ALL_TRACTS.CSV

4 Node CART® Regression: luof_boolean versus Income_10k, NH_WhiteP

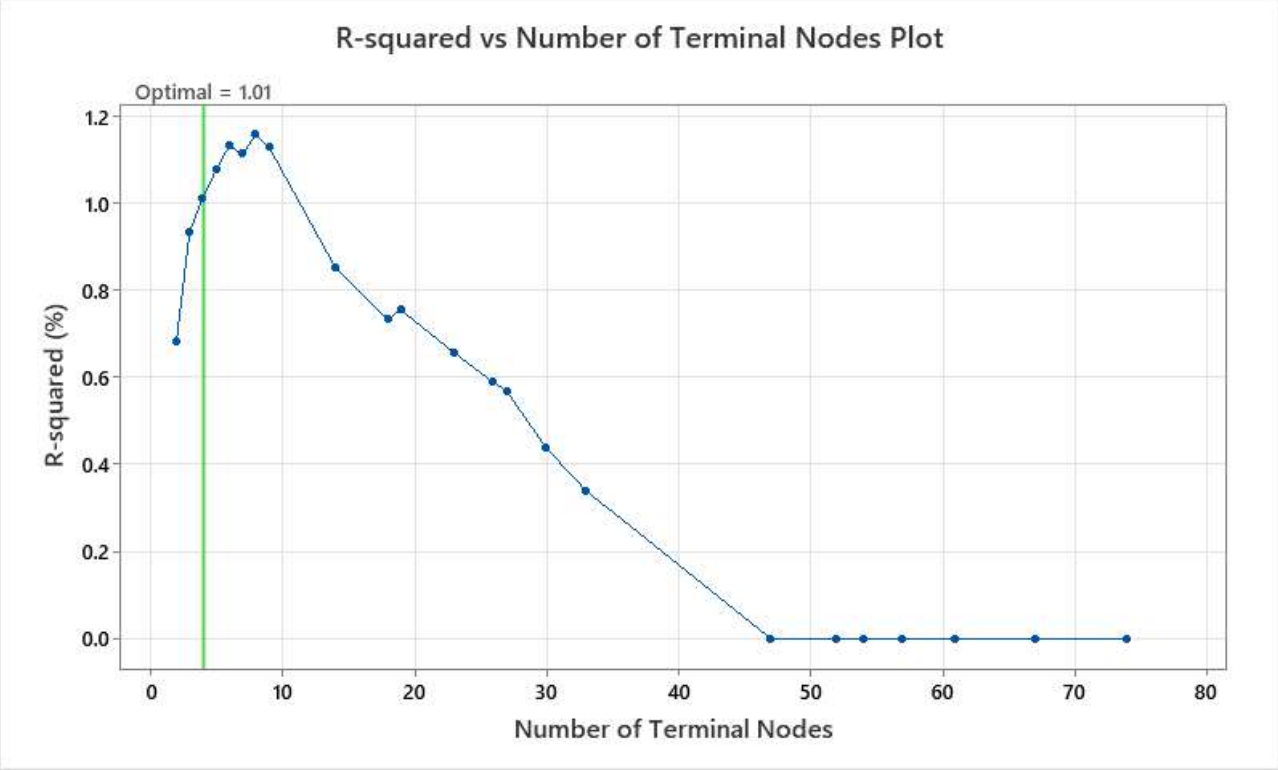
*** WARNING * Minitab stopped growing the tree at 1024 nodes due to default limits.**

