**DATA MINING**

**TUGAS PERTEMUAN KE – 7**



**Dosen Pembimbing**

**JUNTA ZENIARJA M.Kom.**

Disusun Oleh :

Luthfi Kamal Ananda (A11.2020.12586) // A11.4411

|  |
| --- |
|  |

**PROGRAM STUDI S-1 TEKNIK INFORMATIKA**

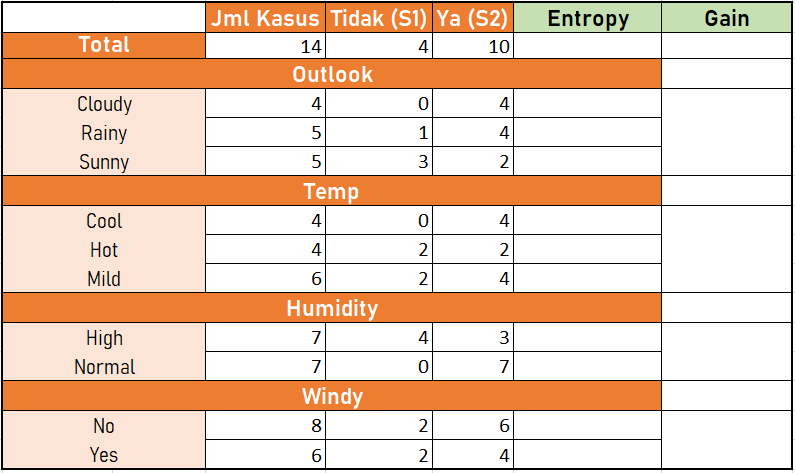
**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS DIAN NUSWANTORO  
SEMARANG  
2023**



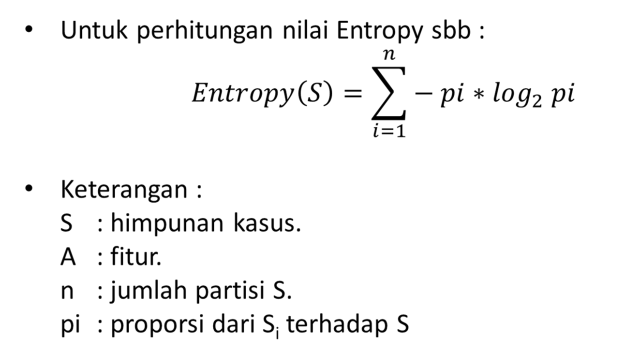
**Jawaban**

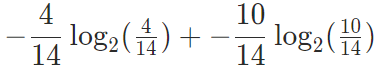
**Langkah 1**

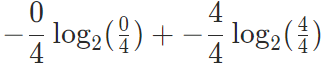
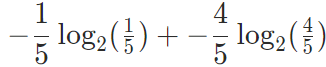
Dengan menggunakan Algoritma C4.5 Berikut adalah tabel yang didapat setelah menghitung jumlah kasus, jumlah kasus untuk keputusan **Play** dan **Don’t Play** berdasarkan atribut **OUTLOOK, TEMPERATURE, HUMIDITY** dan **WINDY**

