

Logit-DecisionTree

luismor

14/4/2021

```
library(readxl)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
library(psych)
library(caret)
```

```
## Loading required package: lattice

## Loading required package: ggplot2

##
## Attaching package: 'ggplot2'

## The following objects are masked from 'package:psych':
##
##   %+%, alpha
```

```
library(tree)
library(randomForest)
```

```
## randomForest 4.6-14

## Type rfNews() to see new features/changes/bug fixes.

##
## Attaching package: 'randomForest'
```

```
## The following object is masked from 'package:ggplot2':
##
##     margin

## The following object is masked from 'package:psych':
##
##     outlier

## The following object is masked from 'package:dplyr':
##
##     combine
```

```
library(parallel)
library(doParallel)
```

```
## Loading required package: foreach
```

```
## Loading required package: iterators
```

```
#Carga de datos
```

```
df <- read_excel("/Users/unimooc/Dropbox/2021/Directorio R/Research/Cooperation/DATOS 2004-2014.xlsx", sheet = "DATOS")
```

```
df <- na.omit(df)
anyNA(df)
```

```
## [1] FALSE
```

```
#Transformamos 1-2 = 1, 3-4 = 0
```

```
df$GI01[df$GI01 == 2] = 1
df$GI02[df$GI02 == 2] = 1
df$GI03[df$GI03 == 2] = 1
df$GI04[df$GI04 == 2] = 1
df$GI05[df$GI05 == 2] = 1
df$GI06[df$GI06 == 2] = 1
df$GI07[df$GI07 == 2] = 1
df$GI08[df$GI08 == 2] = 1
df$GI09[df$GI09 == 2] = 1
```

```
df$GI01[df$GI01 == 3] = 0
df$GI02[df$GI02 == 3] = 0
df$GI03[df$GI03 == 3] = 0
df$GI04[df$GI04 == 3] = 0
df$GI05[df$GI05 == 3] = 0
df$GI06[df$GI06 == 3] = 0
df$GI07[df$GI07 == 3] = 0
df$GI08[df$GI08 == 3] = 0
df$GI09[df$GI09 == 3] = 0
```

```
df$GI01[df$GI01 == 4] = 0
df$GI02[df$GI02 == 4] = 0
df$GI03[df$GI03 == 4] = 0
```

```

df$GI04[df$GI04 == 4] = 0
df$GI05[df$GI05 == 4] = 0
df$GI06[df$GI06 == 4] = 0
df$GI07[df$GI07 == 4] = 0
df$GI08[df$GI08 == 4] = 0
df$GI09[df$GI09 == 4] = 0

#Convertimos variables a factores
df$MDOLOCAL <- factor(df$MDOLOCAL)
df$MDONAC <- factor(df$MDONAC)
df$MDOUE <- factor(df$MDOUE)
df$INNPROD <- factor(df$INNPROD)
df$INNPROC <- factor(df$INNPROC)
df$INNFABRI <- factor(df$INNFABRI)
df$INNOBIEN <- factor(df$INNOBIEN)
df$INNOSERV <- factor(df$INNOSERV)
df$INNLOGIS <- factor(df$INNLOGIS)
df$INNAPOYO <- factor(df$INNAPOYO)
df$COOP <- factor(df$COOP)
df$GI01 <- factor(df$GI01)
df$GI02 <- factor(df$GI02)
df$GI03 <- factor(df$GI03)
df$GI04 <- factor(df$GI04)
df$GI05 <- factor(df$GI05)
df$GI06 <- factor(df$GI06)
df$GI07 <- factor(df$GI07)
df$GI08 <- factor(df$GI08)
df$GI09 <- factor(df$GI09)

#Creamos bases de datos de ciclos económicos
df.gr <- filter(df, TIEMPO <= 2007)
df.cr <- filter(df, TIEMPO %in% c(2008, 2009, 2010))
df.re <- filter(df, TIEMPO >= 2011)

```

```

set.seed(2021)
df.p <- createDataPartition(y = df$INTINN,
                             p = 0.8, list = F)

Training <- df[df.p,]

```

```

## Warning: The 'i' argument of '['()' can't be a matrix as of tibble 3.0.0.
## Convert to a vector.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_warnings()' to see where this warning was generated.

```

```

Test <- df[-df.p,]

df.gr.p <- createDataPartition(y = df.gr$INTINN,
                               p = 0.8, list = F)

Training.gr <- df.gr[df.gr.p,]
Test.gr <- df.gr[-df.gr.p,]

