Note: \*Tulisan ini diketik dengan menggunakan microsoft Word

WAKTU	PAKET	FREKWEKSI	PRIORITAS	GANGGUAN
PENDEK	BESAR	SEDANG	RENDAH	GANGGUAN
PENDEK	KECIL	TINGGI	RENDAH	NORMAL
PENDEK	KECIL	SEDANG	TINGGI	GANGGUAN
PENDEK	KECIL	TINGGI	RENDAH	NORMAL
PENDEK	KECIL	SEDANG	TINGGI	NORMAL
PANJANG	BESAR	SEDANG	RENDAH	NORMAL
PANJANG	KECIL	TINGGI	TINGGI	GANGGUAN
PENDEK	BESAR	SEDANG	RENDAH	NORMAL
PANJANG	KECIL	RENDAH	TINGGI	NORMAL
PENDEK	KECIL	TINGGI	TINGGI	NORMAL
PANJANG	BESAR	TINGGI	TINGGI	NORMAL
PANJANG	KECIL	RENDAH	TINGGI	NORMAL

- 1. Buatlah tree dan rule untuk mendeteksi adanya gangguan pada jaringan komputer menggunakan data di atas
- 2. Berapa persen besarnya error yang terjadi tanpa penyederhanaan (pruning) dan dengan penyederhanaan

#### > Entropy Gangguan

$$P_{gangguan} = 3/12 = 0.25$$

$$P_{normal} = 9 / 12 = 0.75$$

$$Entropy = -(P(i) * log_2(P(i)) + P(i) * log_2(P(i)))$$

$$= - (0.25 * log_2 (0.25) + 0.75 * log_2 (0.75)$$

$$= - (-0.811)$$

$$= 0.811$$

### > Entropy Waktu

• Pendek 7

$$P(Pendek, Gangguan) = 2/7 = 0.285$$

$$P_{(Pendek, Normal)} = 5 / 7 = 0.715$$

Entropy waktu pendek = 
$$-(0.285 * log_2(0.285) + 0.715 * log_2(0.715))$$

$$= -(-0.861)$$

$$= 0.861$$

Panjang 5

P(Panjang, Gangguan) = 1/5 = 0.2

P(Panjang, Normal) = 4/5 = 0.8

Entropy waktu panjang =  $-(0.2 * log_2(0.2) + 0.8 * log_2(0.8))$ 

= - (0.2 \* (-2.321) + 0.8 \* (-0.321))

= - (-0.721)

= 0.721

Entropy waktu = 7/12 \* 0.861 + 5/12 \* 0.721

= 0.502 + 0.300 = 0.802

Information Gain Waktu = 0.811 - 0.802 = 0.009

### > Entropy Paket

• Besar 4

P(Besar, Gangguan) = 1 / 4 = 0.25

P(Besar, Normal) = 3 / 4 = 0.75

Entropy Paket Besar = -  $(0.25 * log_2 (0.25) + 0.75 * log_2 (0.75))$ 

= - (0.25 \* (-2) + 0.75 \* (-0.415))

= - (-0.811)

= 0.811

Kecil 8

P(Kecil, Gangguan) = 2/8 = 0.25

P(Kecil, Normal) = 6/8 = 0.75

Entropy Paket Kecil = -  $(0.25 * log_2 (0.25) + 0.75 * log_2 (0.75))$ 

= - (0.25 \* (-2) + 0.75 \* (-0.415))

= - (-0.811)

