Data Mining

Warih Adi Kusuma

A11.2020.12821

A11.4612

Pertemuan 7


Entropy rumus:

$$Entropy(S) \equiv -p_{\oplus} \log_2 p_{\oplus} - p_{\ominus} \log_2 p_{\ominus}$$

Gain rumus:

$$Gain(S,A) \equiv Entropy(S) - \sum_{i=1}^{n} \frac{|Si|}{|S|} \cdot Entropy(Si)$$


Node 1

| | | Kasus | Don't Play | Play | <mark>Entropy</mark> | <mark>Gain</mark> |
|---|---|---|---|---|---|---|
| Total Kasus | | 14 | 4 | 10 | 0.863120569 | |
| Outlook | | | | | | 0.258521037 |
| | Cloudy | 4 | 0 | 4 | 0 | |
| | Rainy | 5 | 1 | 4 | 0.721928095 | |
| | Sunny | 5 | 3 | 2 | 0.970950594 | |
| Temperature | | | | | | 0.183850925 |
| | Cool | 4 | 0 | 4 | 0 | |
| | Hot | 4 | 2 | 2 | 1 | |
| | Mild | 6 | 2 | 4 | 0.918295834 | |
| <mark>Humidity</mark> | | | | | | <mark>0,370506501</mark> |
| | High | 7 | 4 | 3 | 0.985228136 | |
| | Normal | 7 | 0 | 7 | 0 | |
| Windy | | | | | | 0.005977711 |
| | No | 8 | 2 | 6 | 0.811278124 | |
| | Yes | 6 | 4 | 2 | 0.918295834 | |

Klasifikasi Humdity

Node 1.1

| | | Kasus | Don't Play | Play | | |
|---|---|---|---|---|---|---|
| Humidity High | | 7 | 4 | 3 | 0.985228136 | |
| Outlook | | | | | | 0.69951385 |
| | Cloudy | 2 | 0 | 2 | 0 | |
| | Rainy | 2 | 1 | 1 | 0.721928095 | |
| | Sunny | 3 | 3 | 0 | 0.970950594 | |
| Temperature | | | | | | 0.0202244207 |
| | Cool | 0 | 0 | 0 | 0 | |
| | Hot | 3 | 2 | 1 | 0.918295834 | |
| | Mild | 4 | 2 | 2 | 1 | |
| Windy | | | | | | 0.005977711 |
| | No | 4 | 2 | 2 | 1 | |
| | Yes | 3 | 4 | 1 | 0.918295834 | |

Klasifikasi Rainy

Node 1.1.2

| | | Kasus | Don't Play | Play | Entropy | Gain |
|---|---|---|---|---|---|---|
| Humidity High dan Outlook Rain | | 2 | 1 | 1 | 1 | |
| Temperature | | | | | | 0 |
| | Cool | 0 | 0 | 0 | 0 | |
| | Hot | 0 | 0 | 0 | 0 | |
| | Mild | 2 | 1 | 1 | 1 | |
| Windy | | | | | | 1 |
| | No | 4 | 2 | 2 | 1 | |
| | Yes | 3 | 4 | 1 | 0.918295834 | |

Pohon Keputusan Algoritma: