

TUGAS 6

Machine Learning

Decision Tree



Nama:

Addina Nurul Mawaddah

(1941720240)

Kelas:

TI-2B

PROGRAM STUDI TEKNIK INFORMATIKA

JURUSAN TEKNOLOGI INFORMASI

POLITEKNIK NEGERI MALANG

APRIL 2021

Decision Tree

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. Lakukan penyederhaan (Pruning)
3. Berapa persen besarnya error yang terjadi tanpa penyederhanaan (pruning) dan dengan penyederhanaan

Jawab :

1. Memilih Node Awal :

Node :

- **Waktu**

Waktu	Gangguan	Jumlah
Pendek	Gangguan	2
Pendek	Normal	5
Panjang	Gangguan	1
Panjang	Normal	4

$$q1 = -2/7 \log 2 \frac{2}{7} - 5/7 \log 2 \frac{5}{7} = 0.86$$

$$q2 = -1/5 \log 2 \frac{1}{5} - 4/5 \log 2 \frac{4}{5} = 0.72$$

Entropy untuk waktu :

$$E = 7/12 q1 + 5/12 q2$$

$$= 7/12 (1.04) + 5/12 (0.72) = 0.8$$

- **Paket**

Paket	Gangguan	Jumlah
Besar	Gangguan	1

Besar	Normal	3
Kecil	Gangguan	2
Kecil	Normal	6

$$q_1 = -1/4 \log_2 1/4 - 3/4 \log_2 3/4 = 0.81$$

$$q_2 = -2/8 \log_2 2/8 - 6/8 \log_2 6/8 = 0.81$$

Entropy untuk paket :

$$\begin{aligned} E &= 4/12 q_1 + 8/12 q_2 \\ &= 4/12 (0.81) + 8/12 (0.81) \\ &= 0.81 \end{aligned}$$

- Frekwensi

Frekwensi	Gangguan	Jumlah
Sedang	Gangguan	2
Sedang	Normal	3
Tinggi	Gangguan	1
Tinggi	Normal	4
Rendah	Normal	2

$$q_1 = -2/5 \log_2 2/5 - 3/5 \log_2 3/5 = 0.97$$

$$q_2 = -1/5 \log_2 1/5 - 4/5 \log_2 4/5 = 0.72$$

$$q_3 = -2/2 \log_2 2/2 - 0 \log_2 0 = 0$$

Entropy untuk Frekwensi :

$$\begin{aligned} E &= 5/12 q_1 + 5/12 q_2 + 2/2 q_3 \\ &= 5/12 (0.97) + 5/12 (0.72) + 2/2 (0) \\ &= 0.7 \end{aligned}$$

• Prioritas

Prioritas	Gangguan	Jumlah
Rendah	Gangguan	1
Rendah	Normal	4
Tinggi	Gangguan	2
Tinggi	Normal	5

$$q_1 = -1/5 \log_2 1/5 - 4/5 \log_2 4/5 = 0.72$$

$$q_2 = -2/7 \log_2 2/7 - 5/7 \log_2 5/7 = 0.86$$

Entropy untuk Prioritas :

$$\begin{aligned} E &= 5/12 q_1 + 7/12 q_2 \\ &= 5/12 (0.72) + 7/12 (0.86) \\ &= 0.8 \end{aligned}$$

Leaf Node

Frekwensi = Tinggi

Waktu	Paket	Prioritas	Gangguan
Pendek	Kecil	Rendah	Normal
Pendek	Kecil	Rendah	Normal
Panjang	Kecil	Tinggi	Gangguan
Pendek	Kecil	Tinggi	Normal
Panjang	Besar	Tinggi	Normal