****

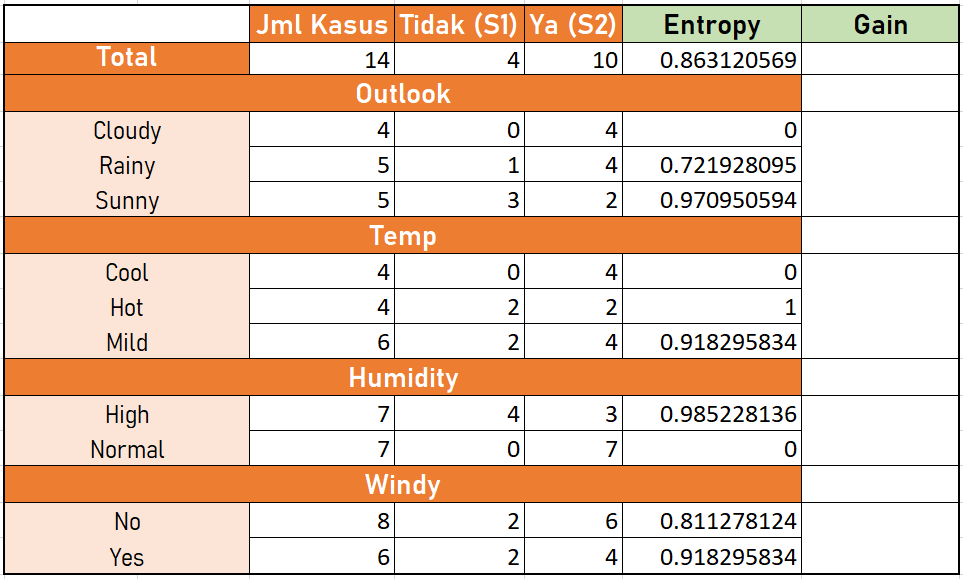
**Perhitungan Node 1 [1]**



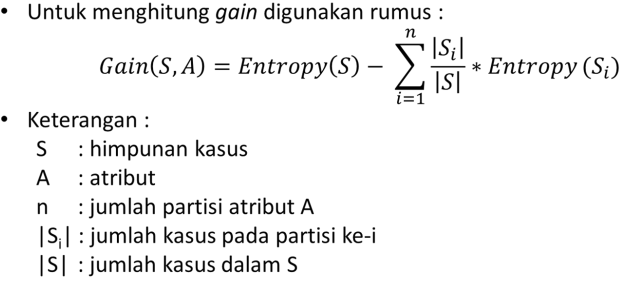
Dengan begitu untuk mencari Entropy dapat menggunakan :

* Entropy(Total) =
* Entropy(Cloudy) =
* Entropy(Raining) =
* dan seterusnya..

Sehingga ditemukan tabel sebagai berikut :

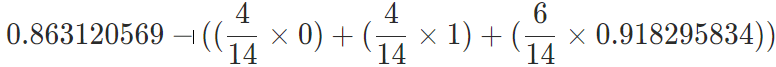


**Perhitungan Node 1 [2]**

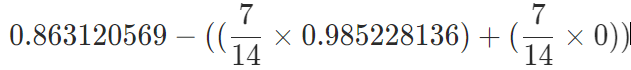
****

Dengan begitu untuk mencari Gain sebagai berikut :