= 0.811

Entropy Paket = 4/12 \* 0.811 + 8/12 \* 0.811

= 0.333 \* 0.811 + 0.667 \* 0.811

=0.811

Information Gain Paket = 0.811 - 0.811 = 0

## > Entropy Frekwensi

• Tinggi 5

$$P(Tinggi, Gangguan) = 1 / 5 = 0.2$$

$$P(Tinggi, Normal) = 4/5 = 0.8$$

Entropy Frekwensi Tinggi = 
$$-(0.2 * log_2 (0.2) + 0.8 * log_2 (0.8))$$
  
=  $-(0.2 * (-2.321) + 0.8 * (-0.321))$ 

$$= -(-0.721)$$

$$= 0.721$$

Sedang 5

$$P(Sedang, Gangguan) = 2/5 = 0.4$$

$$P(Sedang, Normal) = 3/5 = 0.6$$

Entropy Frekwensi Sedang = - 
$$(0.4 * log_2 (0.4) + 0.6 * log_2 (0.6))$$

$$= -(-0.97)$$

$$= 0.97$$

• Rendah 2

$$P(Rendah, Gangguan) = 0 / 2 = 0$$

$$P(Rendah, Normal) = 2/2 = 1$$

Entropy Frekwensi Rendah = - 
$$(0 * log_2(0) + 1 * log_2(1)) = 0$$

$$= 0.300 + 0.404 + 0$$

$$= 0.704$$

Information Gain Frekwensi = 0.811 - 0.704 = 0.107 ~ Root

#### > Entropy Prioritas

• Rendah 5

$$P(Rendah, Gangguan) = 1/5 = 0.2$$

$$P(Rendah, Normal) = 4/5 = 0.8$$

Entropy Prioritas Rendah = - 
$$(0.2 * log_2 (0.2) + 0.8 * log_2 (0.8))$$

$$= -(-0.721)$$

$$= 0.721$$

• Tinggi 7

P(Tinggi, Gangguan) = 2/7 = 0.285

P(Tinggi, Normal) = 5/7 = 0.715

Entropy Prioritas Tinggi = -  $(0.285 * log_2 (0.285) + 0.715 * log_2 (0.715))$ 

= - (0.285 \* (-1.810) + 0.715 \* (-0.483))

= - (-0.861)

= 0.861

Entropy Prioritas = 5/12 \* 0.721 + 7/12 \* 0.861

= 0.300 + 0.502

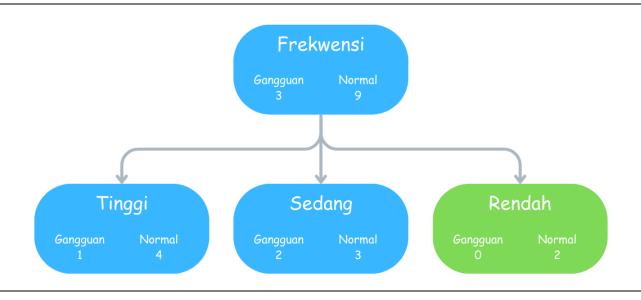
= 0.802

Information Gain Prioritas = 0.811 - 0.802 = 0.009

Nilai IG terbesar adalah Frekwensi, maka ini menjadi Root

# Frekwensi Gangguan Normal 3 9

Karena Frekwensi Sedang dan Frekwensi Tinggi yang memiliki + dan – maka mereka punya leaf node.



- Entropy Frekwensi Tinggi Waktu 5

   Entropy Frekwensi Tinggi Waktu Panjang | 2 P(Tinggi, Panjang, Gangguan) = 1/2 = 0.5 P(Tinggi, Panjang, Normal) = 1/2 = 0.5  $Entropy FT WPan = -(0.5 * log_2(0.5) + 0.5 * log_2(0.5))$  = -(0.5 \* (-1) + 0.5 \* (-1)) = -(-1)
  - Entropy Frekwensi Tinggi Waktu Pendek | 3
     P(Tinggi, Pendek, gangguan) = 0 / 3 = 0
     P(Tinggi, Pendek, Normal) = 3 / 3 = 1
     Entropy FT WPen = 0