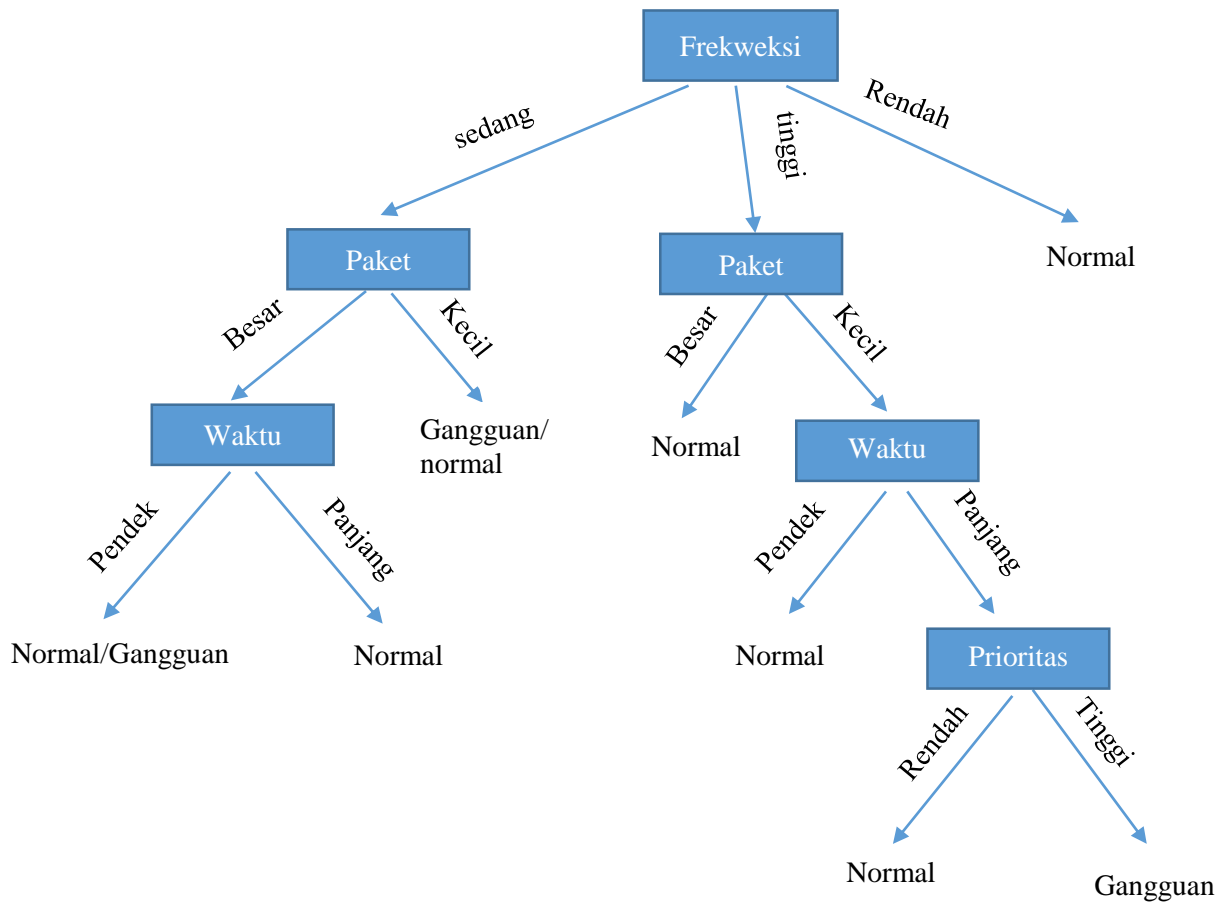
Waktu	Gangguan	Jumlah
Pendek	Gangguan	0
	Normal	3
Panjang	Gangguan	1
	Normal	1
Entropy = 0.89		

Paket	Gangguan	Jumlah
Kecil	Gangguan	1
	Normal	3
Besar	Gangguan	0
	Normal	1
Entropy = 0.81		

Prioritas	Gangguan	Jumlah
Rendah	Gangguan	0
	Normal	2

Tinggi	Gangguan	1
	Normal	2
Entropy = 0.93		

Tree



Hasil data tarining

Waktu	Paket	Frekweksi	Prioritas	Gangguan	Prediksi
Pendek	Besar	Sedang	Rendah	Gangguan	Gangguan
Pendek	Kecil	Tinggi	Rendah	Normal	Normal
Pendek	Kecil	Sedang	Tinggi	Gangguan	Gangguan
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

Rule :

R1: IF frekweksi = “Rendah” THEN gangguan = “normal”

R2: IF frekweksi = “Sedang” and Paket = “Kecil” THEN gangguan = “Gangguan/normal”

R3: IF frekweksi = “Sedang” and Paket = “Besar” and waktu = “pendek” THEN
gangguan = normal/gangguan

R4: IF frekweksi = “Sedang” and Paket = “Besar” and waktu = “panjang” THEN
gangguan = Normal

R5: IF frekweksi = “tinggi” and Paket = “Besar” THEN gangguan = “Normal”

R6: IF frekweksi = “tinggi” and Paket = “kecil” and waktu = “pendek” THEN gangguan
= “Gangguan”

R7: IF frekweksi = “tinggi” and Paket = “kecil” and waktu = “panjang” and prioritas =
“rendah” THEN gangguan = Normal

R8: IF frekweksi = “tinggi” and Paket = “kecil” and waktu = “panjang” and prioritas =
“tinggi” THEN gangguan = Normal

2. Lakukan penyederhaan (purning)

Frekweksi	Gangguan	Jumlah
Sedang	Gangguan	2
Sedang	Normal	3
Tinggi	Gangguan	4
Tinggi	Normal	1
Rendah	Gangguan	0
Rendah	Normal	2

	Sedang	Tinggi	Rendah	Marginal
Gangguan	2	4	0	6
Normal	3	1	2	6
Marginal	5	5	2	12



Uji independensi Chi square (E_{ij}):

	Sedang	Tinggi	Rendah	Marginal
Gangguan	2.5	2.5	1	6
Normal	2.5	2.5	1	6
Marginal	5	5	2	12

O_{ij}	E_{ij}	$O_{ij} - E_{ij}$	$(O_{ij} - E_{ij})^2$	$(O_{ij} - E_{ij})^2 / E_{ij}$
2	2.5	-0.5	0.25	0.1
4	2.5	1.5	2.25	0.9
0	1	-1	1	1
3	2.5	0.5	0.25	0.1
1	2.5	-1.5	2.25	0.9
2	1	1	1	1

3. Berapa persen besarnya error yang terjadi

Waktu	Paket	Frekweksi	Prioritas	Gangguan	Prediksi
Pendek	Besar	Sedang	Rendah	Gangguan	Gangguan
Pendek	Kecil	Tinggi	Rendah	Normal	Normal
Pendek	Kecil	Sedang	Tinggi	Gangguan	Gangguan
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

$$\begin{aligned}
 \text{Kesalahan (e)} &= 1/12 * 100\% \\
 &= 8.3\%
 \end{aligned}$$