

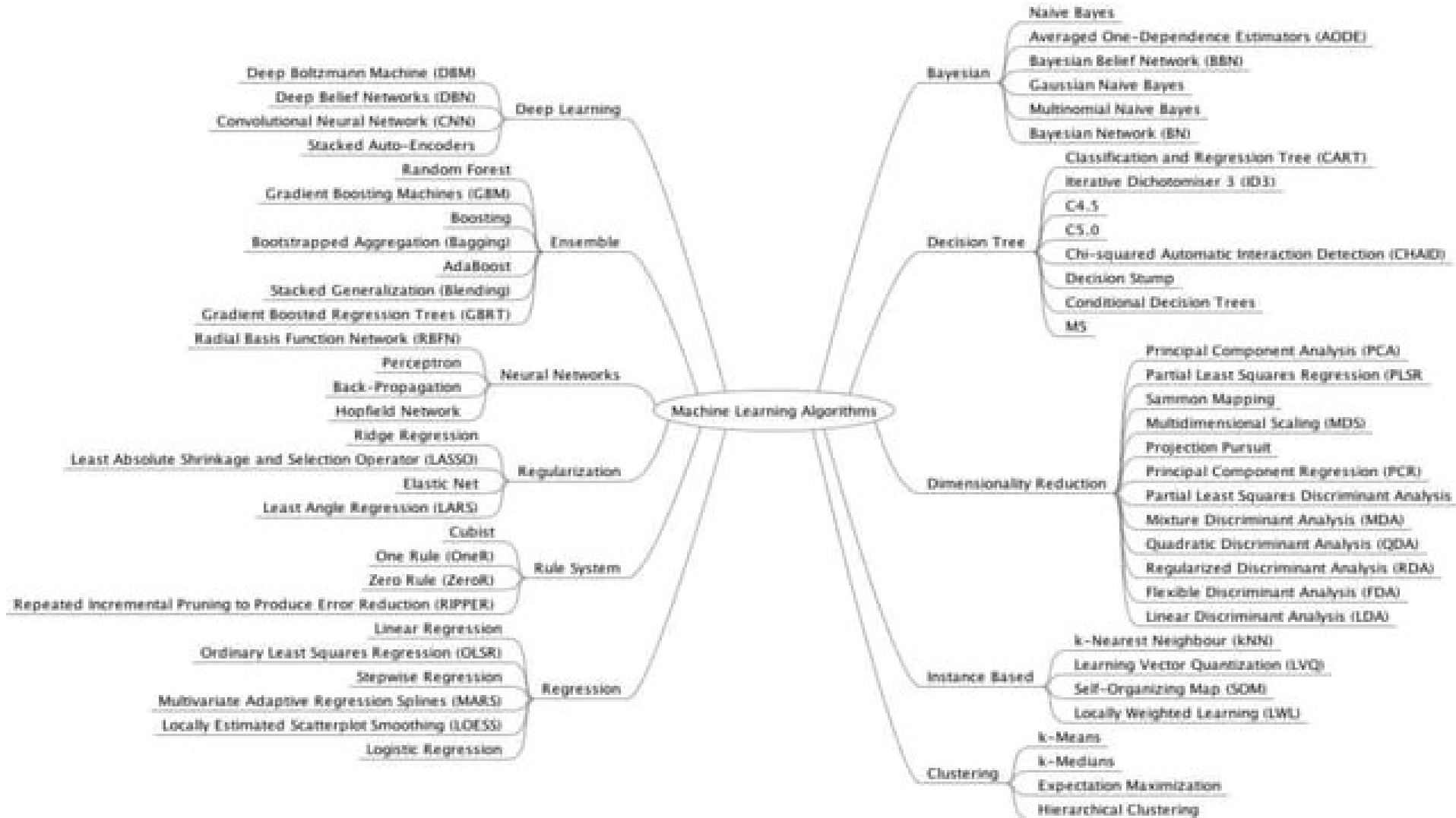
PENERAPAN ***MACHINE LEARNING***

Team Teaching

MACHINE LEARNING MODELS

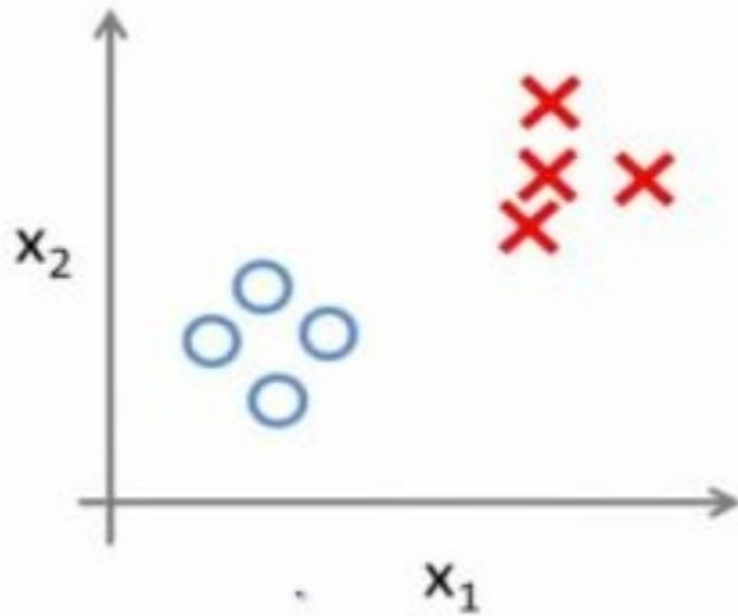


MACHINE LEARNING ALGORITHMS

