

Report

I did not get the output as desired. Especially during the find function and entropy classes as I repeatedly got some errors or the other regarding the number of function calls being out of bound. Hence I am attaching the pseudocode and tried to explain the logic in the best way I can.

Function find

- Combined the attributes and classes into a mx16 cross array
- I separated the positive and negative values to separate array
- Found info gain for each iteration (comparing current attribute to the class)
- Found the maximum to split on..ie the root in first case

Function train

- check to see if zero and majority votes based on number of elements in class
- In the recursive part, I called the find function. Initialised the left and right node. Tested the class for zero or one and recalled the function again sending the corresponding nodes.

Function classify

- If the split has not occurred then it is a leaf node. Hence change the class with the decision attribute.
- If split has occurred, split it into left and right.

Function entropy (P)

$$h = -1 * p(1) * \log_2(p(1)) + -1 * p(2) * \log_2(p(2)) ;$$

Function entropy(class)

- Find the unique number of ones and zeros
- Hence compute probability and entropy using
$$h = -1 * p(1) * \log_2(p(1)) + -1 * p(2) * \log_2(p(2)) ;$$