```

```

df.cr.p <- createDataPartition(y = df.cr$INTINN,
                               p = 0.8, list = F)

Training.cr <- df.cr[df.cr.p,]
Test.cr <- df.cr[-df.cr.p,]

df.re.p <- createDataPartition(y = df.re$INTINN,
                               p = 0.8, list = F)

Training.re <- df.re[df.re.p,]
Test.re <- df.re[-df.re.p,]

lg.gr.01 <- glm(GIO1 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFAFRI + INNLOGIS +
                INNAPOYO,
                family=binomial(link = "logit"),
                data = Training.gr)
summary(lg.gr.01)

##
## Call:
## glm(formula = GIO1 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFAFRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.gr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3176  -0.9259   0.5774   0.7774   2.1947
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.019e+00  7.443e-02 -13.690 < 2e-16 ***
## COOP1       3.620e-01  4.654e-02  7.777 7.42e-15 ***
## TAMANO      -3.806e-05  1.136e-05 -3.351 0.000806 ***
## MDOLOCAL1    6.089e-02  5.890e-02  1.034 0.301239
## MDONAC1     3.273e-01  4.727e-02  6.924 4.39e-12 ***
## MDOUE1      2.099e-01  3.087e-02  6.800 1.05e-11 ***
## INNPROD1    9.232e-01  6.457e-02 14.299 < 2e-16 ***
## INNOBIEN1    6.969e-01  5.637e-02 12.362 < 2e-16 ***
## INNOSERV1    3.440e-01  4.709e-02  7.304 2.80e-13 ***
## INNPROC1    -5.147e-01  5.293e-02 -9.725 < 2e-16 ***
## INNFAFRI1    7.110e-01  4.281e-02 16.606 < 2e-16 ***
## INNLOGIS1    2.015e-01  4.444e-02  4.533 5.80e-06 ***
## INNAPOYO1    3.132e-01  4.105e-02  7.628 2.38e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 36507  on 28159  degrees of freedom
## Residual deviance: 31071  on 28147  degrees of freedom
## AIC: 31097
##
## Number of Fisher Scoring iterations: 4

```

```

probabs.gr.01 <- predict(lg.gr.01, Test.gr, type='response')
preds.gr.01 <- ifelse(probabs.gr.01 > 0.5, 1, 0)

confusionMatrix(factor(preds.gr.01), factor(Test.gr$GIO1))

```

```

## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 1396  783
##           1 1039 3821
##
##           Accuracy : 0.7412
##           95% CI : (0.7308, 0.7514)
##       No Information Rate : 0.6541
##       P-Value [Acc > NIR] : < 2.2e-16
##
##           Kappa : 0.4135
##
##  McNemar's Test P-Value : 2.315e-09
##
##           Sensitivity : 0.5733
##           Specificity : 0.8299
##       Pos Pred Value : 0.6407
##       Neg Pred Value : 0.7862
##           Prevalence : 0.3459
##       Detection Rate : 0.1983
##       Detection Prevalence : 0.3096
##       Balanced Accuracy : 0.7016
##
##       'Positive' Class : 0
##

```

```

lg.gr.02 <- glm(GIO2 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
                INNAPPOYO,
                family=binomial(link = "logit"),
                data = Training.gr)
summary(lg.gr.02)

```

```

##
## Call:
## glm(formula = GIO2 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##     INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
##     INNAPPOYO, family = binomial(link = "logit"), data = Training.gr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0810  -0.9659   0.6372   0.8790   2.1387
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)

```

```
## (Intercept) -1.365e+00  7.252e-02 -18.821 < 2e-16 ***
## COOP1       2.990e-01  4.240e-02   7.052 1.77e-12 ***
## TAMANO     -3.255e-05  1.099e-05  -2.961 0.00307 **
## MDOLOCAL1  -5.308e-02  5.636e-02  -0.942 0.34635
## MDONAC1     5.469e-01  4.666e-02  11.719 < 2e-16 ***
## MDOUE1      3.154e-01  2.926e-02  10.781 < 2e-16 ***
## INNPROD1    9.089e-01  5.915e-02  15.364 < 2e-16 ***
## INNOBIEN1   4.152e-01  5.078e-02   8.176 2.93e-16 ***
## INNOSERV1   1.312e-01  4.063e-02   3.230 0.00124 **
## INNPROC1   -4.728e-01  5.034e-02  -9.394 < 2e-16 ***
## INNFIABRI1  7.660e-01  4.053e-02  18.899 < 2e-16 ***
## INNLOGIS1   1.997e-01  4.133e-02   4.831 1.36e-06 ***
## INNAPYOY01  3.310e-01  3.835e-02   8.632 < 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: 38294 on 28159 degrees of freedom
## Residual deviance: 33731 on 28147 degrees of freedom
## AIC: 33757
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.gr.02 <- predict(lg.gr.02, Test.gr, type='response')
preds.gr.02 <- ifelse(probabs.gr.02 > 0.5, 1, 0)

confusionMatrix(factor(preds.gr.02), factor(Test.gr$GI02))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 1629  769
##           1 1344 3297
##
##           Accuracy : 0.6998
##           95% CI : (0.689, 0.7105)
##           No Information Rate : 0.5776
##           P-Value [Acc > NIR] : < 2.2e-16
##
##           Kappa : 0.3684
##
## Mcnemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.5479
##           Specificity : 0.8109
##           Pos Pred Value : 0.6793
##           Neg Pred Value : 0.7104
##           Prevalence : 0.4224
##           Detection Rate : 0.2314
##           Detection Prevalence : 0.3407
##           Balanced Accuracy : 0.6794
##
```

```
##      'Positive' Class : 0
##
```

```
lg.gr.03 <- glm(GIO3 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFAFRI + INNLOGIS +
                INNAPOYO,
                family=binomial(link = "logit"),
                data = Training.gr)
summary(lg.gr.03)
```

```
##
## Call:
## glm(formula = GIO3 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFAFRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.gr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.6006  -1.0729   0.5357   0.7872   1.3999
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.335e-01  7.662e-02  -5.658 1.53e-08 ***
## COOP1        4.283e-01  5.211e-02   8.221 < 2e-16 ***
## TAMANO      -1.299e-05  1.249e-05  -1.039 0.298620
## MDOLOCAL1    4.441e-02  6.145e-02   0.723 0.469829
## MDONAC1      1.402e-01  4.862e-02   2.883 0.003938 **
## MDOUE1       1.154e-01  3.231e-02   3.571 0.000356 ***
## INNPROD1     8.544e-01  7.385e-02  11.569 < 2e-16 ***
## INNOBIEN1    2.924e-01  6.650e-02   4.396 1.10e-05 ***
## INNOSERV1    4.742e-01  5.294e-02   8.956 < 2e-16 ***
## INNPROC1    -3.180e-01  5.722e-02  -5.558 2.74e-08 ***
## INNFAFRI1    8.875e-01  4.782e-02  18.560 < 2e-16 ***
## INNLOGIS1    2.887e-01  5.047e-02   5.720 1.07e-08 ***
## INNAPOYO1    5.741e-01  4.626e-02  12.411 < 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: 32676  on 28159  degrees of freedom
## Residual deviance: 28754  on 28147  degrees of freedom
## AIC: 28780
##
## Number of Fisher Scoring iterations: 5
```

```
probabs.gr.03 <- predict(lg.gr.03, Test.gr, type='response')
preds.gr.03 <- ifelse(probabs.gr.03 > 0.5, 1, 0)

confusionMatrix(factor(preds.gr.03), factor(Test.gr$GIO3))
```

```
## Confusion Matrix and Statistics
##
```

```
##           Reference
## Prediction    0    1
##           0  592  442
##           1 1268 4737
##
##           Accuracy : 0.7571
##           95% CI : (0.7469, 0.767)
##           No Information Rate : 0.7358
##           P-Value [Acc > NIR] : 2.297e-05
##
##           Kappa : 0.2716
##
## Mcnemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.3183
##           Specificity : 0.9147
##           Pos Pred Value : 0.5725
##           Neg Pred Value : 0.7888
##           Prevalence : 0.2642
##           Detection Rate : 0.0841
##           Detection Prevalence : 0.1469
##           Balanced Accuracy : 0.6165
##
##           'Positive' Class : 0
##
```

```
lg.gr.04 <- glm(GIO4 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
  INNOBIEN + INNOSERV + INNPROC + INNFBARI + INNLOGIS +
  INNAPOYO,
  family=binomial(link = "logit"),
  data = Training.gr)
summary(lg.gr.04)
```

```
##
## Call:
## glm(formula = GIO4 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##       INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBARI + INNLOGIS +
##       INNAPOYO, family = binomial(link = "logit"), data = Training.gr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2225  -0.9861   0.6414   1.0038   1.6125
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -9.319e-01  7.146e-02 -13.040 < 2e-16 ***
## COOP1       1.334e-01  4.123e-02  3.236  0.00121 **
## TAMANO      2.638e-05  1.256e-05  2.100  0.03573 *
## MDOLOCAL1   4.045e-03  5.593e-02  0.072  0.94235
## MDONAC1     8.784e-02  4.596e-02  1.911  0.05599 .
## MDOUE1     -5.026e-02  2.955e-02 -1.700  0.08904 .
## INNPROD1    3.797e-02  5.988e-02  0.634  0.52604
## INNOBIEN1   -8.999e-03  5.238e-02 -0.172  0.86360
## INNOSERV1   4.208e-01  3.964e-02 10.613 < 2e-16 ***
```



