Assignment_5_task_5.2

Machine Learning (WiSe 2025/2026)

Author: Suvansh Shukla
Matriculation No. 256245

Assignment 5 Task 5.2

To calculate the gain ration we need information gain and intrinsic information (a.k.a Split Info).

$Gain\ Ratio = \frac{Information\ Gain}{Intrinsic\ Information}$

$Information\ Gain = Entropy(Parent) - \sum_i \frac{N_i}{N}.Entropy(Child_i)$

$Intrinsic\ Information = -\sum_i \frac{N_i}{N}.log_2\left(\frac{N_i}{N}\right)$

Part A

Since we are only using perfect splits (2.5, 3.5, 5.5, 9.5), this would mean that the Entropy
of all the children after the root would simply be Zero.

This means that the Information Gain for such a tree would be equal to it's root node.

So...

$Information\ Gain = -\frac{10}{15}log_2\frac{10}{15} - \frac{5}{15}log_2\frac{5}{15}$

$Information\ Gain = 0.9182$

$Intrinsic\ Information = -\frac{10}{15}log_2\frac{10}{15} - \frac{5}{15}log_2\frac{5}{15}$

$Intrinsic\ Information = 0.9182$

So, $Gain\ Ratio = 1$

Part B

We see that splitting at 5.5, 9.5, gives us a tree with all leaf nodes with zero entropy.

There is only one other node where there is no pure split. So,

$Information\ Gain = 0.918 - (0 + \left(\frac{4}{15}log_2\frac{4}{15} + \frac{5}{15}log_2\frac{5}{15}\right) + 0 + 0)$

$Information\ Gain = 0.918 - (-0.991) = 1.909$

$Intrinsic\ Information = -\left(\frac{6}{15}log_2\frac{6}{15} + \frac{4}{15}log_2\frac{4}{15} + \frac{5}{15}log_2\frac{5}{15}\right)$

$Intrinsic\ Information = 1.547$

$Gain\ Ratio = \frac{1.909}{1.547} = 1.234$

Part C

Splitting at 9.5 gives us a 2 level decision tree, with 5 instances in one node and 10 in the other.

The leaf node with 10 instances also has 0 entropy!

$Information\ Gain = 0.918 - \left(0 + \left(\frac{5}{15}log_2\frac{5}{15} + \frac{5}{15}log_2\frac{5}{15}\right) + 0\right)$

$Information\ Gain = 0.918 - (-1) = 1.918$

$Intrinsic\ Information = -1$

So, $Gain\ Ratio = -1.918$