Integrating Machine Learning Model
by Rizki Fajar Nugroho

(date of delivery)

Trainer Profile

Rizki Fajar Nugroho
Data Scientist at SaaS Company
LinkedIn - Rizki Fajar Nugroho

Table of Content

Content
Machine Learning Fundamentals
Machine Learning Workflows

Machine Learning Fundamentals

Machine Learning

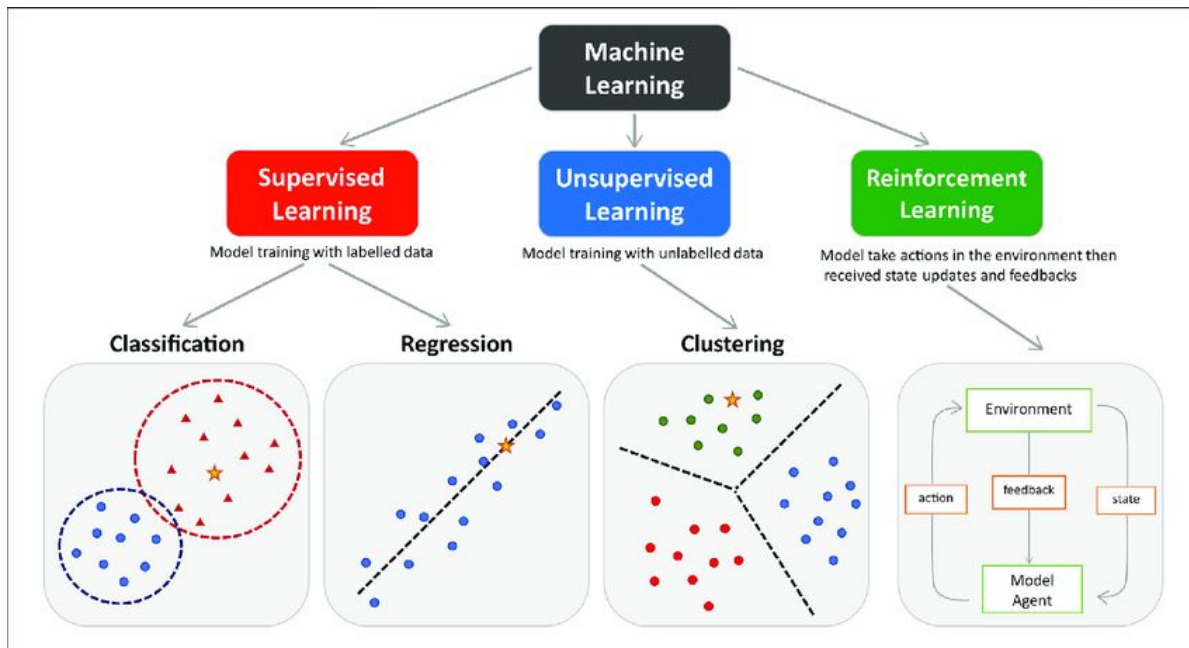Method to give computers ability to **learn from data** without explicitly programmed.

VS

Learn From Experience

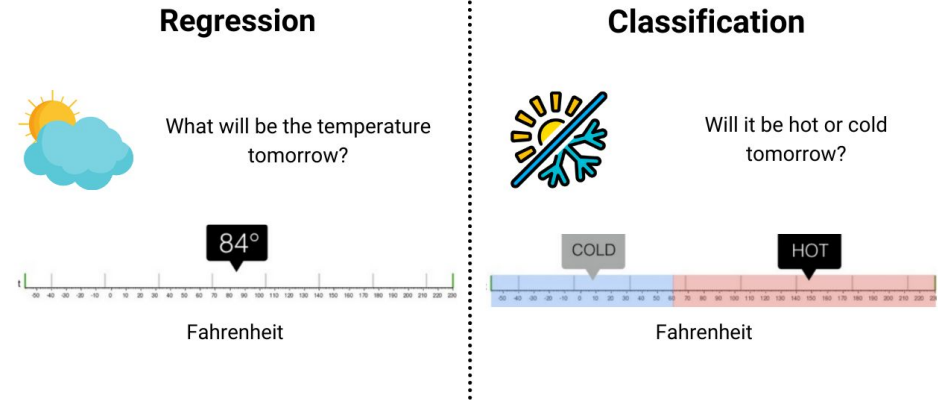Learn From Data

Machine Learning Types

Concept of Supervised Learning

Supervised Learning refers to a class of systems and algorithms that determine a **predictive** model **using data points** with known outcomes.

