Aprendizagem Supervisionada Classificatória

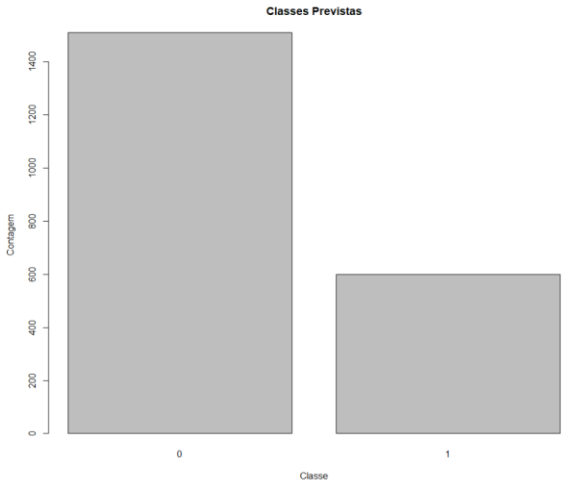
PEDRO LEITE - PEDRO CARVALHO

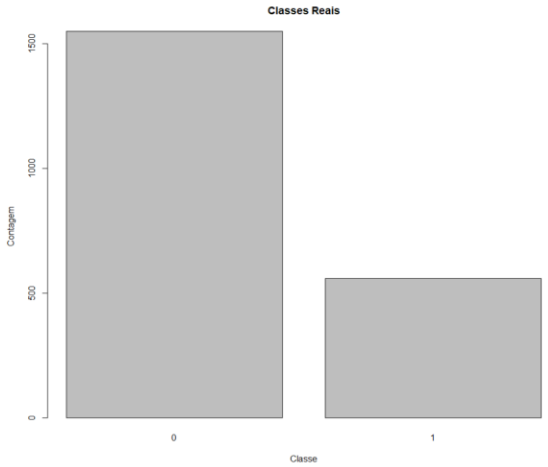
Telco Customer Churn

Focused customer retention programs

7043 clientes
21 variáveis
"Churn" é a variável objetivo

Naive Bayes





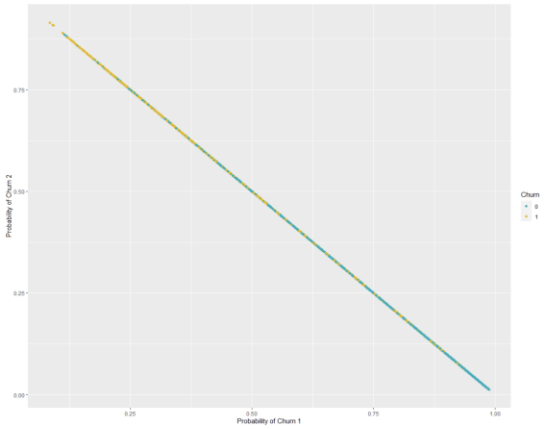

```
> error_rate <- mean(test_data$Churn != test_data$predicted)
> print(error_rate)
[1] 0.2388626
```

