

IS-429 Lab Week#14 Models Build & Comparison 00000054804

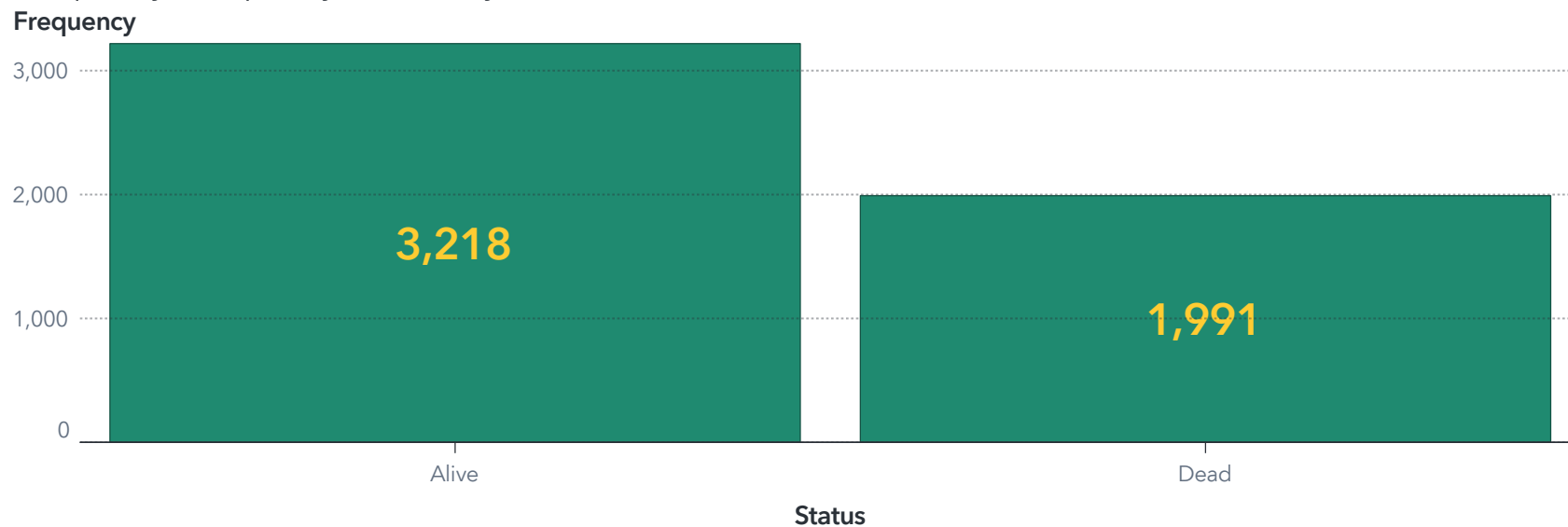
Creation Date: Tuesday, May 30, 2023 03:01:53 PM

