

4 Node CART® Classification: luof_boolean versus Income_10k, NH_W...

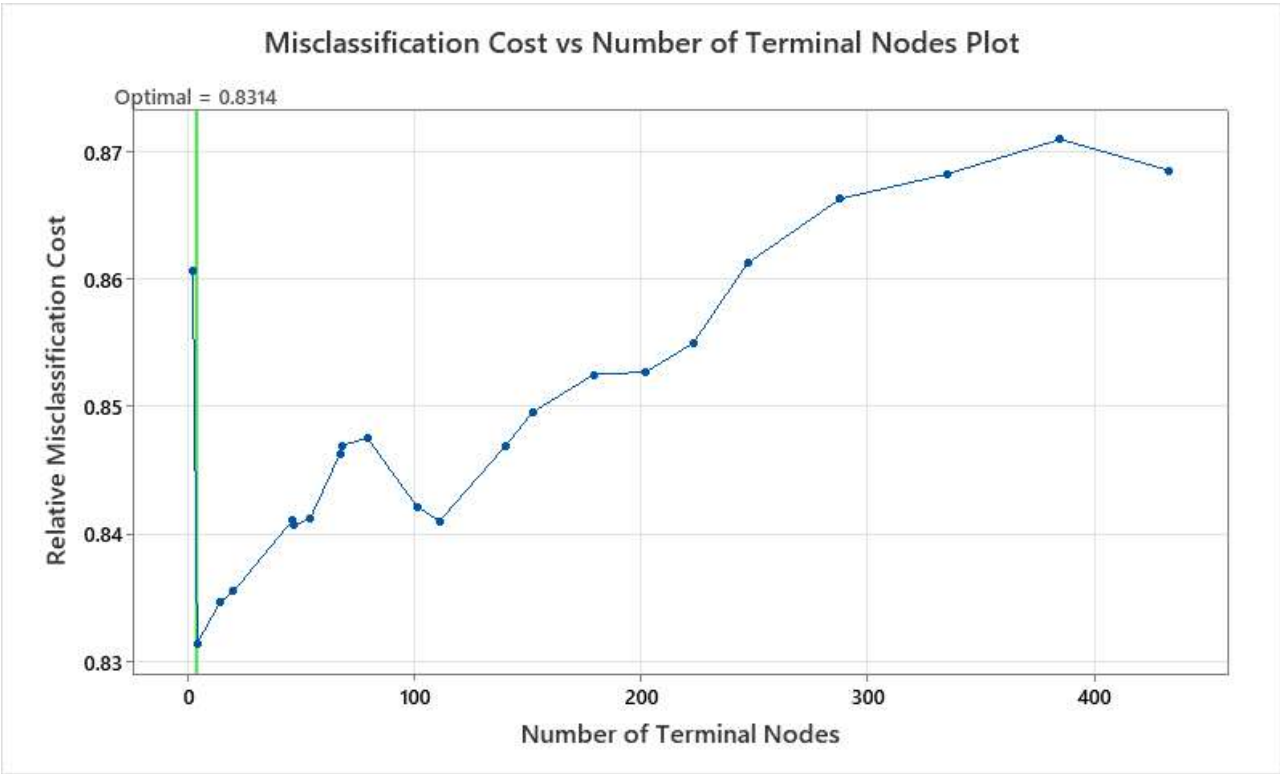
* WARNING * Minitab stopped growing the tree at 1024 nodes due to default limits.

Method

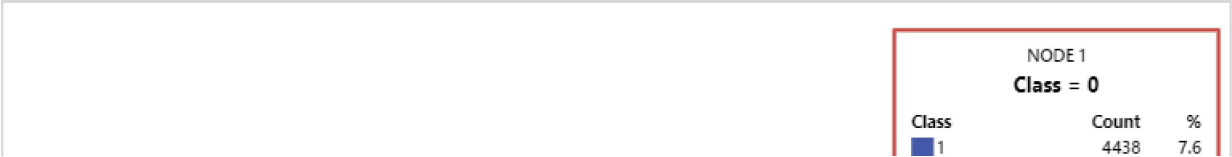
Prior probabilities	Same for all classes
Node splitting	Gini
Optimal tree	Within 1 standard error of minimum misclassification cost
Model validation	70/30% training/test sets
Rows used	82906