```
## INNPROC1      6.271e-01  5.006e-02  12.527 < 2e-16 ***
## INNFBABRI1    6.357e-01  4.075e-02  15.599 < 2e-16 ***
## INNLOGIS1     4.899e-01  4.245e-02  11.541 < 2e-16 ***
## INNAPOYO1     5.347e-01  3.775e-02  14.163 < 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: 38559  on 28159  degrees of freedom
## Residual deviance: 34265  on 28147  degrees of freedom
## AIC: 34291
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.gr.04 <- predict(lg.gr.04, Test.gr, type='response')
preds.gr.04 <- ifelse(probabs.gr.04 > 0.5, 1, 0)
```

```
confusionMatrix(factor(preds.gr.04), factor(Test.gr$GI04))
```

```
## Confusion Matrix and Statistics
##
##              Reference
## Prediction    0    1
##              0 1577  750
##              1 1464 3248
##
##              Accuracy : 0.6855
##              95% CI : (0.6745, 0.6963)
##      No Information Rate : 0.568
##      P-Value [Acc > NIR] : < 2.2e-16
##
##              Kappa : 0.3406
##
##      McNemar's Test P-Value : < 2.2e-16
##
##              Sensitivity : 0.5186
##              Specificity : 0.8124
##              Pos Pred Value : 0.6777
##              Neg Pred Value : 0.6893
##              Prevalence : 0.4320
##              Detection Rate : 0.2240
##      Detection Prevalence : 0.3306
##              Balanced Accuracy : 0.6655
##
##              'Positive' Class : 0
##
```

```
lg.gr.05 <- glm(GI05 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
                INNAPOYO,
                family=binomial(link = "logit"),
                data = Training.gr)
summary(lg.gr.05)
```

```
##
## Call:
## glm(formula = GIO5 ~ COOP + TAMANO + MDOLocal + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.gr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2671  -0.8743   0.6229   0.9242   1.5462
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.8353671  0.0720641 -11.592 < 2e-16 ***
## COOP1        0.2326010  0.0424502   5.479 4.27e-08 ***
## TAMANO       0.0000328  0.0000131   2.504  0.0123 *
## MDOLocal1    0.0029388  0.0566281   0.052  0.9586
## MDONAC1      0.0269692  0.0463207   0.582  0.5604
## MDOUE1       0.0047115  0.0298246   0.158  0.8745
## INNPROD1     -0.0077079  0.0608954  -0.127  0.8993
## INNOBIEN1    0.0256338  0.0533954   0.480  0.6312
## INNOSERV1    0.4633656  0.0406189  11.408 < 2e-16 ***
## INNPROC1     0.6469421  0.0507643  12.744 < 2e-16 ***
## INNFBRI1     0.7648246  0.0414627  18.446 < 2e-16 ***
## INNLOGIS1    0.3986627  0.0432506   9.218 < 2e-16 ***
## INNAPOYO1    0.4536578  0.0389672  11.642 < 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: 38139  on 28159  degrees of freedom
## Residual deviance: 33623  on 28147  degrees of freedom
## AIC: 33649
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.gr.05 <- predict(lg.gr.05, Test.gr, type='response')
preds.gr.05 <- ifelse(probabs.gr.05 > 0.5, 1, 0)

confusionMatrix(factor(preds.gr.05), factor(Test.gr$GIO5))
```

```
## Confusion Matrix and Statistics
##
##              Reference
## Prediction    0    1
##              0 1543  782
##              1 1339 3375
##
##              Accuracy : 0.6987
##              95% CI : (0.6878, 0.7094)
##              No Information Rate : 0.5906
```

```
##      P-Value [Acc > NIR] : < 2.2e-16
##
##              Kappa : 0.3579
##
##      McNemar's Test P-Value : < 2.2e-16
##
##              Sensitivity : 0.5354
##              Specificity : 0.8119
##              Pos Pred Value : 0.6637
##              Neg Pred Value : 0.7160
##              Prevalence : 0.4094
##              Detection Rate : 0.2192
##      Detection Prevalence : 0.3303
##      Balanced Accuracy : 0.6736
##
##      'Positive' Class : 0
##
```

```
lg.gr.06 <- glm(GIO6 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
  INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
  INNAPOYO,
  family=binomial(link = "logit"),
  data = Training.gr)
summary(lg.gr.06)
```

```
##
## Call:
## glm(formula = GIO6 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.gr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9542  -0.9927  -0.7018   1.1282   1.8663
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.397e+00  7.188e-02 -19.431  < 2e-16 ***
## COOP1        1.190e-01  3.882e-02   3.066  0.00217 **
## TAMANO       5.787e-05  1.161e-05   4.982  6.29e-07 ***
## MDOLOCAL1    -1.524e-01  5.467e-02  -2.787  0.00532 **
## MDONAC1      1.508e-01  4.702e-02   3.208  0.00134 **
## MDOUE1       1.257e-01  2.920e-02   4.307  1.66e-05 ***
## INNPROD1     -1.139e-01  5.748e-02  -1.982  0.04752 *
## INNOBIEN1     2.162e-01  4.945e-02   4.373  1.23e-05 ***
## INNOSERV1     2.348e-01  3.701e-02   6.342  2.26e-10 ***
## INNPROC1     2.908e-01  4.898e-02   5.938  2.89e-09 ***
## INNFBABRI1    7.612e-01  3.800e-02  20.030  < 2e-16 ***
## INNLOGIS1     3.288e-01  3.720e-02   8.839  < 2e-16 ***
## INNAPOYO1     4.151e-01  3.465e-02  11.982  < 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: 38095 on 28159 degrees of freedom
## Residual deviance: 35133 on 28147 degrees of freedom
## AIC: 35159
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.gr.06 <- predict(lg.gr.06, Test.gr, type='response')
