

## No. — Decision Tree Algoritma ID3 —

Date: 08 Okt 2024

Waktu	No.	Paket	Frekwensi	Prioritas	Gangguan	
PE	1	Besar	Sedang.	Rendah	G.	
✓ PE	2	✓ Kecil	✓ Tinggi	Rendah ✓	N	
PE	3	Kecil	Sedang.	Tinggi	G	■ PE = Pendek
✓ PE	4	✓ Kecil	✓ Tinggi	Rendah ✓	N	■ PA = Panjang
PE	5	Kecil	Sedang.	Tinggi	N	
PA	6	Besar	Sedang.	Rendah	N	
PA	7	✓ Kecil	✓ Tinggi	✓ Tinggi	G.	■ G = gangguan
PE	8	Besar	Sedang.	Rendah	N	N = normal
PA	9	Kecil	Rendah	Tinggi	N	
✓ PE	10	✓ Kecil	✓ Tinggi	✓ Tinggi	N	
PA	11	✓ Besar	✓ Tinggi	✓ Tinggi	N	
PA	12	Kecil	Rendah	Tinggi	N	

1. Buat decision tree dan rule untuk mendekripsi adanya gangguan pada jaringan komputer!

a. ⇒ Entropy Gangguan

$$P(G) = 3/12 = 1/4 = 0,25$$

$$P(N) = 9/12 = 3/4 = 0,75$$

Maka entropy gangguan:  $- P(\ominus) \log_2 P(\ominus) - P(\oplus) \log_2 P(\oplus)$

$$= -0,25 \log_2 (0,25) - 0,75 \log_2 (0,75)$$

$$= [(-0,25 \times -2)] - (0,75 \times (-0,415))$$

$$= 0,5$$

$$= +0,81125 \rightarrow \text{Entropy parental}$$

## b) Entropy Waktu & Information Gain Waktu

$$P(\text{Pendek}) = 7/12 = 0,583 \rightarrow 0,584$$

$$P(\text{Panjang}) = 5/12 = 0,416$$

$$* P(\text{Pendek, Gangguan} : Y_1) = 2/7 = 0,285$$

$$P(\text{Pendek, Normal} : Y_2) = 5/7 = 0,715$$

$$\begin{aligned}\text{Entropy waktu pendek} &= -0,715 \log_2(0,715) - 0,285 \log_2(0,285) \\ &= (-0,715 \times -0,483) - 0,285 (-1,01) \\ &= 0,861 //\end{aligned}$$

$$* P(\text{Panjang, gangguan}) = 1/5 = 0,2$$

$$P(\text{Panjang, normal}) = 4/5 = 0,8$$

$$\begin{aligned}\text{Entropy waktu panjang} &= -(0,2 \times \log_2(0,2) + 0,8 \times \log_2(0,8)) \\ &= -(0,2(-2,321) + 0,8(-0,321)) \\ &= 0,4642 \quad 0,4642 //\end{aligned}$$

$$* \text{Entropy Waktu} = P(\text{Pendek}) \times E.\text{Pendek} + P(\text{panjang}) \times E.\text{panjang}$$

$$\begin{aligned}&\xrightarrow{7/12 \times 0,861 + 5/12 \times 0,4642} 0,50225 + 0,3 \\ * \text{Information Gain Waktu} &= \text{Entropy gangguan} - \text{Entropy waktu} \\ &= 0,861 - 0,50225 \\ &= 0,359\end{aligned}$$