* Gain(Total, Outlook) =



* ****Gain(Total, Temp) =

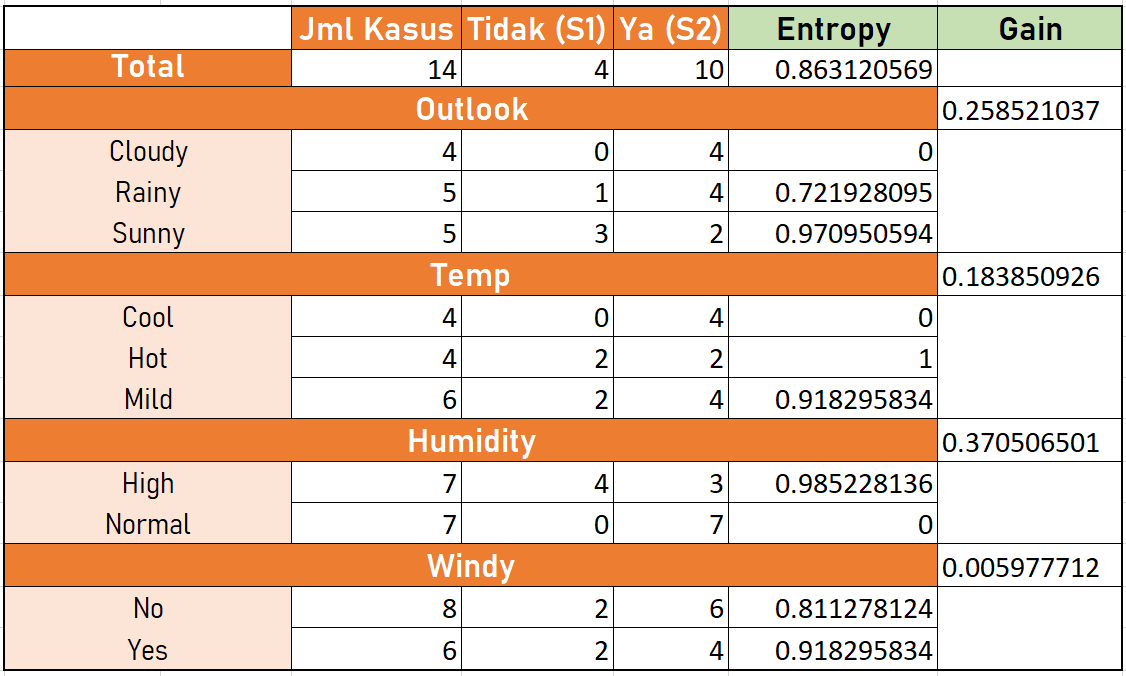


* ****Gain(Total, Humid) =



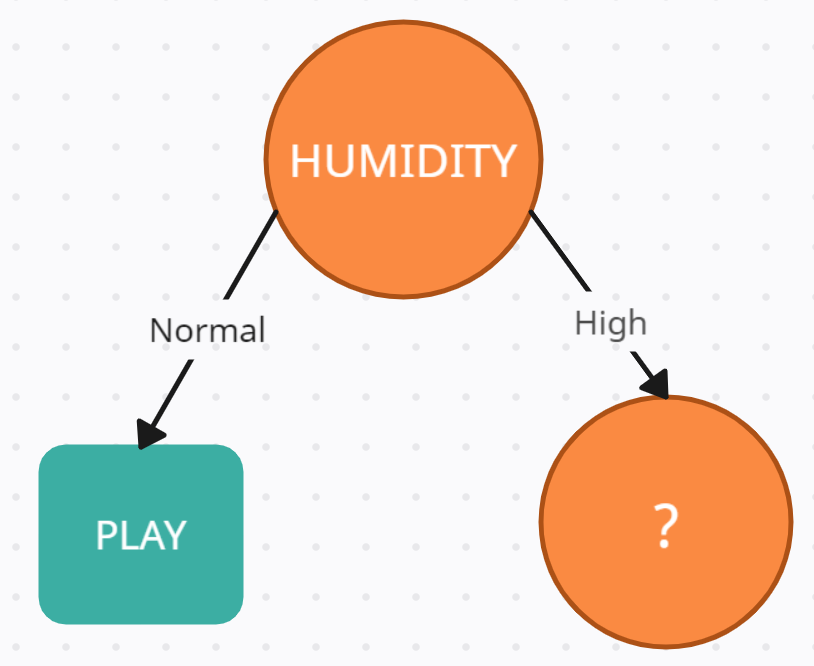
* Gain(Total, Windy) =

Sehingga ditemukan tabel sebagai berikut :



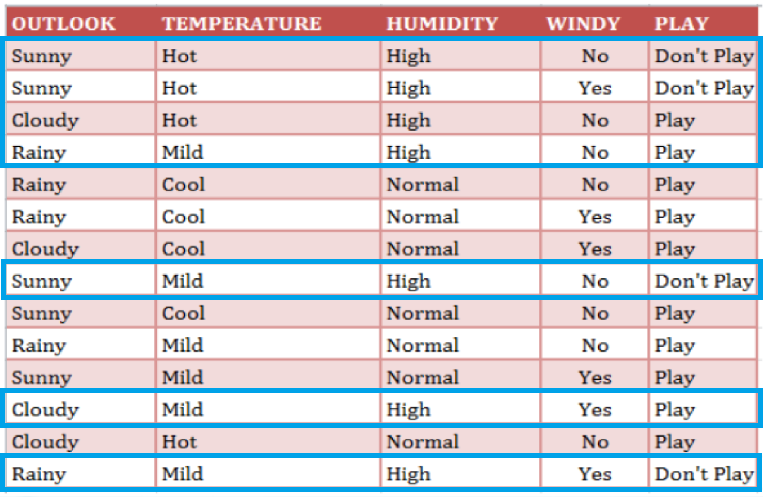
**Perhitungan Node 1 [3]**

* Dapat dilihat bahwa atribut yang memiliki **Gain** tertinggi adalah **HUMIDITY** sehingga **HUMIDITY** dapat menjadi node akar
* Terdapat 2 nilai atribut pada **HUMIDITY** yaitu **HIGH** dan **NORMAL**
* Nilai atribut **NORMAL** sudah mengklasifikasikan kasus menjadi 1 yaitu hanya memiliki keputusan **Play**, sehingga tidak perlu dilakukan perhitungan lebih lanjut
* Tetapi **HIGH** masih perlu dilakukan perhitungan lagi

