Applied AI & Machine Learning CS-333

Dr. Abbas Hussain

PNEC, NUST

Lecture 3

Spring 2026

Machine Learning (ML) Models

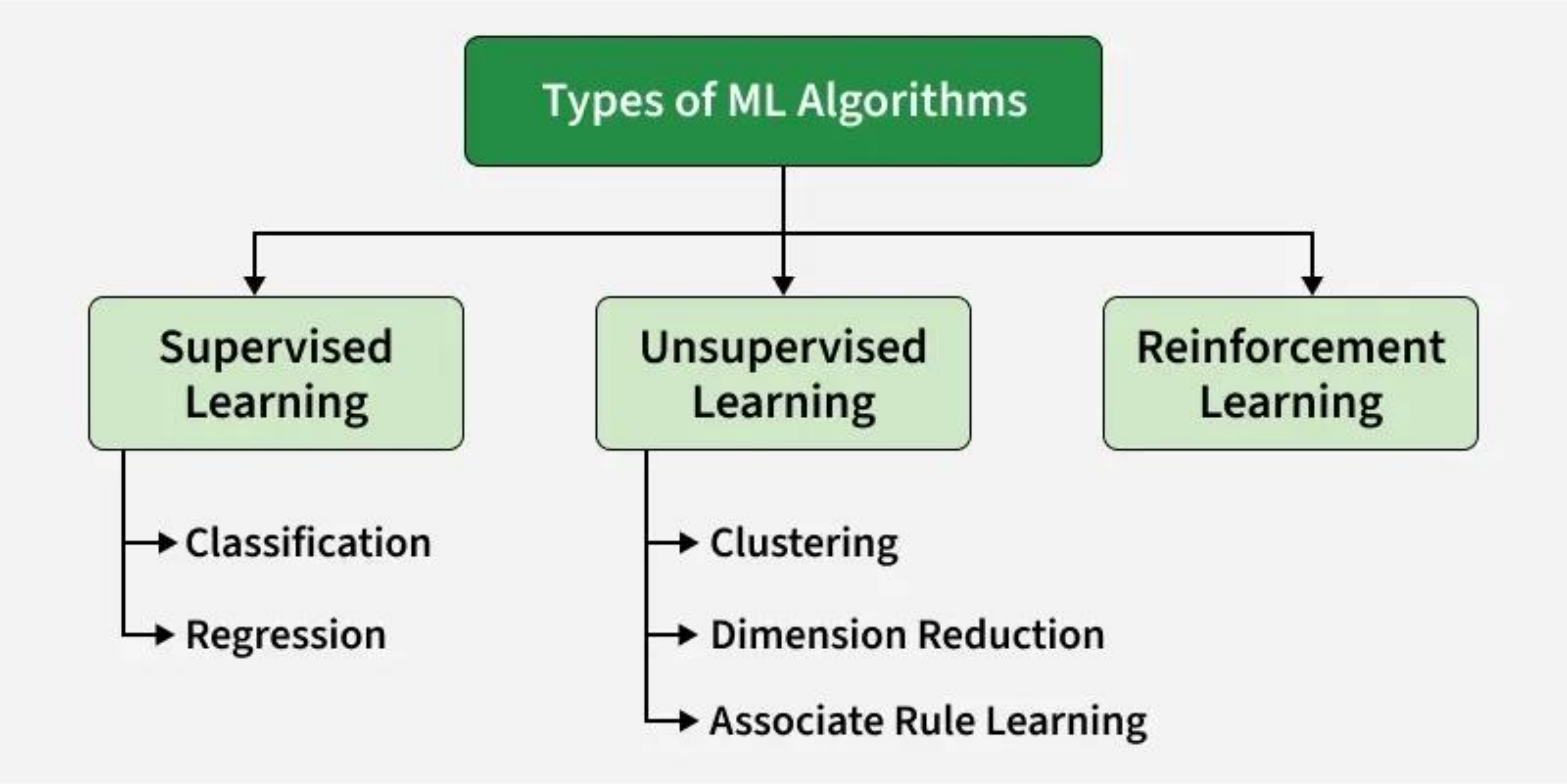