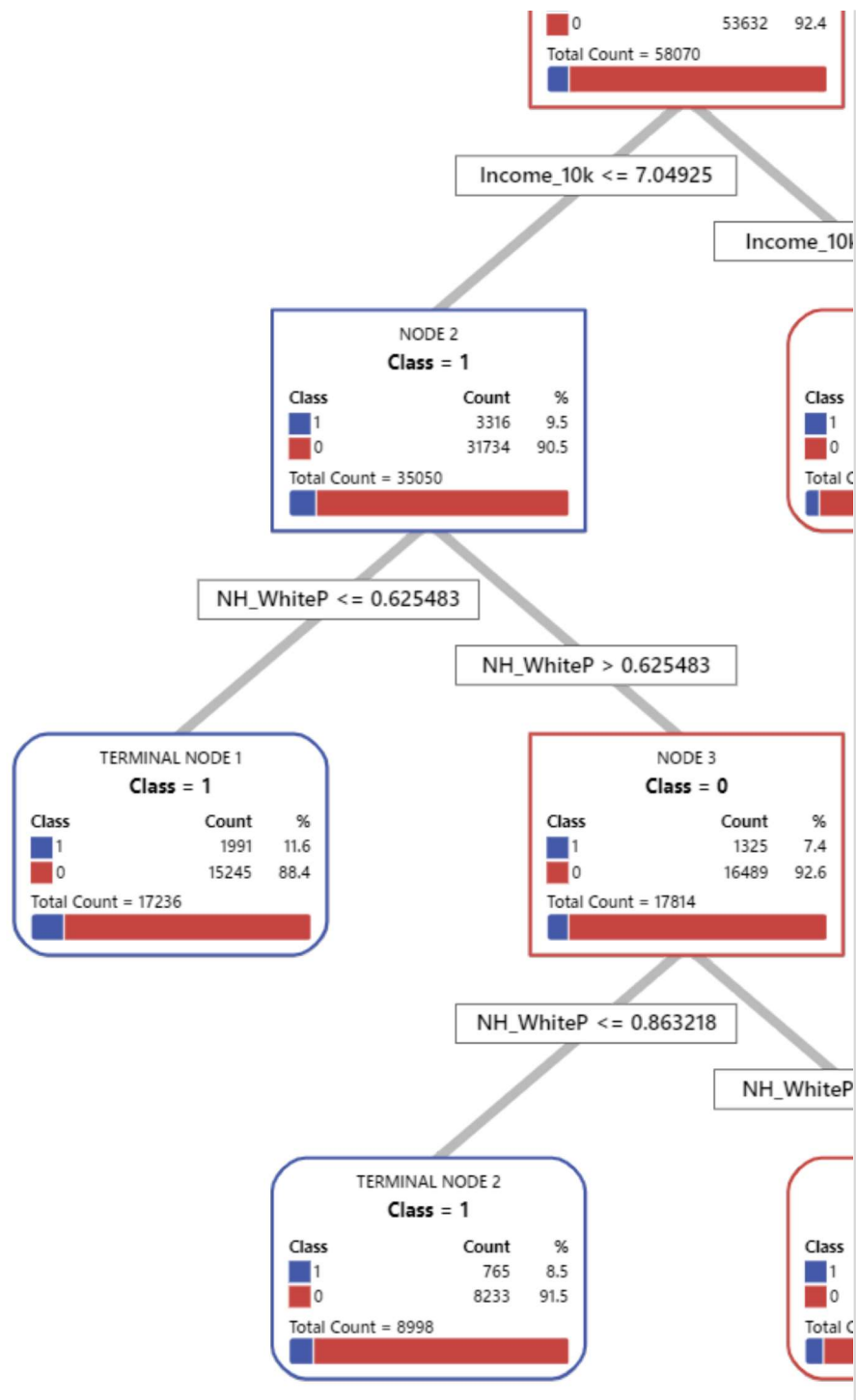
Binary Response Information

Variable	Class	Training		Test	
		Count	%	Count	%
luof_boolean	1 (Event)	4438	7.64	1899	7.65
	0	53632	92.36	22937	92.35
	All	58070	100.00	24836	100.00