= 1

- Entropy Frekwensi Tinggi Waktu = 2/5\*1+3/5\*0=0.4IG Frekwensi Tinggi – Waktu = 0.704-0.4=0.304
- Entropy Frekwensi Sedang Waktu 5
   Entropy Frekwensi Sedang Waktu Panjang | 1
   P(Sedang, Panjang, Gangguan) = 0 / 1 = 0
   P(Sedang, Panjang, Normal) = 1 / 1 = 1
   Entropy FS WPan = 0
  - Entropy Frekwensi Sedang Waktu Pendek | 4 P(Sedang, Pendek, Gangguan) = 2/4 = 0.5 P(Sedang, Pendek, Normal) = 2/4 = 0.5  $Entropy FS WPen = -(0.5 * log_2 (0.5) + 0.5 * log_2 (0.5))$  = -(0.5 \* (-1) + 0.5 \* (-1)) = -(-1)
  - Entropy Frekwensi Sedang Waktu = 1/5\*0+4/5\*1=0.8IG Frekwensi Sedang – Waktu = 0.704-0.8=-0.096

```
> Entropy Frekwensi Tinggi - Paket 5
      • Entropy Frekwensi Tinggi - Paket Besar 1
        P(Tinggi, Besar, Gangguan) = D/1 = D
         P(Tinggi, Besar, Normal) = 1/1 = 1
         Entropy Fringgi - PBesar = 0
      • Entropy Frekwensi Tinggi - Paket Kecil | 4
         P(Tinggi, Kecil, Gangguan) = 1/4 = 0.25
        P(Tinggi, Kecil, Normal) = \frac{3}{4} = 0.75
         Entropy F_{Tinggi} - P_{Besar} = -(0.25 * log_2(0.25) + 0.75 * log_2(0.75))
= -(0.25 * (-2) + 0.75 * (-0.415))
                                  = - (-0.811)
                                  = 0.811
         Entropy Frekwensi Tinggi - Paket = 1/5 * 0 + 4/5 * 0.811
                                                = 0 + 0.648 = 0.648
         IG Frekwensi Tinggi - Paket = 0.704 - 0.648 = 0.056
> Entropy Frekwensi Sedang - Paket 5
      • Entropy Frekwensi Sedang - Paket Besar | 3
        P(Sedang, Besar, Gangguan) = 1/3 = 0.333
        P(Sedang, Besar, Normal) = 2/3 = 0.667
         Entropy Fsedang - PBesar = - (0.333 * log2 (0.333) + 0.667 * log2 (0.667))
                                  = - (0.333 * (-1.586) + 0.667 * (-0.584))
                                  = - ((-0.528) + (-0.389))
                                  = - (-0.917)
                                  = 0.917

    Entropy Frekwensi Sedang – Paket Kecil | 2

         P(Sedang, Kecil, Gangguan) = 1/2 = 0.5
         P(Sedang, Kecil, Normal) = 1/2 = 0.5
         Entropy F_{Sedang} - P_{Kecil} = -(0.5 * log_2 (0.5) + 0.5 * log_2 (0.5))
```