Method

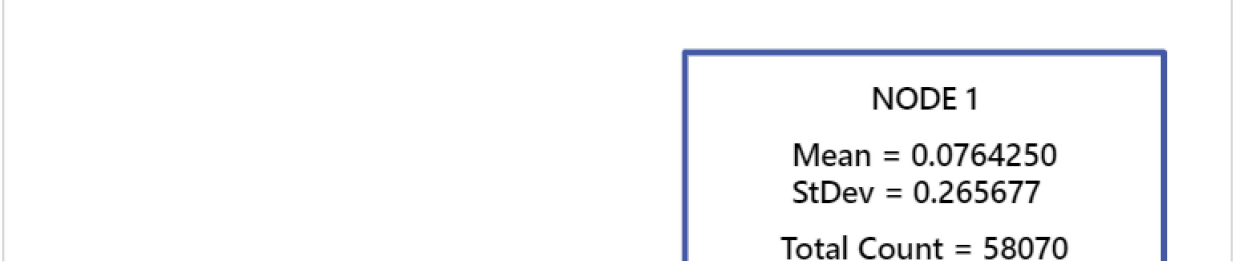
Node splitting	Least squared error
Optimal tree	Within 1 standard error of maximum R-squared
Model validation	70/30% training/test sets
Rows used	82906