Author: christopher.darren@student.umn.ac.id

IS-429 Lab Week#14 00000054804-Christopher Darren

Data Exploration

Frequency, Frequency Percent by Status



Frequency Percent

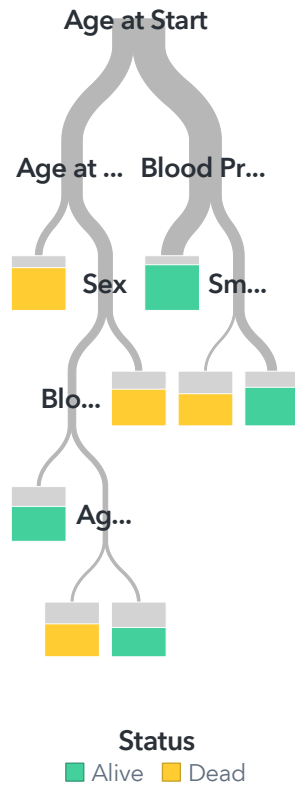


C45 Prediction Status

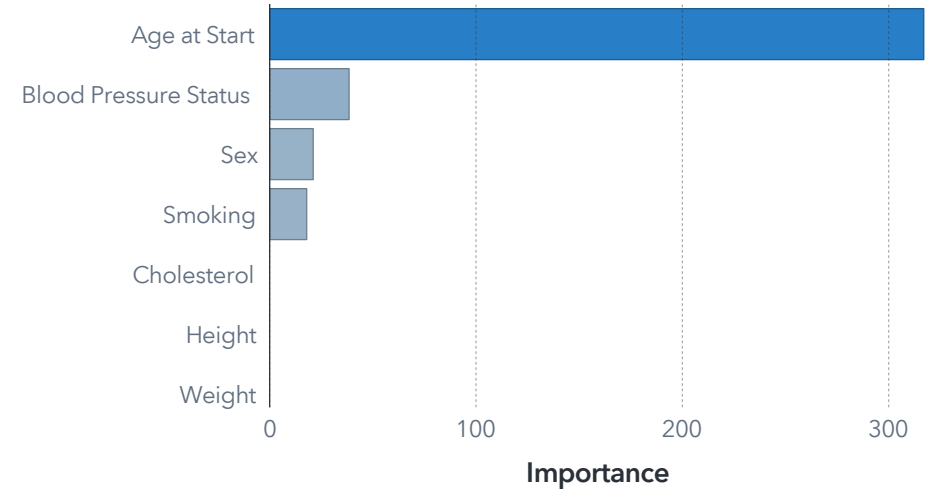
Decision Tree of Status

Decision Tree **Status** (event=Dead) Misclassification Rate **0.2573** Observations Used **5,209**

Tree



Variable Importance



Confusion Matrix

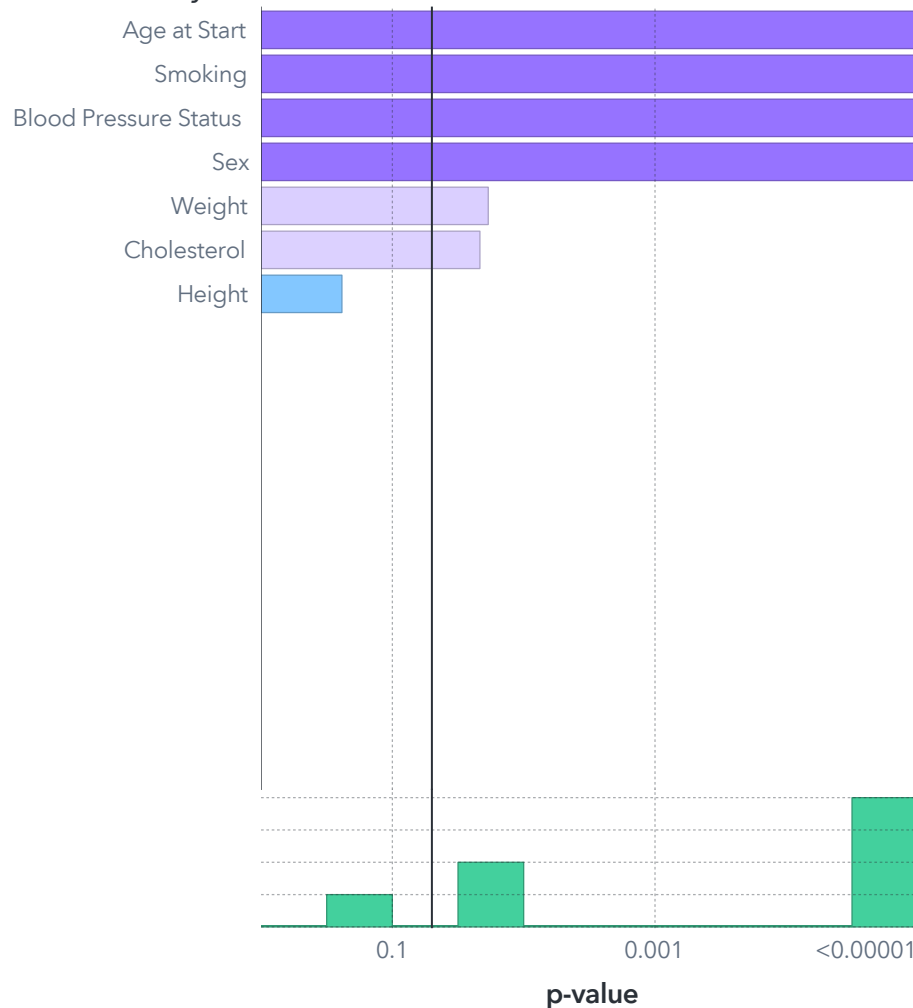
		Observed			
		Training		Validation	
Predicted	Alive	1,898	352	799	169
	Dead	586	810	264	331
Partition		Training		Validation	

