📅 DAY 17 — Linear Regression

Goal : Understand how models learn relationships

1️⃣ What is Linear Regression?

Linear Regression tries to learn a **straight-line relationship:**

$$y = w_1x_1 + w_2x_2 + \cdots + b$$

In simple words:

> "How much does `y` change when `x` changes?"

Example:

- More experience → higher salary
- Bigger house → higher price

2️⃣ Why Linear Regression Is IMPORTANT

Even though it's "simple", it:

- Teaches how models learn
- Explains loss functions
- Introduces optimization
- Is used in real systems

Many advanced models are extensions of this.

3️⃣ Model Components

Term	Meaning
Weight(w)	Feature importance
Bias (b)	Starting point
Prediction	`y_hat`
Error	`y - y_hat`

4️⃣ Cost Function (How the Model Knows It's Wrong)

Mean Squared Error (MSE)

$$MSE = \frac{1}{n}\sum(y - \hat{y})^2$$

Why square ?

- Penalizes large errors
- Differentiable (math-friendly)

> The model's goal is to **minimize this value**.

5️⃣ How does the Model Learn?

Gradient Descent (IMPORTANT)

1. Start with random weights
2. Measure error
3. Adjust weights slightly
4. Repeat until error is small

```
Too high error → adjust weights
Repeat many times → best line found
```

Learning Rate controls step size:

- Too high → overshoot
- Too low → slow learning

6️⃣ Overfitting & Underfitting (REVISIT)

Problem	Meaning
Underfitting	Model too simple
Overfitting	Model too complex

Linear regression usually underfits complex data.

7️⃣ Assumptions of Linear Regression

Linear Regression assumes:

1. Linear relationship

2. No extreme outliers
3. Low multicollinearity
4. Errors are normally distributed

If assumptions break → poor performance.

8️⃣ Implement Linear Regression (Sklearn)

```
from sklearn.linear_model import LinearRegression

model = LinearRegression()
model.fit(X_train, y_train)

y_pred = model.predict(X_test)
```

9️⃣ Interpreting Coefficients (VERY IMPORTANT)

```
print(model.coef_)
print(model.intercept_)
```

Meaning:

- Positive coef → y increases with x
- Larger value → stronger influence

🧠 This is why linear models are **interpretable**.

1️⃣0️⃣ Evaluation Metrics for Regression

Metric	Meaning
MAE	Average absolute error
MSE	Penalizes large errors
RMSE	Same unit as target
R^2	Variance explained