Entropy Frekwensi Sedang - Paket = 
$$3/5 * 0.917 + 2/5 * 1$$
  
=  $0.550 + 0.4 = 0.950$ 

```
Entropy Frekwensi Tinggi – Prioritas 5
      • Entropy Frekwensi Tinggi - Prioritas Tinggi | 3
        P(Tinggi, Tinggi, Gangguan) = 1/3 = 0.333
        P(\text{Tinggi}, \text{Tinggi}, \text{Normal}) = 2/3 = 0.667
        Entropy Fringgi - Pringgi = - (0.333 * log2 (0.333) + 0.667 * log2 (0.667))
                                = - (0.333 * (-1.586) + 0.667 * (-0.584))
                                = - ((-0.528) + (-0.389))
                                = - (-0.917)
                                = 0.917

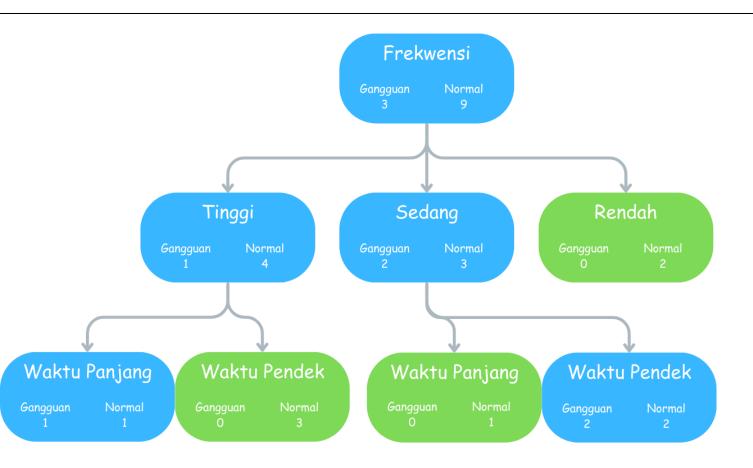
    Entropy Frekwensi Tinggi – Prioritas Rendah | 2

        P(Tinggi, Rendah, Gangguan) = 0/2 = 0
        P(Tinggi, Rendah, Normal) = 2/2 = 1
        Entropy Fringgi - Prendah = 0
        Entropy Frekwensi Tinggi - Prioritas = 3/5 * 0.917 + 2/5 * 0
                                                 = 0.550 + 0 = 0.550
        IG Frekwensi Tinggi - Prioritas = 0.704 - 0.550
                                           = 0.154
> Entropy Frekwensi Sedang – Prioritas 5

    Entropy Frekwensi Sedang – Prioritas Tinggi | 2

        P(Sedang, Tinggi, Gangguan) = 2/5 = 0.4
        P(Sedang, Tinggi, Normal) = 3/5 = 0.6
        Entropy Fsedang - = - (0.4 * log_2 (0.4) + 0.6 * log_2 (0.6))
                          = -(0.4*(-1.321) + 0.6*(-0.736))
                          = -(-0.97)
                          = 0.97
      • Entropy Frekwensi Sedang - Prioritas Rendah | 3
        P(Sedang, Rendah, Gangguan) = 1/3 = 0.333
        P(Sedang, Rendah, Normal) = 2/3 = 0.667
        Entropy Fsedang - Prendah = - (0.333 * log2 (0.333) + 0.667 * log2 (0.667))
                                 = - (0.333 * (-1.586) + 0.667 * (-0.584))
                                 = -((-0.528) + (-0.389))
                                 = - (-0.917)
                                 = 0.917
        Entropy Frekwensi Sedang - Prioritas = 2/5 * 0.97 + 3/5 * 0.917
```

= 0.388 + 0.550 = 0.938



- > Entropy Frekwensi Tinggi Waktu Panjang Paket 2
  - Entropy Frekwensi Tinggi Waktu Panjang Paket Kecil | 1
     P(Tinggi, Panjang, Kecil, Gangguan) = 1/1 = 1
     P(Tinggi, Panjang, Kecil, Normal) = 0/1 = 0

Entropy Fringgi - Wranjang - Pkecil = 0

• Entropy Frekwensi Tinggi – Waktu Panjang – Paket Besar | 1 P(Tinggi, Panjang, Besar, Gangguan) = 0/1 = 0

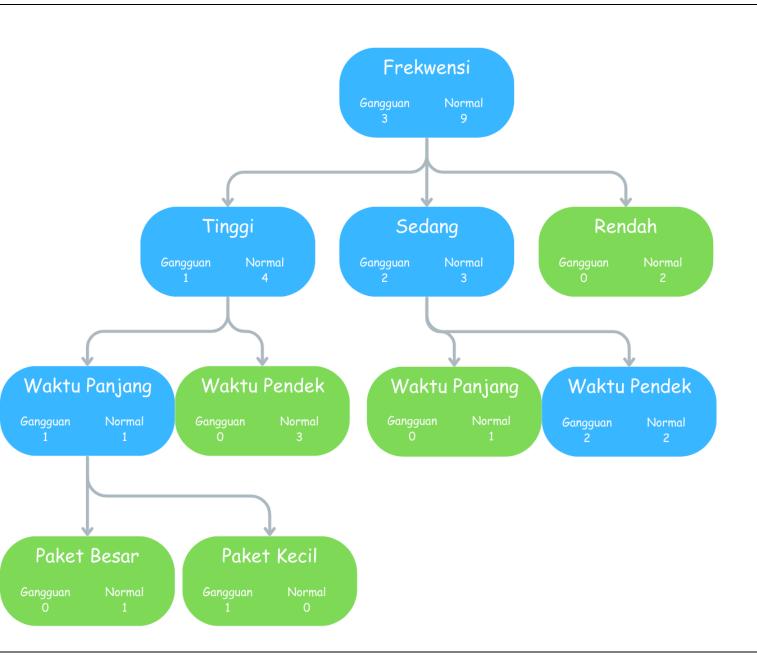
P(Tinggi, Panjang, Besar, Normal) = 1/1 = 1