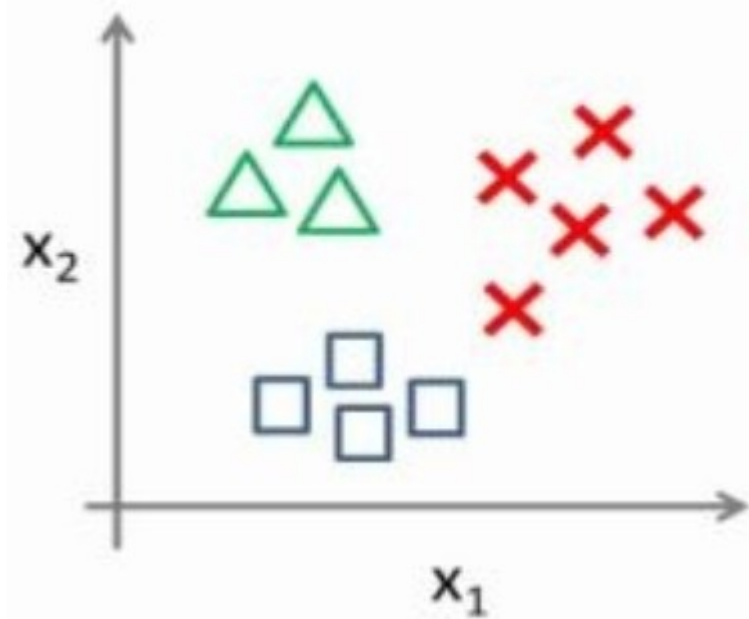


CLASSIFICATION

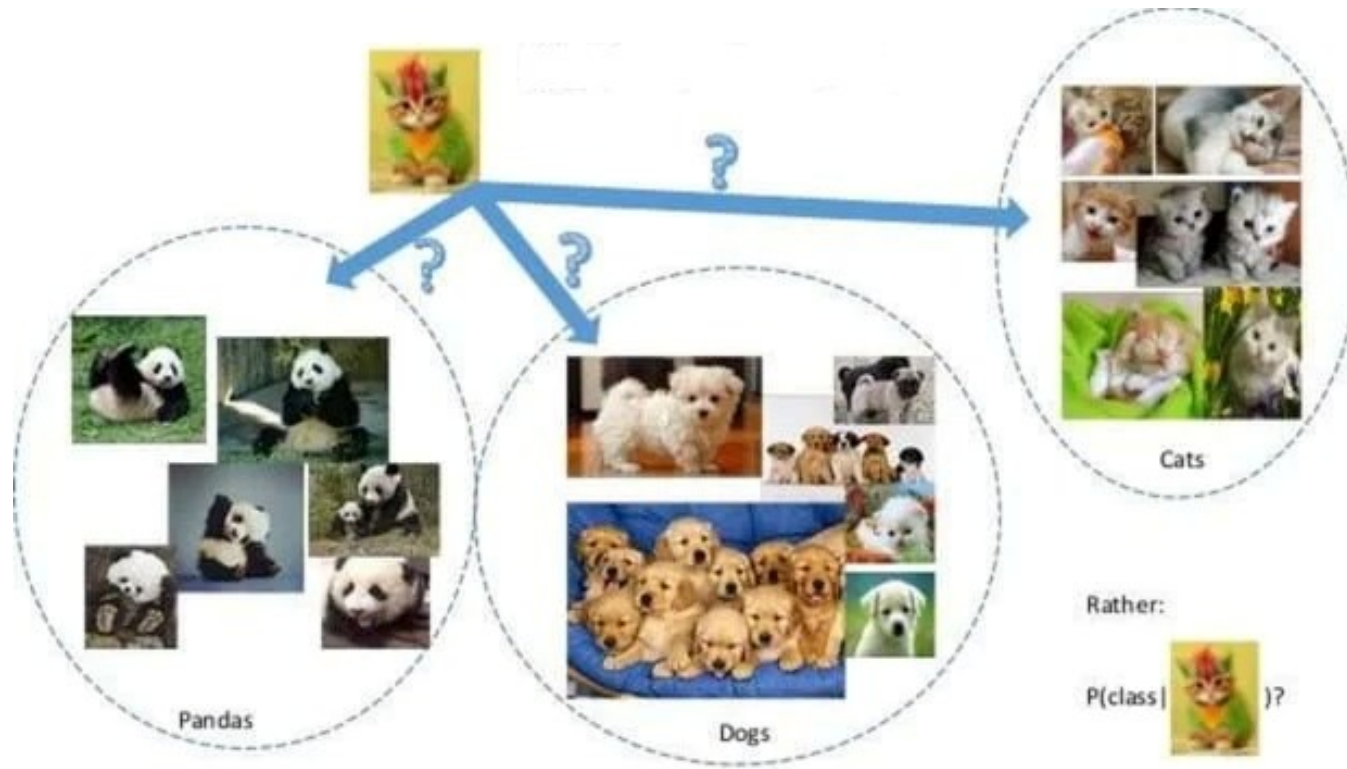
Binary classification:



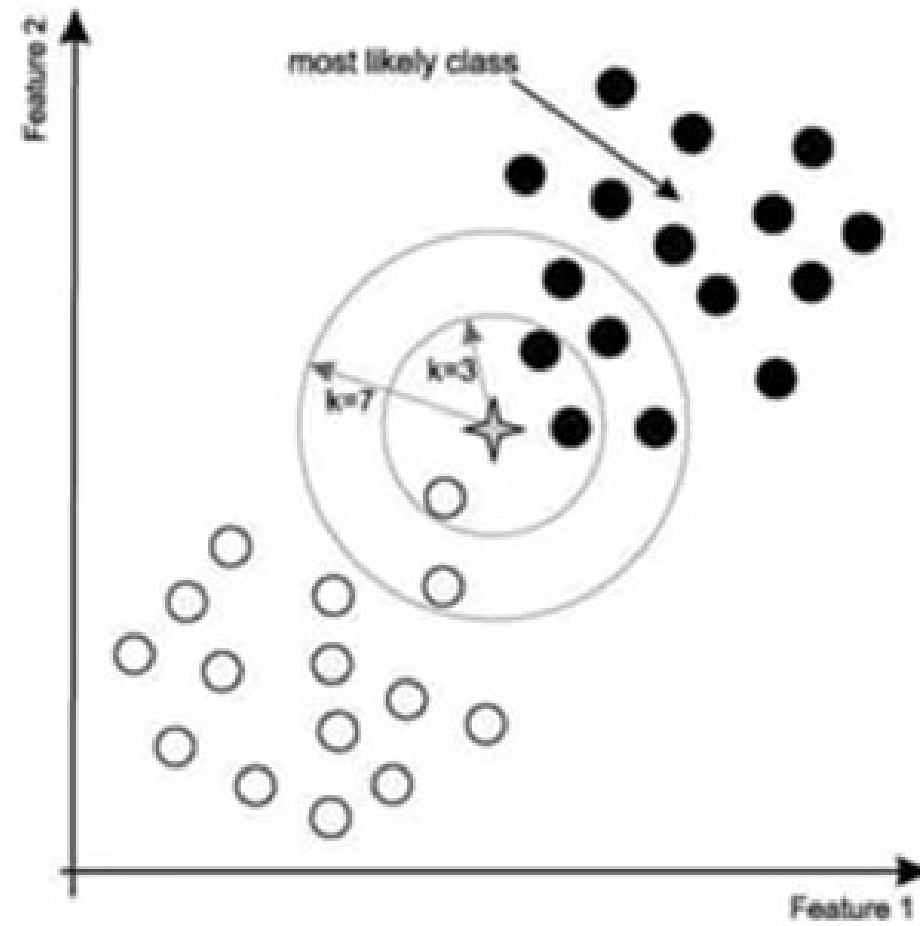
Multi-class classification:



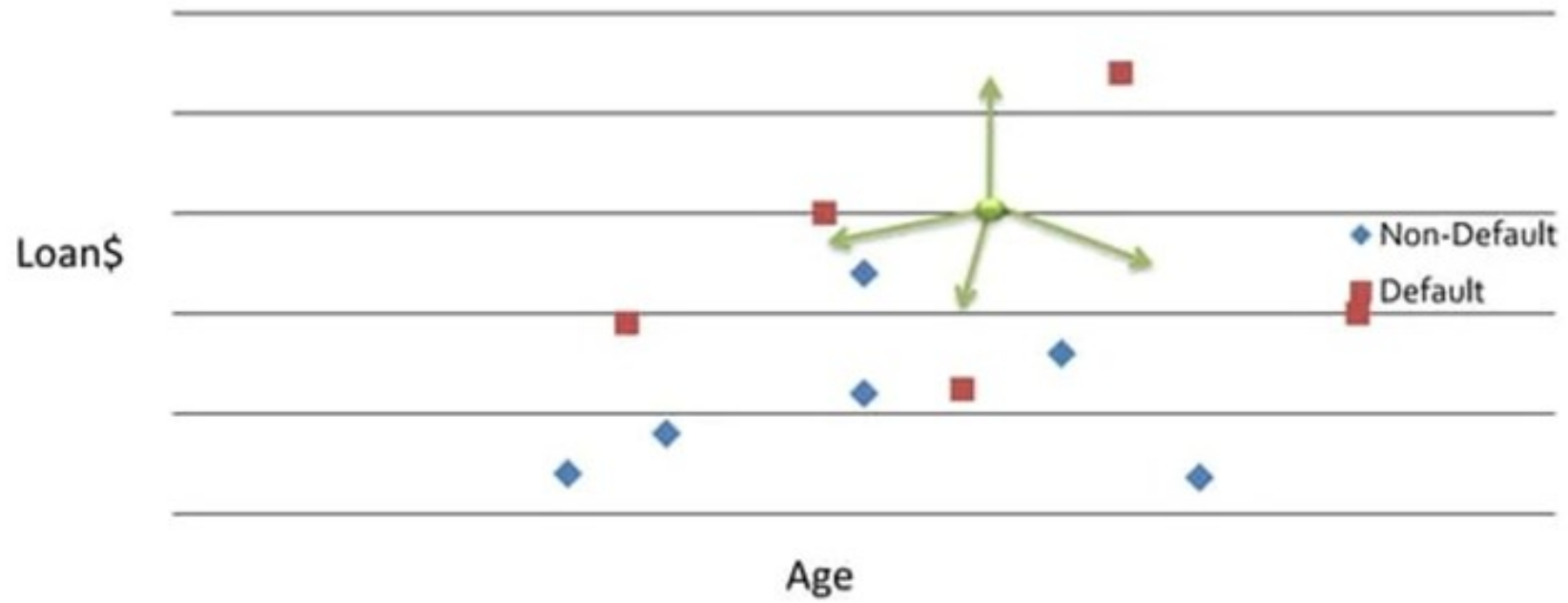
CLASSIFICATION



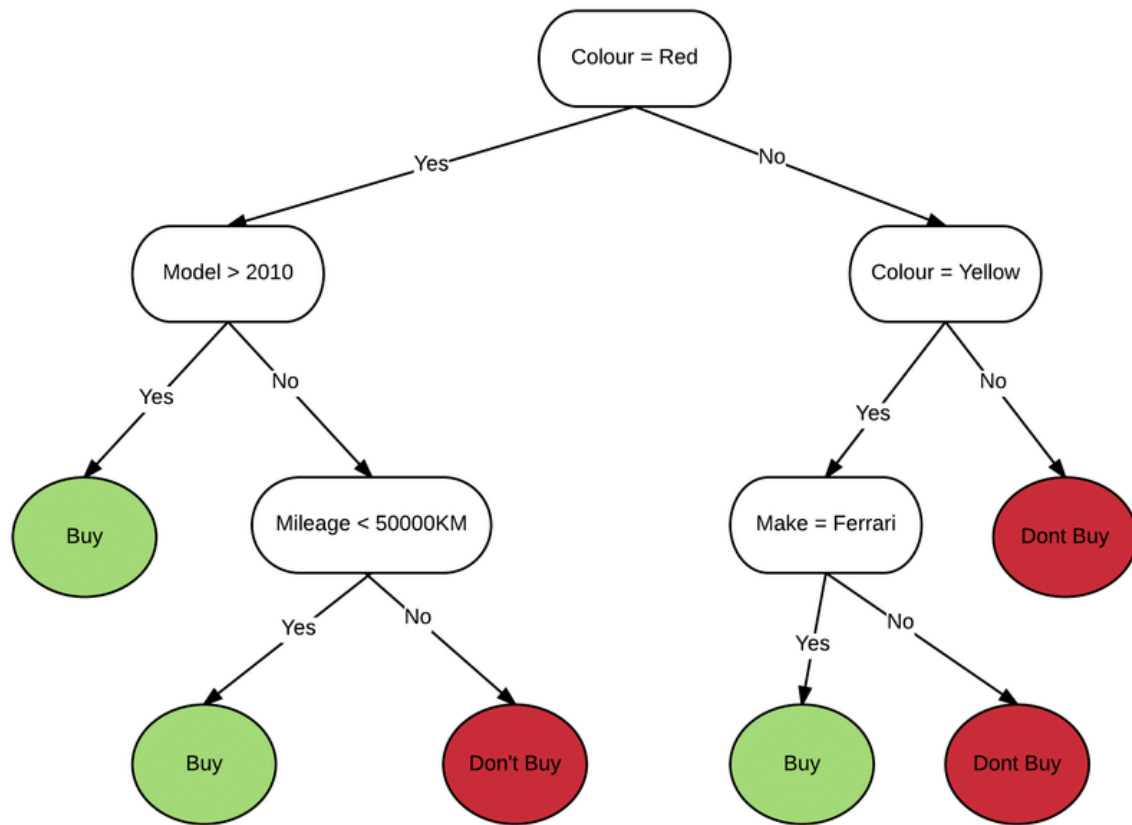
KNN



KNN



DECISION TREE



- Library yang bisa digunakan pada python:

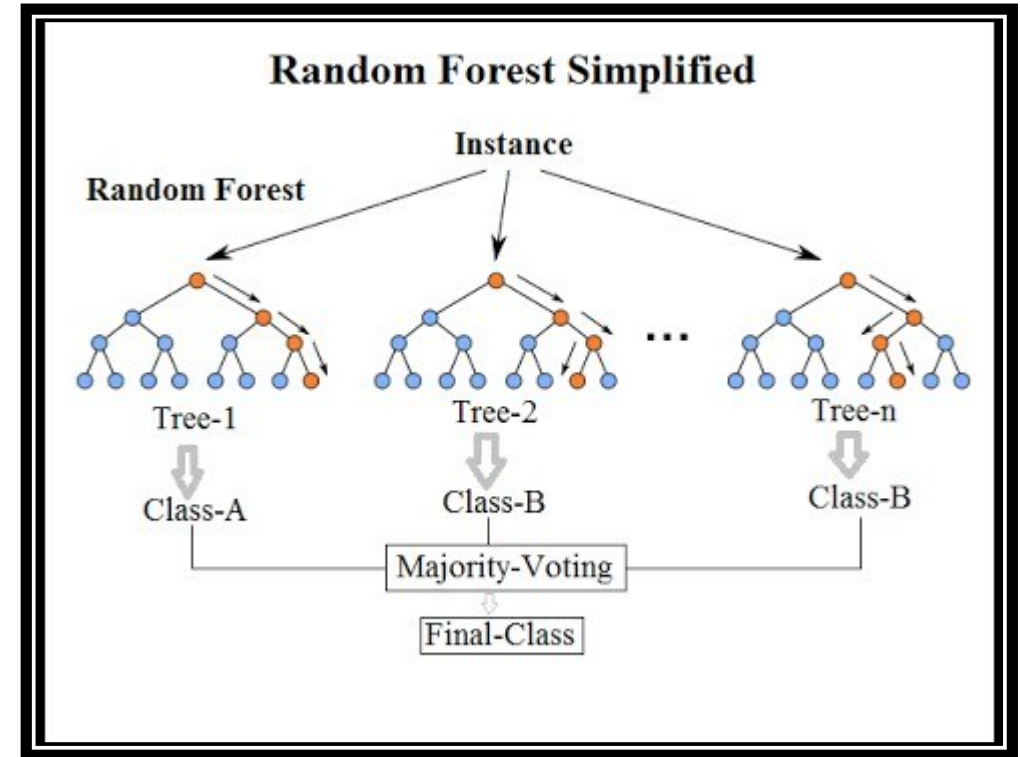
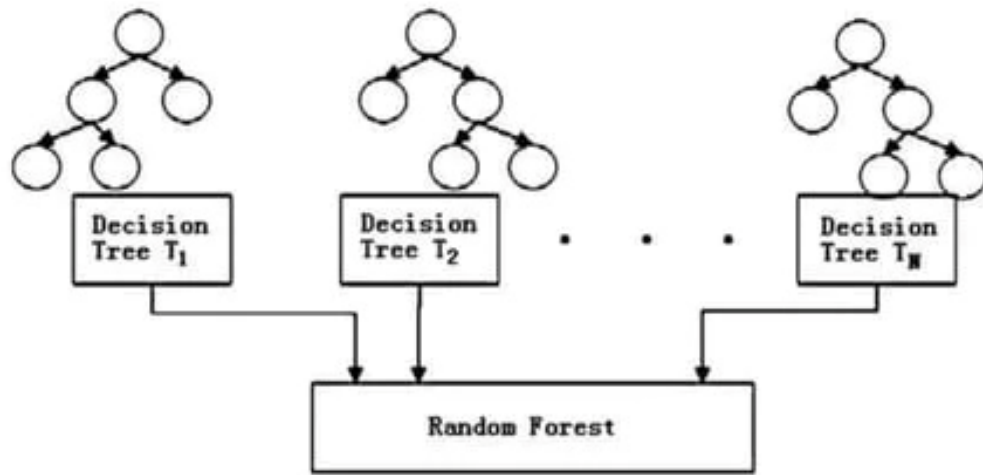
`import pandas`

`from sklearn import tree`

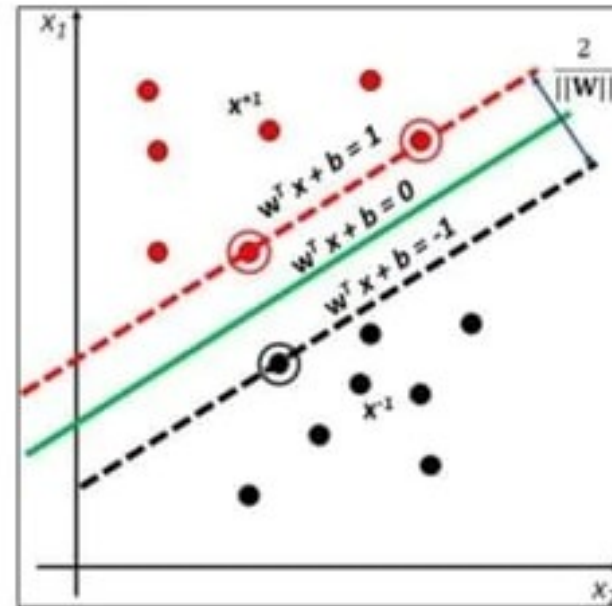
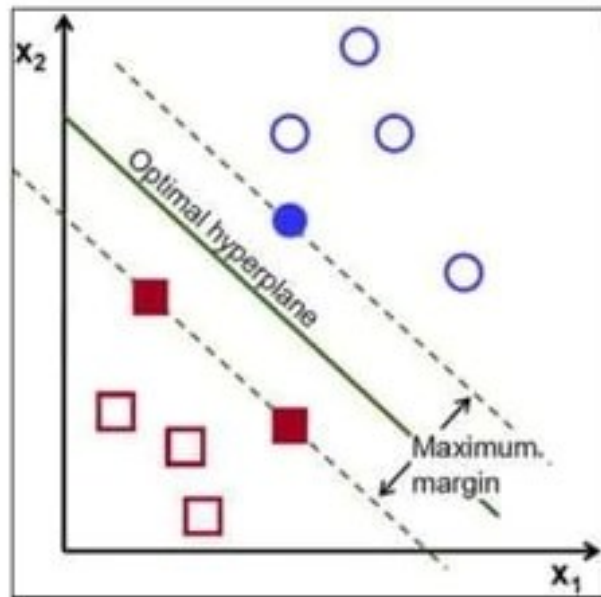
`import pydotplus`