Implementation of Supervised Learning:

- **Regression** - The model finds outputs that are real variables (number which can have decimals.)
- **Classification** - The model finds classes in which to place its inputs.

Regression

What will be the temperature tomorrow?

84°

Fahrenheit

Classification

Will it be hot or cold tomorrow?
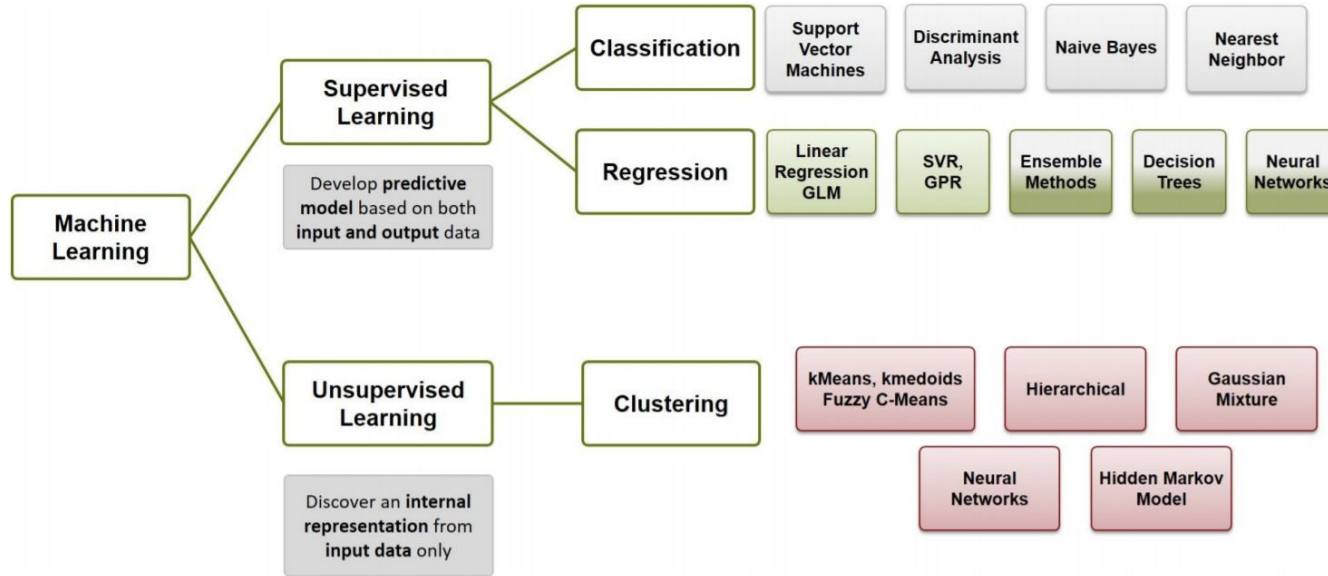
COLD   HOT

Fahrenheit

Concept of Unsupervised Learning

In unsupervised learning, **only input data** is provided in the dataset. There are **no labelled outputs** to aim for. But it may be surprising to know that it is still possible to find **many interesting and complex patterns hidden within data without any labels. The goal is to capture interesting structure / information**
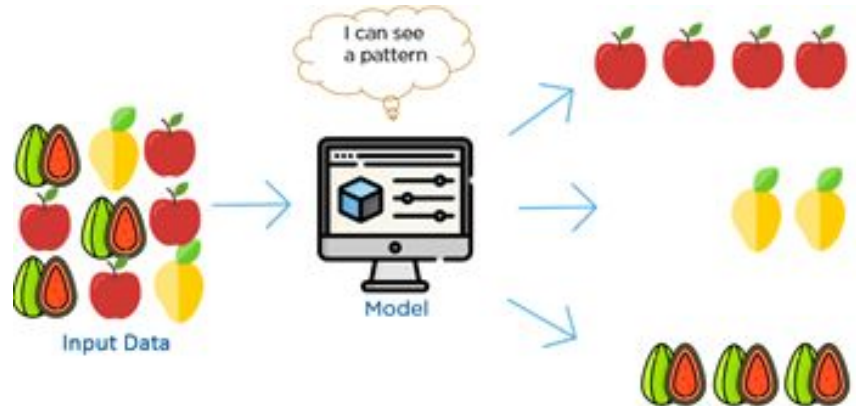
Machine Learning Algorithms