Optimal Tree Diagram



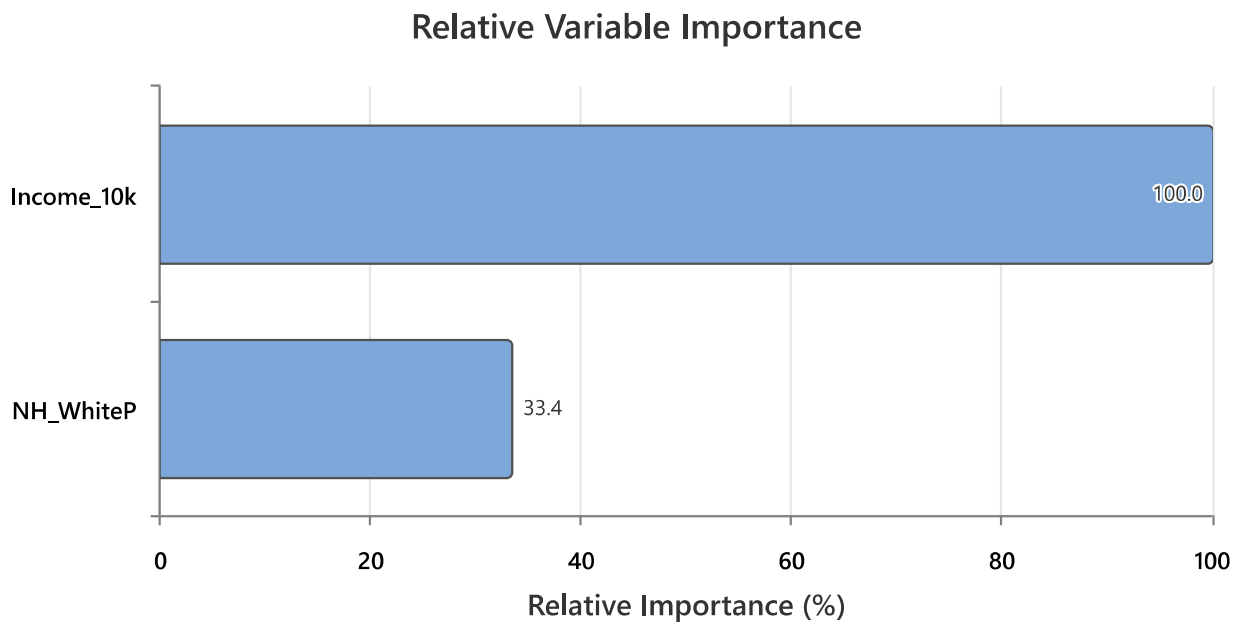


Model Summary

Total predictors	2
Important predictors	2
Number of terminal nodes	4
Minimum terminal node size	8816

Statistics	Training	Test
Average -loglikelihood	0.2644	0.2655
Area under ROC curve	0.6079	0.5983
95% CI	(0.1375, 1)	(0.1277, 1)
Lift	1.5115	1.4742

Misclassification cost 0.8168 0.8314



Variable importance measures model improvement when splits are made on a predictor. Relative importance is defined as % improvement with respect to the top predictor.

Confusion Matrix

Actual Class	Predicted Class (Training)				Predicted Class (Test)			
	Count	1	0	% Correct	Count	1	0	% Correct
1 (Event)	4438	2756	1682	62.1	1899	1148	751	60.5
0	53632	23478	30154	56.2	22937	9999	12938	56.4
All	58070	26234	31836	56.7	24836	11147	13689	56.7

Statistics	Training (%)	Test (%)
True positive rate (sensitivity or power)	62.1	60.5
False positive rate (type I error)	43.8	43.6
False negative rate (type II error)	37.9	39.5
True negative rate (specificity)	56.2	56.4

