CP-March-18-2025

Ensrin Debbarma

A) Yes: 8
No: 4
Total: 12

$P(Yes) = \frac{8}{12} = \frac{2}{3}$

$P(No) = \frac{4}{12} = \frac{1}{3}$

$Gini(parent) = 1 - \left[\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2\right]$

$= 1 - \left[\frac{4}{9} + \frac{1}{9}\right]$

$= 1 - \frac{5}{9}$

$= \frac{4}{9}$

≈ 0.44

Outlook

Sunny: Days: (1, 2, 8, 9, 11)

Decisions: No, No, No, Yes, Yes

$Gini(Sunny) = 1 - \left[\left(\frac{3}{5}\right)^2 + \left(\frac{2}{5}\right)^2\right]$

$= 1 - \left[\left(\frac{9}{25} + \frac{4}{25}\right)\right]$

$= 1 - \frac{13}{25}$

$= 1 - \frac{13}{25} = 0.48$

∴ Weighted by $\frac{5}{12} = \frac{5}{12} \times 0.48$

$= 0.20$

Gini (Overcast): Days: (3, 7, 12) Decisions: Yes, Yes, Yes

$= 1 - \left[\left(\frac{3}{3}\right)^2 + 0^2\right] = 1 - 1 = 0$

weighted by $\frac{3}{12} = \frac{3}{12} \times 0 = 0$.

Rain: Days: (4, 5, 6, 10)

Decisions: Yes, Yes, No, Yes

* $Gini(Rain) = 1 - \left[\left(\frac{3}{4}\right)^2 + \left(\frac{1}{4}\right)^2\right]$

$= 1 - \left[\frac{9+1}{16}\right]$

$= 1 - \frac{10}{16}$

$= \frac{16-10}{16}$

$= \frac{6}{16} = 0.375$

weighted by $\frac{4}{12} = \frac{4}{12} \times 0.375$

$= 0.125$

Total Gini = 0.20 + 0 + 0.125

= 0.325

Gini gain = 0.44 - 0.325

= 0.1194

Gini for temperature ≈ 0.075

Gini for Humidity ≈ 0.0555

Gini for wind ≈ 0.0277

∴ Outlook best root split.

:) Build tree

1. Overcast Branch

Days 3, 7, 12 → all yes

→ pure leaf = Yes.

2. Sunny Branch

Days 1, 2, 8, 9, 11 → decisions (No, No, No, Yes, Yes)

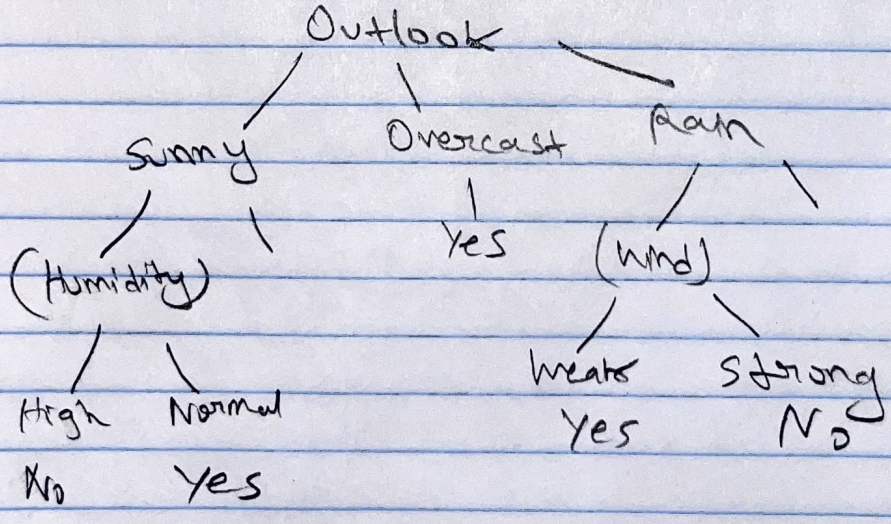
- Humidity:
  - High → Days 1, 2, 8 → all No → leaf = No.
  - Normal → Days 9, 11 → both Yes → leaf = Yes.

3. Rain branch

day 4, 5, 6, 10 → decisions: (Yes, Yes, No, Yes)

Wind

- Weak → Days 4, 5, 10 → all yes → leaf = Yes
- Strong → Day 6 → No → leaf = No



1. Sunny, Mild, High, Strong, No
2. Rain, Hot, Normal, Strong, No
3. Overcast, Mild, Normal, Weak, Yes.

```
Running Naive Bayes classifier on the iris dataset...

Training data: 100 samples with 4 features
Prior probabilities:
  P(C+): 0.34
  P(C-): 0.66
  True Positives: 16
  False Positives: 1
  True Negatives: 33
  False Negatives: 0
  Accuracy: 0.98
  Precision: 0.9411764705882353
  Recall: 1.0
PS C:\Users\nrndb\OneDrive\Desktop\CMPSC 445\Assignments\HW 4> python HW4.py buy
Running Naive Bayes classifier on the buy dataset...

Training data: 14 samples with 4 features
Prior probabilities:
  P(C+): 0.6428571428571429
  P(C-): 0.35714285714285715
  True Positives: 2
  False Positives: 1
  True Negatives: 1
  False Negatives: 0
  Accuracy: 0.75
  Precision: 0.6666666666666666
  Recall: 1.0
```