



Decision Tree

Tim Pengajaran

Mata Kuliah **Machine Learning**

Jurusan Teknologi Informasi Tahun 2022



Disclaimer

- Materi presentasi ini termasuk contoh, gambar, dan referensi diberikan **hanya sebagai informasi dan penunjang pembelajaran**
- Penggunaan merek dan nama perusahaan pada materi ini mungkin mengandung hak kekayaan intelektual yang dilindungi oleh undang-undang dan **sepenuhnya dimiliki pemegang hak kekayaan intelektual**
- Kredit diberikan kepada pemilik lisensi terhadap aset-aset *open-source* yang digunakan pada materi ini sesuai dengan lisensi *open-source* yang digunakan

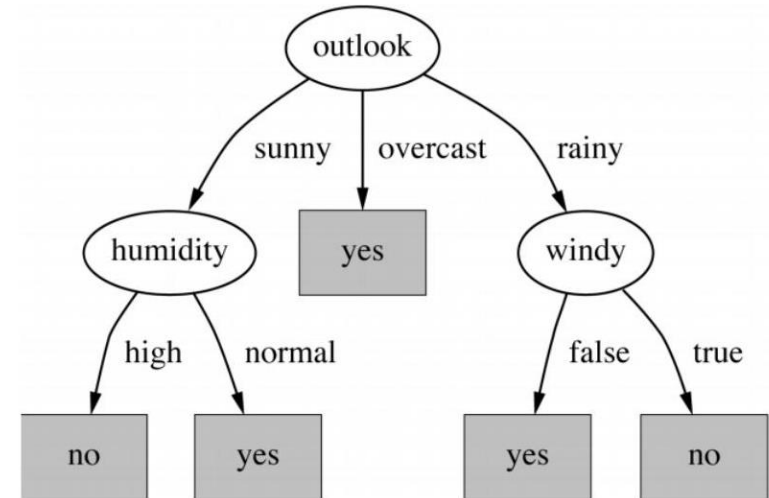


Outlines

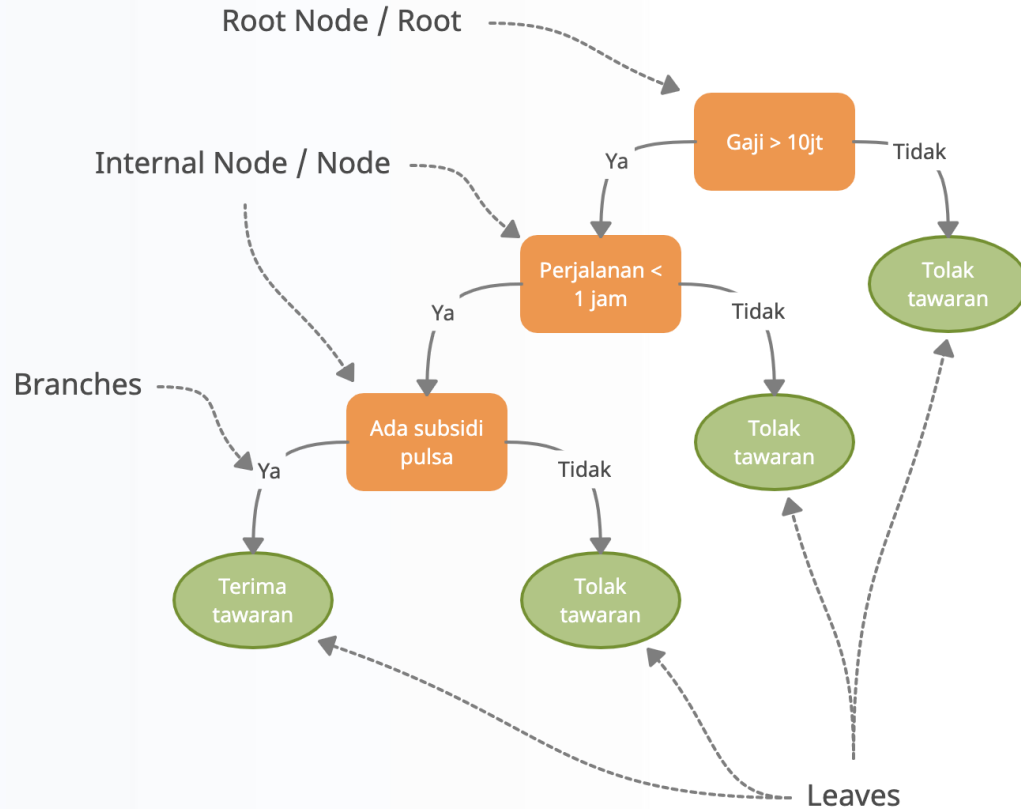
- Apa itu **"Decision Tree"**
- Intuisi Perhitungan "Decision Tree"
- Proses Perhitungan "Decision Tree"
 - **Gini Impurity**
 - **Entropy**
- Implementasi "Decision Tree" dengan Scikit-learn