## c) Entropy Dan Information Gain Paket

$$P(\text{Besar}) = 4/12 = 1/3 = 0,33$$

$$P(\text{Kecil}) = 8/12 = 2/3 = 0,67$$

$$* P(\text{Besar, Gangguan}) = 1/4 = 0,25$$

$$P(\text{Besar, Normal}) = 3/4 = 0,75$$

$$\begin{aligned}\text{Maka entropy paket besar: } & (0,25 \times \log_2 10 + 0,75 \log_2 (0,75)) \\ &= -(0,25(-2) + 0,75(-0,415)) \\ &= 0,81125 //\end{aligned}$$

$$* P(\text{Kecil, Gangguan}) = 2/8 = 1/4 = 0,25$$

$$P(\text{Kecil, Normal}) = 6/8 = 3/4 = 0,75$$

$$\begin{aligned}\text{Maka entropy paket kecil: } & -(0,25(-2) + 0,75(-0,415)) \\ &= 0,81125\end{aligned}$$

## \* Entropy paket (Gabungan)

$$P(\text{Besar}) \times E.\text{Besar} + P(\text{Kecil}) \times E.\text{Kecil}$$

$$= (0,33 \times 0,81125) + (0,67 \times 0,811)$$

$$\approx 0,811$$

## \* Information Gain Paket

$$\text{Entropy Gangguan} - \text{Entropy paket}$$

$$= 0,861 - 0,811$$

$$= 0$$

8.0) Entropy & I(G) Frekwensi

$$P(\text{Tinggi}) = 5/12 = 0,417$$

$$P(\text{Sedang}) = 5/12 = 0,417$$

$$P(\text{Rendah}) = 2/12 = 1/6 = 0,16$$

$$\star P(\text{Tinggi, Gangguan}) = 1/5 = 0,2$$

$$P(\text{Tinggi, Normal}) = 4/5 = 0,8$$

Maka entropy frekwensi tinggi

$$= -(0,2 \log_2(0,2) + 0,8 \log(0,8))$$

$$= 0,721,$$

$$\star P(\text{Sedang, Gangguan}) = 2/5 = 0,4$$

$$P(\text{Sedang, Normal}) = 3/5 = 0,6$$

Maka entropy frekwensi sedang

$$= -(0,4 \log(0,4) + 0,6 \log_2(0,6))$$

$$= -(0,4(-1,321) + 0,6(-0,736))$$

$$= 0,197$$

$$\star P(\text{Rendah, Gangguan}) = 0/2 = 0$$

$$P(\text{Rendah, Normal}) = 2/2 = 1$$

Maka entropy frek. rendah

$$= -(0 \log_2(0) + 1 \log_2(1))$$

$$= 0$$

\* Entropy Frekwensi (Gabungan)

$$= (0,42 \times 0,721) + (0,42 \times 0,197) + (0,16 \times 0)$$

$$= 0,30282 + 0,4074$$

$$= 0,70922 //$$

\* Information Gain Frekwensi

$$= 0,811 - 0,709$$

$$= 0,107$$

e) Entropy dan Information Gain Prioritas

$$P(\text{Tinggi}) = 7/12 = 0,583$$

$$P(\text{Rendah}) = 5/12 = 0,416$$

$$* P(\text{Prioritas Tinggi}, \text{Gangguan}) = 2/7 = 0,285$$

$$P(\text{Tinggi}, \text{Normal}) = 5/7 = 0,715$$

Entropy prioritas tinggi

$$= -(0,285 \log_2(0,285) + 0,715 \log_2(0,715))$$

$$= -(0,285(-1,81) + 0,715(-0,483))$$

$$= 0,861$$

$$* P(\text{Rendah}, \text{Gangguan}) = 1/5 = 0,2$$

$$P(\text{Rendah}, \text{Normal}) = 4/5 = 0,8$$

Entropy prioritas rendah

$$= -(0,2 \log_2(0,2) + 0,8 \log_2(0,8))$$

$$= -(0,2(-2,321) + 0,8(-0,321))$$

$$= 0,721$$

$$* \text{Entropy Prioritas (gabungan)} = 0,802$$

\* Information Gain Prioritas

$$= 0,811 - 0,802$$

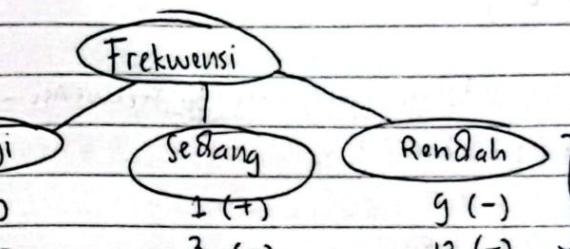
$$= 0,009$$

f) Information Gain seluruh fitur

Information Gain	Nilai
Waktu	0,009
Paket	0
Frekuensi	0,107
Prioritas	0,009

Maka IG bernilai terbesar adalah milik fitur "Frekuensi".