`from sklearn.tree import DecisionTreeClassifier`

RANDOM FOREST



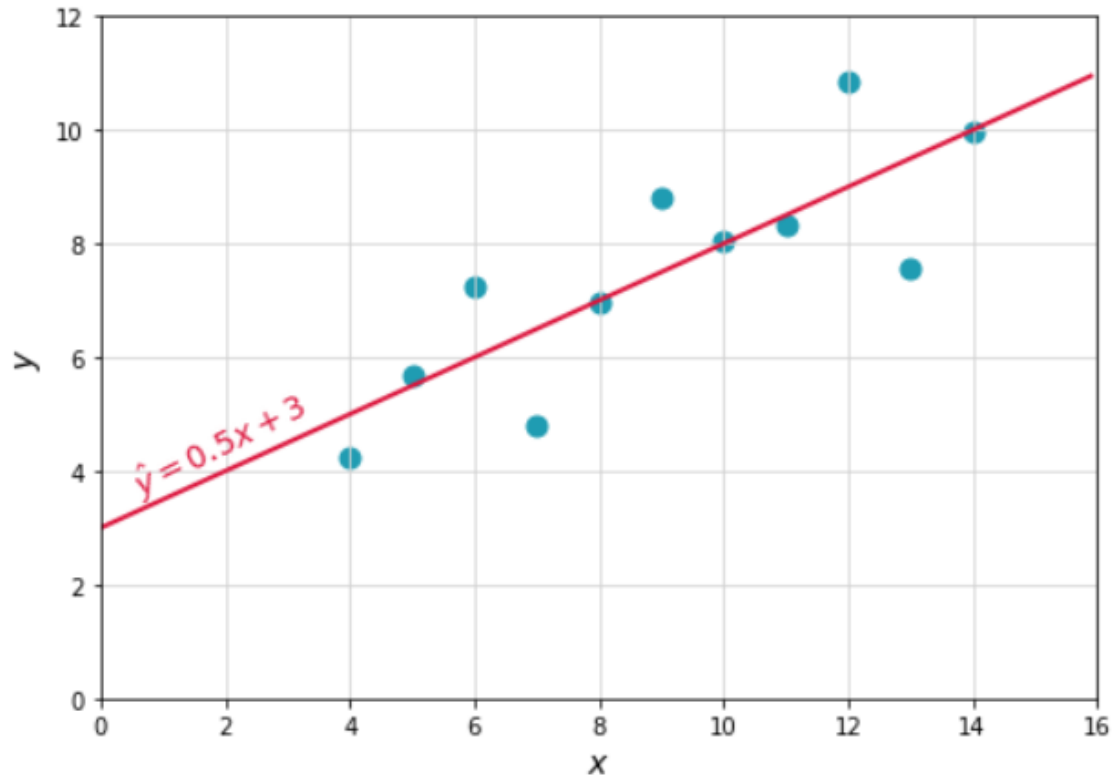
SVM



CONTOH LIBRARY

```
from sklearn import tree
from sklearn import svm
from sklearn import neighbors
from sklearn.metrics import accuracy_score
from sklearn.naive_bayes import GaussianNB
```

LINEAR REGRESION

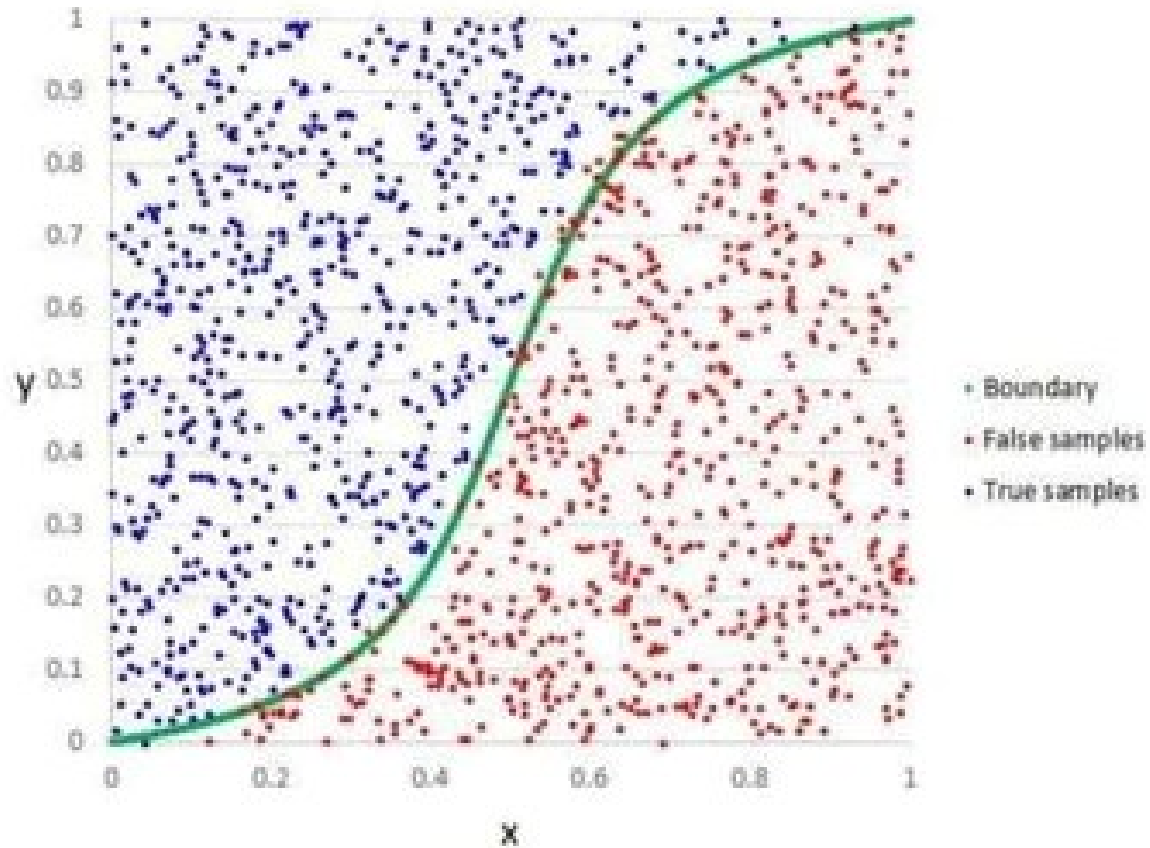


- Library yang bisa digunakan pada python:

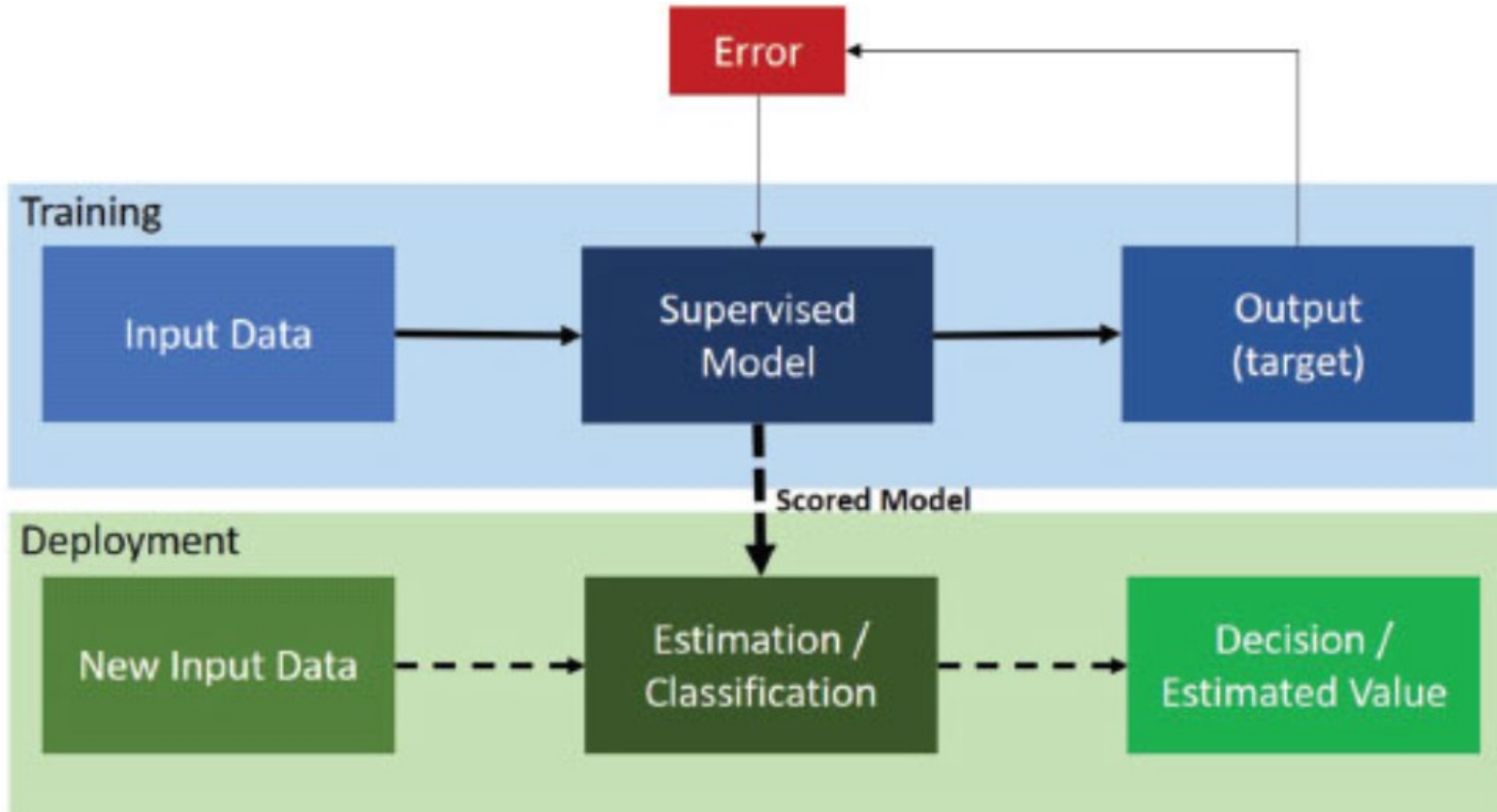
`import matplotlib.pyplot as plt`

`from scipy import stats`

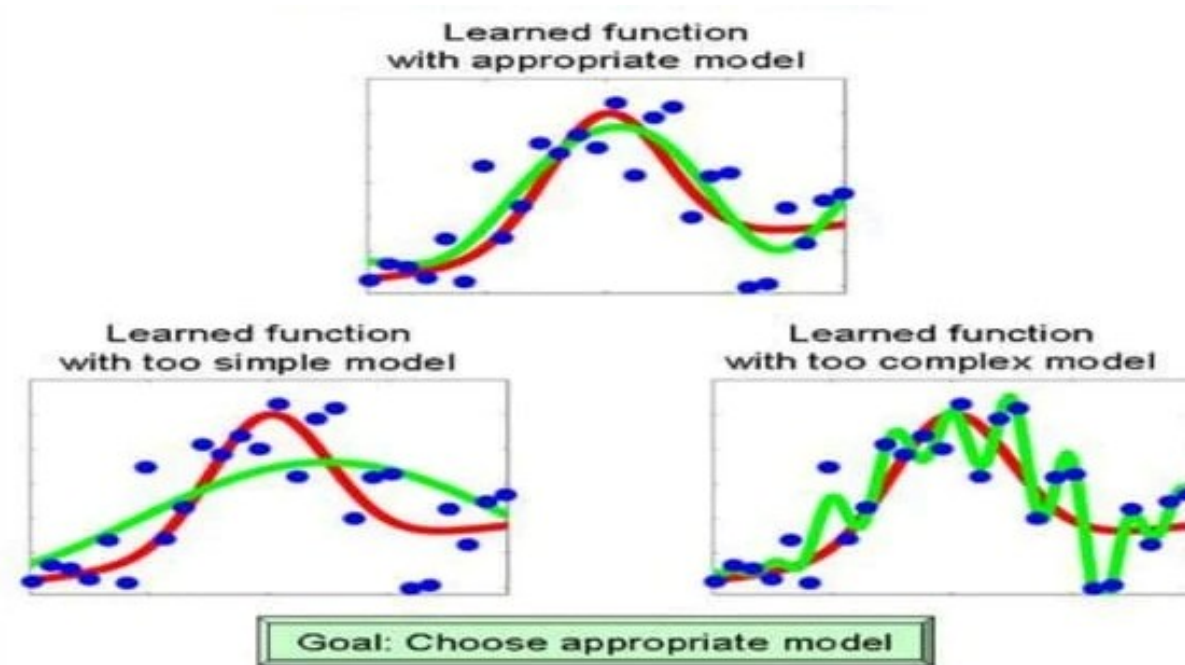
LOGISTIC REGRESION



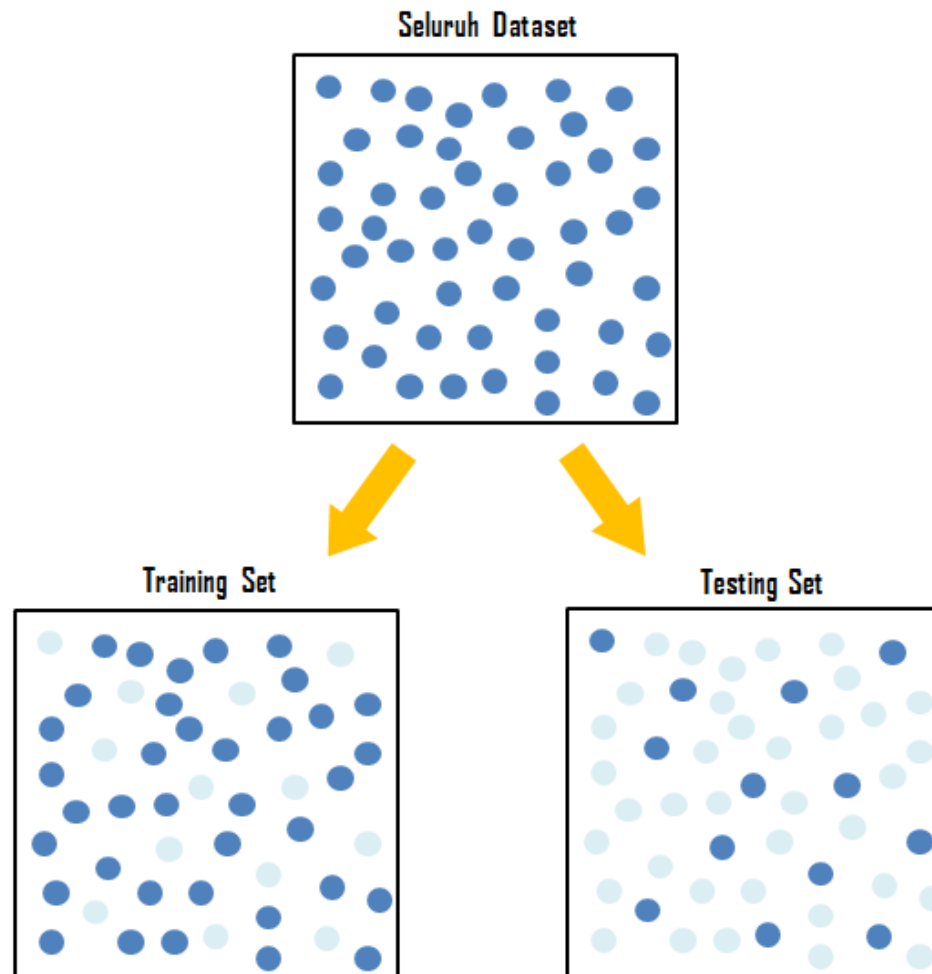
METODE TRAINING SUPERVISED MODEL



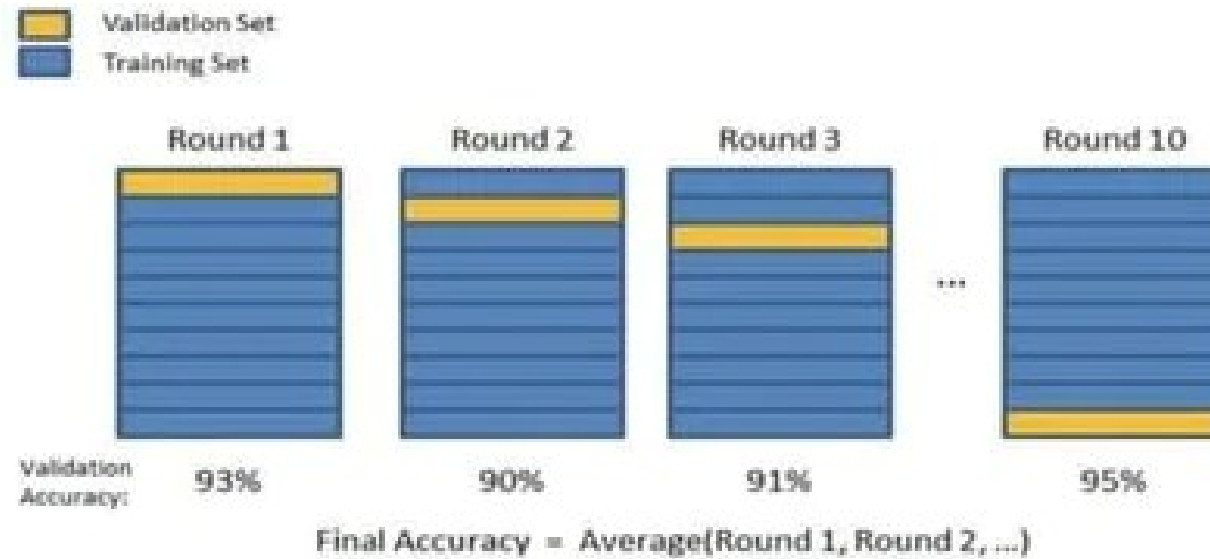