Apa itu Decision Tree?

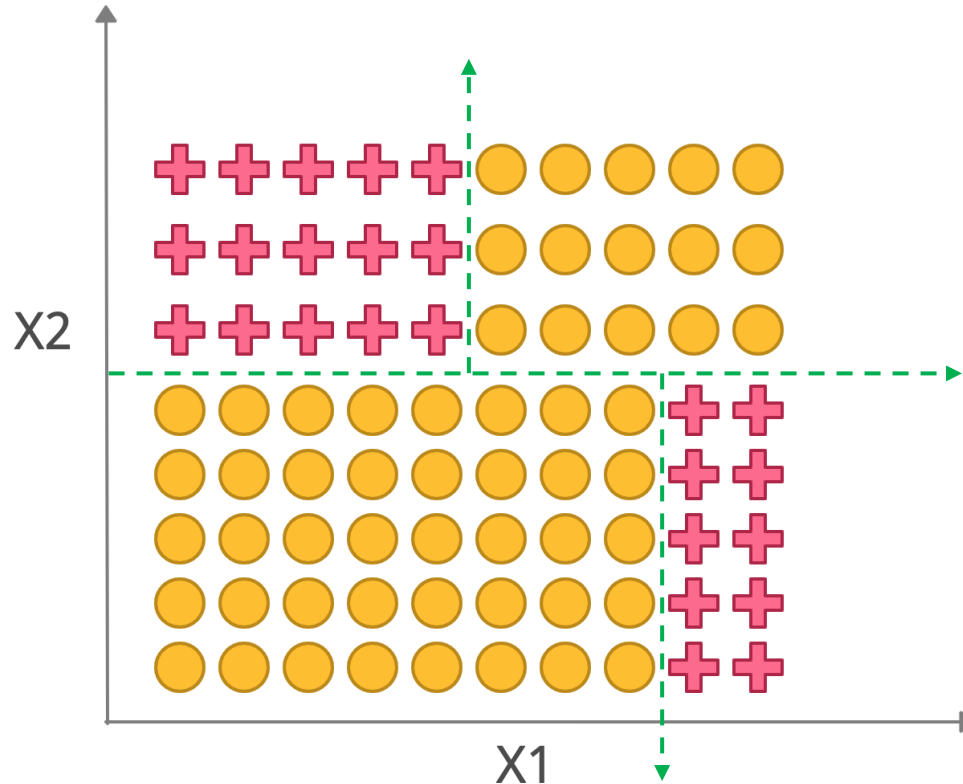
- Model **serupa graf berbentuk pohon** yang memodelkan sebuah keputusan
- Tujuannya untuk memetakan target dengan pembelajaran sederhana berdasarkan nilai fitur
- Bersifat *non-parametrik* / Tidak bergantung pada distribusi probabilitas