Misclassification

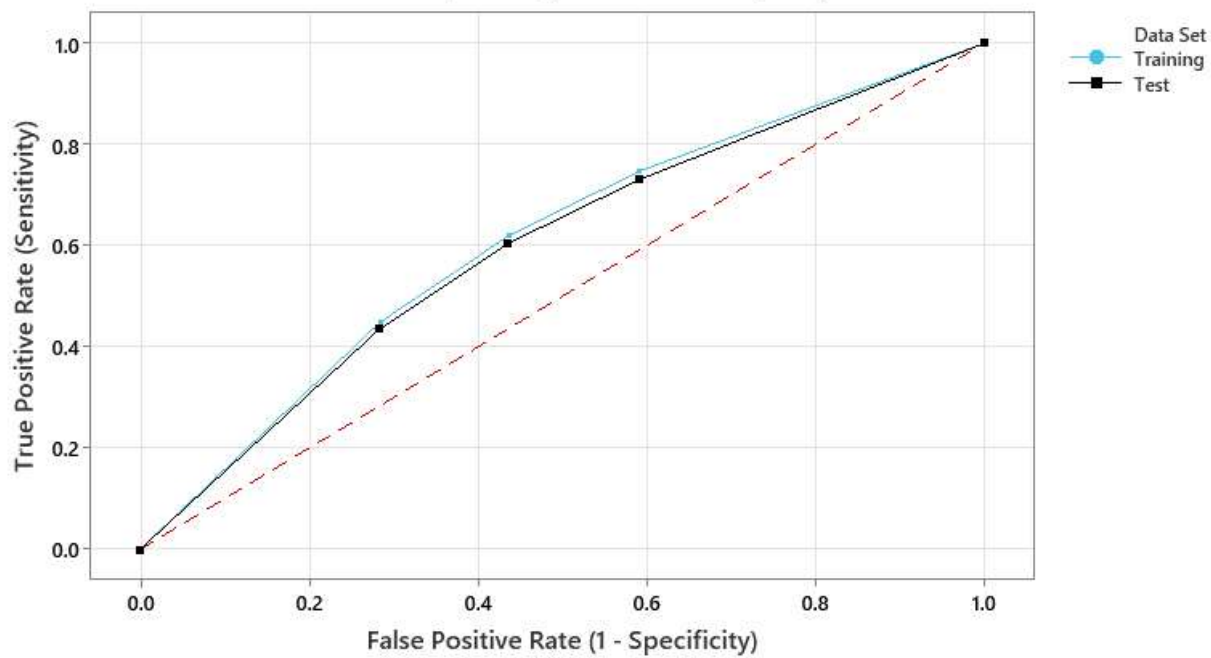
Input

Misclassification Cost	Predicted Class
------------------------	-----------------

Actual Class	1	0
1		1.00
0	1.00	

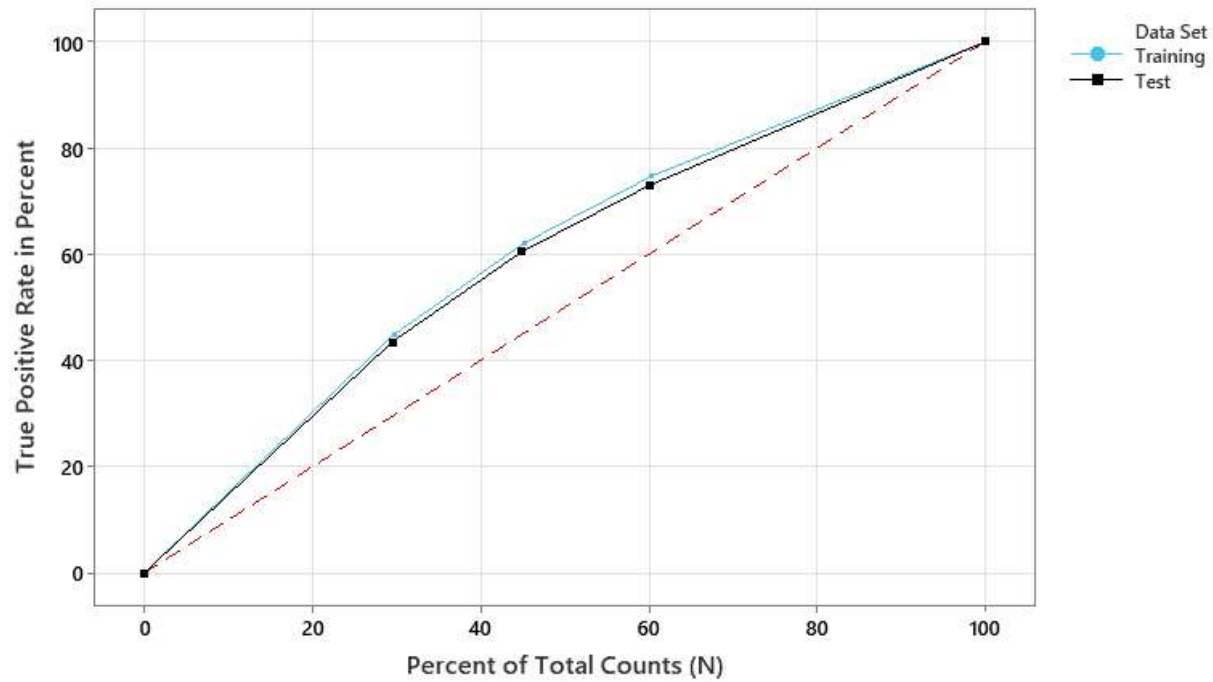
Actual Class	Training				Test			
	Count	Misclassified	% Error	Cost	Count	Misclassified	% Error	Cost
1 (Event)	4438	1682	37.9	0.3790	1899	751	39.5	0.3955
0	53632	23478	43.8	0.4378	22937	9999	43.6	0.4359
All	58070	25160	43.3	0.4084	24836	10750	43.3	0.4157

Receiver Operating Characteristic (ROC) Curve



Area Under Curve: Training = 0.6079, Test = 0.5983

Gain Chart



Lift Chart