Kemudian fitur "Frekuensi" akan menjadi "Root".



Tidak relevan.

Karena "Tinggi" dan "Sedang" memiliki (+) dan (-) maka memiliki leaf node.

g) Menghitung lagi entropy "Frekuensi Tinggi" & "Frekuensi Sedang".

Frekuensi,

Total data  
frekuensi sedang

\*\* Entropy

- Waktu

Banyaknya data  
waktu panjang sedang  
Banyaknya data  
waktu singkat sedang

\*\* Entropy Frekuensi Sedang - Panjang  $\rightarrow P = 1/5$

$P(\text{Sedang, Panjang, Gangguan}) = 0$   $\rightarrow$  Entropy frekuensi  
 $P(\text{Sedang, Panjang, Normal}) = 1$   $\rightarrow$  tinggi-waktu panjang

$\ominus$  = tidak = normal  
 $\oplus$  = ya = gangguan

No:

Date:

\*\* Entropy sedang - pendek  $\rightarrow P = 4/5$

$$P(\text{Sedang, Pendek, Gangguan}) = 2/4 = 1/2 = 0,5$$

$$P(\text{Sedang, pendek, normal}) = 2/4 = 0,5$$

Entropy Frekuensi Sedang - Waktu Pendek

$$= -(0,5 \log_2(0,5) + 0,5 \log_2(0,5))$$

$$= -(0,5(-1) + 0,5(-1))$$

$$= 1,$$

\*\* Entropy Total (Frekuensi = Waktu)

$$= \left(\frac{1}{5} \times 0\right) + \left(\frac{4}{5} \times 1\right)$$

$$= 0 + 4/5$$

$$= 0 + 0,8$$

$$= 0,8,,$$

\*\* Information Gain (Frekuensi - Waktu)

$$= \text{Entropy (Frekuensi)} - \text{Entropy Total (Frekuensi - Waktu)}$$

$$= 0,704 - 0,8$$

$$= -0,006,,$$

No:

Date:

\*\*\* Entropy Frekuensi - Paket

$$\text{*** Entropy Sedang - Besar} \rightarrow P = 3/5 = 0,6$$

$$P(\text{Sedang, Besar, Gangguan}) = 1/3 = 0,33$$

$$P(\text{Sedang, Besar, Normal}) = 2/3 = 0,67$$

Maka Entropy (Sedang, Besar)

$$= -(0,67 \log_2(0,67) + 0,33 \log_2(0,33))$$

$$= -(0,67(-0,57) + 0,33(-1,59))$$

$$= 0,9066,,$$

\*\*\* Entropy sedang - Kecil  $\rightarrow P = 2/5 = 0,4$

$$P(\text{Sedang, Kecil, Gangguan}) = 1/2 = 0,5$$

$$P(\text{Sedang, Kecil, Normal}) = 1/2 = 0,5$$

Maka Entropy (Sedang, Normal)

$$= -(0,5 \log_2(0,5) + 0,5 \log_2(0,5))$$

$$= 1,,$$

\*\*\* Entropy Total (Frekuensi, Paket)

$$= (0,6 \times 0,9066) + (0,4 \times 1)$$

$$= 0,54.. + 0,4$$

$$= 0,943,,$$

\*\*\* Information Gain (Frekuensi, Paket)

$$= \text{Entropy (Frekuensi)} - \text{Entropy Total (Frekuensi, Paket)}$$

$$= 0,704 - 0,943$$

$$= -0,236$$

OKAY

OKAY

### \*\*\*\* Entropy Frekuensi - Prioritas

$$**** \text{ Entropy Sedang, Tinggi} \rightarrow P = 2/5 = 0,4$$

- $P(\text{Sedang, Tinggi, Gangguan}) = 1/2 = 0,5$
- $P(\text{Sedang, Tinggi, Normal}) = 1/2 = 0,5$