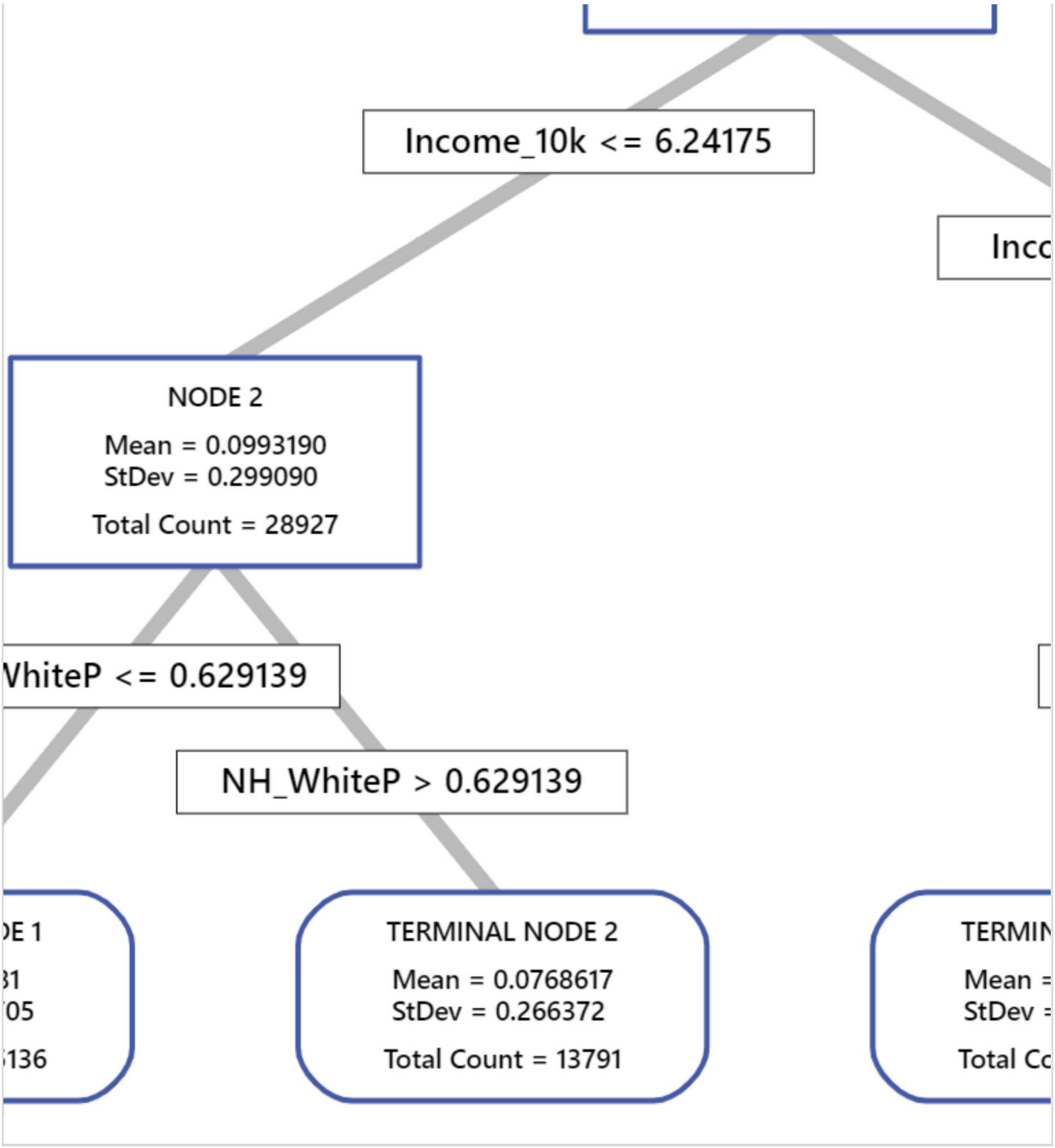
Response Information

Data Set	N	% of N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Training	58070	70.0	0.0764250	0.265679	0	0	0	0	1
Test	24836	30.0	0.0764616	0.265741	0	0	0	0	1