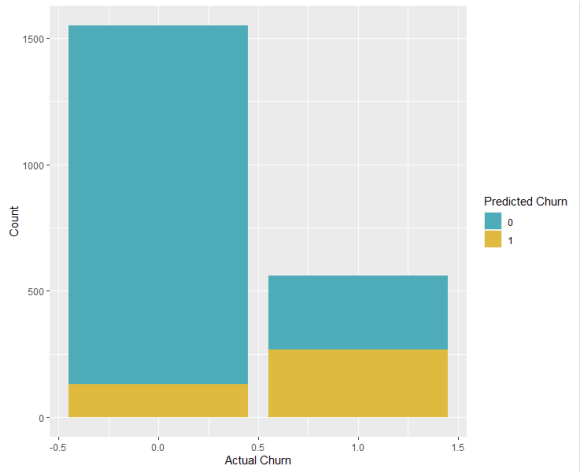
LDA

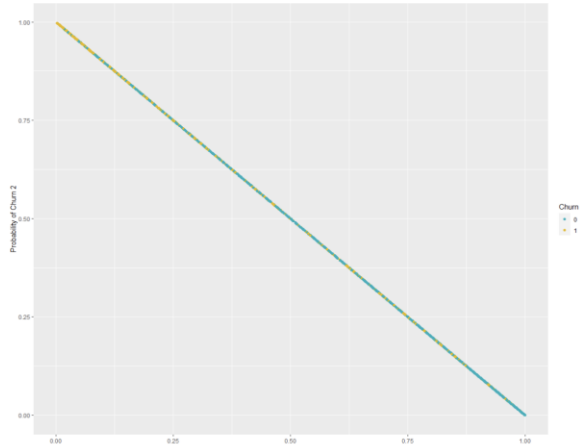
QDA

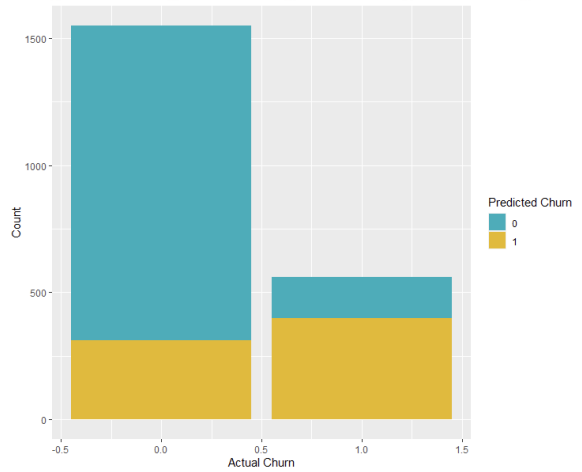
Separação das variáveis

Contagem das observações





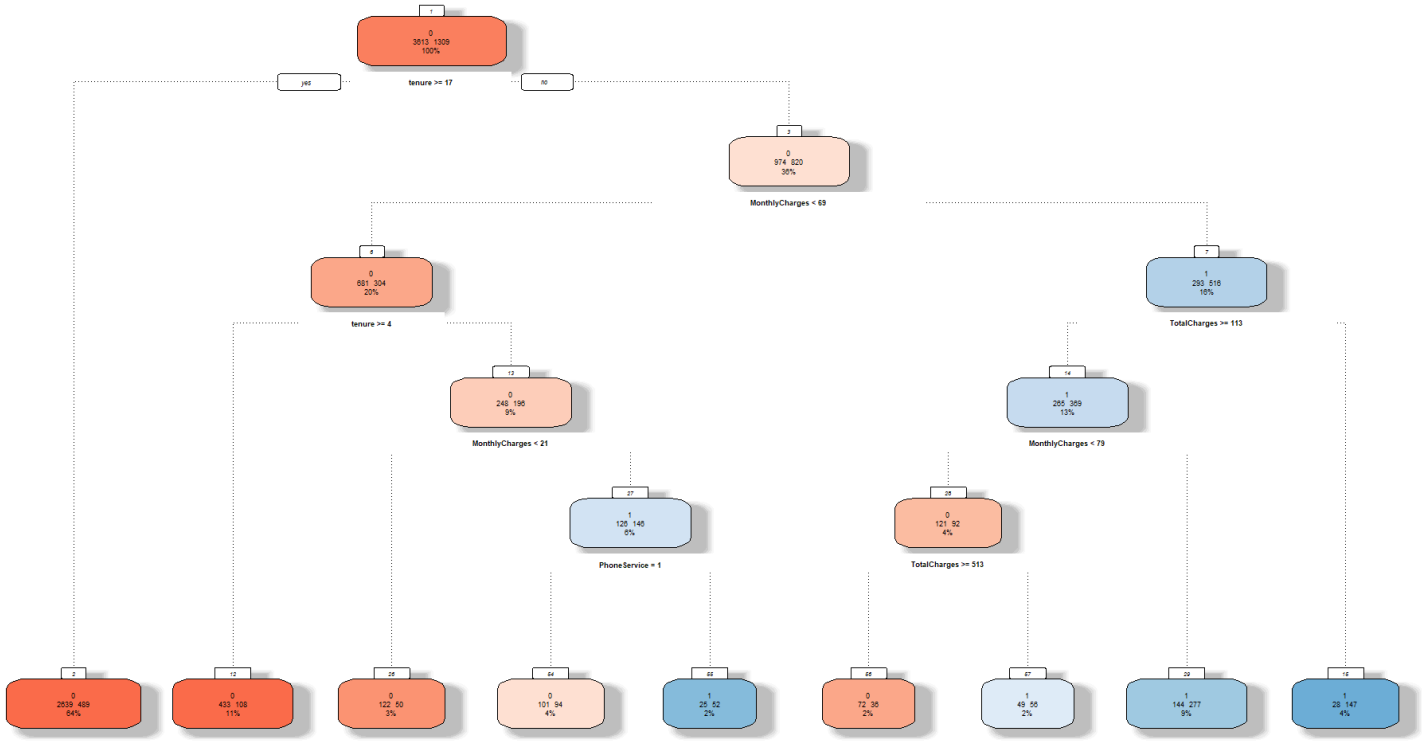




Regressão Logística

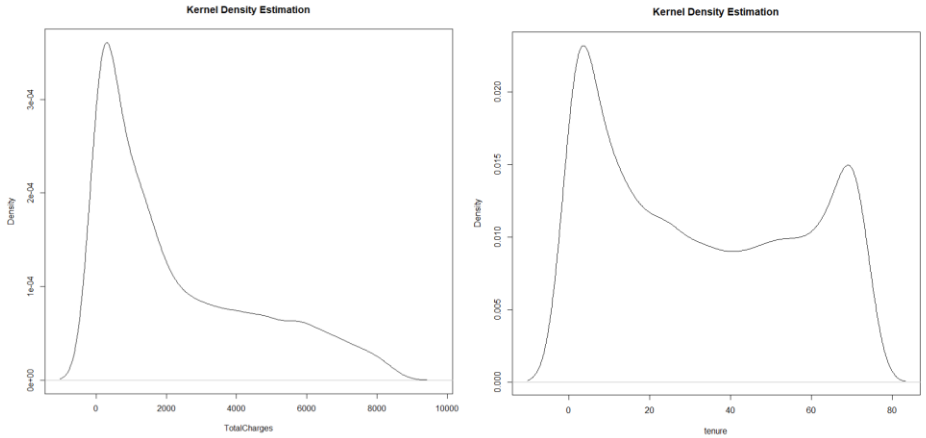
```
Call:
glm(formula = Churn ~ gender + Partner + SeniorCitizen + Dependents +
    tenure + PhoneService + PaperlessBilling + TotalCharges +
    MonthlyCharges, family = binomial, data = df)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.9420  -0.6895  -0.3716   0.7355   3.2678

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)     -1.055e+00  1.500e-01  -7.033 2.02e-12 ***
gender           3.406e-03  6.294e-02   0.054 0.956850
Partner          2.349e-02  7.541e-02   0.311 0.755440
SeniorCitizen    4.879e-01  8.177e-02   5.967 2.41e-09 ***
Dependents      -3.155e-01  8.636e-02  -3.653 0.000259 ***
tenure          -7.038e-02  5.631e-03 -12.499  < 2e-16 ***
PhoneService    -8.803e-01  1.173e-01  -7.505 6.14e-14 ***
PaperlessBilling 5.321e-01  7.121e-02   7.473 7.86e-14 ***
TotalCharges     1.811e-04  6.277e-05   2.885 0.003911 **
MonthlyCharges   2.845e-02  1.911e-03  14.886  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 8143.4  on 7031  degrees of freedom
Residual deviance: 6174.5  on 7022  degrees of freedom
AIC: 6194.5

Number of Fisher Scoring iterations: 6
```