Maka Entropy (Sedang, Tinggi)

$$= -(0,5 \log_2(0,5) + 0,5 \log_2(0,5))$$

$$= 1$$

$$**** \text{ Entropy Sedang, Rendah} \rightarrow P = 3/5 = 0,6$$

$$P(\text{Sedang, Rendah, Gangguan}) = 1/3 = 0,33$$

$$P(\text{Sedang, Rendah, Normal}) = 2/3 = 0,67$$

Maka Entropy (Sedang Rendah)

$$= -(0,67 \log_2(0,67) + 0,33 \log_2(0,33))$$

$$= -(0,67(-0,57) + 0,33(-1,59))$$

$$= 0,9066$$

### \*\*\*\* Entropy Total (Frekuensi, Prioritas)

$$= (0,4 \times 1) + (0,6 \times 0,9066)$$

$$= 0,4 + 0,543$$

$$= 0,943$$

### \*\*\*\* Information Gain (Frekuensi, Prioritas)

$$= 0,704 - 0,943$$

$$= -0,239$$

### \*\* Entropy Frekuensi - Waktu

$$** \text{ Frekuensi Tinggi - Waktu Panjang} (P = \frac{2}{5}) \rightarrow 0,4$$

- $\text{Peluang (Frekuensi Tinggi, Waktu Panjang, Gangguan)} = 1/2 = 0,5$

- $\text{Peluang (Frekuensi Tinggi, Waktu Panjang, Normal)} = 1/2 = 0,5$

Maka entropy Frekuensi Tinggi - Waktu Panjang adalah

$$= -(0,5 \log_2(0,5) + 0,5 \log_2(0,5))$$

$$= 1$$

### \*\* Frekuensi Tinggi - Waktu Pendek ( $P = \frac{3}{5}$ ) $\rightarrow 0,6$

- $\text{Peluang (Frekuensi Tinggi, Waktu Pendek, Gangguan)} = 0/3 = 0$

- $\text{Peluang (Frekuensi Tinggi, Waktu Pendek, Normal)} = 3/3 = 1$

Entropy Frekuensi Tinggi - Waktu Pendek

$$= -(1 \log(1) + 0 \log_2(0))$$

$$= 0$$

### \*\* Entropy Total (Frekuensi - Waktu)

$$= (0,4 \times 1) + (0,6 \times 0)$$

$$= 0,4$$

### \*\* Information Gain (Frekuensi - Waktu)

$$= 0,704 - 0,4$$

$$= 0,304$$

### \*\*\* Entropy Frekuensi - Paket

\*\*\* Frekuensi Tinggi - Paket Besar ( $P = 1/5 = 0,2$ )

$$\circ \text{Peluang (Frek. Tinggi, Paket Besar, Gangguan)} = 0,2 = 0$$

$$\circ \text{Peluang (Frek. Tinggi, Paket Besar, Normal)} = 2/5 = 0,4$$

◦ Entropy Frekuensi Tinggi - Paket Besar

$$= -(1 \log_2(1) + 0 \log_2(0))$$

$$= 0$$

\*\*\* Frekuensi Tinggi - Paket Kecil ( $P = 4/5 = 0,8$ )

$$\circ \text{Peluang (Frek. Tinggi, Paket Kecil, Gangguan)} = 1/4 = 0,25$$

$$\circ \text{Peluang (Frek. Tinggi, Paket Kecil, Normal)} = 3/4 = 0,75$$

◦ Entropy Frekuensi Tinggi - Paket Kecil

$$= -(0,75 \log_2(0,75) + 0,25 \log_2(0,25))$$

$$= -(0,75(-0,415) + 0,25(-2))$$

$$= 0,8112$$

\*\*\* Entropy Total (Frekuensi - Paket)