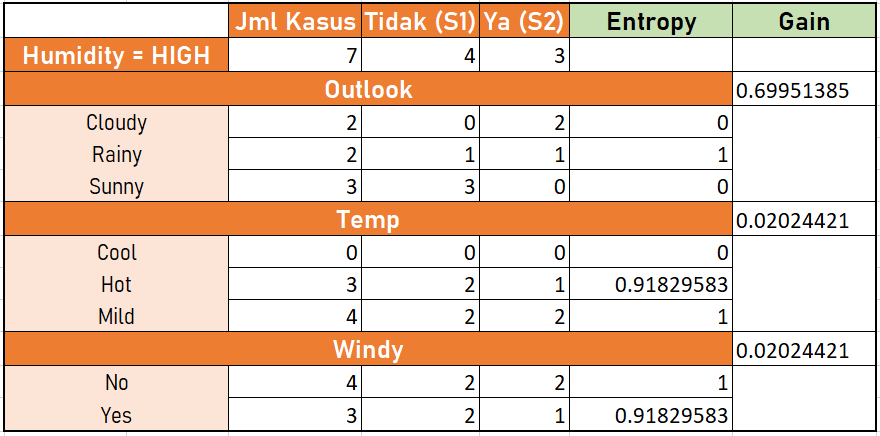


**Perhitungan Node 1.1**

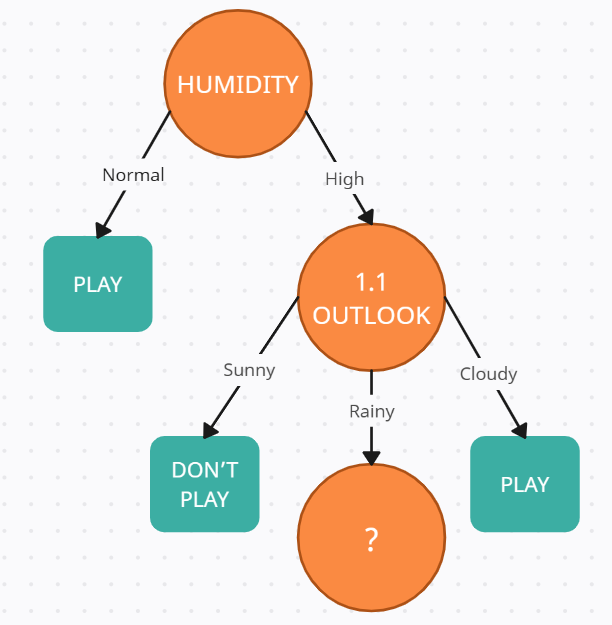
* Setelah itu, melakukan perhitungan **Gain** lagi pada tiap – tiap atribut yang dapat menjadi node akar dari nilai **HUMIDITY = HIGH**



* Sehingga dilakukan lagi perhitungan dengan data yang memiliki **HUMIDITY = HIGH** dengan atribut **OUTLOOK, TEMPERATURE** dan **WINDY**
* Dengan menggunakan rumus sebelumnya untuk mencari **Entropy** dan **Gain** dengan data yang memiliki **HUMIDITY = HIGH** dapat ditemukan tabel baru sebagai berikut

