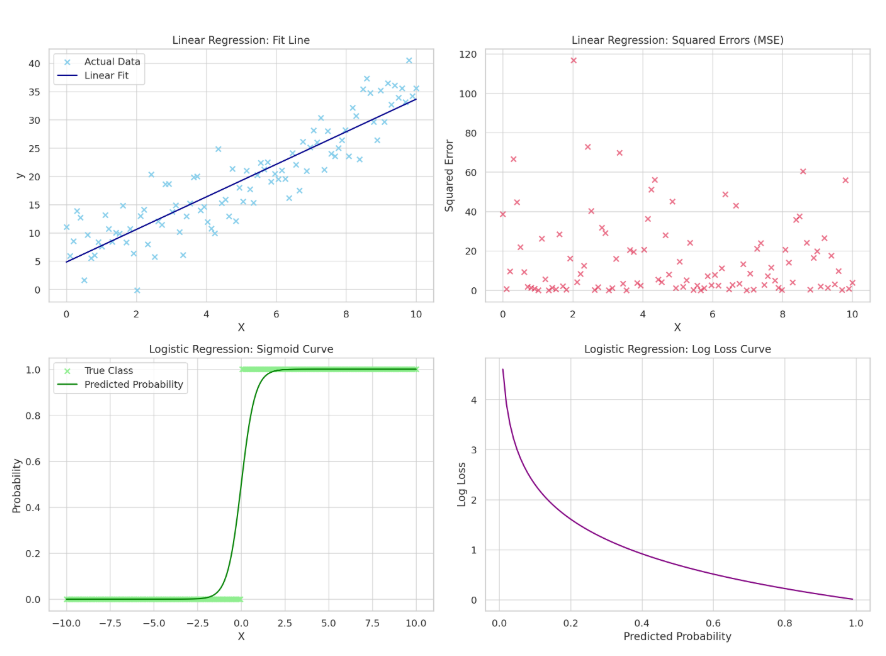
For multi-class classification:



Where:



* Here’s a visual breakdown of how loss functions behave in linear and logistic regression:



🔹 Top Left (Linear Regression – Fit Line):

* Shows the actual data points (blue dots) and the best-fit line learned by linear regression (dark blue).
* The goal is to minimize the distance between the line and the actual points.

🔹 Top Right (Linear Regression – Squared Errors):

* Displays the squared errors for each data point.
* These are the values used in the Mean Squared Error (MSE) loss function.
* Larger vertical errors (residuals) are penalized more heavily.

🔹 Bottom Left (Logistic Regression – Sigmoid Curve):

* Logistic regression maps inputs (X) to probabilities using the sigmoid function.
* The green curve shows the model’s predicted probability of class 1 (positive class).
* Actual class labels are shown as dots.

🔹 Bottom Right (Logistic Regression – Log Loss Curve):

* This shows how the log loss behaves for a positive class (label = 1).
* If the predicted probability is close to 1, the loss is small.
* If it’s close to 0 (wrong prediction), the log loss is very high.