Optimal Tree Diagram

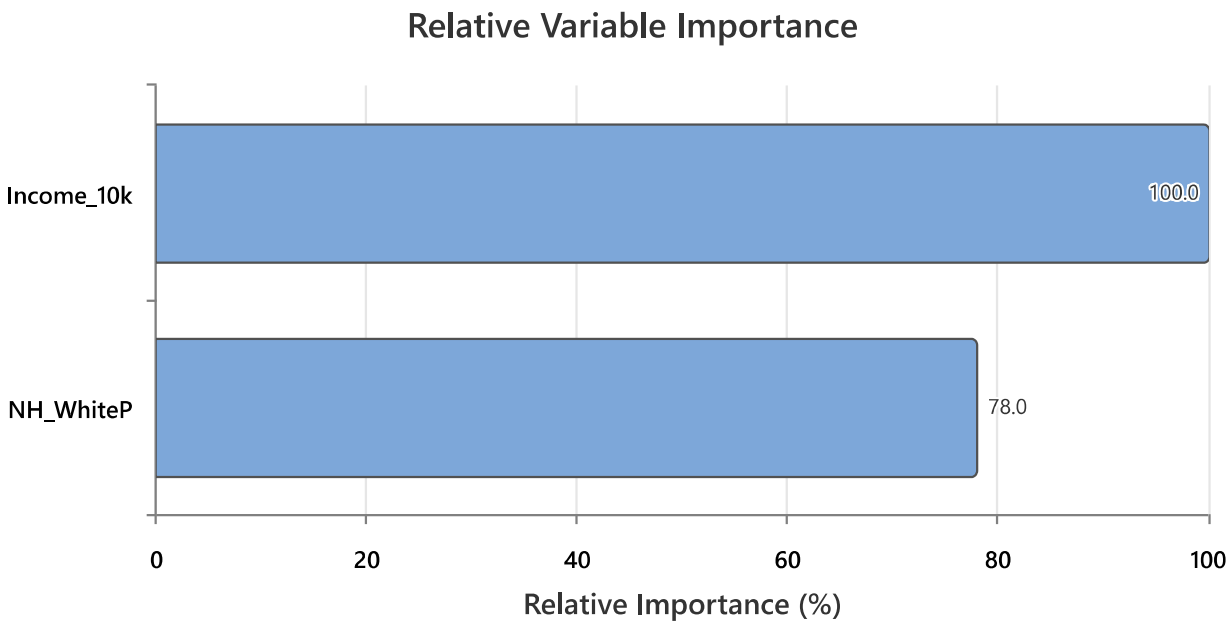




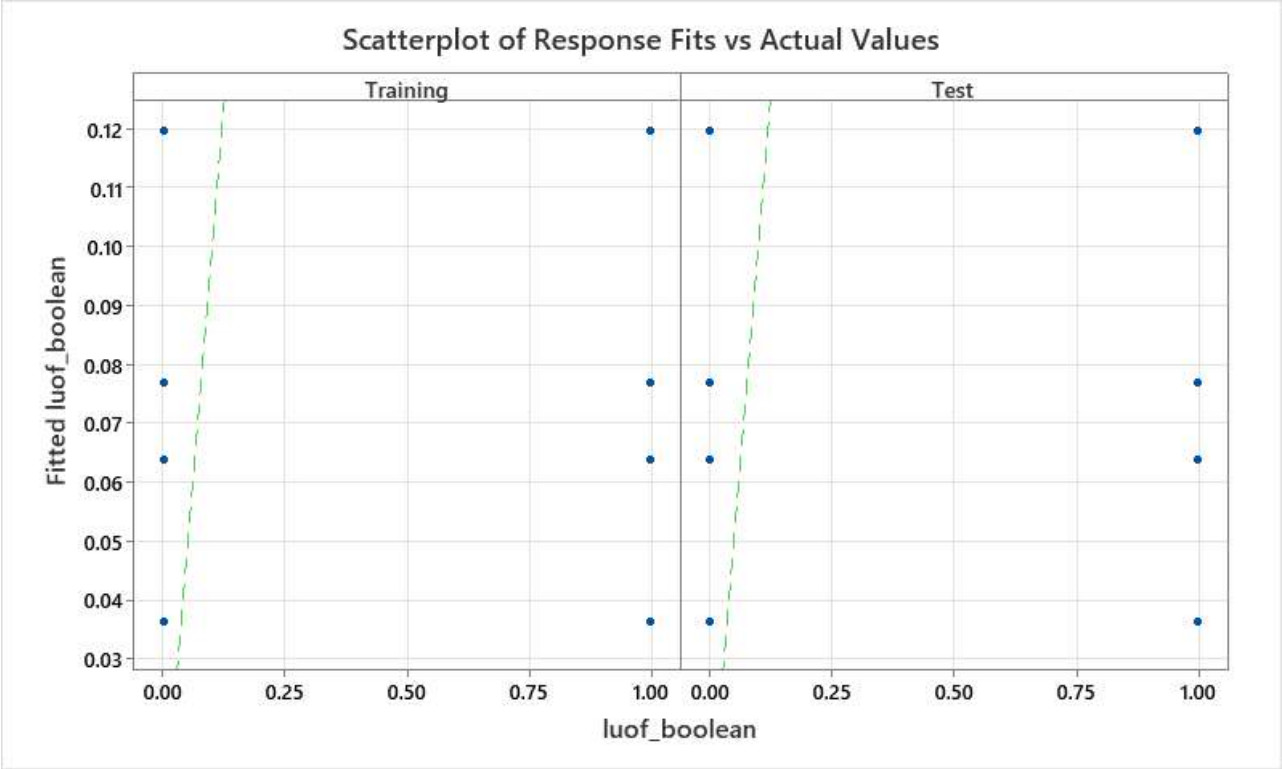
Model Summary

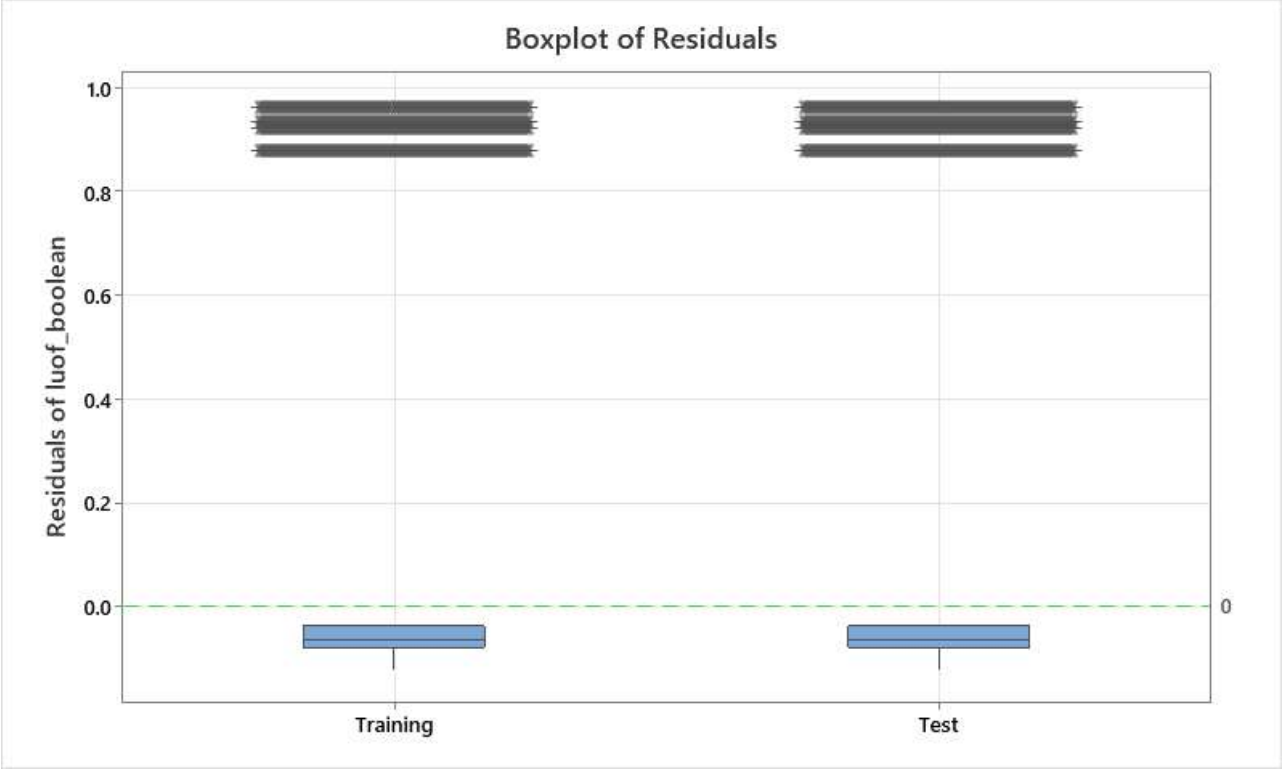
Total predictors	2