Logistic Regression

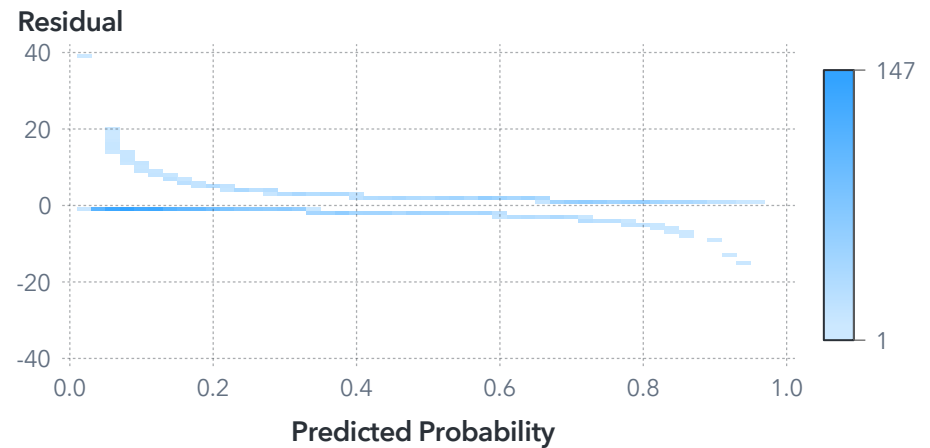
A logistic regression model requires that there be no missing values in any of the selected predictors; any patient with a missing values is excluded. In this case, that meant we had to throw away 170 of the patients. That's a small number here, but it is important to consider when we are looking at other datasets.

Logistic Regression **Status** (event=Dead) Misclassification Rate (Event) **0.2568** Observations Used **5,039** Unused **170**

