Models

Matt Dube

December 15, 2018

library(readr)  
library(dplyr)  
library(here)  
library(caret)  
library(gbm)  
library(car)  
library(nnet)  
library(earth)  
library(pROC)  
library(vip)  
library(tibble)

Load data

customer\_clean <- read\_csv(here("00\_Data/clean", "customer\_clean.csv"))

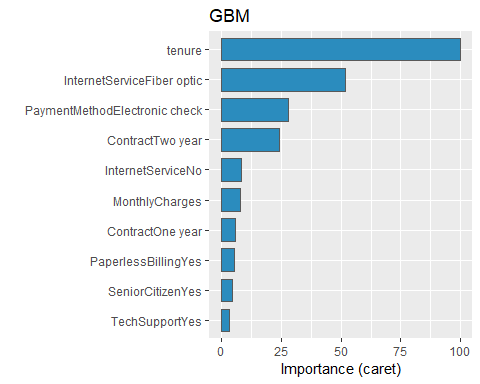
split the data

set.seed(5432)  
index <- createDataPartition(customer\_clean$Churn, p=.75, list = FALSE)  
trainData <- customer\_clean[index,]  
testData <- customer\_clean[-index,]

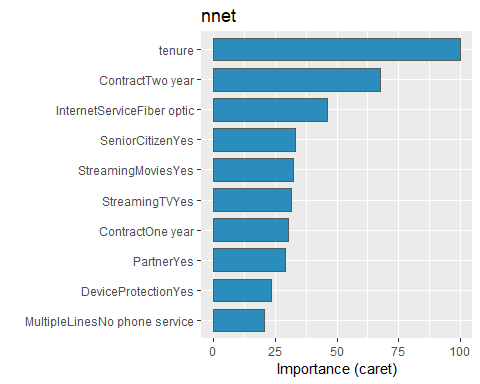
Load models

gbm.fit <- readRDS(here("03\_Models","gbm\_model.rda"))  
nnet.fit <- readRDS(here("03\_Models","nnet\_model.rda"))  
mars.fit <- readRDS(here("03\_Models","mars\_model.rda"))  
xgb.fit <- readRDS(here("03\_Models","xgb\_model.rda"))

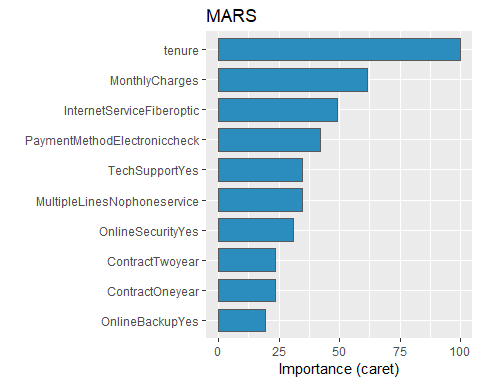
vip(gbm.fit, fill = "#2b8cbe") + ggtitle("GBM")



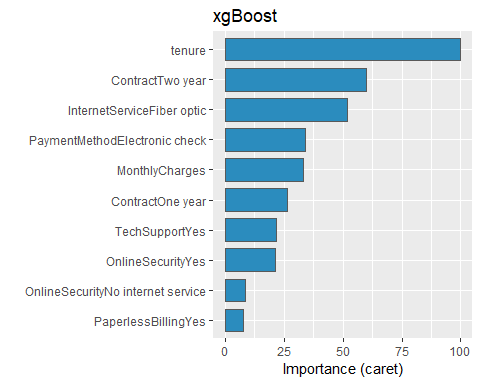
vip(nnet.fit, fill = "#2b8cbe") + ggtitle("nnet")



vip(mars.fit, fill = "#2b8cbe") + ggtitle("MARS")



vip(xgb.fit, fill = "#2b8cbe") + ggtitle("xgBoost")



make predictions with each model

gbm.predict <- predict(gbm.fit, newdata = testData)  
nnet.predict <- predict(nnet.fit, newdata = testData)  
mars.predict <- predict(mars.fit, newdata = testData)  
xgb.predict <- predict(xgb.fit, newdata = testData)

varImp(mars.fit)

## earth variable importance  
##   
## only 20 most important variables shown (out of 27)  
##   
## Overall  
## tenure 100.000  
## MonthlyCharges 61.467  
## InternetServiceFiberoptic 49.090  
## PaymentMethodElectroniccheck 42.083  
## MultipleLinesNophoneservice 34.618  
## TechSupportYes 34.618  
## OnlineSecurityYes 30.976  
## ContractOneyear 23.546  
## ContractTwoyear 23.546  
## OnlineBackupYes 19.532  
## PaperlessBillingYes 17.495  
## SeniorCitizenYes 10.910  
## MultipleLinesYes 6.322  
## StreamingTVNointernetservice 0.000  
## PaymentMethodMailedcheck 0.000  
## OnlineSecurityNointernetservice 0.000  
## PaymentMethodCreditcard(automatic) 0.000  
## StreamingMoviesYes 0.000  
## PartnerYes 0.000  
## OnlineBackupNointernetservice 0.000

test\_X <- testData %>% select(-Churn)

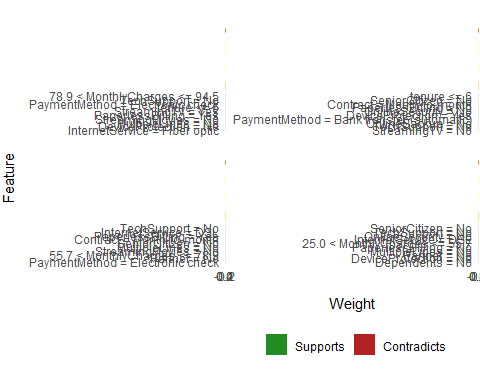
library(lime)

##   
## Attaching package: 'lime'

## The following object is masked from 'package:dplyr':  
##   
## explain

library(vip)  
  
set.seed(5431)  
local\_obs\_1 <- test\_X[sample(1:nrow(test\_X),4, replace = FALSE),]  
  
set.seed(6541)  
local\_obs\_2 <- test\_X[sample(1:nrow(test\_X),4, replace = FALSE),]  
  
  
explainer\_mars <- lime(test\_X, mars.fit, n\_bins = 5)

explanation\_mars <- explain(  
 x = local\_obs\_1,  
 explainer = explainer\_mars,  
 n\_permutations = 5000,  
 dist\_fun = "manhattan",  
 kernel\_width = .75,  
 n\_features = 10,  
 feature\_select = "auto",  
 labels = "Yes"  
)  
  
plot\_features(explanation\_mars)



explanation\_mars <- explain(  
 x = local\_obs\_2,  
 explainer = explainer\_mars,  
 n\_permutations = 5000,  
 dist\_fun = "manhattan",  
 kernel\_width = .75,  
 n\_features = 10,  
 feature\_select = "lasso\_path",  
 labels = "Yes"  
)  
  
plot\_features(explanation\_mars)