preds.gr.06 <- ifelse(probabs.gr.06 > 0.5, 1, 0)

confusionMatrix(factor(preds.gr.06), factor(Test.gr$GIO6))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 3360 1654
##           1  819 1206
##
##           Accuracy : 0.6487
##           95% CI : (0.6374, 0.6598)
##       No Information Rate : 0.5937
##       P-Value [Acc > NIR] : < 2.2e-16
##
##           Kappa : 0.2366
##
##  McNemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.8040
##           Specificity : 0.4217
##           Pos Pred Value : 0.6701
##           Neg Pred Value : 0.5956
##           Prevalence : 0.5937
##           Detection Rate : 0.4773
##       Detection Prevalence : 0.7123
##           Balanced Accuracy : 0.6128
##
##           'Positive' Class : 0
##
```

```
lg.gr.07 <- glm(GIO7 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFIABRI + INNLOGIS +
                INNAPPOYO,
                family=binomial(link = "logit"),
                data = Training.gr)
summary(lg.gr.07)
```

```
##
## Call:
## glm(formula = GIO7 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFIABRI + INNLOGIS +
##      INNAPPOYO, family = binomial(link = "logit"), data = Training.gr)
##
```

```
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4731  -0.9071  -0.6828   1.2239   1.9812
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.581e+00  7.478e-02 -21.147  < 2e-16 ***
## COOP1        3.077e-01  3.877e-02   7.938 2.06e-15 ***
## TAMANO       3.418e-05  1.031e-05   3.315 0.000916 ***
## MDOLocal1    -1.217e-01  5.639e-02  -2.158 0.030955 *
## MDONAC1     -1.769e-02  4.965e-02  -0.356 0.721546
## MDOUE1       2.562e-01  3.060e-02   8.371  < 2e-16 ***
## INNPROD1    -2.947e-01  6.053e-02  -4.869 1.12e-06 ***
## INNOBIEN1    4.736e-01  5.218e-02   9.076  < 2e-16 ***
## INNOSERV1    2.041e-01  3.714e-02   5.496 3.89e-08 ***
## INNPROC1     1.785e-01  5.161e-02   3.458 0.000544 ***
## INNFIABRI1   7.412e-01  4.005e-02  18.505  < 2e-16 ***
## INNLOGIS1    1.768e-01  3.754e-02   4.710 2.48e-06 ***
## INNAPYOY01   2.147e-01  3.527e-02   6.088 1.15e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 35479  on 28159  degrees of freedom
## Residual deviance: 33335  on 28147  degrees of freedom
## AIC: 33361
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.gr.07 <- predict(lg.gr.07, Test.gr, type='response')
preds.gr.07 <- ifelse(probabs.gr.07 > 0.5, 1, 0)

confusionMatrix(factor(preds.gr.07), factor(Test.gr$GI07))
```

```
## Confusion Matrix and Statistics
##
##              Reference
## Prediction    0    1
##              0 4441 1799
##              1  364  435
##
##              Accuracy : 0.6927
##              95% CI : (0.6818, 0.7035)
##      No Information Rate : 0.6826
##      P-Value [Acc > NIR] : 0.03523
##
##              Kappa : 0.1436
##
## Mcnemar's Test P-Value : < 2e-16
##
##              Sensitivity : 0.9242
##              Specificity : 0.1947
##              Pos Pred Value : 0.7117
```

```
##          Neg Pred Value : 0.5444
##          Prevalence : 0.6826
##          Detection Rate : 0.6309
##          Detection Prevalence : 0.8865
##          Balanced Accuracy : 0.5595
##
##          'Positive' Class : 0
##
```

```
lg.gr.08 <- glm(GIO8 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFAFRI + INNLOGIS +
                INNAPOYO,
                family=binomial(link = "logit"),
                data = Training.gr)
summary(lg.gr.08)
```

```
##
## Call:
## glm(formula = GIO8 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFAFRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.gr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.6858  -0.9376  -0.7362   1.1760   1.8639
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.276e+00  7.164e-02 -17.817  < 2e-16 ***
## COOP1        5.061e-01  3.828e-02  13.221  < 2e-16 ***
## TAMANO       2.410e-05  1.033e-05   2.333   0.0197 *
## MDOLOCAL1    -4.038e-02  5.510e-02  -0.733   0.4636
## MDONAC1      -1.203e-01  4.687e-02  -2.567   0.0102 *
## MDOUE1       2.698e-01  2.932e-02   9.203  < 2e-16 ***
## INNPROD1     -1.091e-01  5.861e-02  -1.862   0.0626 .
## INNOBIEN1     5.041e-01  5.063e-02   9.957  < 2e-16 ***
## INNOSERV1     2.850e-03  3.651e-02   0.078   0.9378
## INNPROC1      2.073e-02  4.933e-02   0.420   0.6743
## INNFAFRI1     7.508e-01  3.893e-02  19.289  < 2e-16 ***
## INNLOGIS1     3.741e-01  3.730e-02  10.030  < 2e-16 ***
## INNAPOYO1     1.562e-01  3.482e-02   4.486  7.25e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 37618  on 28159  degrees of freedom
## Residual deviance: 35187  on 28147  degrees of freedom
## AIC: 35213
##
## Number of Fisher Scoring iterations: 4
```

```

probabs.gr.08 <- predict(lg.gr.08, Test.gr, type='response')
preds.gr.08 <- ifelse(probabs.gr.08 > 0.5, 1, 0)

confusionMatrix(factor(preds.gr.08), factor(Test.gr$GI08))