$$= (0,2 \times 0) + (0,8 \times 0,8112)$$

$$= 0 + (0,64896)$$

$$= 0,64896 \approx 0,648$$

\*\*\* Information Gain (Frekuensi - Paket)

$$< 0,704 - 0,648$$

$$> 0,064 \\ 0,056 //$$

### \*\*\* Entropy Frekuensi - Prioritas Tinggi

\*\*\* Frekuensi Tinggi - Prioritas ( $P = 3/5 = 0,6$ )

$$\circ \text{Peluang (Frek. Tinggi, Prioritas Tinggi, Gangguan)} = 1/3 = 0,33$$

$$\circ \text{Peluang (Frek. Tinggi, Prioritas Tinggi, Normal)} = 2/3 = 0,67$$

◦ Entropy Frekuensi Tinggi - Prioritas Tinggi

$$= -(0,67 \log_2(0,67) + 0,33 \log_2(0,33))$$

$$= -(0,67(-0,57) + 0,33(-1,59))$$

$$= 0,90$$

\*\*\* Frekuensi Tinggi - Prioritas Rendah ( $P = 2/5 = 0,4$ )

$$\circ \text{Peluang (Frek. Tinggi, Prioritas Rendah, Gangguan)} = 0/2 = 0$$

$$\circ \text{Peluang (Frek. Tinggi, Prioritas Rendah, Normal)} = 2/2 = 1$$

◦ Entropy Frekuensi Prioritas Rendah - Prioritas Tinggi

$$= -(1 \log_2(1) + 0 \log_2(0))$$

$$= 0$$

\*\*\* Entropy Total (Frekuensi - Prioritas)

$$= (0,6 \times 0,90) + (0,4 \times 0)$$

$$= 0,54 //$$

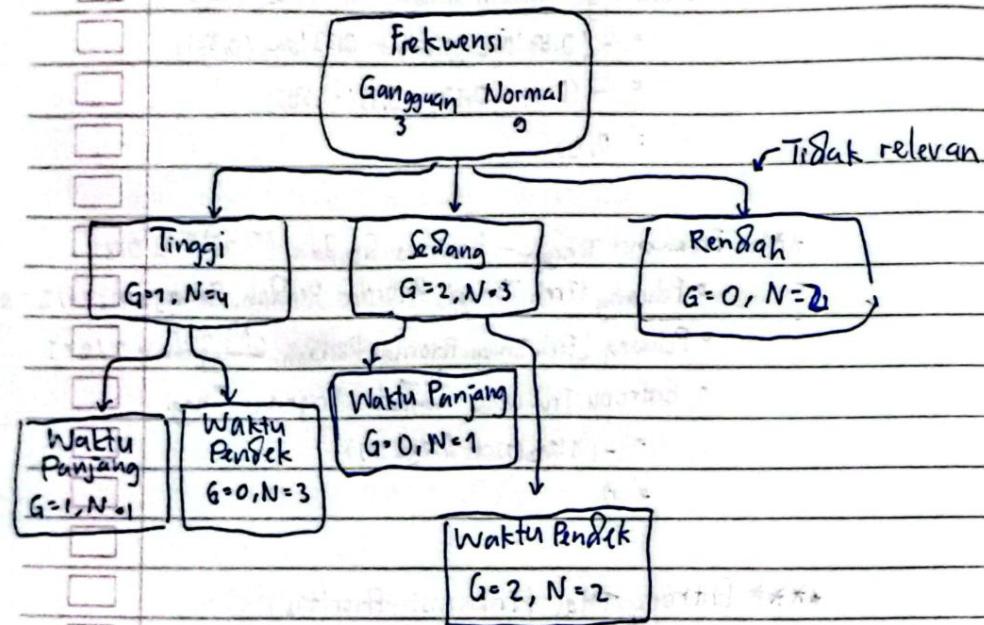
\*\*\* Information Gain (Frekuensi - Prioritas)

$$= 0,704 - 0,54$$

$$= 0,164 //$$

## b) Information Gain Frekuensi

Frekuensi Tinggi			Frekuensi Sedang		
Waktu	Paket	Prioritas	Waktu	Paket	Prioritas
0,304	0,056	0,09	-0,09	-0,2406	-0,234