Istilah dalam Decision Tree



Intuisi Model Decision Tree



Proses Perhitungan Decision Tree #1



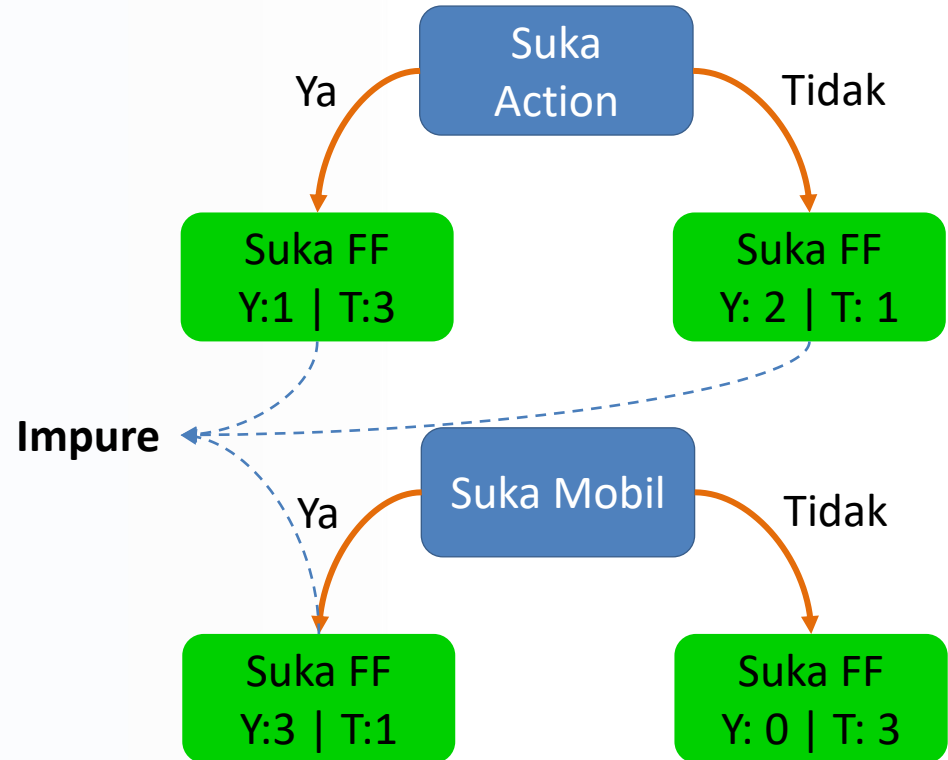
Suka Action	Suka Mobil	Umur	Suka FF
Ya	Ya	7	Tidak
Ya	Tidak	12	Tidak
Tidak	Ya	18	Ya
Tidak	Ya	35	Ya
Ya	Ya	38	Ya
Ya	Tidak	50	No
Tidak	Tidak	83	No

Langkah pertama – Think! Variabel apa yang harus diletakkan pada root?

Proses Perhitungan Decision Tree #2

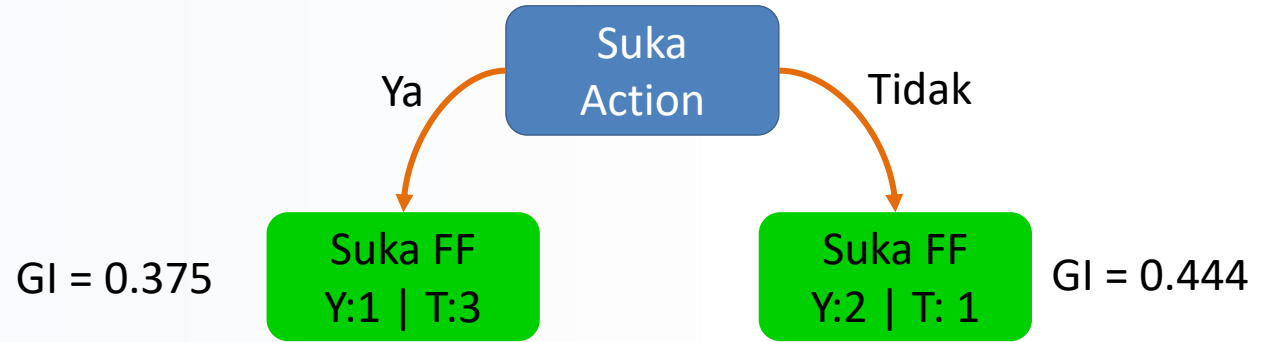
Cek seberapa baik variabel memisahkan target berdasarkan nilai variabel

Suka Action	Suka Mobil	Umur	Suka FF
Ya	Ya	7	Tidak
Ya	Tidak	12	Tidak
Tidak	Ya	18	Ya
Tidak	Ya	35	Ya
Ya	Ya	38	Ya
Ya	Tidak	50	Tidak
Tidak	Tidak	83	Tidak



Proses Perhitungan Decision Tree #3

Kalkulasi *impurity* dengan “Gini Impurity”

