Table of Contents

[The Evaluation Tab 2](#_Toc67865631)

[1. Validation Dataset – Default Settings 2](#_Toc67865632)

[1.1.1. Error Matrix 2](#_Toc67865633)

[1.1.2. Measures of Performance 2](#_Toc67865634)

[1.1.3. Risk Chart Decision Tree 3](#_Toc67865635)

[2. Validation Dataset – Reduced Number of Splits 4](#_Toc67865636)

[2.1.1. Error Matrix 4](#_Toc67865637)

[2.1.2. Measures of Performance 4](#_Toc67865638)

[2.1.3. Comparison/Findings for Error Matrix and Measures of Performance 4](#_Toc67865639)

[2.1.4. Risk Chart Decision Tree 5](#_Toc67865640)

[2.1.5. Comparison/Findings for Risk Chart to Default 5](#_Toc67865641)

[3. Validation Dataset –Default Settings, Input Variable Modification 6](#_Toc67865642)

[3.1.1. Error Matrix 6](#_Toc67865643)

[3.1.2. Measures of Performance 6](#_Toc67865644)

[3.1.3. Comparison/Findings for Error Matrix and Measures of Performance 6](#_Toc67865645)

[3.1.4. Risk Chart Decision Tree 7](#_Toc67865646)

[3.1.5. Comparison/Findings for Risk Chart to Default 7](#_Toc67865647)

[References 8](#_Toc67865648)

# The Evaluation Tab

## Validation Dataset – Default Settings

## Error Matrix

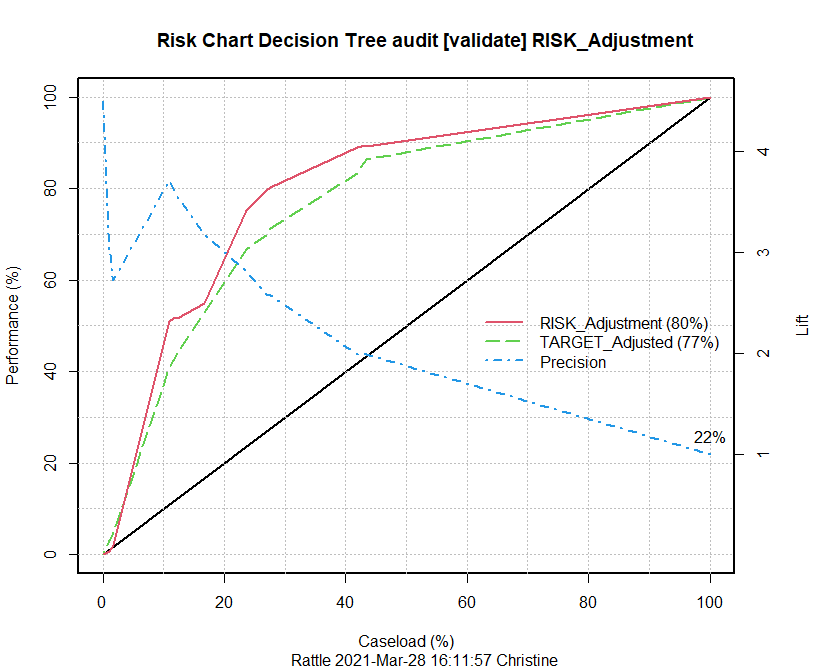
Graphical user interface, text, application, email

Description automatically generated

## Measures of Performance

|  |  |
| --- | --- |
| Precision | TP/(TP+FP) = 35/(35+15) = 35/50 = .70 = **70%** |
| Sensitivity | TP/(TP+FN) = 35/(35+31) = 35/66 = .53 = **53%** |
| Specificity | TN/(TN+FP) = 219/(219+15) = 219/234 = .94 = **94%** |

## Risk Chart Decision Tree



## Validation Dataset – Reduced Number of Splits

## Error Matrix

Graphical user interface, text, application, email

Description automatically generated

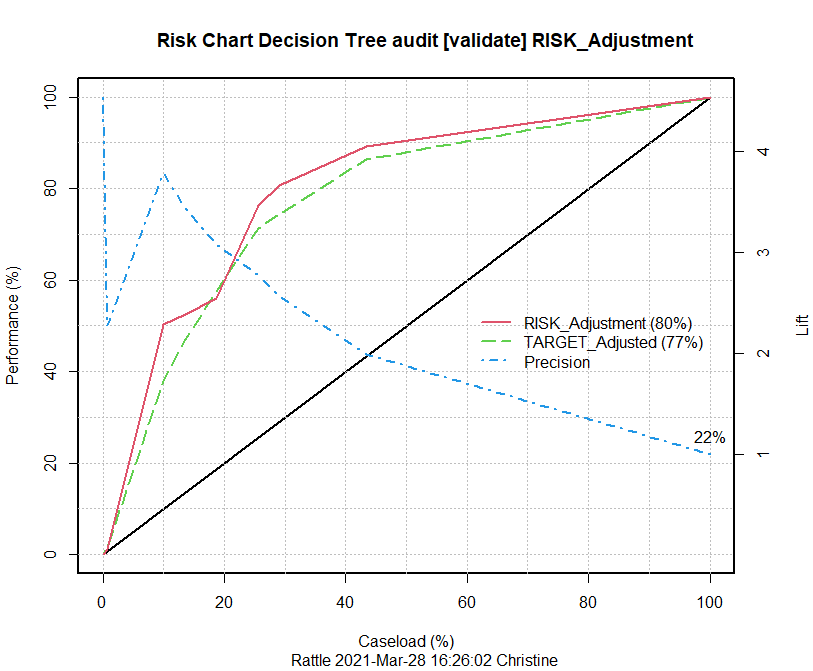
## Measures of Performance

|  |  |
| --- | --- |
| Precision | TP/(TP+FP) = 38/(38+18) = 38/56 = .68 = **68%** |
| Sensitivity | TP/(TP+FN) = 38/(38+28) = 38/66 = .58 = **58%** |
| Specificity | TN/(TN+FP) = 216/(216+18) = 216/234 = .92 = **92%** |