From Data to Prediction

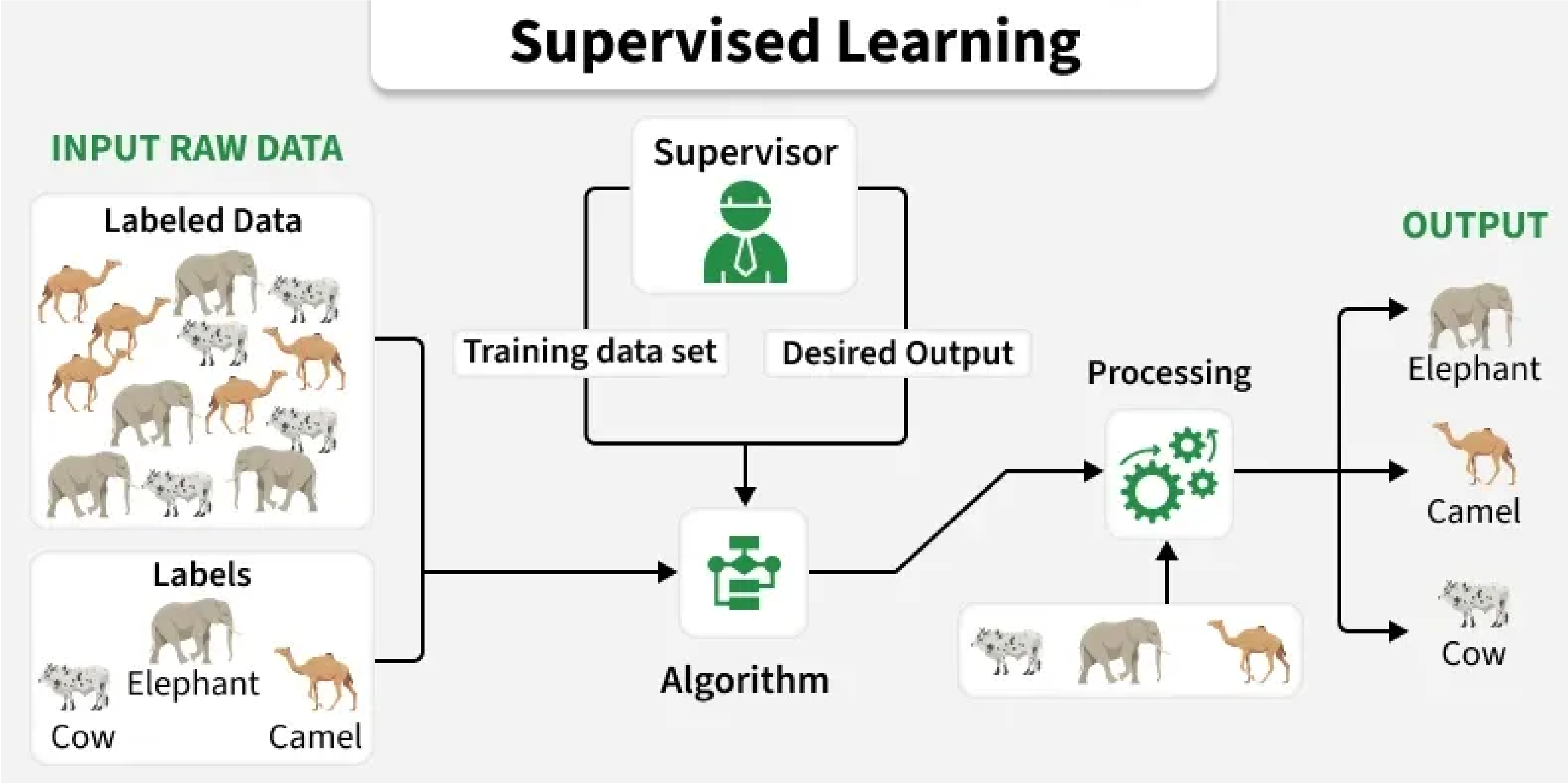
Data → Model → Loss → Optimization → Prediction

Learning Outcomes

By the end of this lecture, students will be able to:

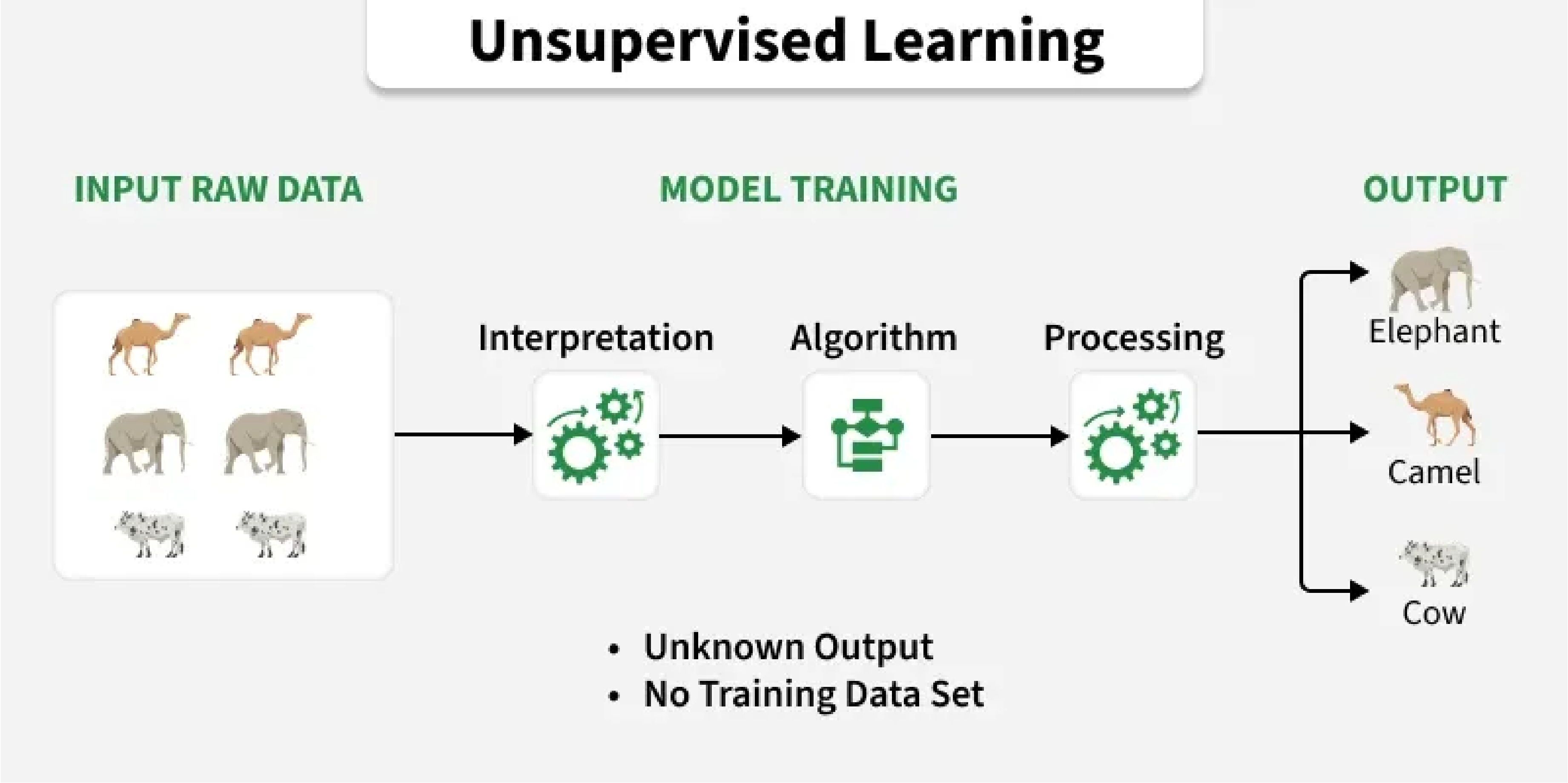
1. **Differentiate** between regression and classification problems.
2. **Explain** the mathematical formulation of linear regression.
3. **Interpret** model parameters (slope and intercept).
4. **Describe** the role of loss (cost) functions in model training.
5. **Compute and interpret** MSE, RMSE, and R^2.
6. **Select appropriate evaluation metrics** based on application needs.
7. **Analyze** model performance and identify underfitting and overfitting.