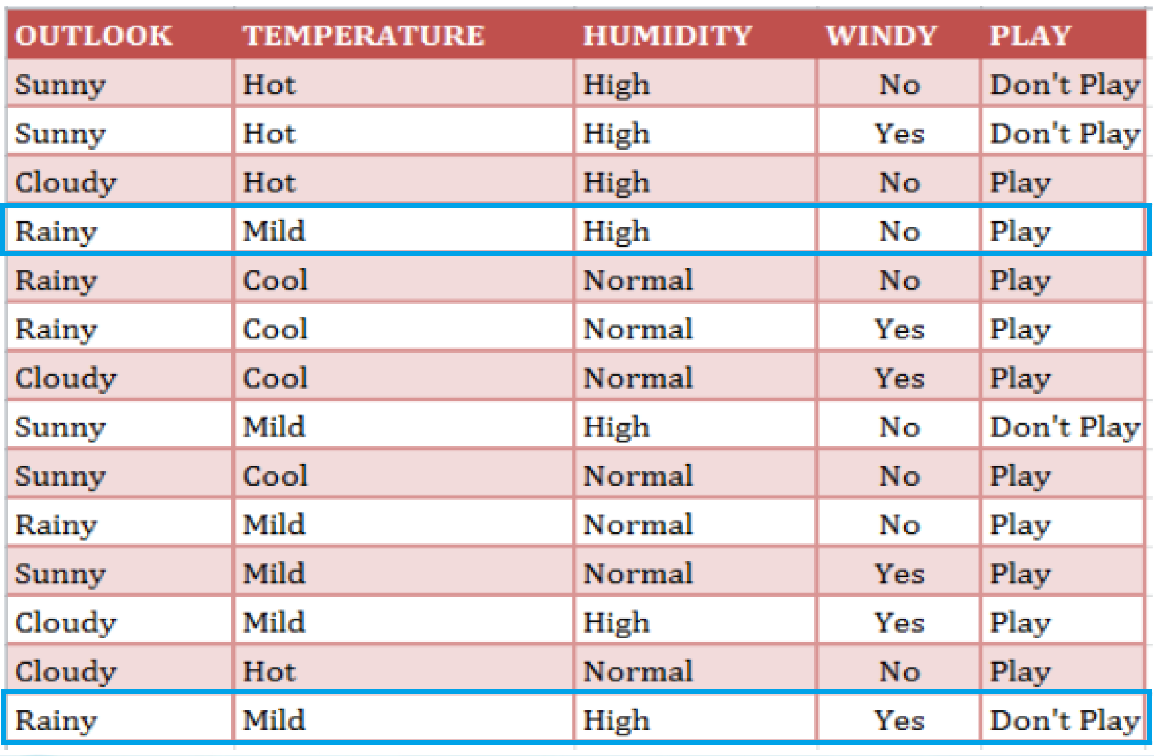


* Dapat dilihat bahwa atribut **OUTLOOK** adalah atribut yang memiliki **GAIN** tertinggi sebesar 0.69951385
* Sehingga **OUTLOOK** dapat dijadikan menjadi node cabang dari nilai atribut **HIGH**
* Ada tiga nilai dari atribut **OUTLOOK**  yaitu :
  + **CLODY 🡺** Klasifikasi Kasus 1 (**PLAY**)
  + **SUNNY 🡺** Klasifikasi Kasus 2 (**DON’T PLAY**)
  + **RAINY 🡺** Perlu perhitungan lagi

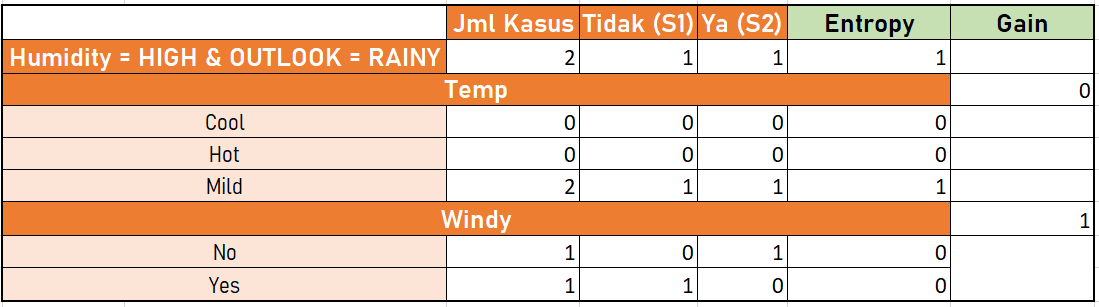


**Perhitungan Node 1.2**

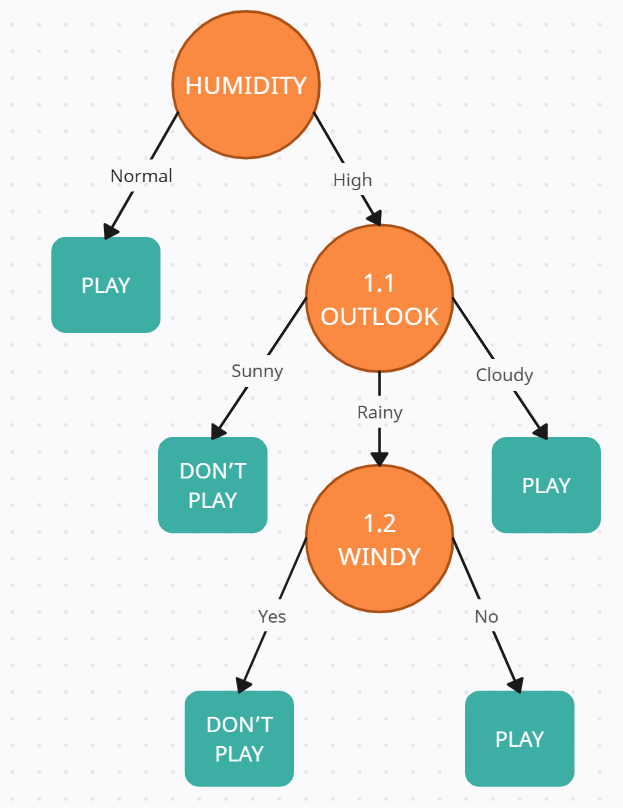
* Setelah itu lakukan perhitungan **GAIN** lagi pada tiap – tiap atribut yang dapat menjadi node akar dari nilai **HUMIDITY = HIGH DAN OUTLOOK = RAINY**

