

# MACHINE LEARNING

Nama : Hanif Ridal Warits

NPM : 41155050210060

Kelas : Informatika A2 – 2021

## Tugas Pertemuan 3

### 1.0. Lakukan praktik dari

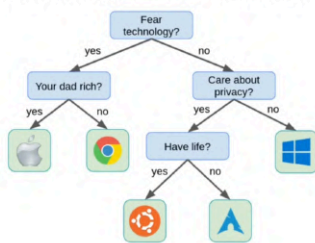
[https://youtu.be/5wwXKtLkyqs?si=fn88eveu\\_qbCC6b3](https://youtu.be/5wwXKtLkyqs?si=fn88eveu_qbCC6b3),

buat screenshot dengan nama kalian pada coding, kumpulkan dalam bentuk pdf, dari kegiatan ini:

### 1.1. Pengenalan komponen Decision Tree: root, node, leaf

#### Konsep Dasar

Terminology: root node, internal node, leaf node



Model machine learning decision tree terdiri dari root, node & leaf, dari kasus diatas :



**CART**  
Classification And Regression Tree merupakan default model yang disediakan sklearn

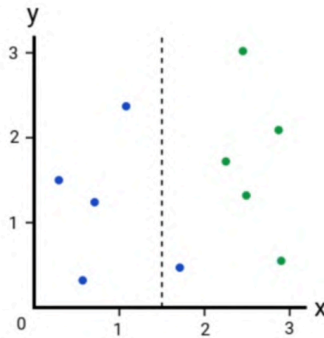
Algoritma Decision Tree lainnya:

- ID3
- C4.5
- C5.0

## 1.2. Pengenalan Gini Impurity

Impurity measure(pengukuran ketidakmurnian)

### Gini Impurity



Ruas Kiri:

$$\begin{aligned} G &= 1 - \sum_i P_i^2 \\ &= 1 - P(\text{biru})^2 \\ &= 1 - \left(\frac{4}{10}\right)^2 = 0 \end{aligned}$$

Ruas Kanan:

$$\begin{aligned} G &= 1 - \sum_i P_i^2 \\ &= 1 - (P(\text{biru})^2 + P(\text{hijau})^2) \\ &= 1 - \left(\left(\frac{1}{6}\right)^2 + \left(\frac{5}{6}\right)^2\right) = 0.278 \end{aligned}$$

Average Gini Impurity:

$$\begin{aligned} G &= \frac{4}{4+6} \times 0 + \frac{6}{4+6} \times 0.278 \\ &= 0.1668 \end{aligned}$$

Gini impurity adalah acuan CART, yang diterapkan

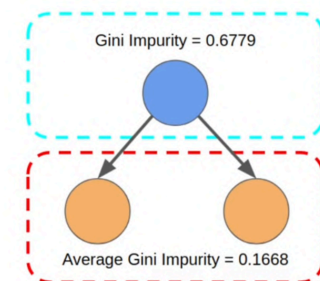
jangkauan nilai 0 & 1

0 = murni yang sempurna

1 = tidak murni

## 1.3. Pengenalan Information Gain

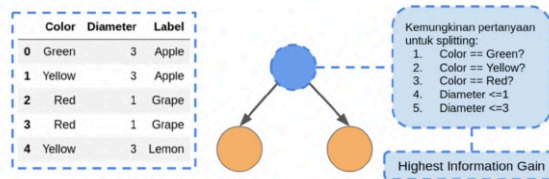
### Information Gain



$$\text{Information Gain} = 0.6779 - 0.1668 = 0.51$$

## 1.4. Membangun Decision Tree

## Membangun Decision Tree



$$\begin{aligned} G &= 1 - (P(\text{apple})^2 + P(\text{grape})^2 + P(\text{lemon})^2) \\ &= 1 - ((\frac{2}{5})^2 + (\frac{2}{5})^2 + (\frac{1}{5})^2) \\ &= 0.63 \end{aligned}$$

## 1.5. Persiapan dataset: Iris Dataset

```
15 SKLearn - DecisionTree.ipynb ☆
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

Decision Tree Classification

Dataset

[1] Suggested code may be subject to a license [AdityaSingh17/MMLAB] IssaHassan/Intro-MachineLearning-with-Python | 15chenst/Sampling_102001647
from sklearn.datasets import load_iris

X, y = load_iris(return_X_y = True)

print(f'Dimensi Features : {X.shape}')
print(f'Class : {set(y)}')
```

Dimensi Features : (150, 4)  
Class : {0, 1, 2}

```
[2] Suggested code may be subject to a license [200009063/Machine_Learning]
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.3, random_state = 0)
```

## 1.6. Training model Decision Tree Classifier

```
Classification dengan DecisionTreeClassifier

[3] from sklearn.tree import DecisionTreeClassifier

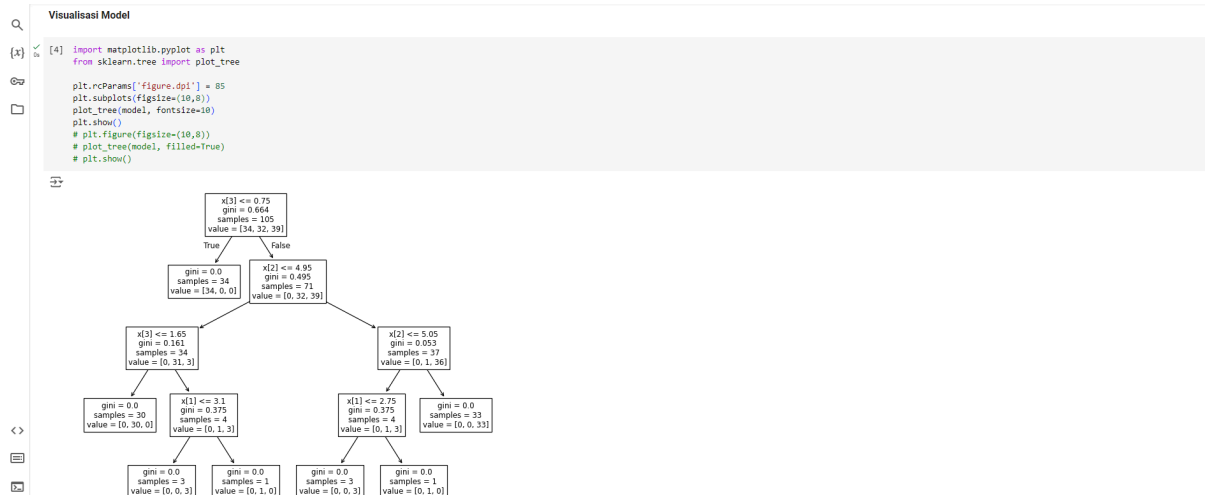
model = DecisionTreeClassifier(max_depth=4)

model.fit(X_train, y_train)
```

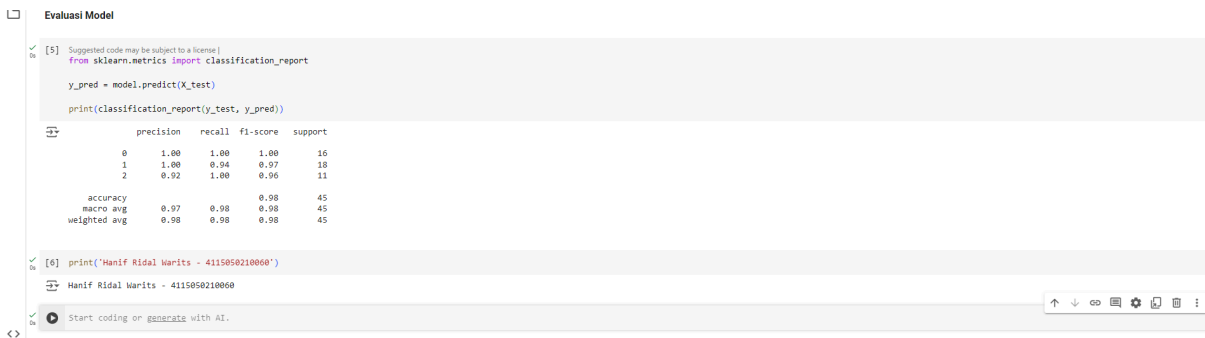
DecisionTreeClassifier

DecisionTreeClassifier(max\_depth=4)

## 1.7. Visualisasi model Decision Tree



## 1.8. Evaluasi model Decision Tree



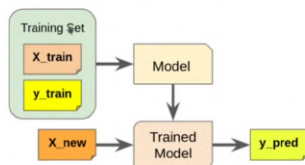
## 2.0. Lakukan praktik dari

[https://youtu.be/yKovaQ6tyV8?si=HnHG6kcoCsDwvo\\_0](https://youtu.be/yKovaQ6tyV8?si=HnHG6kcoCsDwvo_0),

buat screenshot dengan nama kalian pada coding, kumpulkan dalam bentuk pdf, dari kegiatan ini:

## 2.1. Proses training model Machine Learning secara umum

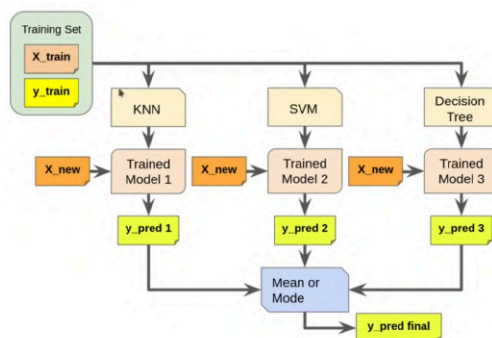
### General ML Model Training



## 2.2. Pengenalan Ensemble Learning

### Ensemble Learning: heterogeneous & homogeneous

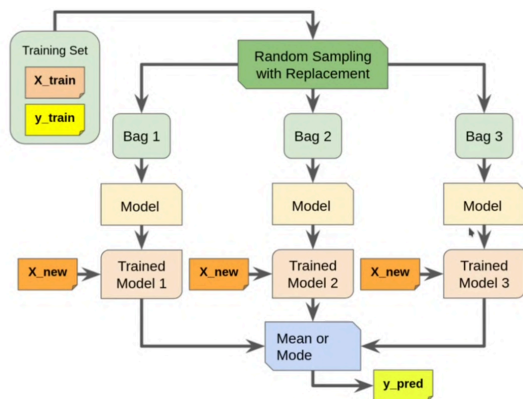
Referensi: [https://en.wikipedia.org/wiki/Ensemble\\_learning](https://en.wikipedia.org/wiki/Ensemble_learning)



## 2.3. Pengenalan Bootstrap Aggregating | Bagging

## Bagging: Bootstrap Aggregating

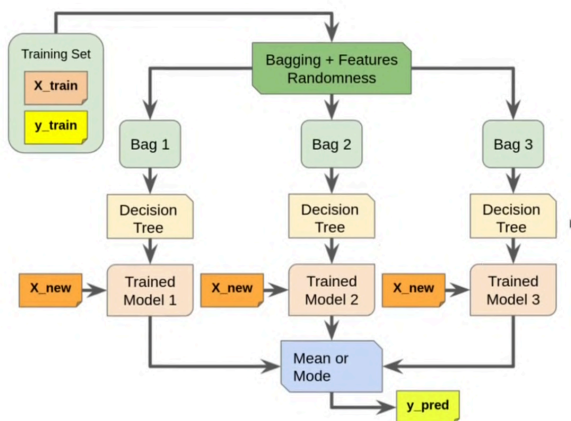
Referensi: [https://en.wikipedia.org/wiki/Bootstrap\\_aggregating](https://en.wikipedia.org/wiki/Bootstrap_aggregating)



## 2.4. Pengenalan Random Forest | Hutan Acak

### Random Forest

Referensi: [https://en.wikipedia.org/wiki/Random\\_forest](https://en.wikipedia.org/wiki/Random_forest)



## 2.5. Persiapan dataset | Iris Flower Dataset

```
16 SKLearn - Random Forest.ipynb
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text
[+]
{x}
Dataset

[3] Suggested code may be subject to a license | TKVuelhaAI/CampoVIP-KNN
from sklearn.datasets import load_iris
X, y = load_iris(return_X_y = True)
print(f'Dimensi Features : {X.shape}')
print(f'Class : {set(y)}')

Dimensi Features : (150, 4)
Class : {0, 1, 2}

[4] from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.3, random_state = 0)
```

## 2.6. Implementasi Random Forest Classifier dengan Scikit Learn

### Classification dengan RandomForestClassifier

```
[5] Suggested code may be subject to a license
from sklearn.ensemble import RandomForestClassifier

model = RandomForestClassifier(n_estimators = 100, random_state = 0)
model.fit(X_train, y_train)
```

```
RandomForestClassifier
RandomForestClassifier(random_state=0)
```

## 2.7. Evaluasi model dengan Classification Report

### Evaluasi Model

```
[6] from sklearn.metrics import classification_report

y_pred = model.predict(X_test)

print(classification_report(y_test, y_pred))
```

	precision	recall	f1-score	support
0	1.00	1.00	1.00	16
1	1.00	0.94	0.97	18
2	0.92	1.00	0.96	11
accuracy			0.98	45
macro avg	0.97	0.98	0.98	45
weighted avg	0.98	0.98	0.98	45

```
[7] print('Hanif Ridal Warits - 41155050210000')
```

```
Hanif Ridal Warits - 41155050210000
```