MODEL SELECTION



TRAINING VS CROSS VALIDATION (TEST)



TRAINING VS CROSS VALIDATION (TEST)



TRAINING VS CROSS VALIDATION (TEST)

- Contoh program:

```
from sklearn.model_selection import train_test_split
```

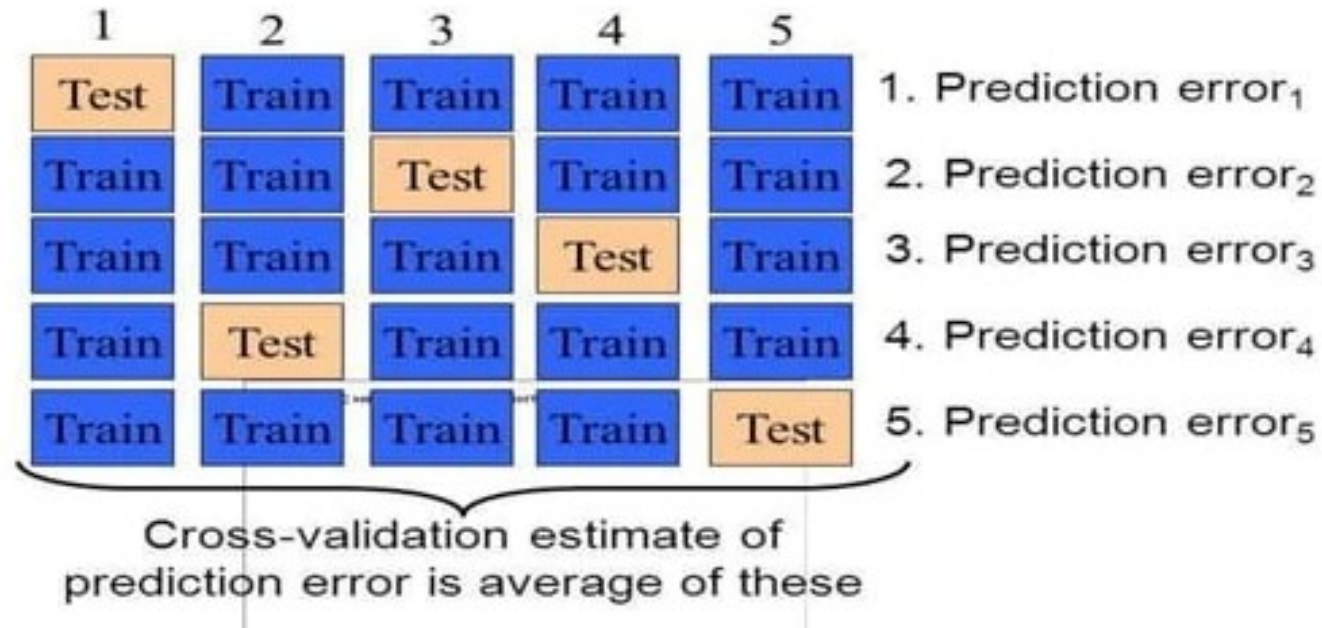
→ train/test split data langsung:

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
```

→ train/test split data menggunakan parameter random state:

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=3)
```

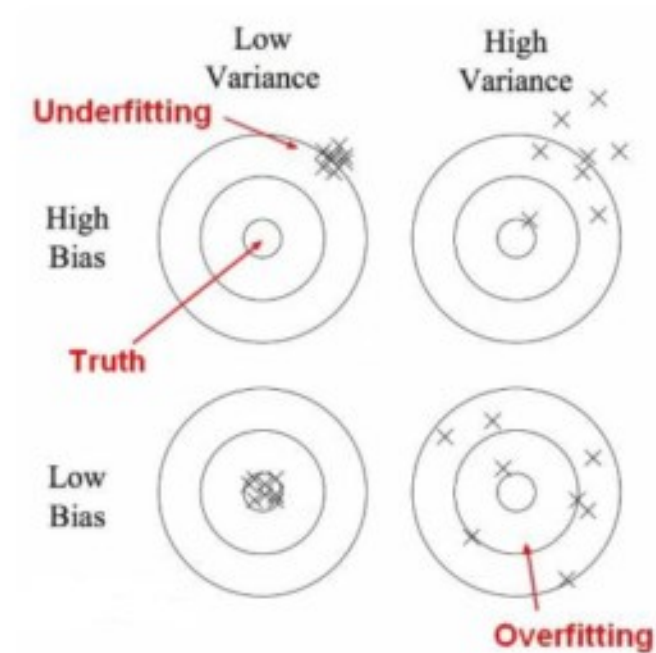
CROSS VALIDATION (TEST)



BIAS DAN VARIANCE

- Bias adalah perbedaan antara rata rata hasil prediksi dari model ML dengan data nilai yang sebenarnya.
- Variance adalah variabel dari prediksi model untuk data tertentu dengan memberikan informasi persebaran data.

BIAS DAN VARIANCE



- *Underfitting* merupakan kondisi saat nilai akurasi data training rendah diikuti dengan nilai akurasi data testing yang rendah.
- *Overfitting* merupakan kondisi dimana nilai akurasi pada data training tinggi dengan nilai akurasi data testing rendah

BIAS DAN VARIANCE

- Contoh algoritma *machine learning* dengan *low bias*:

Decision Tree, K-Nearest Neighbors (KNN), dan Support Vector Machine (SVM).

- Contoh algoritma *machine learning* dengan *high bias*:

Linear Regression, Logistic Regression, dan Linear Discriminant Analysis.

- Contoh algoritma *machine learning* dengan *low variance*:

Linear Regression, Logistic Regression, dan Linear Discriminant Analysis.

- Contoh algoritma *machine learning* dengan *high variance*:

Decision Tree, KNN, dan SVM.

BIAS DAN VARIANCE

- Solusi:

mengkonfigurasi algoritma *machine learning* yang digunakan dengan cara parameterisasi yang bertujuan untuk menyeimbangkan *bias* dan *variance*.

→ Dengan mengurangi *variance* akan meningkatkan *bias*, dan mengurangi *bias* akan meningkatkan *variance*. Anda dapat mengatur nilai parameter yang digunakan untuk menemukan keseimbangan *bias-variance* pada model Anda sehingga membuat akurasi menjadi lebih baik.

EVALUATION METRIC

$$accuracy = \frac{\textit{jumlah data prediksi yang benar}}{\textit{jumlah seluruh data prediksi}}$$

CONFUSION MATRIX

Actual Values			
		1 (Positive)	0 (Negative)
Predicted Values	1 (Positive)	TP (True Positive)	FP (False Positive) Type I Error
	0 (Negative)	FN (False Negative) Type II Error	TN (True Negative)

$$accuracy = \frac{TN + TP}{TN + FP + FN + TP}$$

$$precision = \frac{TP}{FP + TP}$$

$$\text{Sensitivity (recall)} = \frac{TP}{TP + FN}$$

$$specificity = \frac{TN}{TN + FP}$$

$$f1 \text{ score} = 2 \frac{1}{\frac{1}{precision} + \frac{1}{recall}}$$



$$f1 \text{ score} = 2 \frac{precision * recall}{precision + recall}$$

CONFUSION MATRIX

- Contoh:

n=165	Predicted: NO	Predicted: YES
Actual: NO	50	10
Actual: YES	5	100

- Tentukan:
 - a. accuracy,
 - b. precision,
 - c. recall,
 - d. specificity,
 - e. f1 score

CONFUSION MATRIX

- Contoh program

```
from sklearn.metrics import classification_report
print (classification_report(y_test, y_pred))
```

ouput

#		precision	recall	f1-score	support
	0	0.98	0.97	0.97	98
	1	0.93	0.96	0.95	45
	accuracy			0.97	143
	macro avg	0.96	0.96	0.96	143
	weighted avg	0.97	0.97	0.97	143

PENGUKURAN PERFORMANSI REGRESI

- MSE (MEAN SQUARE ERROR)
- RMSE (ROOT MEAN SQUARE ERROR)
- MAE (MEAN ABSOLUTE ERROR)
- MBE (MEAN BIAS ERROR)
- MAPE (MEAN ABSOLUTE PERCENTAGE ERROR)
- Dan sebagainya

TUGAS 1

- Kumpulkan program python yang sudah anda praktekan selama perkuliahan machine learning dengan minimal 4 program yang berbeda beserta outputnya.
- Tugas maksimal dikumpulkan Rabu tanggal 23 September 2026 Pukul 23.59.
- Link Pengumpulan pada Ilmu2
- **Aturan penamaan folder:**

NPM_Nama_T1

TUGAS 1

- Aturan penamaan file:

NPM_program1

NPM_program2 dan seterusnya

NPM_output1

NPM_output2 dan seterusnya

TERIMA KASIH