## Comparison/Findings for Error Matrix and Measures of Performance

When comparing the Error Matrix and Performance Measure numbers to the default settings model (Section 6.1.1), this model, using a Max Depth of 5, is only a few percentage points less than the original model which was set at a Max Depth of 30. The Overall error for both models is 15.3% and the Averaged class error is a point lower than the default settings value. Although the ‘Reduced Number of Splits’ model has slightly lower values for Precision, Sensitivity, and Specificity, it may be worth using the smaller, less complicated model.

## Risk Chart Decision Tree



## Comparison/Findings for Risk Chart to Default

When comparing the visual representation of this model and the Default Settings (Original) model they are again very similar in performance and lift; however, the risk chart above (Section 7.1.4) confirms that this is a better model. Stated another way, when our caseload (or percentage of taxpayers that we sampled) is between 20% and 50% the use of this model on the Validation Dataset will provide twice the performance.

## Validation Dataset –Default Settings, Input Variable Modification

## Error Matrix

Graphical user interface, text, application, email

Description automatically generated

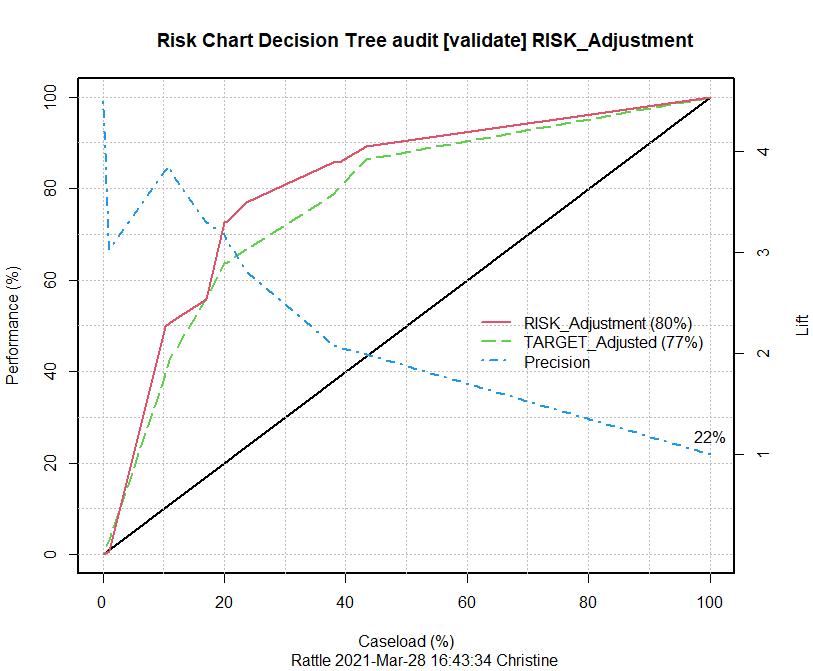
## Measures of Performance

|  |  |
| --- | --- |
| Precision | TP/(TP+FP) = 37/(37+14) = 37/51 = .73 = **73%** |
| Sensitivity | TP/(TP+FN) = 37/(37+29) = 37/66 = .56 = **56%** |
| Specificity | TN/(TN+FP) = 220/(220+14) = 220/234 = .94 = **94%** |

## Comparison/Findings for Error Matrix and Measures of Performance

When comparing this model’s Performance Measures to those of the Default Settings model, we find that the Specificity measurement is the same at 94% which shows that both models can find 94% of the true negatives within the Validation Dataset. When reviewing the Sensitivity performance measure, this model is finding 56% of the actual positives versus the Default Settings model which is finding 53%. The Precision performance measure for this model comes in at 73% and the Default Settings model has a 3% lower accuracy in predicting positive targets. Lastly, the Overall Error and Averaged Class Error are lower for this model, 14.4% and 24.95% respectively. As a result, the recommendation would be to use this model over the Default Settings model.

## Risk Chart Decision Tree



## Comparison/Findings for Risk Chart to Default

The visual representations of this model and the Default Settings (Original) model are again very similar in performance and lift; however, the risk chart above (Section 7.1.4) confirms that this is a better model. Stated another way, when our caseload (or percentage of taxpayers that we sampled) is between 20% and 50% the use of this model on the Validation Dataset will provide twice the performance.

# References

R Core Team (2021). R: A language and environment for statistical  
computing. R Foundation for Statistical Computing, Vienna, Austria.   
URL https://www.R-project.org/.

Berry, G.S.L.M. J. (2010). *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management*. [devry]. Retrieved from https://devry.vitalsource.com/#/books/9781118275603/