\* Entropy Frekuensi - Waktu Panjang - Paket

\* Frek. Tinggi - Waktu Panjang - Paket Kecil ( $P = \frac{3}{2}$ )  $\approx 0,15$

- Peluang (Tinggi, Panjang, Kecil, Gangguan) =  $\frac{3}{4} = 0,75$

- Peluang ( " , " , " , Normal) =  $\frac{0}{4} = 0$

- Entropy  $F_{Tinggi} - W_{panjang} - P_{kecil} = 0$

$$* F_{Tinggi} - W_{panjang} - P_{Besar} (P = \frac{1}{2}) = 0,15$$

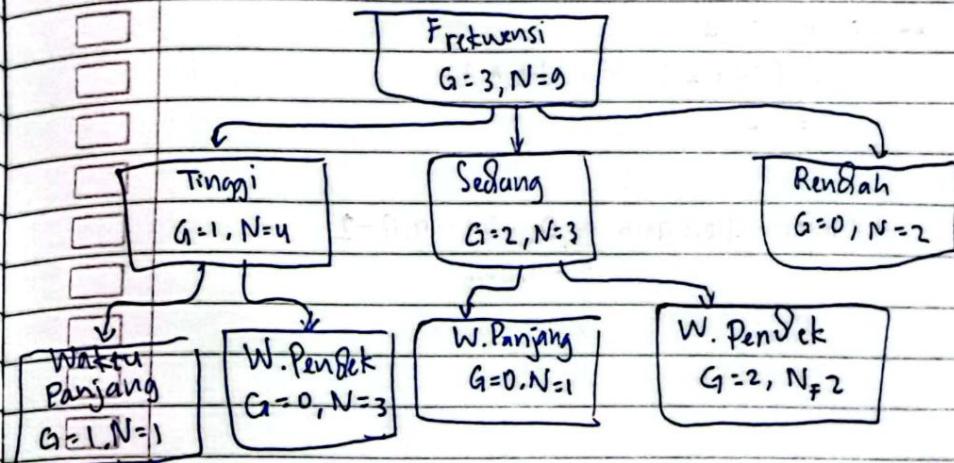
- $P(Tinggi, Panjang, Besar, Gangguan) = 0/4 = 0$

- $P(Tinggi, Panjang, Besar, Normal) = 1/4 = 0,25$

- Entropy  $F_{Tinggi} - W_{panjang} - P_{Besar} = 0$

$$* \text{Entropy Total } F_T - W_P - P = 0$$

$$* \text{Information Gain} = 0,704 - 0,11 - 0 = 0,594$$



\*\*\*  $E_{Fs} - W_{pendek} - P_{kecil}$

\*\*  $E_{Fsedang} - W_{pendek} - P_{kecil} (P = 2/4 = 1/2 = 0,5)$

- $P(Fs - Wp - Pk, Gangguan) = 1/2 = 0,5$

- $P(Fs - Wp - Pk, Normal) = 1/2 = 0,5$

• Entropy  $F_S - W_{P_E} - P_F$

$$= -(0,5 \log_2(0,5) + 0,5 \log_2(0,5))$$

= 1

\*\* Entropy  $F_S - W_{P_E} - P_B$  besar ( $P_{2/4} = 1/2 = 0,5$ )

