Nama : Triansyah Amarullah Ahmad Prayoga

NPM : 41155050210034

Kelas : TIF-A2

Matkul : Machine Learning

TUGAS 3

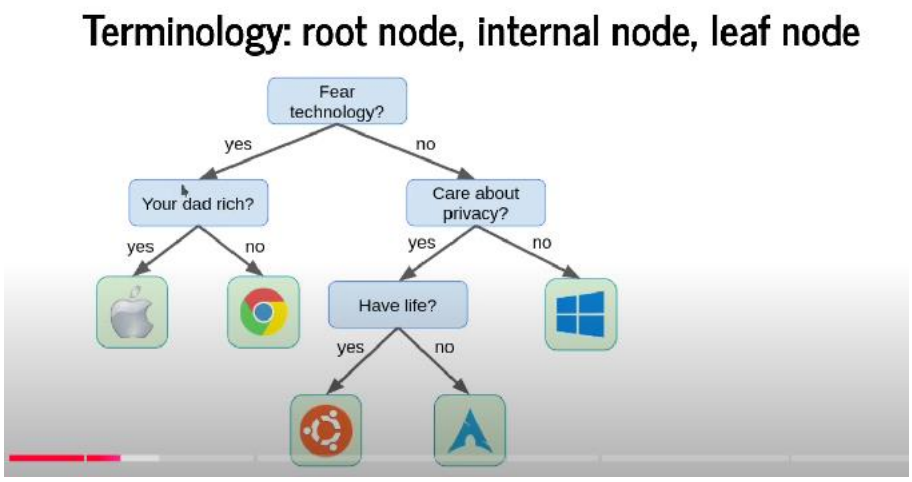
Tugas Pertemuan 3

1.0. Lakukan praktik dari https://youtu.be/5wwXKtLkyqs?si=fn88eveu_qbCC6b3 , buat screenshot dengan nama kalian pada coding, kumpulkan dalam bentuk pdf, dari kegiatan ini:

```
print('Triansyah Amarullah Ahmad Prayoga', '41155050210034')
```

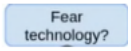
```
Triansyah Amarullah Ahmad Prayoga 41155050210034
```

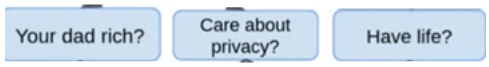
1.1. Pengenalan komponen Decision Tree: root, node, leaf



Gambar diatas merupakan pohon yang terbalik, root node sebagai akar yang paling atas bertuliskan ‘fear technology?’ internal node berisi pertanyaan lanjutan jika pertanyaan dijawab ya menuju ke kiri sedangkan dijawab no menuju ke kanan. Diatas internal node ada 3 yaitu ‘Your dad rich?’ , ‘Care about privacy’, dan ‘Have life?’. Sedangkan leaf node adalah daunnya seperti apple, google, linux, windows.

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

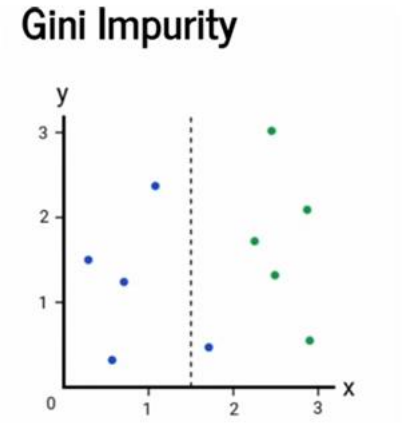
root : 

node : 

leaf : 

Algoritma Decision Tree lainnya:
- ID3
- C4.5
- C5.0

1.2. Pengenalan Gini Impurity



Ruas Kiri:

$$G = 1 - \sum_{i}^{n} P_i^2$$
$$= 1 - P(biru)^2$$
$$= 1 - (\frac{4}{4})^2 = 0$$

Ruas Kanan:

$$G = 1 - \sum_{i}^{n} P_i^2$$
$$= 1 - (P(biru)^2 + P(hijau)^2)$$
$$= 1 - ((\frac{1}{6})^2 + (\frac{5}{6})^2) = 0.278$$