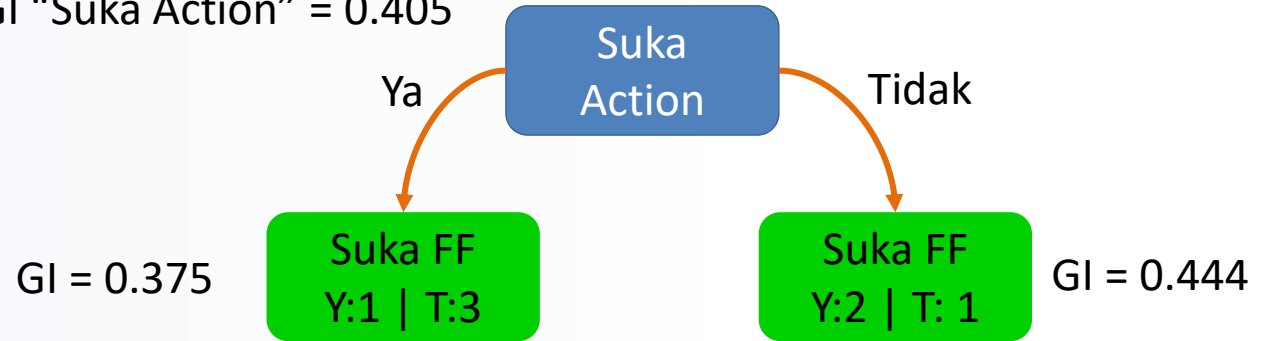


$$\begin{aligned} \text{Gini Impurity leaf} &= 1 - (\text{Probabilitas "ya"})^2 - (\text{Probabilitas "tidak"})^2 \\ &= 1 - \left(\frac{1}{1+3}\right)^2 - \left(\frac{1}{2+1}\right)^2 \\ &= 0.375 \end{aligned}$$

Proses Perhitungan Decision Tree #4

Membobotkan "GI" setiap leaf untuk mendapatkan "GI" Suka Action

GI "Suka Action" = 0.405

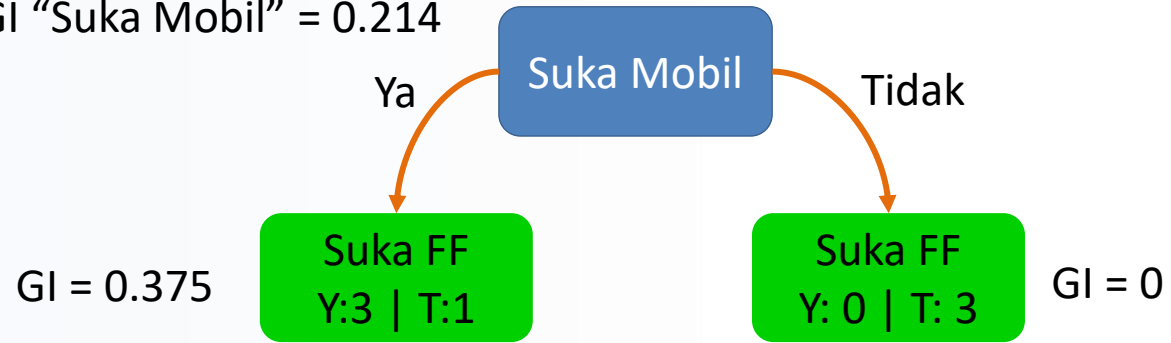


$$GI \text{ "Suka Action"} = \left(\frac{4}{4 + 3} \right) 0.375 - \left(\frac{3}{4 + 3} \right) 0.444$$

Proses Perhitungan Decision Tree #5

Lakukan teknik yang sama untuk "Suka Mobil"

GI "Suka Mobil" = 0.214



$$\begin{aligned} GI \text{ "Suka Mobil"} &= \left(\frac{4}{4 + 3} \right) 0.375 - \left(\frac{3}{4 + 3} \right) 0 \\ &= 0.214 \end{aligned}$$

Proses Perhitungan Decision Tree #6

Bagaimana dengan “Umur”?

