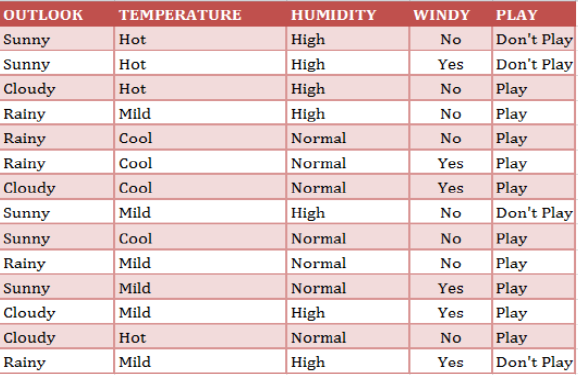
**TUGAS DATA MINING**

**PERTEMUAN MINGGU KE – 7**

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



**Penyelesaian :**

**Langkah 1**

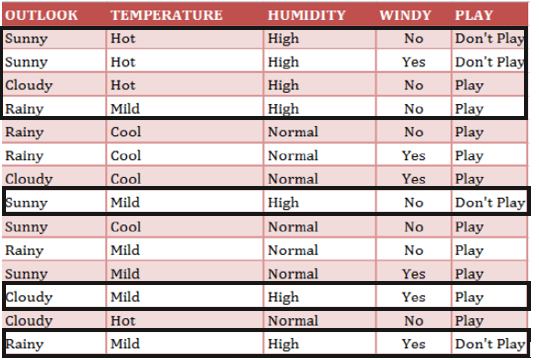
Menghitung jumlah kasus, jumlah kasus untuk keputusan **Play** dan jumlah kasus untuk keputusan **Don’t Play**. Entropy dari semua kasus dan kasus yang dibagi berdasarkan atribut Outlook Temperature Humidity dan Windy. Setelah itu lakukan perhitungan Gain untuk setiap atribut.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Cases (S) | Don’t Play (S1) | Play (S2) | Entropy | Gain |
| Total |  | 14 | 4 | 10 | 0,86312 |  |
| Outlook |  |  |  |  |  | 0,26 |
|  | Cloudy | 4 | 0 | 4 | 0 |  |
|  | Rainy | 5 | 1 | 4 | 0,72193 |  |
|  | Sunny | 5 | 3 | 2 | 0,97095 |  |
| Temperature |  |  |  |  |  | 0,18 |
|  | Cool | 4 | 0 | 4 | 0 |  |
|  | Hot | 4 | 2 | 2 | 1 |  |
|  | Mild | 6 | 2 | 4 | 0,9183 |  |
| Humidity |  |  |  |  |  | 0,37 |
|  | High | 7 | 4 | 3 | 0,98523 |  |
|  | Normal | 7 | 0 | 7 | 0 |  |
| Windy |  |  |  |  |  | 0,01 |
|  | No | 8 | 2 | 6 | 0,81128 |  |
|  | Yes | 6 | 4 | 2 | 0,9183 |  |

Dari hasil tersebut di atas, dapat diketahui bahwa atribut dengan gain tertinggi adalah Humidity yaitu sebesar 0 37 sehingga Humidity dapat menjadi node akar. Ada dua nilai atibut dari Humidity yaitu High dan Normal. Untuk nilai atribut Normal sudah mengklasifikasikan kasus menjadi 1 yaitu keputusannya Play, sehingga tidak perlu dilakukan perhitungan lebih lanjut. Akan tetapi, untuk nilai High masih perlu dilakukan perhitungan lagi.

**Langkah 2**

Menghitung jumlah kasus, jumlah kasus untuk keputusan **Play** jumlah kasus untuk keputusan **Don’t Play**. Entropy dari semua kasus dan kasus yang dibagi berdasarkan atribut Outlook Temperature dan Windy yang dapat menjadi node akar dari nilai atribut High. Setelah itu lakukan perhitungan Gain, untuk tiap tiap atribut.