Average Gini Impurity:

$$G = \frac{4}{4+6} \times 0 + \frac{6}{4+6} \times 0.278$$
$$= 0.1668$$

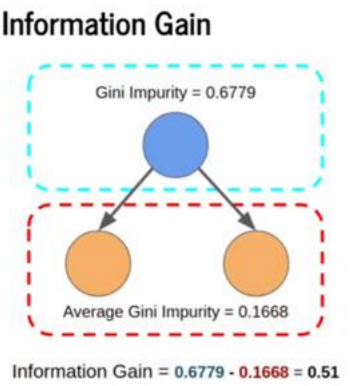
Gini impurity adalah acuan CART, yang diterapkan

jangkauan nilai 0 & 1

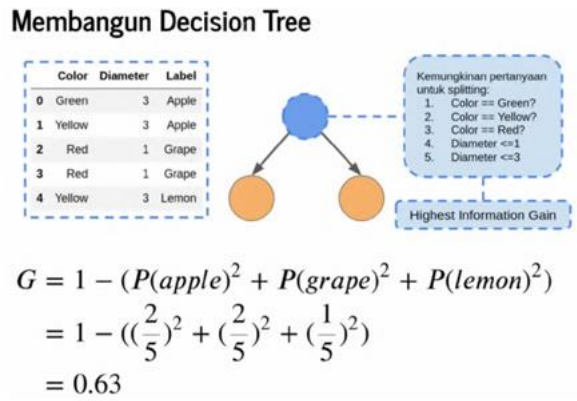
0 =murni yang sempurna

1 = tidak murni

1.3. Pengenalan Information Gain



1.4. Membangun Decision Tree



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

1.5. Persiapan dataset: Iris Dataset

```
[1]  Suggested code may be subject to a license | AdityaSingh17/MLLAB | IssaHassan/Intro-MachineLearning-with-Python | 15cherish/Sampling_102003647
     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 | 2000090063/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

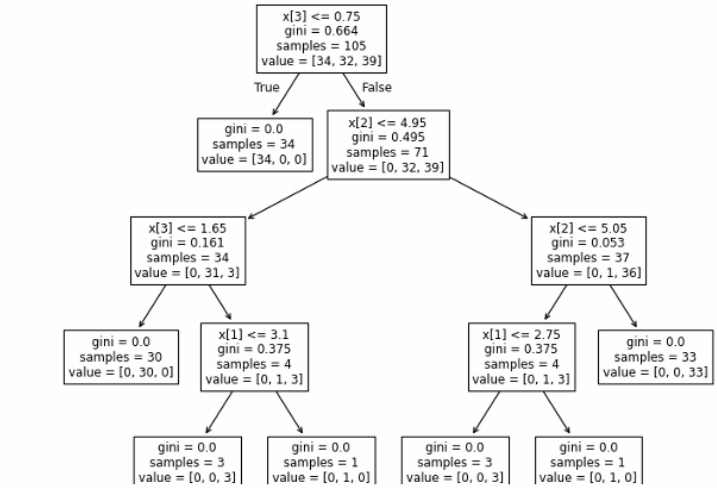
Classification dengan DecisionTreeClassifier

```
[3]  from sklearn.tree import DecisionTreeClassifier

     model = DecisionTreeClassifier(max_depth=4)

     model.fit(X_train, y_train)
```

```
DecisionTreeClassifier(max_depth=4)
```

1.7. Visualisasi model Decision Tree

```
[4]  import matplotlib.pyplot as plt
     from sklearn.tree import plot_tree

     plt.rcParams['figure.dpi'] = 85
     plt.subplots(figsize=(10,8))
     plot_tree(model, fontsize=10)
     plt.show()
     # plt.figure(figsize=(10,8))
     # plot_tree(model, filled=True)
     # plt.show()
```



1.8. Evaluasi model Decision Tree

```
print('Triansyah Amarullah Ahmad Prayoga', '41155050210034')
```

```
Triansyah Amarullah Ahmad Prayoga 41155050210034
```

```
[5]  Suggested code may be subject to a license |
     from sklearn.metrics import classification_report

     y_pred = model.predict(X_test)

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

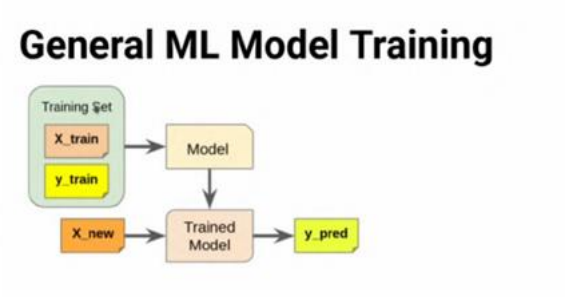
```
              precision    recall  f1-score   support

           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
```

2.0. Lakukan praktik dari https://youtu.be/yKovaQ6tyV8?si=HnHG6kcoCsDwvo_0 , buat screenshot dengan nama kalian pada coding, kumpulkan dalam bentuk pdf, dari kegiatan ini:

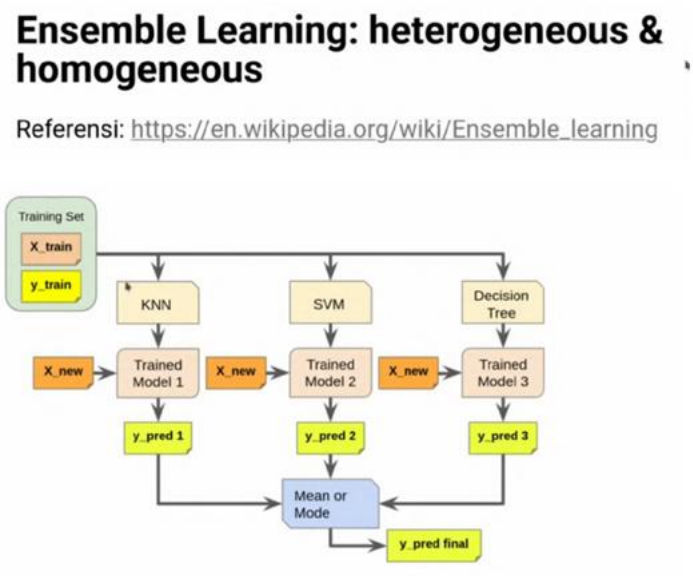
```
print('Triansyah Amarullah Ahmad Prayoga', '41155050210034')
```

```
Triansyah Amarullah Ahmad Prayoga 41155050210034
```

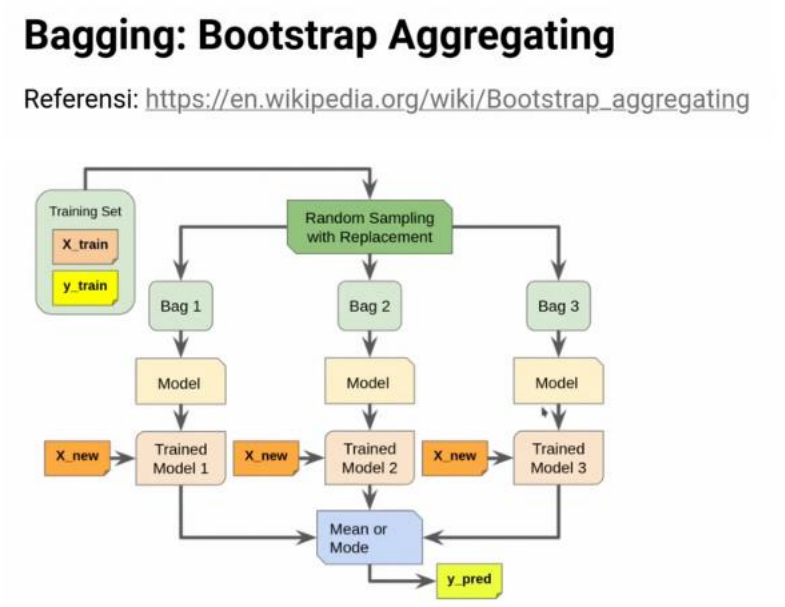
2.1. Proses training model Machine Learning secara umum



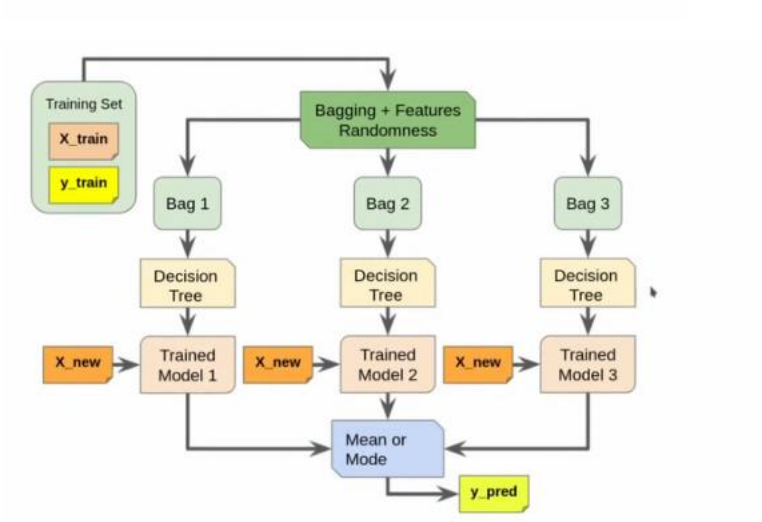
2.2. Pengenalan Ensemble Learning



2.3. Pengenalan Bootstrap Aggregating | Bagging



2.4. Pengenalan Random Forest | Hutan Acak



2.5. Persiapan dataset | Iris Flower Dataset

```
[3] Suggested code may be subject to a license | TKVueltasAlCampo/VIP-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

```
[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

```
[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
```