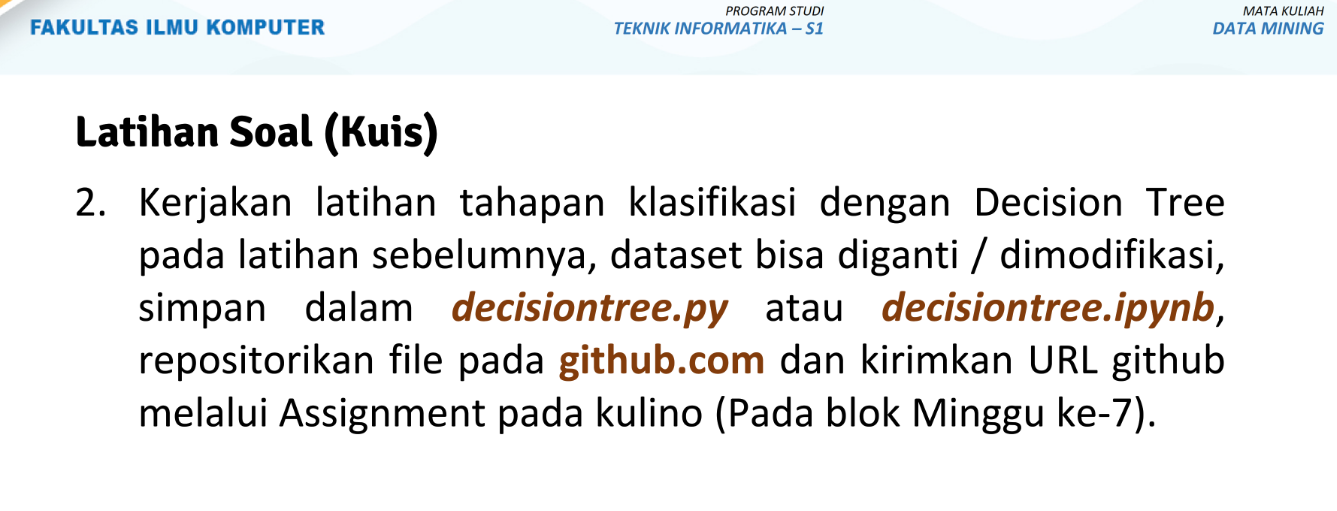


* Sehingga dilakukan lagi perhitungan dengan data yang memiliki **HUMIDITY = HIGH DAN OUTLOOK = RAINY** dengan atribut **TEMPERATURE** dan **WINDY**
* Dengan menggunakan rumus sebelumnya untuk mencari **Entropy** dan **Gain** dengan data yang memiliki **HUMIDITY = HIGH** dan **OUTLOOK = RAINY** dapat ditemukan tabel baru sebagai berikut



* Atribut dengan **GAIN** tertinggi adalah **WINDY** sebesar 1
* Sehingga **WINDY** dapat menjadi node cabang dari nilai atribut **RAINY**
* Ada dua nilai dari atribut **WINDY** yaitu **NO** dan **YES**
  + Nilai **NO** sudah mengklasifikasikan kasus menjadi 1 (**PLAY)**
  + Nilai **YES** sudah mengklasifikasikan kasus menjadi 1(**DON’T PLAY)**
  + Sehingga tidak perlu dilakukan perhitungan lagi
* Sehingga berikut ini adalah decision tree akhir





**Jawaban :**

**LINK GITHUB :** <https://github.com/luthfikamalananda/Decision-Tree>