Entropy Fringgi - Wranjang - Presar = 0

Entropy Frekwensi Tinggi - Waktu Panjang - Paket = 0

$$IG F_{\text{Tinggi}} - W_{\text{Panjang}} - Paket = 0.704 - 0.4 - 0$$

$$= 0.304$$



- Entropy Frekwensi Sedang Waktu Pendek Paket 4
  - Entropy Frekwensi Sedang Waktu Pendek Paket Kecil | 2

$$P(Sedang, Pendek, Kecil, Gangguan) = 1/2 = 0.5$$

$$P(Sedang, Pendek, Kecil, Normal) = 1/2 = 0.5$$

• Entropy Frekwensi Sedang - Waktu Pendek - Paket Besar | 2

$$P(Sedang, Pendek, Besar, Gangguan) = 1/2 = 0.5$$

$$P(Sedang, Pendek, Besar, Normal) = 1/2 = 0.5$$

- > Entropy Frekwensi Sedang Waktu Pendek Prioritas 4
  - Entropy Frekwensi Sedang Waktu Pendek Prioritas Rendah | 2 P(Sedang, Pendek, Rendah, Gangguan) = 1/2 = 0.5

P(Sedang, Pendek, Rendah, Normal) = 1/2 = 0.5

Entropy Fsedang - Wpendek - PBesar = 
$$-(0.5 \log_2(0.5) + 0.5 \log_2(0.5))$$
  
= 1

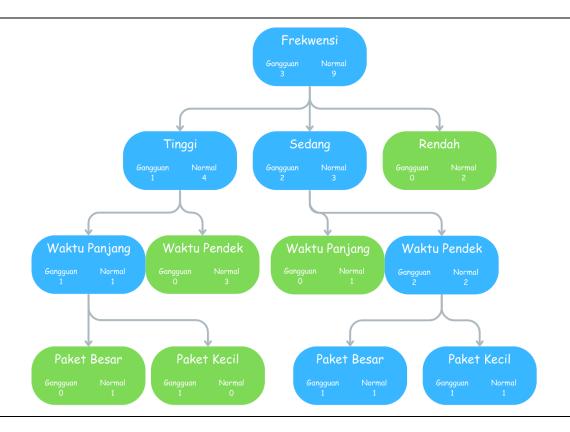
• Entropy Frekwensi Sedang – Waktu Pendek – Prioritas Tinggi | 2 P(Sedang, Pendek, Rendah, Gangguan) = 1/2 = 0.5