Important predictors	2
Number of terminal nodes	4
Minimum terminal node size	10830

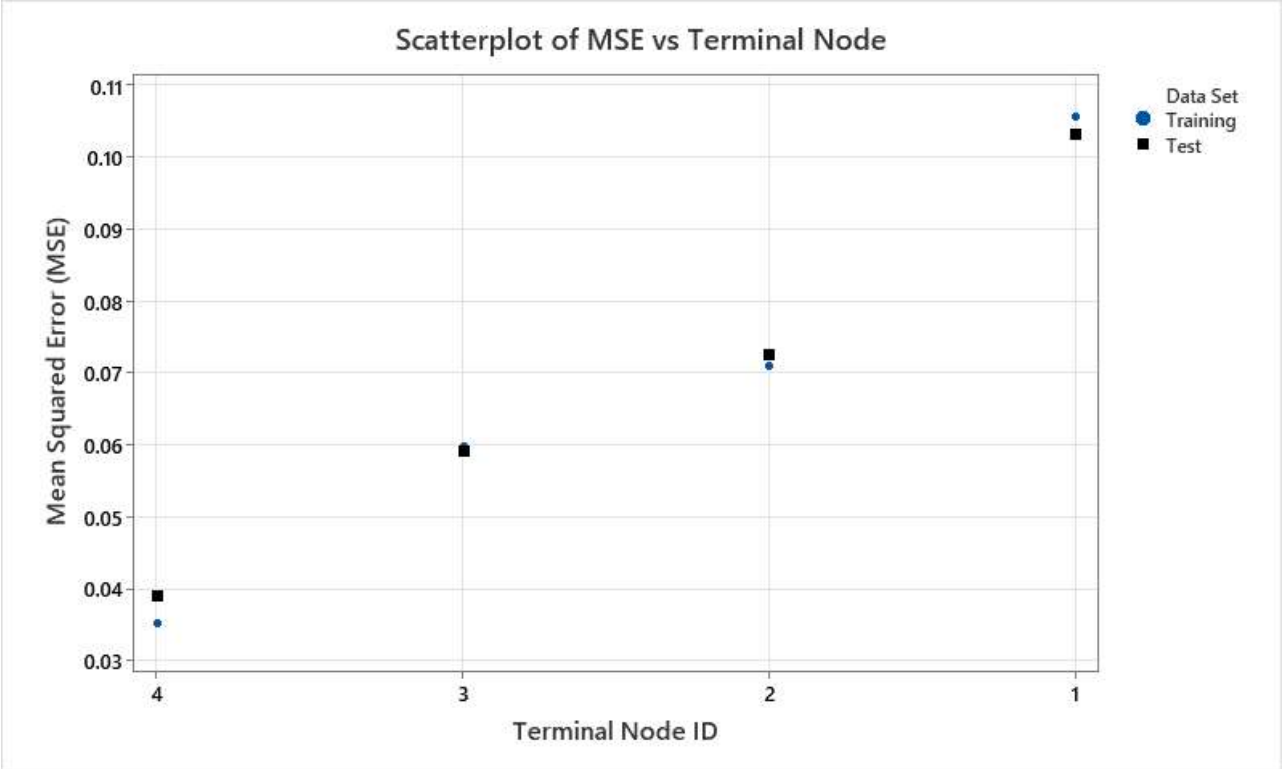
Statistics	Training	Test
R-squared	1.19%	1.01%
Root mean squared error (RMSE)	0.2641	0.2644
Mean squared error (MSE)	0.0697	0.0699
Mean absolute deviation (MAD)	0.1395	0.1396
Mean absolute percent error (MAPE)	0.9126	0.9135