Types of ML Algorithms
Supervised Learning
Unsupervised Learning
Reinforcement Learning
Classification
Regression
Clustering
Dimension Reduction
Associate Rule Learning



Supervised learning problems are mainly divided into:

- **Classification**: predicting categories
- **Regression**: predicting continuous values



Unsupervised learning problems are mainly divided into:

- **Clustering** (group data points into clusters based on their similarities or differences)
- **Dimensionality Reduction** (reduces the number of features while keeping important information)

Examples of ML algorithms

Supervised

Classification and Regression

- Linear Regression
- Logistic Regression
- Decision Trees
- Support Vector Machines (SVM)
- k-Nearest Neighbors (k-NN)
- Naive Bayes
- Random Forest
- Gradient Boosting
- Neural Networks

Unsupervised

Clustering

- K-means clustering
- Gaussian mixture models (GMMs)
- DBSCAN

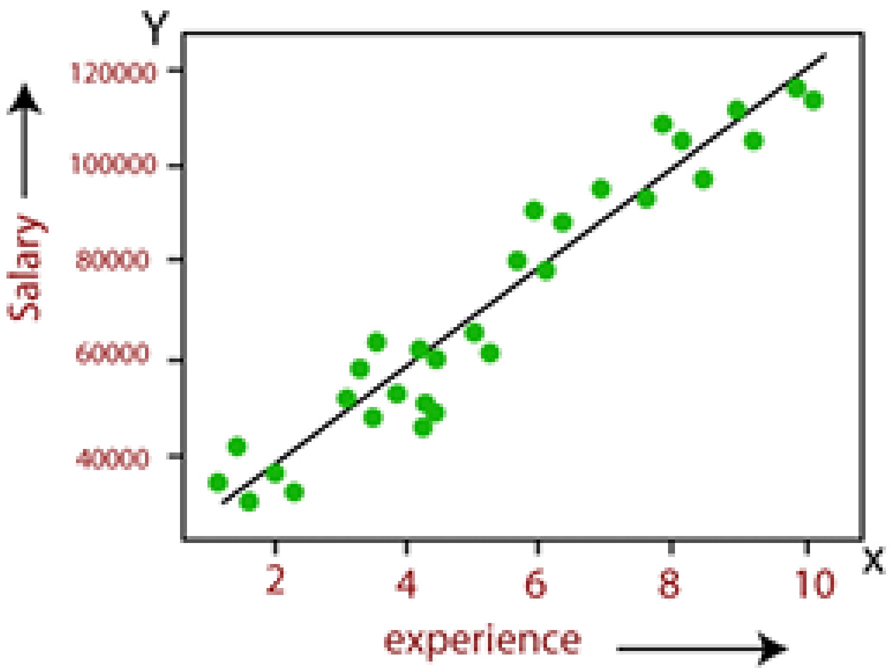
Dimensionality Reduction

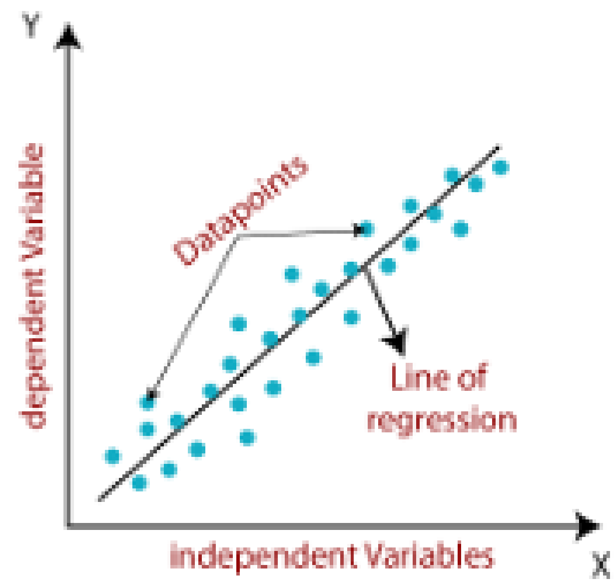
- PCA
- t-SNE

Linear Regression

Predicts a continuous value by fitting a straight line between input and output variables.
Model finds the best fit linear line between the independent and dependent variable

Example: Predicting house prices based on area.





