

NAMA : ABDURRAZZAQ ILHAM AZIZ

KELAS : A11.4508

NIM : A11.2022.14301

TUGAS 7 DATA MINING

Latihan Soal (Kuis)

1. Hitung Entropy dan Gain serta tentukan pohon keputusan yang terbentuk dari contoh kasus keputusan bermain tenis dibawah ini :

OUTLOOK	TEMPERATURE	HUMIDITY	WINDY	PLAY
Sunny	Hot	High	No	Don't Play
Sunny	Hot	High	Yes	Don't Play
Cloudy	Hot	High	No	Play
Rainy	Mild	High	No	Play
Rainy	Cool	Normal	No	Play
Rainy	Cool	Normal	Yes	Play
Cloudy	Cool	Normal	Yes	Play
Sunny	Mild	High	No	Don't Play
Sunny	Cool	Normal	No	Play
Rainy	Mild	Normal	No	Play
Sunny	Mild	Normal	Yes	Play
Cloudy	Mild	High	Yes	Play
Cloudy	Hot	Normal	No	Play
Rainy	Mild	High	Yes	Don't Play

36

1. Menghitung Entropy:

Total data = 14 kasus

Play = 9 kasus

Don't Play = 5 kasus

$$\text{Entropy}(S) = -P \log_2(P) - N \log_2(N)$$

$$= -(9/14)\log_2(9/14) - (5/14)\log_2(5/14)$$

$$= -(0.643 \times -0.637) - (0.357 \times -1.485)$$

$$= 0.408 + 0.532$$

$$= 0.940$$

A. Outlook (Sunny, Cloudy, Rainy):

Sunny (5 kasus): 2 Play, 3 Don't Play

$$\text{Entropy}(\text{Sunny}) = -(2/5)\log_2(2/5) - (3/5)\log_2(3/5) = 0.971$$

Cloudy (4 kasus): 4 Play, 0 Don't Play

Entropy(Cloudy) = 0 (karna semua bermain)

Rainy (5 kasus): 3 Play, 2 Don't Play

$$\text{Entropy(Rainy)} = -(3/5)\log_2(3/5) - (2/5)\log_2(2/5) = 0.971$$

$$\text{Gain(Outlook)} = \text{Entropy(S)} - \sum(|S_v|/|S|) \times \text{Entropy}(S_v)$$

$$= 0.940 - (5/14 \times 0.971 + 4/14 \times 0 + 5/14 \times 0.971)$$

$$= 0.940 - (0.347 + 0 + 0.347)$$

$$= 0.246$$

B. Temperature (Hot, Mild, Cool):

$$\text{Entropy (Hot)} = -(2/6)*\log_2(2/6) - (4/6)*\log_2(4/6) = 0.918$$

$$\text{Entropy (Mild)} = -(4/6)*\log_2(4/6) - (2/6)*\log_2(2/6) = 0.918$$

$$\text{Entropy (Cool)} = 0 \text{ (karna semua bermain)}$$

$$\text{Entropy Temperature} = (6/14)*0.918 + (6/14)*0.918 + (2/14)*0 = 0.849$$

C. Humidity (High, Normal):

$$\text{Entropy (High)} = -(7/10)*\log_2(7/10) - (3/10)*\log_2(3/10) = 0.881$$

$$\text{Entropy (Normal)} = 0 \text{ (karna semua bermain)}$$

$$\text{Entropy Humidity} = (10/14)*0.881 + (4/14)*0 = 0.629$$

D. Windy (Yes, No):

$$\text{Entropy (True)} = -(3/9)*\log_2(3/9) - (6/9)*\log_2(6/9) = 0.918$$

$$\text{Entropy (False)} = -(5/5)*\log_2(5/5) = 0$$

$$\text{Entropy Windy} = (9/14)*0.918 + (5/14)*0 = 0.605$$

2. Menghitung Gain

$$\text{Gain(Outlook)} = \text{Entropy(Total)} - \text{Entropy(Outlook)} = 0.940 - 0.694 = 0.246$$

$$\text{Gain(Temperature)} = \text{Entropy(Total)} - \text{Entropy(Temperature)} = 0.940 - 0.849 = 0.091$$

$$\text{Gain(Humidity)} = \text{Entropy(Total)} - \text{Entropy(Humidity)} = 0.940 - 0.629 = 0.311$$

$$\text{Gain(Windy)} = \text{Entropy(Total)} - \text{Entropy(Windy)} = 0.940 - 0.605 = 0.335$$

3. Hasil Akar Pohon Keputusan

