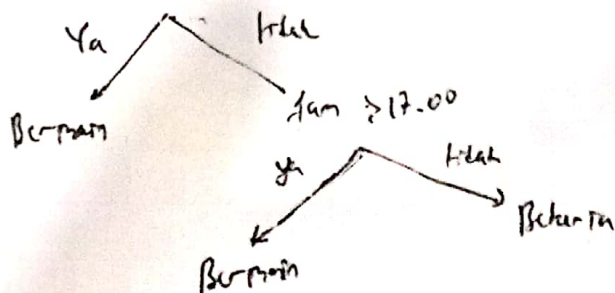


Decision Tree Classifier

- Salah satu metode untuk klasifikasi. Metode untuk mengambil keputusan

ex Hari libur?



Gini Impurity

$$Gini = 1 - \sum_{i=1}^n (P_i)^2, \quad n = \text{banyak kelas}$$

$$Gini \text{ impurity} = 1 - Pr(ya)^2 - Pr(tidak)^2$$

↳ Index untuk menentukan suatu fitur sudah baik atau belum.

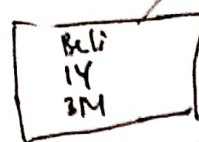
#	Contoh Data	Study	Price	Beli
	Primary Needs	Study		
	Y	Y	7k	N
	Y	N	1k	N
	Y	Y	18k	Y
	N	Y	35k	Y
	N	Y	38k	Y
	Y	N	50k	N
	Y	N	83k	N
	N	N		

Study gini $G_i = 0.22$

Uraikan data angka/numerik:

- Urutkan
- Ambil rata-rata untuk setiap 2 data yg berurutan
- Hitung GI untuk semua rata-rata (n-1)
- Ambil yg terbaik.

kebutuhan primer
Ya = 1
Tidak = 3



$$G_i = 1 - \left(\frac{4}{7}\right)^2 - \left(\frac{3}{7}\right)^2 = 0.38$$

$$G_i = 1 - \left(\frac{2}{3}\right)^2 - \left(\frac{1}{3}\right)^2 = 0.44$$

$$\text{Total } G_i = \frac{1}{7} \times 0.38 + \frac{3}{7} \times 0.44 = 0.4$$

Ambil GI yg terkecil untuk dijadikan fitur.