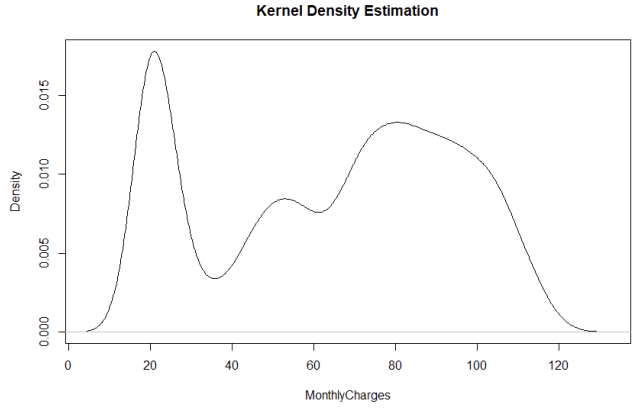
Árvore de Decisão



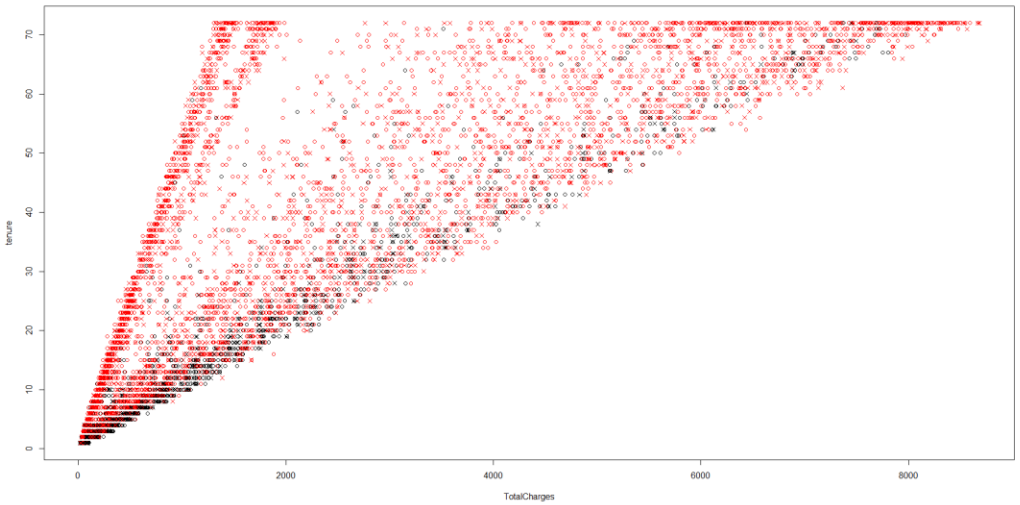
previsoes	0	1
0	1460	339
1	90	221

Método do Núcleo



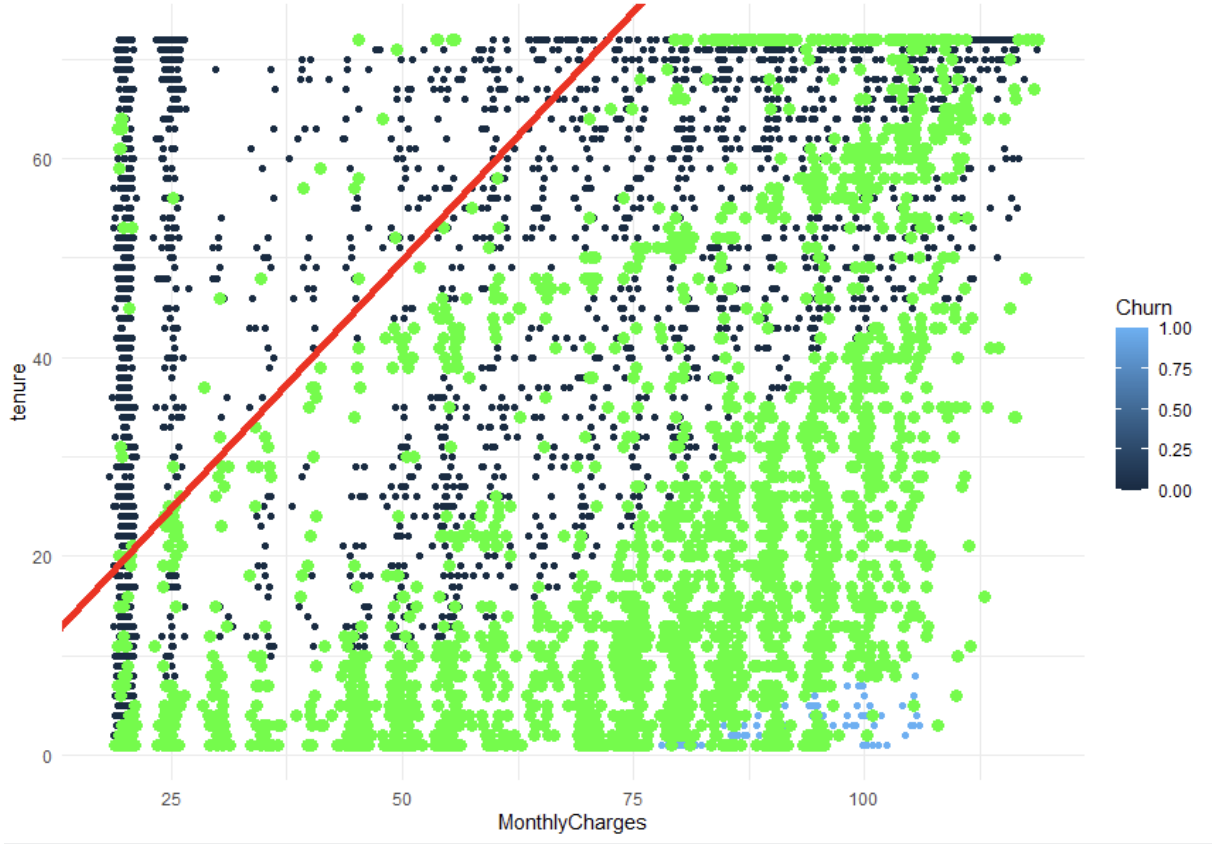


Método dos K-Vizinhos mais Próximos

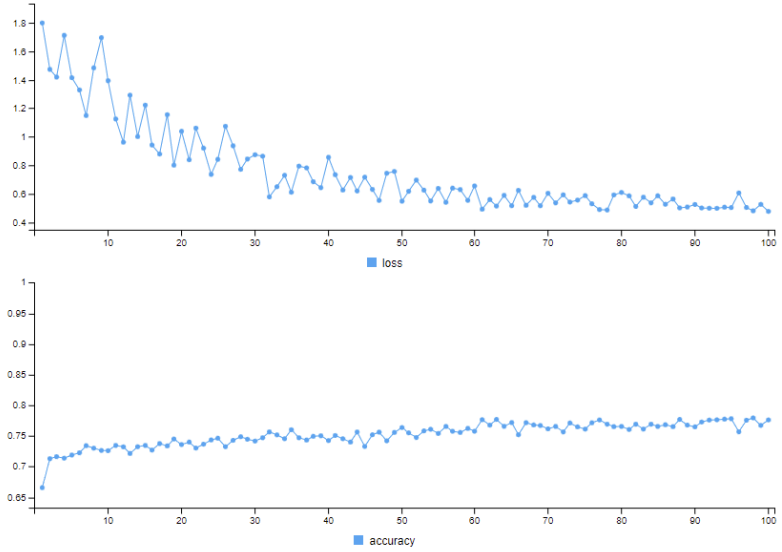


Accuracy = 76%

SVM

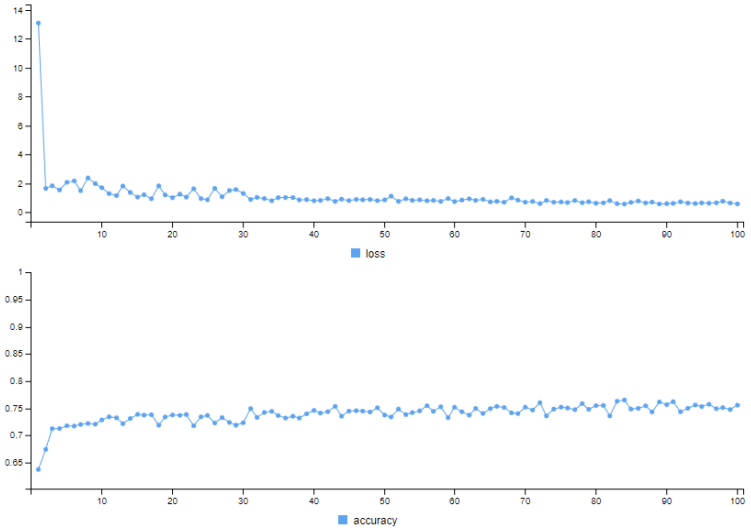

tenure
60
40
20
0
25
50
75
100
MonthlyCharges
Churn
1.00
0.75
0.50
0.25
0.00

Redes Neuronais sem Weight Decay




```
> metrics <- model %>% evaluate(x_test, y_test)
66/66 [==============================] - 0s 909us/step - loss: 0.4230 - accuracy: 0.8033
> print(metrics)
     loss  accuracy
0.4230382 0.8033175
```

Redes Neuronais com Weight Decay




```
> metrics <- model %>% evaluate(x_test, y_test)
66/66 [==============================] - 0s 939us/step - loss: 0.5089 - accuracy: 0.7668
> print(metrics)
     loss  accuracy
0.5089459 0.7668247
```