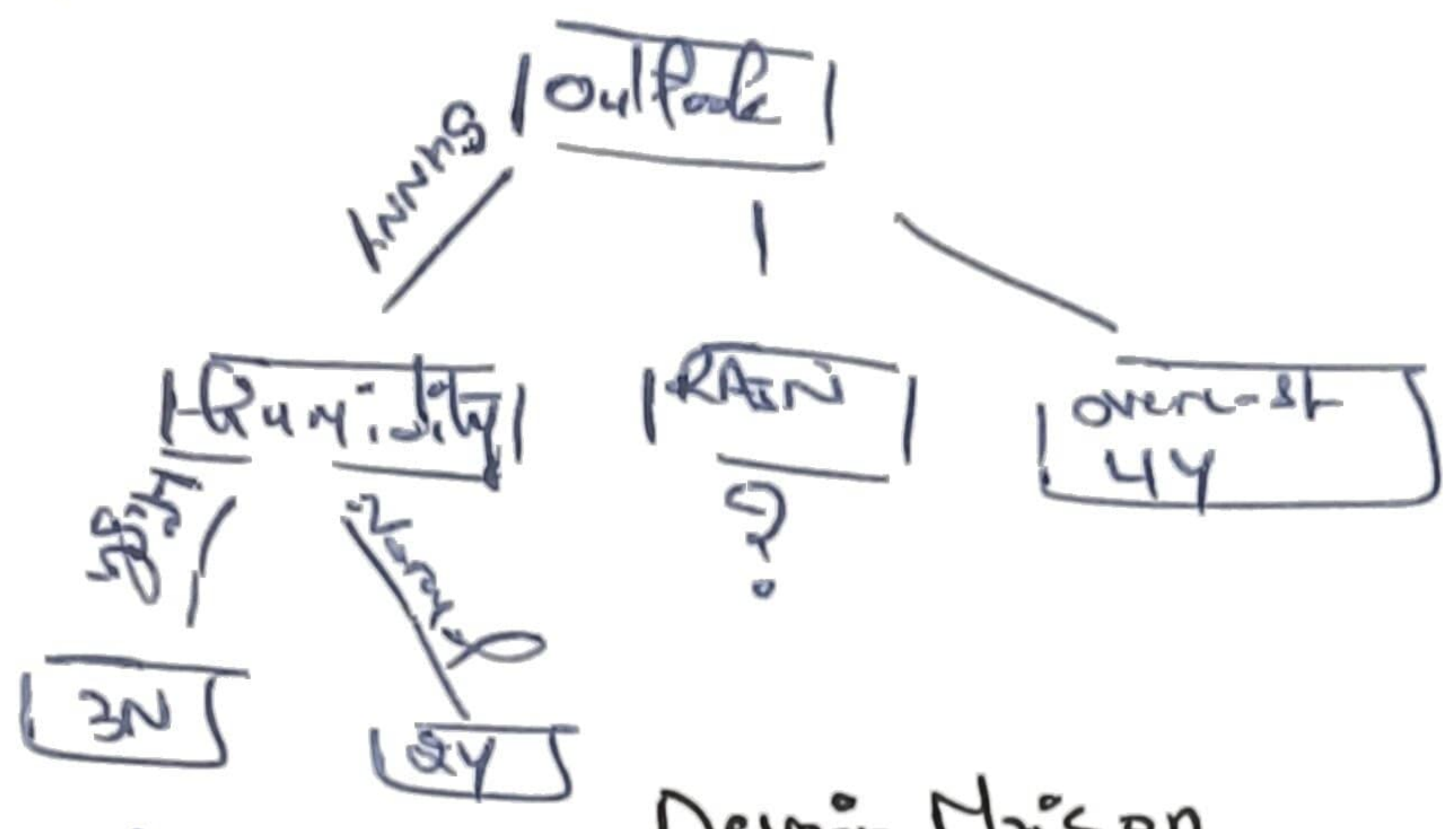


Et donc :



Devoir Maison

on aura : le nouveau tableau de données :

RAIN :

Temp	Humid	wind	play
Mild	High	weak	yes
Cool	Normal	weak	yes
Cool	Normal	strong	no
Mild	Normal	weak	yes
Mild	High	strong	no

on a :

$$S_2 = 3Y \text{ et } 2N$$

et donc $E(S_2) = -\frac{3}{5} \log_2(3/5) - \frac{2}{5} \log_2(2/5) = 0,971$

* Temp : $E(\text{Mild} | S_2) = -\frac{2}{3} \log_2(2/3) - \frac{1}{3} \log_2(1/3) = 0,9182$

$$E(\text{Cool} | S_2) = -\frac{1}{2} \log_2(1/2) - \frac{1}{2} \log_2(1/2) = 1$$

$$G(\text{Temp} | S_2) = 0,971 - 3/5(0,9182) - 2/5(1) = 0,020$$

* Humid : $E(\text{High} | S_2) = -\frac{1}{2} \log_2(1/2) - \frac{1}{2} \log_2(1/2) = 1$

$$E(\text{Normal} | S_2) = -\frac{2}{3} \log_2(2/3) - \frac{1}{3} \log_2(1/3) = 0,9182$$

$$G(\text{Humid} | S_2) = 0,971 - 2/5(1) - 3/5(0,9182) = 0,020$$

* wind : $E(\text{weak} | S_2) = 0 \text{ (3yes)}$

$$E(\text{strong} | S_2) = 0 \text{ (2no)} \text{ et donc}$$

$$G(\text{wind} | S_2) = E(S_2) = 0,971$$

le gain de wind est le plus grand \Rightarrow wind est la variable la plus pertinente on aura comme resultat l'arbre suivant :

