Rohith Ramesh
www.linkedin.com/in/rohithr24

Linear Regression
Supervised machine learning (training on known data and predicting unknown data).

Types:
- Simple Linear Regression
- Multiple Linear Regression
- Polynomial Regression
- Regularization

Simple Linear Regression
This involves 1 input column and 1 output column (2 columns in total).

Multiple Linear Regression involves multiple input columns and 1 output column.

Simple Linear Regression
It is mainly used for numeric prediction based on the best fit line. Simple Linear Regression predicts a number (like marks, price, etc.) based on another value using a straight line.

We start with two columns of known data — one input and one output.

We split the data into:

- 80% for training (to learn the pattern)
- 20% for testing (to check accuracy)

The model draws a straight line through the training data.
This is called the best fit line because it tries to be as close as possible to all points.

We then use this line to predict the test data and compare the predictions with the real values.
The model's accuracy is measured using a score from 0 to 1.
The closer to 1, the better the prediction.

Finally, we use the same line to predict future or unknown values.
A model showing 100% accuracy might be overfitting — memorizing the training data instead of learning general patterns.

Slope of the line
The equation of the line is:

$$y = mx + b$$

where:

- m = slope

- b = y-intercept (theta)

This represents the best fit line.

What is the slope of the line and why do we use it?

Check the program to understand