#1

> table(test$Class, PredC)

PredC

0 1

0 18 4

1 1 42

Bra prediction. Only 5 misclassification out of 65.

> print(m1)

Call: glm(formula = Class ~ ., family = binomial, data = combined)

Coefficients:

(Intercept) RI Na Mg Al Si K

8080.723 -4129.884 -13.395 -4.662 -36.192 -20.826 -8.160

Ca Ba Fe

-4.467 -18.878 27.583

Degrees of Freedom: 148 Total (i.e. Null); 139 Residual

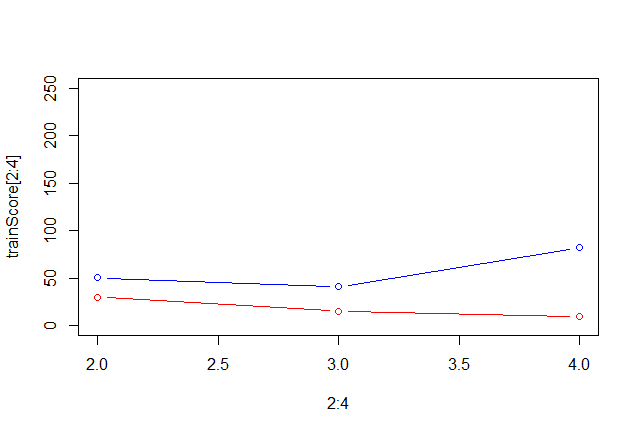
Null Deviance: 146.9

Residual Deviance: 14.74 AIC: 34.74

Probabilistic model:

Decision boundary:

#2



3 leaves, smallest validation error.

> table(test$Class, PredC)

PredC

0 1

0 18 4

1 1 42

> table(test$Class, PredC2)

PredC2

0 1

0 11 11

1 1 42

> PredC3=as.numeric((Pred2[,2]+Pred1)/2>0.5)

> table(test$Class, PredC3)

PredC3

0 1

0 18 4

1 1 42

Tree has a worse test error than Logistic and the combined classifier has same as Logistic. Combining two classifiers eliminated the error of the worse classifier.

#3

> poster(df,"Grades")

[1] 0.001540619

> poster(df,"Popular")

[1] 0.003183166

> poster(df,"Sports")

[1] 0.00297536

More likely that being popular is most important. Probabilities appr 0.2 0.42 0.38