$$\circ P(F_S - W_{P_E} - P_B, \text{Gangguan}) = 1/2 = 0,5$$

$$\circ P(" " " , \text{Normal}) = 1/2 = 0,5$$

•  $E F_S - W_{P_E} - P_B$

$$= -(0,5 \log_2(0,5) + 0,5 \log_2(0,5))$$

= 1

\*\* Entropy Total

$$= (2/4 \times 1) + (2/4 \times 1)$$

= 1

\*\* Information Gain  $\approx 0,704 - 0,8 - 1$

$$= -1,096$$

\*\*\*  $F_S$  sedang -  $W_{pendek}$  - Prioritas

\*\*\*  $F_S$  sedang -  $W_{pendek}$  - Prioritas Rendah ( $P=2/4 = 0,5$ )

$$\circ P(F_S - W_{P_E} - P_F, \text{Gangguan}) = 1/2 = 0,5$$

$$\circ P(F_S - W_{P_E} - P_F, \text{Normal}) = 1/2 = 0,5$$

• Entropy  $F_S - W_{P_E} - P_R$

$$= -(0,5 \log_2(0,5) + 0,5 \log_2(0,5))$$

= 1

\*\*\*  $F_S - W_{P_E} - P_T$  ( $P_{2/4} = 0,5$ )