Linear Regression

Mathematical Equation: Y=aX+b

or in ML form:

$$\hat{y} = wx + b$$

Where:

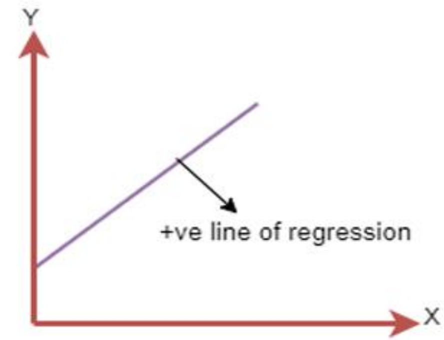
- w → slope
- b → intercept
- ŷ → predicted value (dependent variable)
- X → Independent variables

Visual:

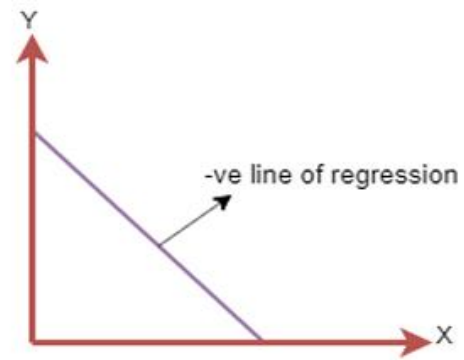
Best-fit line minimizing vertical error.

Interpretation:

- Positive slope → Positive relationship
- Negative slope → Negative relationship



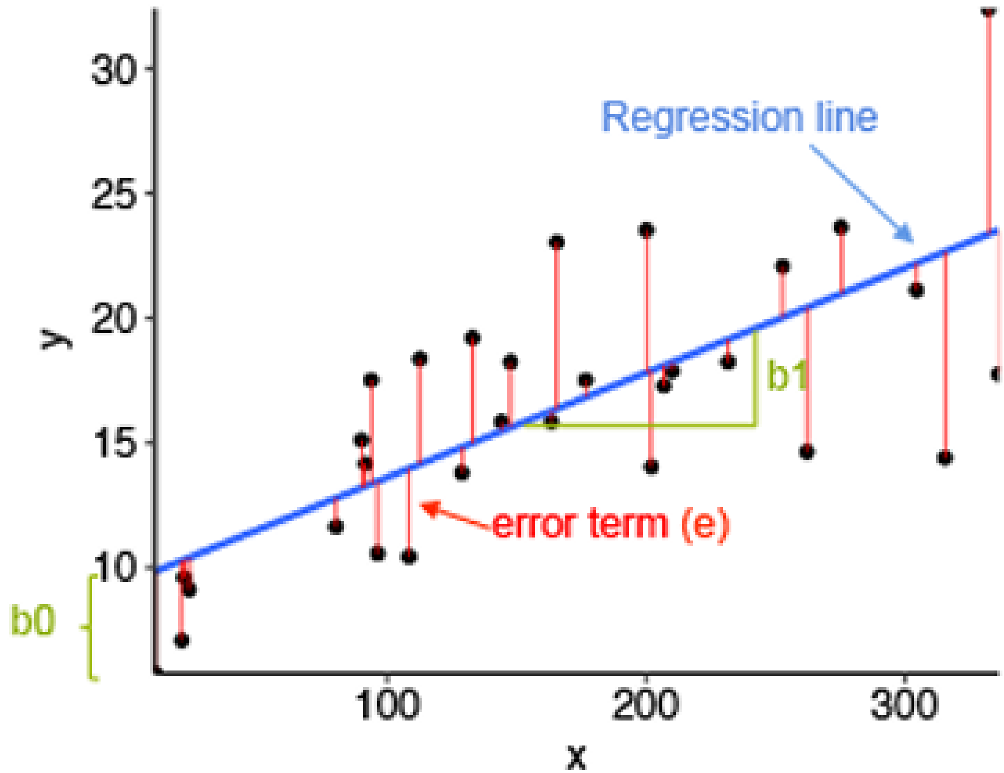
The line equation will be: **$Y= a_0+a_1x$**