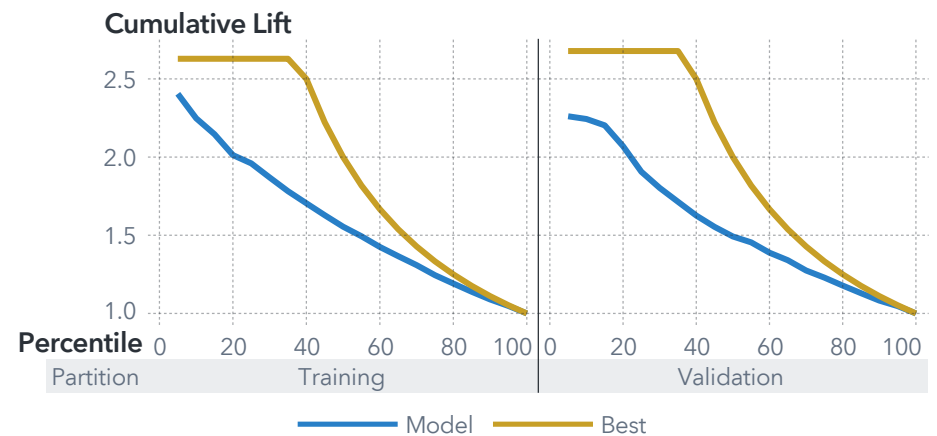
Fit Summary



Residual Plot



Lift

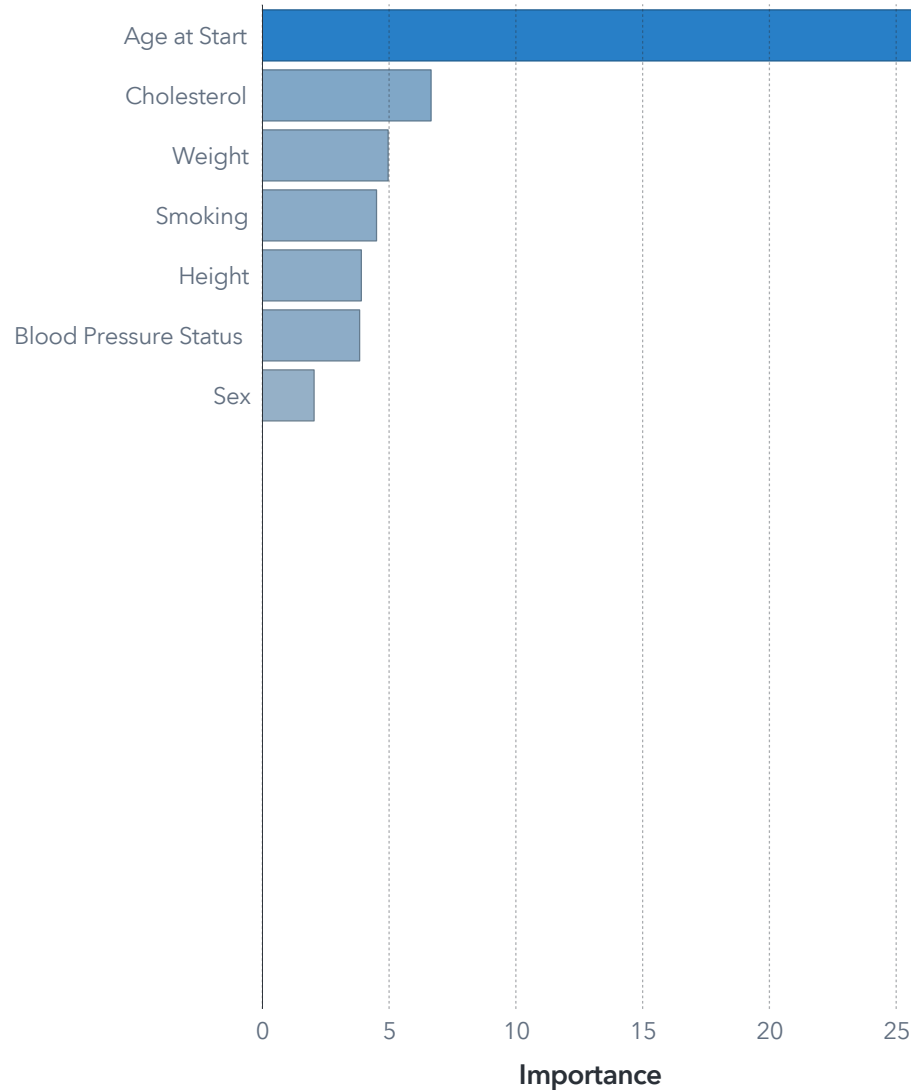


Gradient Boost

Gradient Boosting of Status

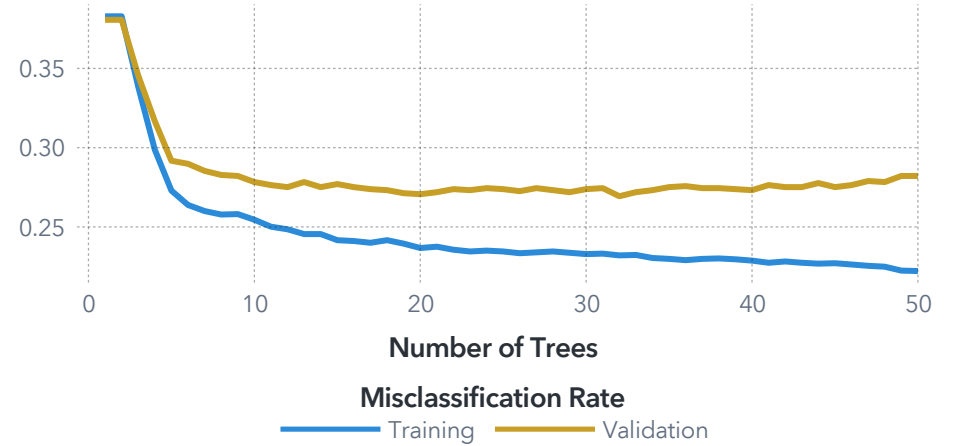
Gradient Boosting **Status** (event=Dead) Misclassification Rate **0.2222** Observations Used **5,209**