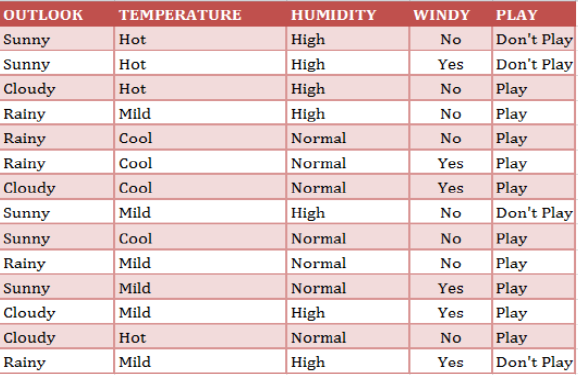


|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Cases (S) | Don’t Play (S1) | Play (S2) | Entropy | Gain |
| Humidity High |  | 7 | 4 | 3 | 0,985228 |  |
| Outlook |  |  |  |  |  | 0,7 |
|  | Cloudy | 2 | 0 | 2 | 0 |  |
|  | Rainy | 2 | 1 | 1 | 1 |  |
|  | Sunny | 3 | 3 | 0 | 0 |  |
| Temperature |  |  |  |  |  | 0,02 |
|  | Cool | 0 | 0 | 0 | 0 |  |
|  | Hot | 3 | 2 | 1 | 0,918296 |  |
|  | Mild | 4 | 2 | 2 | 1 |  |
| Windy |  |  |  |  |  | 0,02 |
|  | No | 4 | 2 | 2 | 1 |  |
|  | Yes | 3 | 2 | 1 | 0,918296 |  |

Dari hasil tersebut di atas, dapat diketahui bahwa atribut dengan gain tertinggi adalah Outlook yaitu sebesar 0.7 sehingga Outlook dapat menjadi node cabang dari nilai atribut High. Ada tiga nilai dari atribut Outlook yaitu Cloudy Rainy dan Sunny. Cloudy untuk klasifikasi kasus 1 (Play). Sunny klasifikasi kasus 1 (Don’t Play). Sedangkan Rainy masih perlu perhitungan lagi.

**Langkah 3**

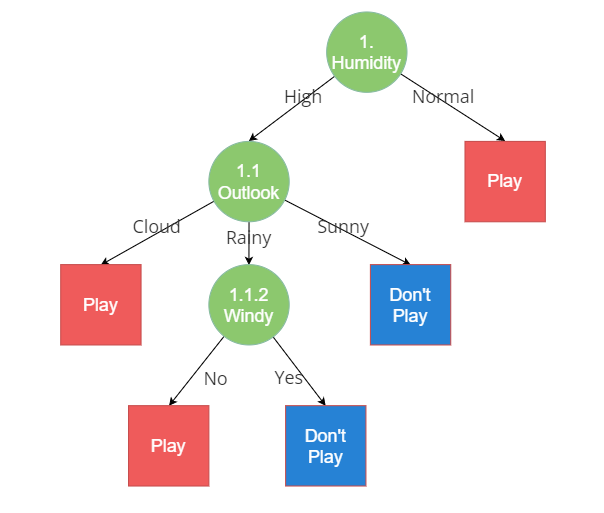
Menghitung jumlah kasus, jumlah kasus untuk keputusan **Play** jumlah kasus untuk keputusan **Don’t Play**. Entropy dari semua kasus dan kasus yang dibagi berdasarkan atribut Temperature dan Windy yang dapat menjadi node cabang dari nilai atribut Rainy. Setelah itu lakukan perhitungan Gain, untuk tiap tiap atribut.



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Cases (S) | Don’t Play (S1) | Play (S2) | Entropy | Gain |
| Humidity High and Outlook Rainy |  | 2 | 1 | 1 | 1 |  |
| Temperature |  |  |  |  |  | 0 |
|  | Cool | 0 | 0 | 0 | 0 |  |
|  | Hot | 0 | 0 | 0 | 0 |  |
|  | Mild | 2 | 1 | 1 | 1 |  |
| Windy |  |  |  |  |  | 1 |
|  | No | 1 | 0 | 1 | 0 |  |
|  | Yes | 1 | 1 | 0 | 0 |  |

Dari hasil tersebut di atas, dapat diketahui bahwa atribut dengan gain tertinggi adalah Windy, yaitu sebesar 1 sehingga Windy dapat menjadi node cabang dari nilai atribut Rainy. Ada dua nilai dari atribut Windy, yaitu No dan Yes. Nilai atribut No sudah mengklasifikasikan kasus menjadi 1 **Play**. Nilai atribut Yes sudah mengklasifikasikan kasus menjadi 1 **Don’t Play** sehingga tidak perlu dilakukan perhitungan lagi.

**Pohon Keputusan Yang Terbentuk**

****