GINI INDEX

Proportion of observations in the m^{th} leaf of K^{th} class.

$$G = \sum_{K=1}^{K} \hat{p}_{mK}\left(1 - \hat{p}_{mK}\right)$$

leaf Class leaf Class

The smaller the value of G the more purity there is in the node.

Used at each node to decide which feature is best to split on.

Measure of purity in tree based methods

BY CHRIS ALBON