Machine Learning - Model Evaluation Metrics

Regression	Classification
- Mean Absolute Error (MAE) - Root Mean Squared Error (RMSE) - R-Squared and Adjusted R-Squared	- Recall - Precision - F1-Score - Accuracy - Area Under the Curve (AUC)

Machine Learning - Regression Model Evaluation Metrics

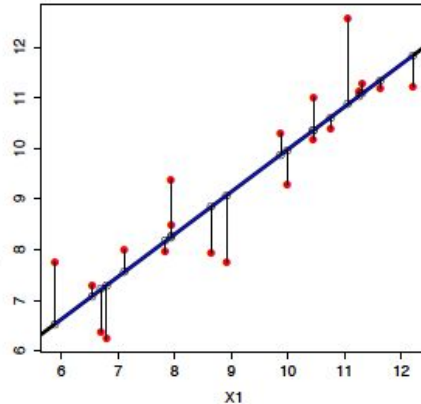|	
Mean squared error	$\text{MSE} = \dfrac{1}{n} \sum_{t=1}^{n} e_t^2$
Root mean squared error	$\text{RMSE} = \sqrt{\dfrac{1}{n} \sum_{t=1}^{n} e_t^2}$
Mean absolute error	$\text{MAE} = \dfrac{1}{n} \sum_{t=1}^{n} |e_t|$
Mean absolute percentage error	$\text{MAPE} = \dfrac{100\%}{n} \sum_{t=1}^{n} \left| \dfrac{e_t}{y_t} \right|$

Machine Learning - RMSE

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2}$$

y_i : actual outcome for obs. i

$\hat{y}_i$: predicted outcome for obs. i

N : Number of observations

- Root Mean Squared Error (**RMSE**)
- Average distance between actual and regression line

Machine Learning - Classification Model Evaluation Metrics - Accuracy

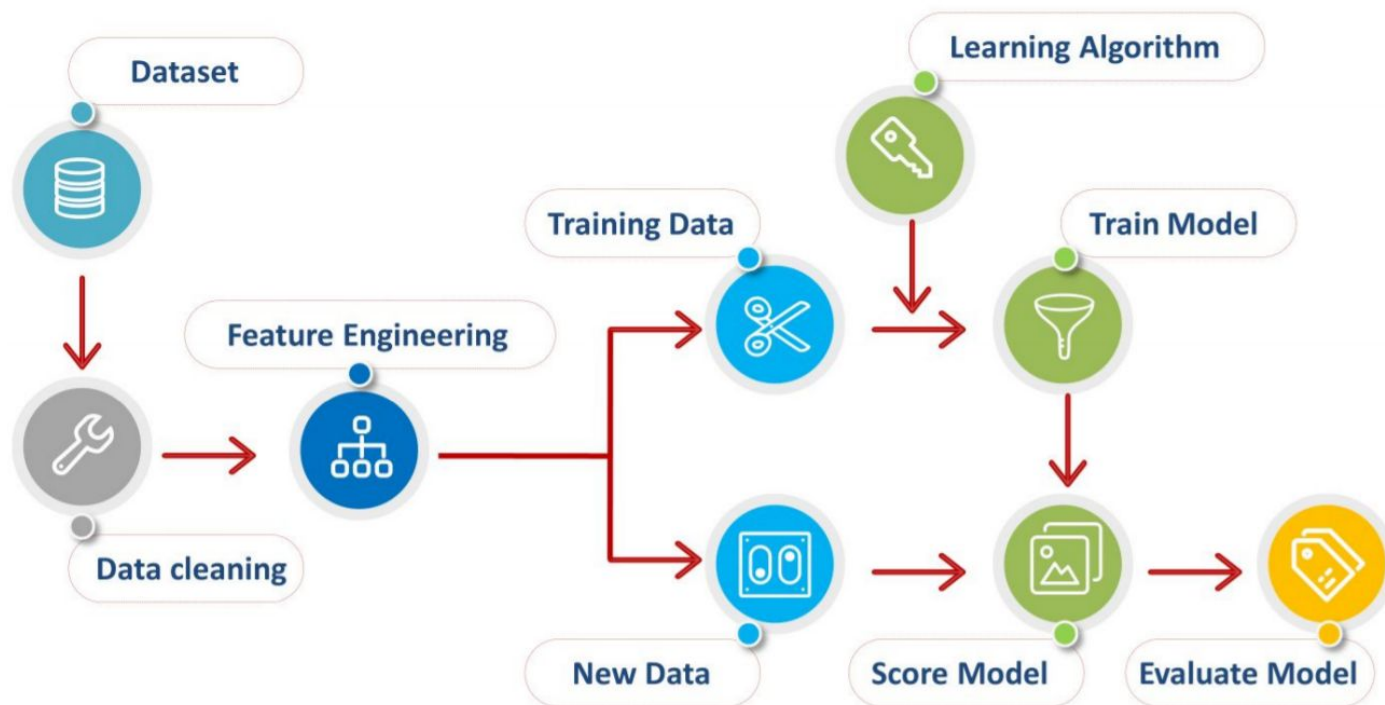$$\text{Accuracy} = \frac{\text{Correct predictions}}{\text{All predictions}}$$