P(Sedang, Pendek, Rendah, Normal) = 1/2 = 0.5

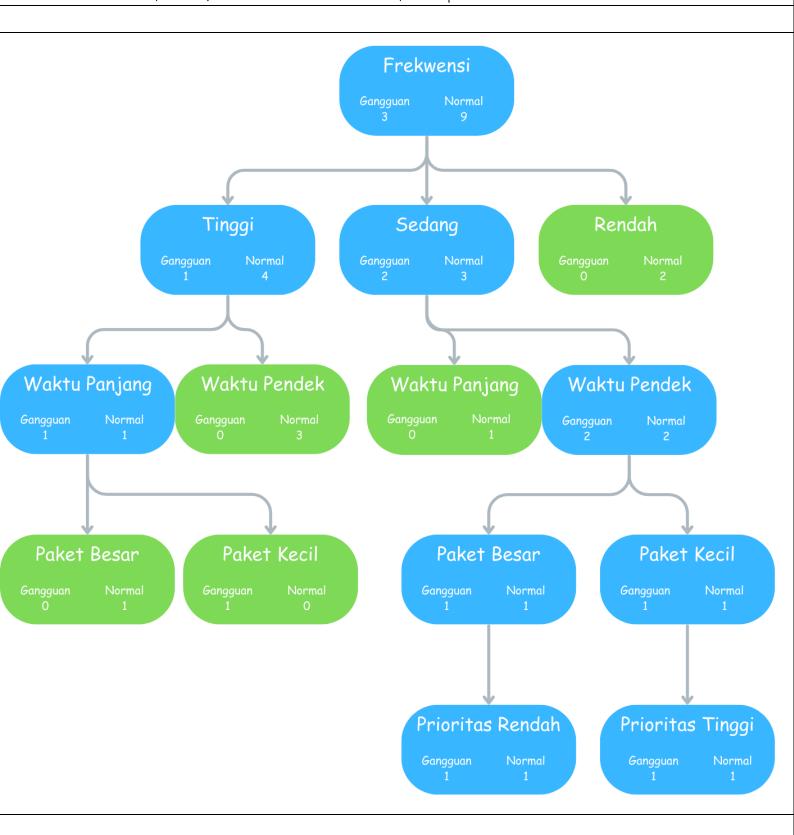
Entropy Fsedang - Wrendek - PTinggi =  $-(0.5 \log_2(0.5) + 0.5 \log_2(0.5))$ = 1

Entropy Fsedang - Wpendek - Prioritas = 2/4 \* 1 + 2/4 \*1 = 1

Karena nilai IG dari entropy Fsedang - Wpendek - Paket sama dengan nilai IG dari Fsedang - Wpendek - Prioritas, maka disini saya ambil fitur Paket



Pada gambar tersebut, karena node Paket Besar dan Paket kecil masih belum Memiliki leaf node, maka ditambahkan fitur prioritas



Meskipun sudah ditambahkan fitur Prioritas sebagai cabang baru, tetapi Pembagian node belum dapat menghasilkan satu nilai pasti

- Membuat Rule (Aturan)
  - R1: IF Frekwensi = Rendah THEN Gangguan = Normal
  - R2: IF Frekwensi = Tinggi AND Waktu = Pendek THEN Gangguan = Normal
  - R3: IF Frekwensi = Tinggi AND Waktu = Panjang AND Paket = Besar THEN Gangguan = Normal
  - R4: IF Frekwensi = Tinggi AND Waktu = Panjang AND Paket = Kecil THEN Gangguan = Gangguan
  - R5: IF Frekwensi = Sedang AND Waktu = Panjang THEN Gangguan = Normal
  - R6: IF Frekwensi = Sedang AND Waktu = Pendek AND Paket = Besar AND Prioritas = Rendah THEN Gangguan = Gangguan
  - R7: IF Frekwensi = Sedang AND Waktu = Pendek AND Paket = Kecil AND Prioritas = Tinggi THEN Gangguan = Normal
- Catatan: Untuk Rule 6 dan 7 diperlukan fitur tambahan atau pandangan expert terkait dengan hal tersebut untuk membagi ke dalam hasil yang pasti sehingga saya define untuk Rule 6 menghasilkan prediksi Gangguan = Gangguan, dan Rule 7 menghasilkan prediksi Gangguan = Normal
- > Melakukan prediksi dengan model decision tree classifiers yang telah dibuat