```

```

## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 3495 1663
##           1   793 1088
##
##           Accuracy : 0.6511
##           95% CI : (0.6398, 0.6622)
##      No Information Rate : 0.6092
##      P-Value [Acc > NIR] : 2.173e-13
##
##           Kappa : 0.2232
##
##  McNemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.8151
##           Specificity : 0.3955
##           Pos Pred Value : 0.6776
##           Neg Pred Value : 0.5784
##           Prevalence : 0.6092
##           Detection Rate : 0.4965
##      Detection Prevalence : 0.7328
##           Balanced Accuracy : 0.6053
##
##           'Positive' Class : 0
##

```

```

lg.gr.09 <- glm(GI09 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
                INNAPOYO,
                family=binomial(link = "logit"),
                data = Training.gr)
summary(lg.gr.09)

```

```

##
## Call:
## glm(formula = GI09 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.gr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7257  -1.0019  -0.7953   1.1563   1.6929
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)

```

```
## (Intercept) -1.152e+00  6.943e-02 -16.596 < 2e-16 ***
## COOP1       3.253e-01  3.812e-02  8.535 < 2e-16 ***
## TAMANO      9.436e-06  1.017e-05  0.928 0.353446
## MDOLocal1   -7.950e-03  5.357e-02 -0.148 0.882035
## MDONAC1     5.476e-02  4.504e-02  1.216 0.224031
## MDOUE1      1.165e-01  2.825e-02  4.123 3.74e-05 ***
## INNPROD1    5.035e-03  5.606e-02  0.090 0.928440
## INNOBIEN1   4.732e-01  4.824e-02  9.810 < 2e-16 ***
## INNOSERV1   1.358e-01  3.615e-02  3.756 0.000173 ***
## INNPROC1    9.046e-02  4.759e-02  1.901 0.057336 .
## INNFIABRI1  4.698e-01  3.769e-02 12.465 < 2e-16 ***
## INNLOGIS1   2.990e-01  3.711e-02  8.057 7.82e-16 ***
## INNAPYOY01  3.492e-01  3.465e-02 10.079 < 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: 38726 on 28159 degrees of freedom
## Residual deviance: 36737 on 28147 degrees of freedom
## AIC: 36763
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.gr.09 <- predict(lg.gr.09, Test.gr, type='response')
preds.gr.09 <- ifelse(probabs.gr.09 > 0.5, 1, 0)

confusionMatrix(factor(preds.gr.09), factor(Test.gr$GI09))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 2861 1703
##           1  988 1487
##
##           Accuracy : 0.6177
##           95% CI : (0.6062, 0.6291)
##           No Information Rate : 0.5468
##           P-Value [Acc > NIR] : < 2.2e-16
##
##           Kappa : 0.2136
##
## Mcnemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.7433
##           Specificity : 0.4661
##           Pos Pred Value : 0.6269
##           Neg Pred Value : 0.6008
##           Prevalence : 0.5468
##           Detection Rate : 0.4064
##           Detection Prevalence : 0.6484
##           Balanced Accuracy : 0.6047
##
```



```
##      'Positive' Class : 0
##
```

```
lg.cr.01 <- glm(GI01 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFAFRI + INNLOGIS +
                INNAPOYO,
                family=binomial(link = "logit"),
                data = Training.cr)
summary(lg.cr.01)
```

```
##
## Call:
## glm(formula = GI01 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFAFRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.cr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3622  -0.9797   0.6172   0.8096   1.7482
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -9.085e-01  9.470e-02  -9.593  < 2e-16 ***
## COOP1        7.023e-01  5.553e-02  12.646  < 2e-16 ***
## TAMANO       -3.154e-05  9.936e-06  -3.174  0.00150 **
## MDOLOCAL1     1.009e-01  7.559e-02   1.334  0.18206
## MDONAC1       3.255e-01  5.863e-02   5.552  2.83e-08 ***
## MDOUE1       2.748e-01  3.742e-02   7.343  2.09e-13 ***
## INNPROD1     9.680e-01  7.490e-02  12.923  < 2e-16 ***
## INNOBIEN1    4.619e-01  6.431e-02   7.182  6.85e-13 ***
## INNOSERV1    1.639e-01  5.301e-02   3.093  0.00198 **
## INNPROC1    -6.225e-01  6.267e-02  -9.932  < 2e-16 ***
## INNFAFRI1    6.210e-01  4.929e-02  12.599  < 2e-16 ***
## INNLOGIS1    2.988e-01  5.188e-02   5.759  8.47e-09 ***
## INNAPOYO1    3.415e-01  4.732e-02   7.218  5.28e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 24599  on 19270  degrees of freedom
## Residual deviance: 21515  on 19258  degrees of freedom
## AIC: 21541
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.cr.01 <- predict(lg.cr.01, Test.cr, type='response')
preds.cr.01 <- ifelse(probabs.cr.01 > 0.5, 1, 0)

confusionMatrix(factor(preds.cr.01), factor(Test.cr$GI01))
```

```
## Confusion Matrix and Statistics
##
```

```