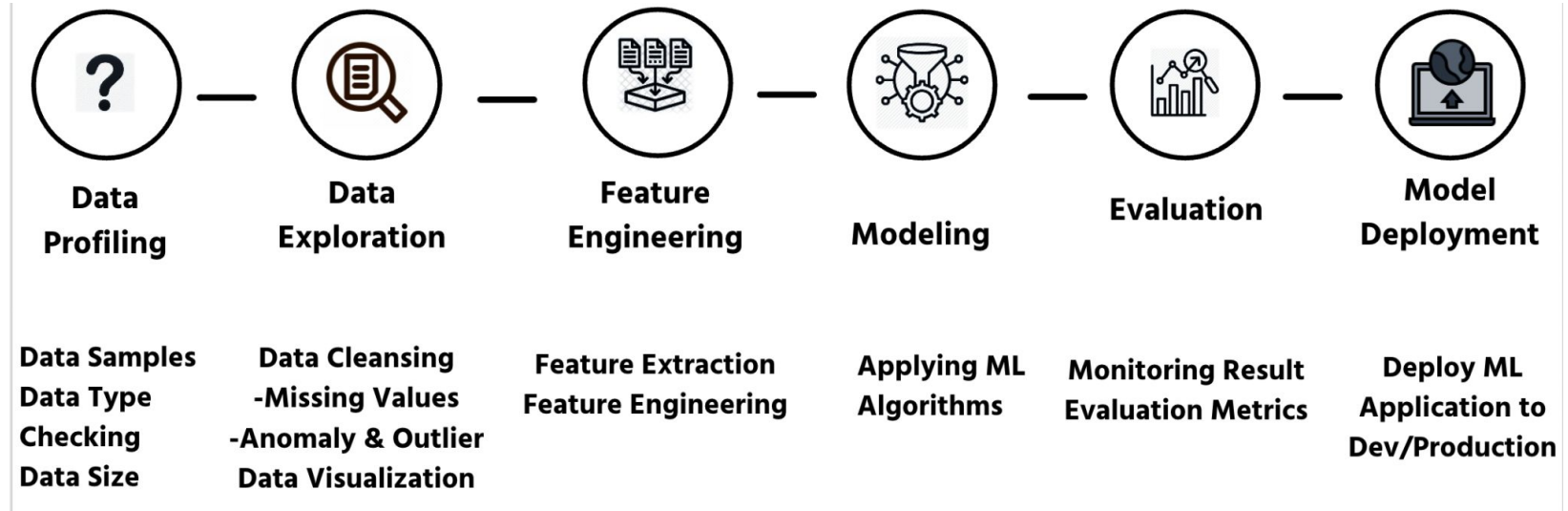# Machine Learning - Classification Model Evaluation Metrics - Confusion Matrix

Machine Learning Workflows

Machine Learning General Steps

Machine Learning Workflow

Data Profiling	Data Exploration	Feature Engineering	Modeling	Evaluation	Model Deployment
Data Samples Data Type Checking Data Size	Data Cleansing -Missing Values -Anomaly & Outlier Data Visualization	Feature Extraction Feature Engineering	Applying ML Algorithms	Monitoring Result Evaluation Metrics	Deploy ML Application to Dev/Production