Variable Importance



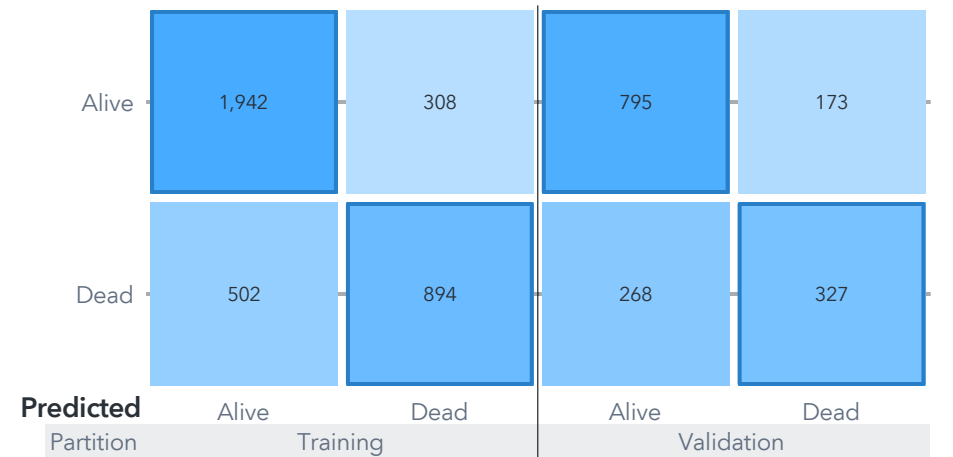
Iteration Plot

Misclassification Rate



Confusion Matrix

Observed

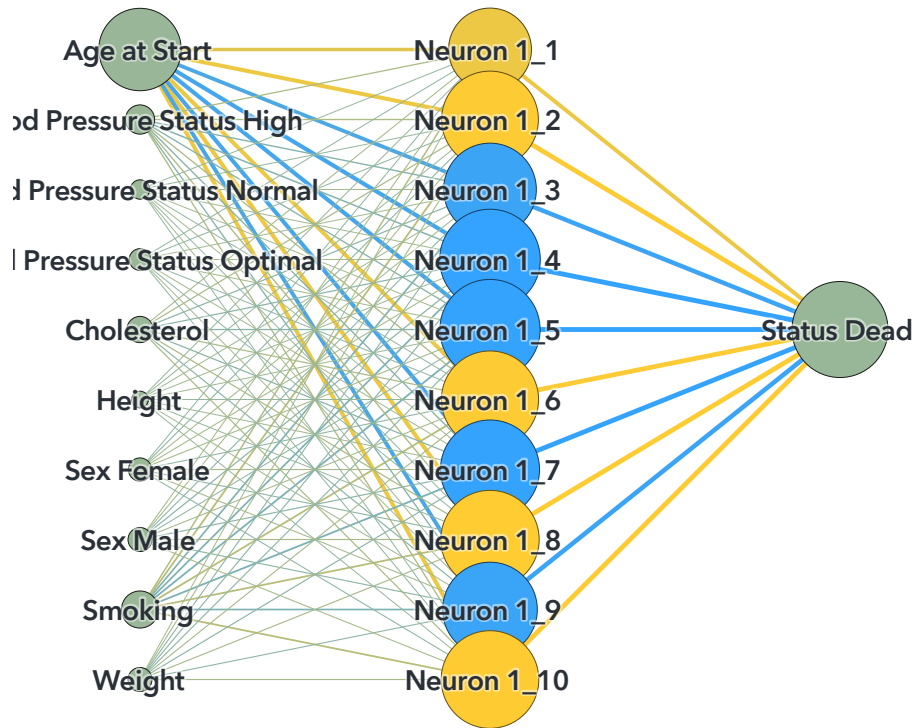


Neural Network

Neural Network of Status

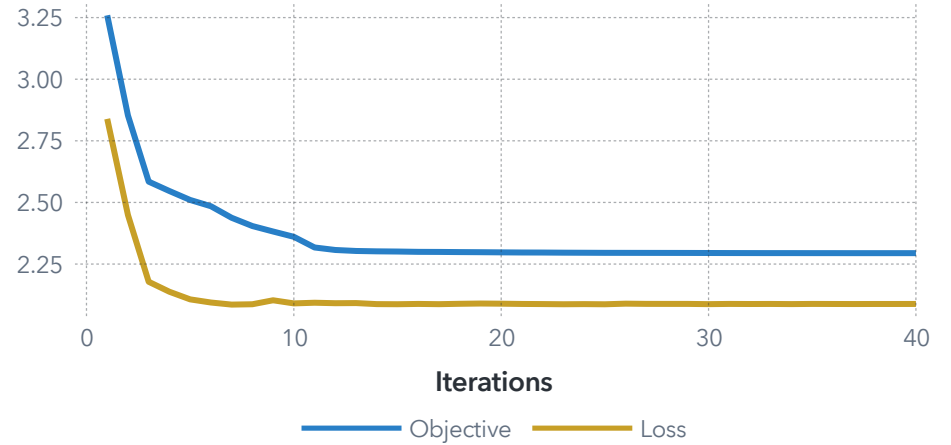