$$\circ P(\text{Sedang-Pendek, Rendah, Gangguan}) = 1/2 = 0,5$$

$$\circ P(" " " " , \text{Normal}) = 1/2 = 0,5$$

•  $E F_S - W_{P_E} - P_T$

$$= -(0,5 \log_2(0,5) + 0,5 \log_2(0,5))$$

= 1

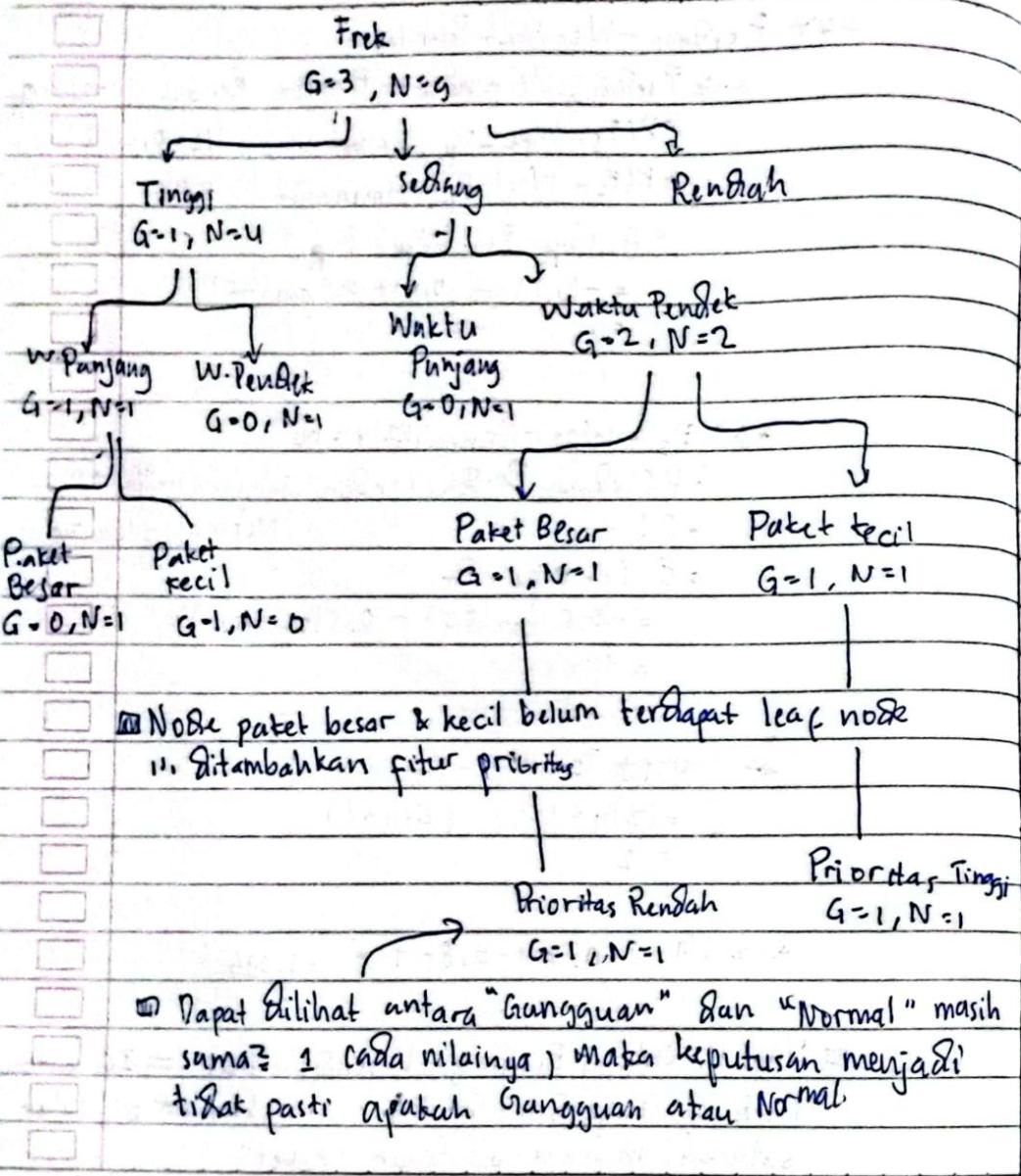
\*\* Entropy  $F_S - W_{P_E} - P_T$

$$= (2/4 \times 1) + (2/4 \times 1)$$

= 1

\*\*\*  $IG = 0,704 - 0,8 - 1 = -1,096$

■ Nilai  $IG$  entropy  $F_S$  sedang -  $W_{pendek}$  - Paket =  $IG$  entropy  $F_S$  sedang -  $W_{pendek}$  - Prioritas. Jadi bisa ambil salah satunya misalnya pilih fitur "Paket".



### Membuat Rules (Output)

NO

## Rules

- 1 IF Frek = Rendah, THEN Gangguan=Normal
- 2 IF Frek = Tinggi, AND Waktu=Pendek, THEN Gangguan=Normal
- 3 IF Frek=Tinggi, AND Waktu=Panjang, AND Paket=Besar, THEN G=N
- 4 IF Frek = T, AND Waktu=Panjang, AND Paket=K, THEN G=G
- 5 IF Frek=S, AND W=Pak, THEN G=N
- 6 IF Frek=S, AND W=Pendek, AND Paket=Besar, AND Prioritas=R, THEN G
- 7 IF Frek=S, AND W=Pendek AND Paket=Kecil AND Prioritas=T, THEN G=N

### Menreditsi Sgn model DT ID3 (Hasil)

- Terjadi ketidakakuratan pada kolom Gangguan. Prediksi di baris 3 (Gangguan, Normal) dan baris 8 (N, G).

### Menghitung nilai error

$$E = \frac{2}{12} \times 100\% \\ = 16,6\%$$

### Pruning Model Akhir

- Lakukan pruning leaf node Frek. Sedang - Waktu Pendek - Paket Besar - Prioritas Rendah & leaf node Frek. Sedang - Waktu Pendek - Paket Kecil - Prioritas Tinggi.
- Frek. Sedang - Waktu Pendek → menjadi leaf node (di klasifikasikan sbg 'Normal').

► Hasil Prediksi Pasca Pruning

- Ketidakakuratan tetap terjadi di dua baris (baris ke-1, baris ke-3).

↓      ↓

G,N      G,N

- Error pasca pruning

$$E = \frac{2}{12} * 100\%$$

$$= 16,6\%$$

► Kesimpulan

Error sebelum dan sesudah pruning sama :

sebesar 16,6% (tidak terjadi perubahan berarti).