Suka Action	Suka Mobil	Umur	Suka FF
Ya	Ya	7	Tidak
Ya	Tidak	12	Tidak
Tidak	Ya	18	Ya
Tidak	Ya	35	Ya
Ya	Ya	38	Ya
Ya	Tidak	50	No
Tidak	Tidak	83	No

9.5
15
26.5
36.5
44
66.5

Adjacent
Average

$GI < 9.5 = 0.429$

Umur < 9.5

Ya

Tidak

$GI = 0$

Suka FF
Y:0 | T:1

$GI = 0.5$

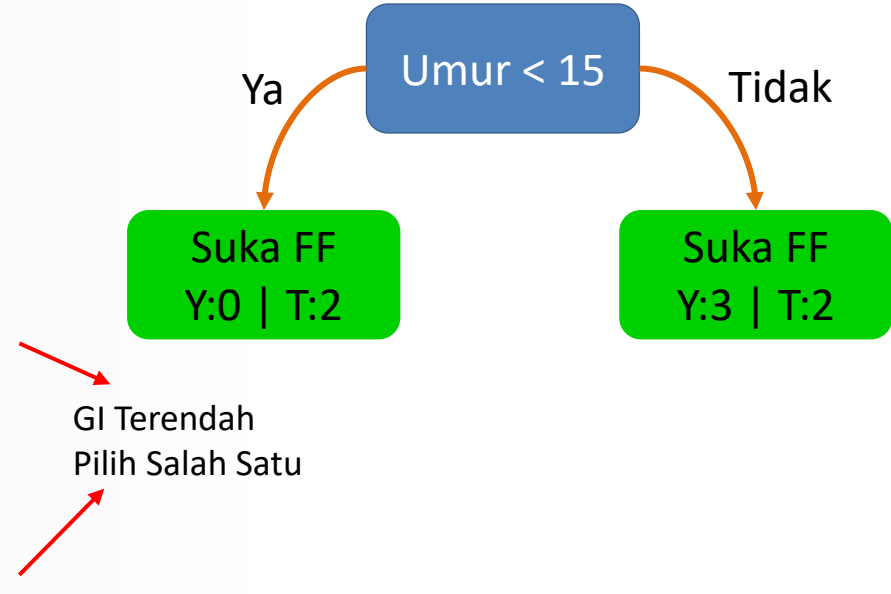
Suka FF
Y:3 | T: 3

Hitung semua GI untuk
setiap Adjacent Average

Proses Perhitungan Decision Tree #7

GI Setiap Adjacent “Umur”?

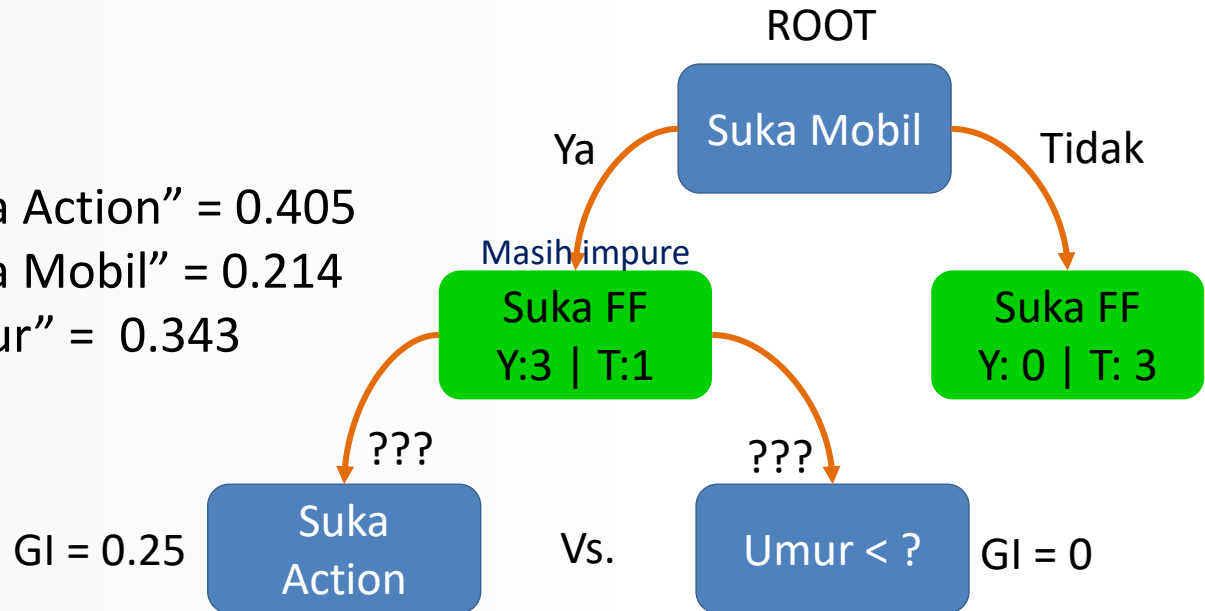
Suka Action	Suka Mobil	Umur	Suka FF	GI Adj
Ya	Ya	7	Tidak	0.429
Ya	Tidak	12	Tidak	0.343
Tidak	Ya	18	Ya	0.476
Tidak	Ya	35	Ya	0.476
Ya	Ya	38	Ya	0.343
Ya	Tidak	50	No	0.429
Tidak	Tidak	83	No	



