Type	Description	Examples
1 Supervised Learning	Model learns from labeled data (data with known outputs)	Regression, Classification
2 Unsupervised Learning	Model learns from unlabeled data (no known outputs)	Clustering, Dimensionality Reduction

---

Inside Supervised Learning, we have:

Subtype	Output Type	Examples
Regression	Continuous output	Linear Regression, Polynomial Regression
Classification	Categorical output	Logistic Regression, Decision Tree, Random Forest, SVM, etc.

```
Artificial Intelligence
 └── Machine Learning
      ├── Supervised Learning
      │    ├── Regression
      │    └── Classification
      └── Unsupervised Learning
           ├── Clustering
           └── Dimensionality Reduction
```

Linear Regression

Linear Regression is a method used to find a linear relationship between independent variable(s) (X) and a dependent variable (Y).

Formula
Y = mx + c

🧠 If there are multiple variables (multiple linear regression):

$$y = b_0 + b_1x_1 + b_2x_2 + ... + b_nx_n$$

Symbol	**Meaning**	**Analogy**
Y	Dependent variable (the value you're predicting)	e.g., Life Expectancy
X	Independent variable (the input)	e.g., GDP
b_0	Intercept → value of Y when X = 0	Like "c" in y = mx + c
b_1	Slope → how much Y changes when X increases by 1 unit	Like "m" in y = mx + c
ε	Error term → difference between actual and predicted Y	The leftover or noise

B0 = Intercept(c)
B1 = the slope(m)

Type	Description
Simple Linear Regression	One independent variable
Multiple Linear Regression	More than one independent variable
Polynomial Linear Regression	Uses higher powers of X (like X^2, X^3) to model curved relationships

m = coefficient(s) → how much Y changes when X changes

c = intercept → value of Y when X = 0

MSE (Mean Squared Error): Measures average of squared prediction errors — lower is better.

RMSE (Root Mean Squared Error): Average error in actual units — shows how far predictions are from real values.

MAE (Mean Absolute Error): Average absolute difference between predicted and actual values.

R^2 (R-Square): Shows how much of the variation in the target is explained by the model.

Adjusted R^2: R^2 adjusted for number of predictors — penalizes adding useless variables.

Adjusted R square = $1 - (1-R2) * (N-1) / N - p -1$

Where:

Symbol	Meaning
R^2	Normal R-squared value
N	Total number of observations (rows)
p	Number of independent variables (features/predictors)

□ Why we need Adjusted R^2

You said it perfectly —
R^2 always increases when you add more independent variables,
even if those variables are **not actually useful** in prediction.

□ So, R^2 can **mislead you** — it'll look like your model is improving,
but in reality, the model might just be becoming more complex and overfitted.

□ How Adjusted R^2 fixes that

- Adjusted R^2 **penalizes unnecessary variables** (via the term *p*).
- If you add a new variable that **actually helps**, Adjusted R^2 will **increase**.
- If you add a variable that **doesn't help**, Adjusted R^2 will **decrease**.

This makes it a **better, fairer measure** of model performance when multiple predictors are used.

Variable	R^2	Adjusted R^2
1 Predictor	0.75	0.74
3 Predictors	0.80	0.77
6 Predictors	0.82	0.76

□ Notice: even though R^2 keeps rising,
Adjusted R^2 starts dropping because not all predictors are useful.

□ In short:

Metric	Meaning	Problem / Solution
R^2	% of variation in Y explained by X	Always increases with new variables
Adjusted R^2	Penalizes unnecessary predictors	Only increases if the new variable adds real value

□ Case 1 – Few useful variables

Model	Independent Variables (p)	What they represent	R^2	Adjusted R^2

Model 1	1 → `engine_size`	Bigger engine → higher price	0.70	0.69
Model 2	2 → `engine_size`, `mileage`	Mileage also affects price	0.82	0.81
Model 3	3 → `engine_size`, `mileage`, `brand_rating`	Brand also important	0.88	0.87

✅Here every new p (variable) **adds real information**,
so **R^2** ↑ and **Adjusted R^2** ↑.
The model truly got better.

Case 2 – Adding useless variables

Now you start adding random columns like the color of the dashboard,
number of cup holders, or serial number.

Model	Independent Variables (p)	Are they useful?	R^2	Adjusted R^2
Model 4	+ `dashboard_color`	❌No relation to price	0.885	0.86
Model 5	+ `cup_holders`, `serial_number`	❌Still no relation	0.89	0.84

R^2 **keeps increasing a little** (because adding any variable can always fit the data a tiny bit more),
but **Adjusted R^2 drops** — it's punishing you for adding useless features.

When we add more independent variables (**p increases**),
the denominator $N - p - 1$ **decreases**.
As a result, the fraction value **increases**.

Since this entire fraction is **subtracted from 1**,
the overall Adjusted R^2 **decreases** — unless the new variable actually improves R^2 a lot.

This means Adjusted R^2 **penalizes** the model for adding too many variables that don't truly help in prediction.

Case 2 = If p decreases denominator N - p - 1 increases and the fraction gets smaller, hence R square goes up.

□ In short:

Change in p	Effect on Denominator	Effect on Fraction	Effect on Adjusted R^2
p increases	Denominator ↓	Fraction ↑	Adjusted R^2 ↓
p decreases	Denominator ↑	Fraction ↓	Adjusted R^2 ↑