> source('~/.active-rstudio-document', echo=TRUE)

> y = glm(ACTION ~ .,data=train, family=binomial)

> summary(y)

Call:

glm(formula = ACTION ~ ., family = binomial, data = train)

Deviance Residuals:

Min 1Q Median 3Q Max

-2.9710 0.3393 0.3458 0.3502 0.5809

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) 3.672e-01 1.057e+00 0.347 0.72837

RESOURCE -3.821e-08 6.918e-07 -0.055 0.95596

MGR\_ID -4.873e-07 6.473e-07 -0.753 0.45153

ROLE\_ROLLUP\_1 -6.269e-06 2.560e-06 -2.449 0.01434

ROLE\_ROLLUP\_2 4.066e-06 4.513e-06 0.901 0.36756

ROLE\_DEPTNAME 1.436e-07 1.268e-06 0.113 0.90981

ROLE\_TITLE -1.636e-06 7.029e-07 -2.327 0.01996

ROLE\_FAMILY\_DESC 2.464e-07 3.555e-07 0.693 0.48827

ROLE\_FAMILY 1.838e-07 2.452e-07 0.750 0.45337

ROLE\_CODE 2.346e-05 7.241e-06 3.240 0.00119

(Intercept)

RESOURCE

MGR\_ID

ROLE\_ROLLUP\_1 \*

ROLE\_ROLLUP\_2

ROLE\_DEPTNAME

ROLE\_TITLE \*

ROLE\_FAMILY\_DESC

ROLE\_FAMILY

ROLE\_CODE \*\*

---

Signif. codes:

0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 14492 on 32768 degrees of freedom

Residual deviance: 14464 on 32759 degrees of freedom

AIC: 14484

Number of Fisher Scoring iterations: 6

> predictX = predict(y, type="response")

> table(train$ACTION, predictX > 0.5)

TRUE

0 1897

1 30872

> library("ROCR")

> ROCRpred = prediction(predictX, train$ACTION)

> auc = as.numeric(performance(ROCRpred, "auc")@y.values)

> ROCRperf = performance(ROCRpred, "tpr", "fpr")

> plot(ROCRperf)

> predictTest = predict(ResourceLog1, type="response", newdata=test)

> summary(predictTest)

Min. 1st Qu. Median Mean 3rd Qu. Max.

0.8504 0.9405 0.9412 0.9420 0.9427 0.9977

>