##           Reference
## Prediction    0    1
##           0  795  477
##           1  832 2712
##
##           Accuracy : 0.7282
##           95% CI : (0.7154, 0.7407)
##           No Information Rate : 0.6622
##           P-Value [Acc > NIR] : < 2.2e-16
##
##           Kappa : 0.3582
##
## Mcnemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.4886
##           Specificity : 0.8504
##           Pos Pred Value : 0.6250
##           Neg Pred Value : 0.7652
##           Prevalence : 0.3378
##           Detection Rate : 0.1651
##           Detection Prevalence : 0.2641
##           Balanced Accuracy : 0.6695
##
##           'Positive' Class : 0
##
```

```
lg.cr.02 <- glm(GIO2 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
  INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
  INNAPOYO,
  family=binomial(link = "logit"),
  data = Training.cr)
summary(lg.cr.02)
```

```
##
## Call:
## glm(formula = GIO2 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.cr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2395  -1.0435   0.6388   0.8783   1.9443
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.414e+00  9.276e-02 -15.243  < 2e-16 ***
## COOP1        5.097e-01  5.062e-02  10.070  < 2e-16 ***
## TAMANO       -3.643e-05  1.002e-05  -3.637  0.000276 ***
## MDOLOCAL1    9.633e-02  7.281e-02   1.323  0.185861
## MDONAC1      6.635e-01  5.823e-02  11.396  < 2e-16 ***
## MDOUE1       3.316e-01  3.606e-02   9.196  < 2e-16 ***
## INNPROD1     6.554e-01  7.153e-02   9.163  < 2e-16 ***
## INNOBIEN1    5.033e-01  6.098e-02   8.254  < 2e-16 ***
## INNOSERV1    2.491e-01  4.960e-02   5.023  5.08e-07 ***
```

```
## INNPROC1      -4.966e-01  6.047e-02  -8.212  < 2e-16 ***
## INNFBABRI1    6.464e-01  4.762e-02  13.575  < 2e-16 ***
## INNLOGIS1     2.959e-01  4.978e-02   5.944  2.77e-09 ***
## INNAPYOY01    3.820e-01  4.539e-02   8.417  < 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: 25611  on 19270  degrees of freedom
## Residual deviance: 22812  on 19258  degrees of freedom
## AIC: 22838
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.cr.02 <- predict(lg.cr.02, Test.cr, type='response')
preds.cr.02 <- ifelse(probabs.cr.02 > 0.5, 1, 0)
```

```
confusionMatrix(factor(preds.cr.02), factor(Test.cr$GI02))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  897  533
##           1  963 2423
##
##           Accuracy : 0.6894
##           95% CI : (0.6761, 0.7024)
##      No Information Rate : 0.6138
##      P-Value [Acc > NIR] : < 2.2e-16
##
##           Kappa : 0.3155
##
##  Mcnemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.4823
##           Specificity : 0.8197
##           Pos Pred Value : 0.6273
##           Neg Pred Value : 0.7156
##           Prevalence : 0.3862
##           Detection Rate : 0.1863
##      Detection Prevalence : 0.2969
##           Balanced Accuracy : 0.6510
##
##           'Positive' Class : 0
##
```

```
lg.cr.03 <- glm(GI03 ~ COOP + TAMANO + MDOLocal + MDONAC + MDOUE + INNPROD +
  INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
  INNAPYOY,
  family=binomial(link = "logit"),
  data = Training.cr)
summary(lg.cr.03)
```

```
##
## Call:
## glm(formula = GIO3 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.cr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.5465  -1.1023   0.5784   0.7586   1.3819
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.830e-01  9.584e-02  -6.083 1.18e-09 ***
## COOP1        7.567e-01  6.132e-02  12.340 < 2e-16 ***
## TAMANO      -1.851e-05  1.016e-05  -1.821  0.0686 .
## MDOLOCAL1    1.621e-01  7.694e-02   2.106  0.0352 *
## MDONAC1      2.436e-01  5.922e-02   4.114 3.89e-05 ***
## MDOUE1       2.120e-01  3.871e-02   5.477 4.33e-08 ***
## INNPROD1     7.127e-01  8.143e-02   8.752 < 2e-16 ***
## INNOBIEN1    3.551e-01  7.154e-02   4.964 6.92e-07 ***
## INNOSERV1    3.416e-01  5.808e-02   5.881 4.09e-09 ***
## INNPROC1    -3.388e-01  6.579e-02  -5.150 2.61e-07 ***
## INNFBRI1     6.021e-01  5.283e-02  11.396 < 2e-16 ***
## INNLOGIS1    3.012e-01  5.602e-02   5.376 7.60e-08 ***
## INNAPOYO1    4.376e-01  5.083e-02   8.608 < 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: 22535  on 19270  degrees of freedom
## Residual deviance: 20365  on 19258  degrees of freedom
## AIC: 20391
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.cr.03 <- predict(lg.cr.03, Test.cr, type='response')
preds.cr.03 <- ifelse(probabs.cr.03 > 0.5, 1, 0)

confusionMatrix(factor(preds.cr.03), factor(Test.cr$GIO3))
```

```
## Confusion Matrix and Statistics
##
##              Reference
## Prediction    0    1
##              0 298 223
##              1 1048 3247
##
##              Accuracy : 0.7361
##              95% CI : (0.7234, 0.7485)
##              No Information Rate : 0.7205
```

```
