

To show that the maximum Gini index of a node is  $1 - \frac{1}{n_c}$  when all records are equally distributed among  $n_c$  different classes, we can start by recalling the formula for the Gini index.

The Gini index for a node is calculated as:

$$Gini = 1 - \sum_{i=1}^{n_c} p_i^2$$

where  $p_i$  is the proportion of records in the node that belong to class  $i$ , and  $n_c$  is the number of different classes.

### Step 1: Set up equal distribution

When all records are equally distributed among the  $n_c$  classes, each class will have the same proportion of records. Therefore, for each class  $i$ :

$$p_i = \frac{1}{n_c}$$

### Step 2: Substitute the proportions into the Gini index formula

Now substitute  $p_i = \frac{1}{n_c}$  into the Gini index formula:

$$Gini = 1 - \sum_{i=1}^{n_c} \left( \frac{1}{n_c} \right)^2$$

### Step 3: Simplify the summation

Since all  $p_i$ 's are the same, the summation simplifies to:

$$Gini = 1 - n_c \times \left( \frac{1}{n_c^2} \right)$$

$$Gini = 1 - \frac{n_c}{n_c^2}$$

$$Gini = 1 - \frac{1}{n_c}$$

### Conclusion

Thus, when all records are equally distributed among the  $n_c$  classes, the Gini index is:

$$Gini = 1 - \frac{1}{n_c}$$

This shows that the maximum Gini index of a node is  $1 - \frac{1}{n_c}$  when the records are equally distributed among  $n_c$  classes.