Proses Perhitungan Decision Tree #8

Bandingkan GI setiap variabel untuk menentukan root

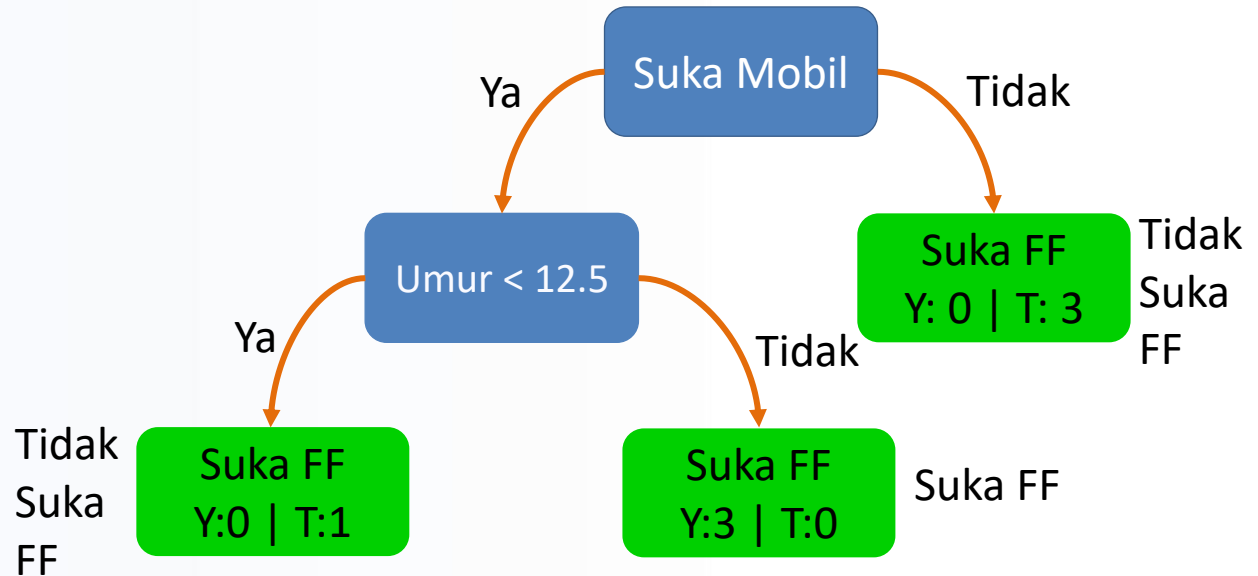
GI "Suka Action" = 0.405
GI "Suka Mobil" = 0.214
GI "Umur" = 0.343



Ulangi proses pengecekan GI untuk "Suka Action" dan "Umur" dengan konteks suka dengan mobil