Neural Network **Status** (event=Dead) Misclassification Rate **0.2526** Observations Used **5,039** Unused **170**

Network



Iteration Plot

Objective / Loss



Confusion Matrix

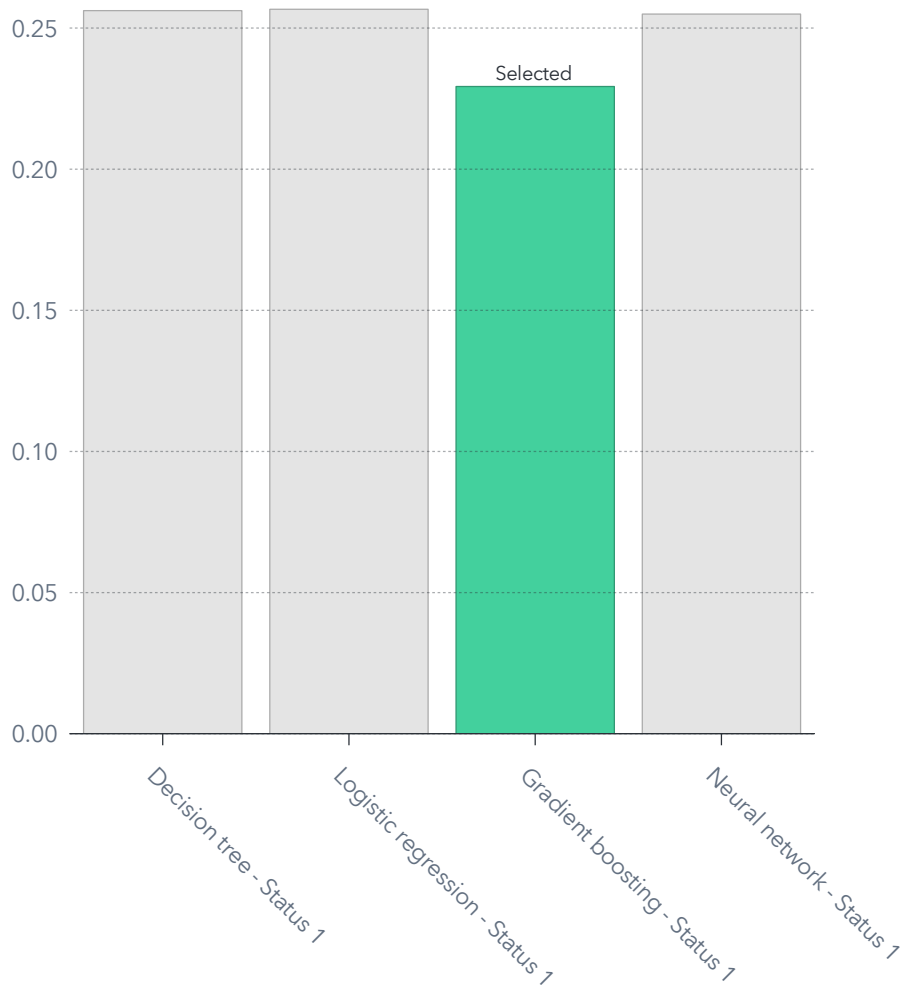
		Observed			
Predicted	Partition	Training		Validation	
		Alive	Dead	Alive	Dead
Alive		1,832	345	783	165
Dead		543	796	258	317