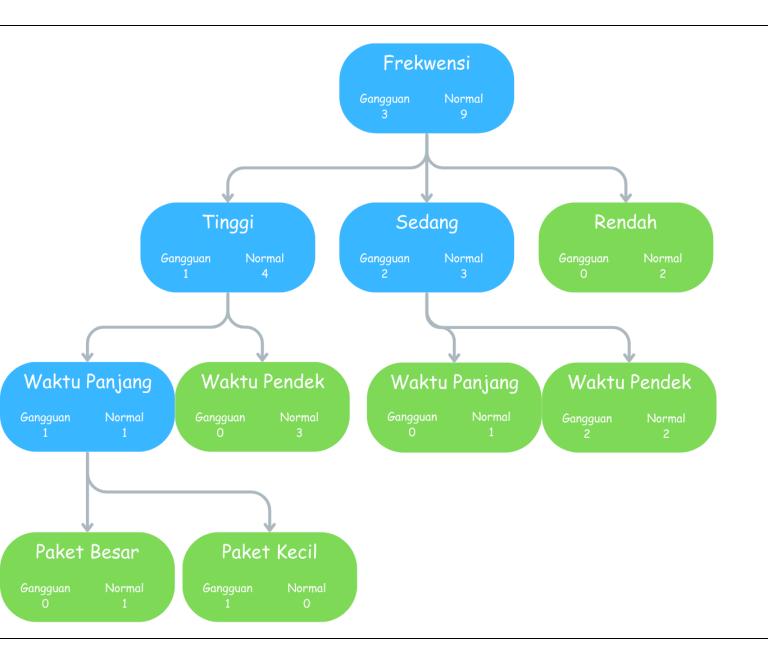
Waktu	Paket	Frekwensi	<b>Prioritas</b>	Gangguan	Prediksi
Pendek	Besar	Sedang	Rendah	Gangguan	Gangguan
Pendek	Kecil	Tinggi	Rendah	Normal	Normal
Pendek	Kecil	Sedang	Tinggi	Gangguan	Normal
Pendek	Kecil	Tinggi	Rendah	Normal	Normal
Pendek	Kecil	Sedang	Tinggi	Normal	Normal
Panjang	Besar	Sedang	Rendah	Normal	Normal
Panjang	Kecil	Tinggi	Tinggi	Gangguan	Gangguan
Pendek	Besar	Sedang	Rendah	Normal	Gangguan
Panjang	Kecil	Rendah	Tinggi	Normal	Normal
Pendek	Kecil	Tinggi	Tinggi	Normal	Normal
Panjang	Besar	Tinggi	Tinggi	Normal	Normal
Panjang	Kecil	Rendah	Tinggi	Normal	Normal

Menghitung nilai kesalahan (Error)

E = 2/12 \* 100%

= 16.690

- Melakukan Pruning (Pemangkasan) pada model decision tree classifiers yang dibuat
  - Terdapat redundansi pada leaf node Frekwensi Sedang Waktu Pendek
  - Melakukan pemangkasan pada leaf node Frekwensi Sedang Waktu Pendek
     Paket Besar Prioritas Rendah dan leaf node Frekwensi Sedang Waktu
     Pendek Paket Kecil Prioritas Tinggi
  - Menjadikan node Frekwensi Sedang Waktu Pendek sebagai leaf node



 Catatan: Untuk node Frekwensi Sedang - Waktu Pendek, diklasifikasikan sebagai Normal

	Hasil	Prediksi	Setelah	Pruning
--	-------	----------	---------	---------

Waktu	<b>Paket</b>	Frekwensi	<b>Prioritas</b>	Status	Prediksi
Pendek	Besar	Sedang	Rendah	Gangguan	Normal
Pendek	Kecil	Tinggi	Rendah	Normal	Normal
Pendek	Kecil	Sedang	Tinggi	Gangguan	Normal
Pendek	Kecil	Tinggi	Rendah	Normal	Normal
Pendek	Kecil	Sedang	Tinggi	Normal	Normal
Panjang	Besar	Sedang	Rendah	Normal	Normal
Panjang	Kecil	Tinggi	Tinggi	Gangguan	Gangguan
Pendek	Besar	Sedang	Rendah	Normal	Normal
Panjang	Kecil	Rendah	Tinggi	Normal	Normal
Pendek	Kecil	Tinggi	Tinggi	Normal	Normal
Panjang	Besar	Tinggi	Tinggi	Normal	Normal
Panjang	Kecil	Rendah	Tinggi	Normal	Normal

## > Menghitung Error setelah Pruning

E = 2/12 \* 100%

= 16.670

Berdasarkan hasil prediksi model decision tree classifiers sebelum dan sesudah pruning, tidak ada perubahan signifikan. Baik sebelum dan sesudah pruning, nilai error nya sama sama 16.6%