Proses Perhitungan Decision Tree #9

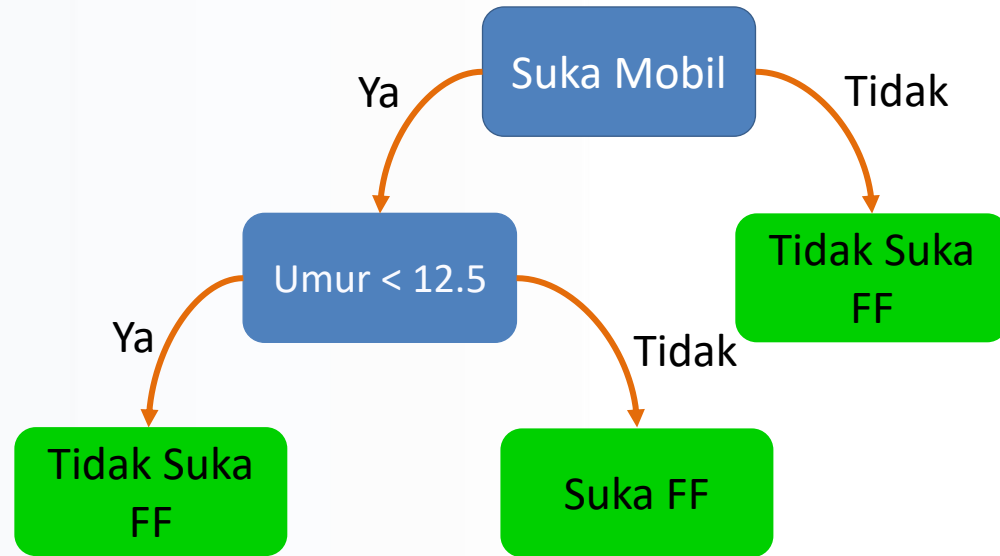
GI “Umur” lebih kecil dibandingkan “Suka Action”



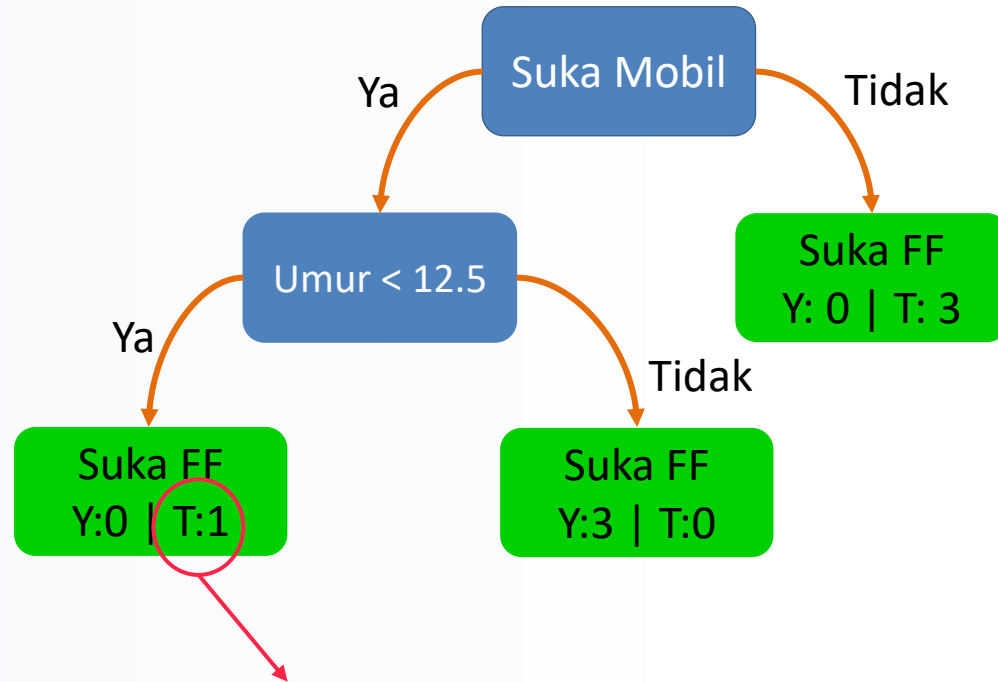
Sudah tidak ditemukan "impurity", tidak perlu di split lagi

Proses Perhitungan Decision Tree #10

Kesimpulan



Permasalahan Pada Decision Tree



Hanya ada 1 sampel → Lack of confidence / lead to overfit → Solusi: Pruning atau menentukan min sampel per leaf



Gini Impurity: Math

$$Gini(t) = 1 - \sum_{i=1}^j P(i|t)^2$$

- j adalah jumlah kelas.
- t subset dari node.
- $P(i|t)$ probabilitas dari kelas i



Thank You for Today
Always Keep Your Spirit!