MODEL SELECTION & SCORING

Model Comparison **Status** (event=Dead)

Fit Statistic

Misclassification Rate (Event)

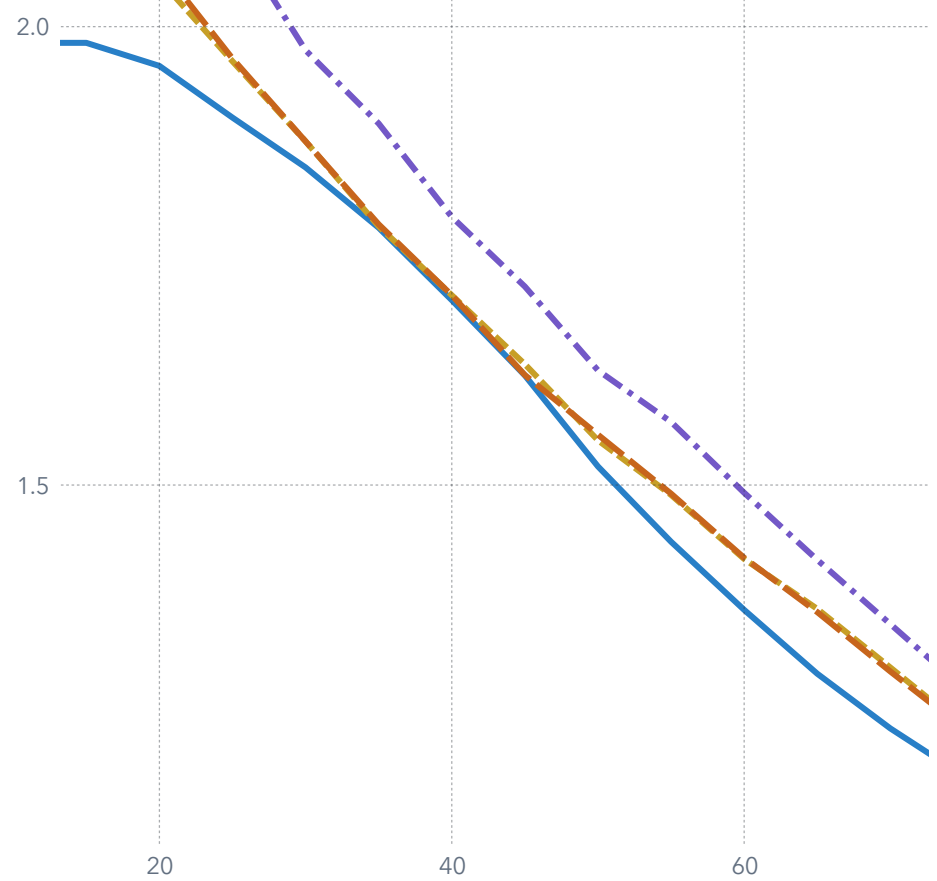


Model

▲ A1.1

Lift

Training: Cumulative Lift



AUTO

- Decision tree - Status 1
- Logistic regression - Status 1
- Gradient boosting - Status 1
- Neural network - Status 1

Appendix

A1.1 Fit Statistic

Warnings: Number of observations for all models do not match.