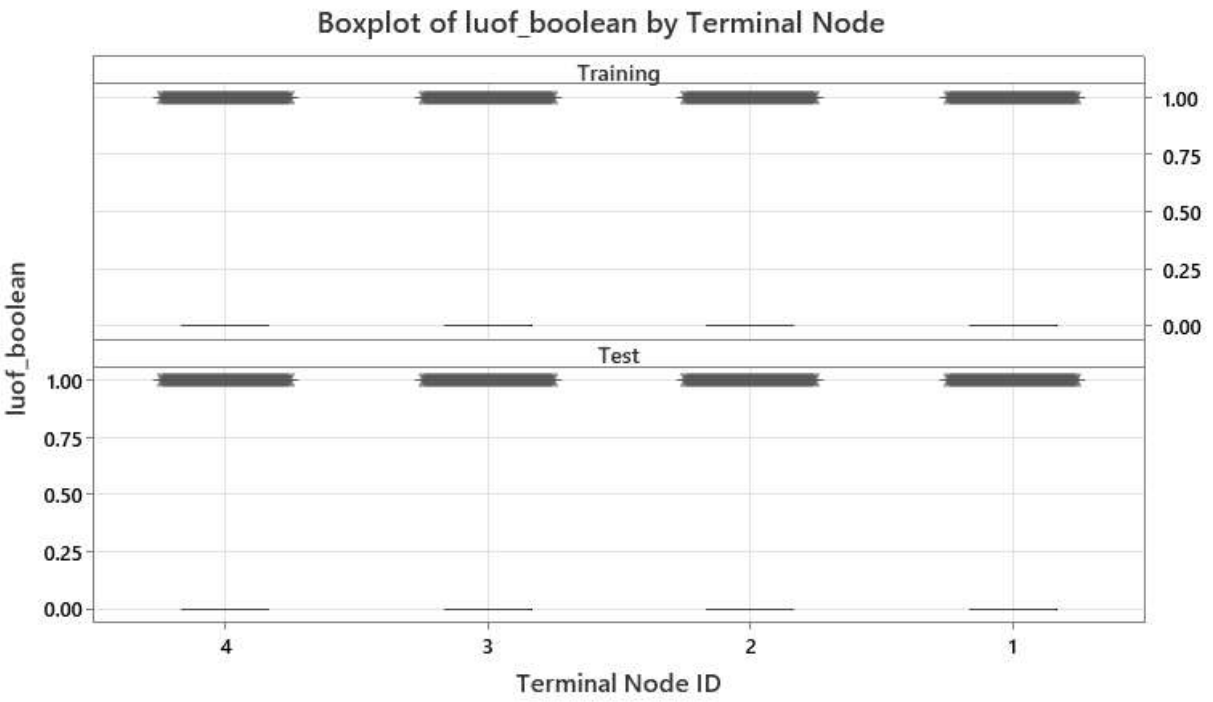


Variable importance measures model improvement when splits are made on a predictor.
Relative importance is defined as % improvement with respect to the top predictor.

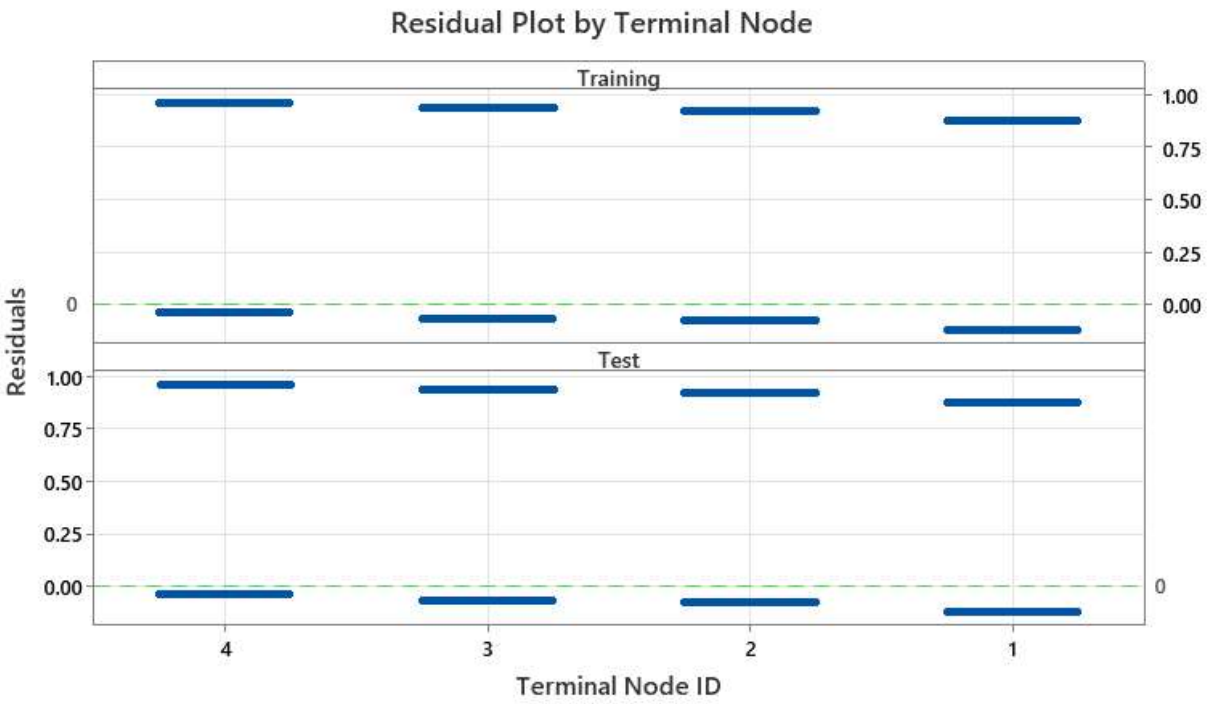



Boxplot of Residuals
Residuals of luof_boolean
1.0
0.8
0.6
0.4
0.2
0.0
0
Training
Test


Scatterplot of MSE vs Terminal Node
Mean Squared Error (MSE)
0.11
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03
4
3
2
1
Terminal Node ID
Data Set
Training
Test



Terminal node sorted in ascending order of MSE for training sample



Terminal node sorted in ascending order of MSE for training sample