The line of equation will be: **$Y= -a_0+a_1x$**

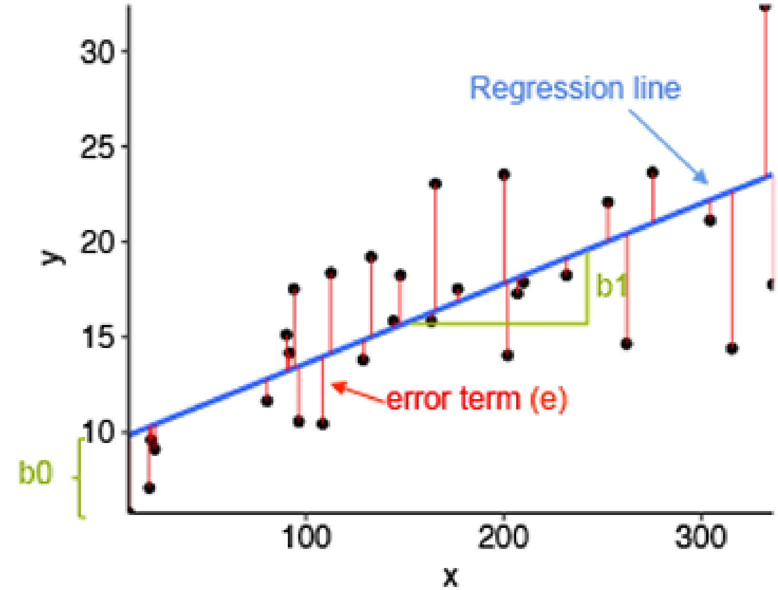
Linear Regression

- A Linear Regression model's main aim is to find the best fit linear line and the optimal values of intercept and coefficients such that the error is minimized.
- Error is the difference between the actual value and Predicted value, and the goal is to reduce this difference.



Linear Regression

Statistical tools for high-throughput data analysis



- x is our dependent variable which is plotted on the x-axis and y is the dependent variable which is plotted on the y-axis.
- Black dots are the data points i.e the actual values.
- bo is the intercept which is 10 and b1 is the slope of the x variable.
- The blue line is the best fit line predicted by the model i.e the predicted values lie on the blue line.
- The vertical distance between the data point and the regression line is known as error or residual.
- Each data point has one residual, and the sum of all the differences is known as the Sum of Residuals/Errors.

Linear Regression

How Do We Find the Best Line?

Residual:

$$Residual = Actual - Predicted$$

Objective:

Minimize sum of squared residuals.

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

Cost Function

Some popular applications of linear regression are:

- Analyzing trends and sales estimates
- Salary forecasting
- Real estate prediction
- Arriving at ETAs in traffic.

Linear Regression

Types of Linear Regression

- **Simple Linear Regression:** If a single independent variable is used to predict the value of a numerical dependent variable, then such a Linear Regression algorithm is called Simple Linear Regression.

- **Multiple Linear regression:** If more than one independent variable is used to predict the value of a numerical dependent variable, then such a Linear Regression algorithm is called Multiple Linear Regression.

$$y = b_0 + b_1x_1 + b_2x_2 + \ldots + b_nx_n$$

Multiple linear regression example

You are a public health researcher interested in social factors that influence heart disease. You survey 500 towns and gather data on the percentage of people in each town who smoke, the percentage of people in each town who bike to work, and the percentage of people in each town who have heart disease.

Linear Regression

Model Evaluation Metrics

1) **MSE** (Mean Squared Error) $$\frac{1}{n}\sum(y-\hat{y})^2$$

Etracted by squared the average difference over the data set.

2) **RMSE** (Root Mean Squared Error) $$\sqrt{MSE}$$

It is the error rate by the square root of MSE.

3) **MAE** (Mean absolute error) $$\frac{1}{n}\sum|y-\hat{y}|$$

Extracted by averaged the absolute difference over the data set.

4) **R^2** (Coefficient of determination)

It represents the coefficient of how well the values fit compared to the original values. The value from 0 to 1 interpreted as percentages. The higher the value is, the better the model is.

How Linear Regression Actually Learns?

Model:

$$\hat{y} = wx + b$$

Training process:

1. Make prediction
2. Calculate error
3. Compute loss (MSE)
4. Update parameters (w, b)
5. Repeat until error is minimized

Linear regression learns optimal weights and bias by minimizing a cost function.

MSE vs R^2 (Conceptual Difference)

MSE	R^2
Measures prediction error	Measures explained variance
Lower is better	Closer to 1 is better
Scale dependent	Scale independent
In squared units	Unitless

Interpretation:

- MSE → "How wrong are predictions?"
- R^2 → "How much variation is explained?"

Terminologies Related to the Regression Analysis

Underfitting and Overfitting:

- If our algorithm works well with the training dataset but not well with test dataset, then such problem is called Overfitting.

And

- if our algorithm does not perform well even with training dataset, then such problem is called underfitting.