Assignment - 4

Develop a decision tree using Gini index.

Day	outlook	Temperature	Humidity	Wind	Play Tennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	yes
4	Rain	Mild	High	Weak	yes
5	Rain	Cool	Normal	Weak	yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	yes
10	Rain	Mild	Normal	Weak	yes
11	Sunny	Mild	Normal	Strong	yes
12	Overcast	Mild	High	Strong	yes
13	Overcast	Hot	Normal	Weak	yes.
14	Rain.	Mild.	High	Strong	No

Outlook : Sunny | Overcast | Rain.

[2+, 3-] | [4+, 0] | [3+, 2-].

Temp : Hot | Cool | Mild.

[2+, 2-] | [3+, 1-] | [4+, 2-].

Humidity : High | Normal

[3+, 4-] | [6+, 1-]

Wind : Weak | Strong.

[6+, 2-] | [3+, 2-].

$Gini(S) = 1 - \sum_{i=1}^{n} p_i^2$

$1 - \left(\left(\frac{9}{14}\right)^2 + \left(\frac{5}{14}\right)^2\right) = 1 - (0.642)^2 + (0.357)^2$

$1 - 0.4132 + 0.1275$

$= 0.4607 = 0.4592$

Gini(Outlook) $= (5/14 \times 0.48) + (4/14 \times 0) + (5/14 \times 0.48) = 0.3428$

Gini(Sunny) $= 1 - ((2/5)^2 + (3/5)^2) = 1 - 0.16 - 0.36 \Rightarrow 0.48$

Gini(Overcast) $= 1 - ((4/4)^2 + (0/4)^2) = 1 - 1 \Rightarrow 0.$

Gini(Rain) $= 1 - ((3/5)^2 + (2/5)^2) = 1 - 0.16 - 0.36 \Rightarrow 0.48$

Gini(Temp) $(4/14 \times 0.5) + (4/14 \times 0.38) + (6/14 \times 0.45) = 0.4407$

Gini(Hot) $= 1 - ((2/4)^2 + (2/4)^2) \Rightarrow 1 - 0.5 = 0.5$

Gini(Cool) $= 1 - ((3/4)^2 + (1/4)^2) \Rightarrow 1 - 0.56 + 0.06 = 0.38 = 062$

Gini(Mild) $= 1 - ((4/6)^2 + (2/6)^2) = 1 - 0.44 + 0.11 = 0.45$

Gini(Humidity) $= (7/14 \times 0.489) + (7/14 \times 0.244) = 0.366$

Gini(High) $= 1 - ((3/7)^2 + (4/7)^2) = 0.489$

Gini(Normal) $= 1 - ((6/7)^2 + (1/7)^2) = 0.244$

Gini(Wind) $= (8/14 \times 0.375) + (5/14 \times 0.5) = 0.428$

Gini(Weak) $= 1 - ((6/8)^2 + (2/8)^2) = 0.375$

Gini(Strong) $= 1 - ((3/6)^2 + (3/6)^2) = 0.5$

Gini = 0.3428, 0.4407, 0.366, 0.428

set the outlook as the root node and build the feature tree

Gini Index of outlook..

Gini(Sunny) = $1-(2/5)^2-(3/5)^2 = 0.48$

Gini(Overcast) = $1-(4/4)^2-(0/4)^2 = 0$

Gini(Rain) = $1-(3/5)^2-(2/5)^2 = 0.48$

Sunny look for other Attributes.

Sunny - Temp		Humidity		Wind	
Hot	mild · cool	High	Normal.	weak	strong
[2-, 0+]	[1+, 1-] [1+, 0-]	[3-, 0+]	[2+, 0-]	[2-, 1+]	[1+, 1-]

Rain. - Temp		Humidity		Wind.	
~~cool~~ mild.		High	Normal	weak	strong
[1+, 1-]	[2+, 1-]	[1+, 1-]	[2+, 1-]	[3+, 0-]	[2-, 0+]

Gini(S, Temp, H) = $1-((0/2)^2+(2/2)^2+0) = 0$

(S, Temp, M) = $1-(1/2)^2+(1/2)^2 = 0.5$

(S, Temp, Cool) = $1-((0/1)^2+(1/1)^2) = 0$ $\frac{2}{5}\times 0+\frac{2}{5}\times 0.5+\frac{1}{5}\times 0$

Gini(S, Hum, H) = $1-((3/3)^2+(0/3)^2) = 0$ $\frac{3}{5}\times 0+\frac{2}{5}\times 0 = 0$ ✓

Gini(S, Hum, N) = $1-((2/2)^2+(0/2)^2) = 0$

Gini(S, Wind Weak) = $1-((2/3)^2+(1/3)^2) = 0.44$ $\frac{3}{5}\times 0.44+\frac{2}{5}\times 0.5$

Gini(S, Wind strong) = $1-((1/2)^2+(1/2)^2) = 0.5$

Gini(R, Temp Cool) = $1-((1/2)^2+(1/2)^2) = 0.5$ $\frac{2}{5}\times 0.5+\frac{3}{5}\times 0.44$

Gini(R, Temp mild) = $1-((2/3)^2+(1/3)^2) = 0.44$

Gini(R, Hum, H) = $1-((1/2)^2+(1/2)^2) = 0.5$ $\frac{2}{5}\times 0.5+\frac{3}{5}\times 0.44$

Gini(R, Hum, N) = $1-((2/3)^2+(1/3)^2) = 0.44$

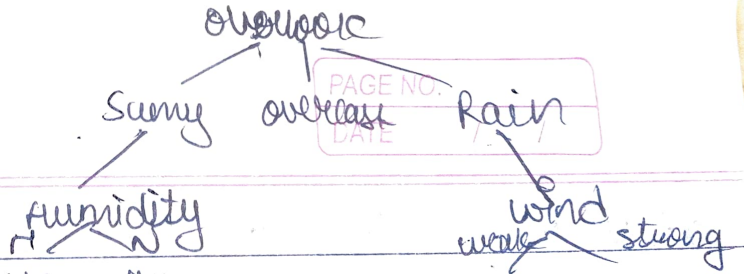
Gini(R, Wind, Weak) = $1-((3/3)^2+(0/3)^2) = 0$ $\frac{3}{5}\times 0+\frac{2}{5}\times 0 = 0$ ✓

Gini(R, Wind, Strong) = $1-((0/2)^2+(2/2)^2) = 0$

sunny → Humidity →

overcast → yes

Rain → wind →



Assignment-5