##      P-Value [Acc > NIR] : 0.008113
##
##              Kappa : 0.1934
##
## Mcnemar's Test P-Value : < 2.2e-16
##
##      Sensitivity : 0.22140
##      Specificity : 0.93573
##      Pos Pred Value : 0.57198
##      Neg Pred Value : 0.75600
##      Prevalence : 0.27949
##      Detection Rate : 0.06188
##      Detection Prevalence : 0.10818
##      Balanced Accuracy : 0.57857
##
##      'Positive' Class : 0
##
```

```
lg.cr.04 <- glm(GIO4 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
  INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
  INNAPOYO,
  family=binomial(link = "logit"),
  data = Training.cr)
summary(lg.cr.04)
```

```
##
## Call:
## glm(formula = GIO4 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.cr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1194  -1.0839   0.6680   0.9679   1.6090
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -9.711e-01  9.074e-02 -10.703  < 2e-16 ***
## COOP1        2.493e-01  4.736e-02   5.264  1.41e-07 ***
## TAMANO       8.243e-06  1.058e-05   0.779  0.436105
## MDOLOCAL1    -3.018e-03  7.185e-02  -0.042  0.966491
## MDONAC1      2.382e-01  5.675e-02   4.198  2.69e-05 ***
## MDOUE1       6.581e-02  3.593e-02   1.832  0.067019 .
## INNPROD1     -2.344e-02  7.088e-02  -0.331  0.740878
## INNOBIEN1     2.043e-01  6.113e-02   3.342  0.000832 ***
## INNOSERV1     4.693e-01  4.736e-02   9.910  < 2e-16 ***
## INNPROC1      5.794e-01  5.970e-02   9.705  < 2e-16 ***
## INNFBABRI1    4.244e-01  4.748e-02   8.939  < 2e-16 ***
## INNLOGIS1     4.277e-01  4.986e-02   8.577  < 2e-16 ***
## INNAPOYO1     4.680e-01  4.426e-02  10.576  < 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: 25851 on 19270 degrees of freedom
## Residual deviance: 23760 on 19258 degrees of freedom
## AIC: 23786
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.cr.04 <- predict(lg.cr.04, Test.cr, type='response')
preds.cr.04 <- ifelse(probabs.cr.04 > 0.5, 1, 0)

confusionMatrix(factor(preds.cr.04), factor(Test.cr$GIO4))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  812  443
##           1 1110 2451
##
##           Accuracy : 0.6775
##           95% CI : (0.6641, 0.6907)
##       No Information Rate : 0.6009
##       P-Value [Acc > NIR] : < 2.2e-16
##
##           Kappa : 0.2861
##
## Mcnemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.4225
##           Specificity : 0.8469
##       Pos Pred Value : 0.6470
##       Neg Pred Value : 0.6883
##           Prevalence : 0.3991
##       Detection Rate : 0.1686
##       Detection Prevalence : 0.2606
##       Balanced Accuracy : 0.6347
##
##       'Positive' Class : 0
##
```

```
lg.cr.05 <- glm(GIO5 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
  INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
  INNAPOYO,
  family=binomial(link = "logit"),
  data = Training.cr)
summary(lg.cr.05)
```

```
##
## Call:
## glm(formula = GIO5 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##     INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
##     INNAPOYO, family = binomial(link = "logit"), data = Training.cr)
##
```

```
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1923  -1.0915   0.6549   0.9464   1.5586
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -8.589e-01  9.110e-02 -9.428  < 2e-16 ***
## COOP1        3.724e-01  4.851e-02  7.677 1.63e-14 ***
## TAMANO       1.401e-05  1.106e-05  1.267  0.2053
## MDOLocal1    -3.489e-03  7.242e-02 -0.048  0.9616
## MDONAC1      2.376e-01  5.693e-02  4.174 3.00e-05 ***
## MDOUE1       3.703e-03  3.619e-02  0.102  0.9185
## INNPROD1     -4.043e-02  7.147e-02 -0.566  0.5716
## INNOBIEN1    1.397e-01  6.172e-02  2.264  0.0236 *
## INNOSERV1    4.537e-01  4.778e-02  9.497  < 2e-16 ***
## INNPROC1     4.239e-01  6.000e-02  7.065 1.60e-12 ***
## INNFIABRI1   6.687e-01  4.787e-02 13.969  < 2e-16 ***
## INNLOGIS1    4.190e-01  5.048e-02  8.301  < 2e-16 ***
## INNAPYOY01   4.811e-01  4.512e-02 10.662  < 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: 25691  on 19270  degrees of freedom
## Residual deviance: 23505  on 19258  degrees of freedom
## AIC: 23531
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.cr.05 <- predict(lg.cr.05, Test.cr, type='response')
preds.cr.05 <- ifelse(probabs.cr.05 > 0.5, 1, 0)

confusionMatrix(factor(preds.cr.05), factor(Test.cr$GI05))
```

```
## Confusion Matrix and Statistics
##
##              Reference
## Prediction    0    1
##              0  798  450
##              1 1090 2478
##
##              Accuracy : 0.6802
##              95% CI : (0.6669, 0.6934)
##      No Information Rate : 0.608
##      P-Value [Acc > NIR] : < 2.2e-16
##
##              Kappa : 0.2862
##
##      McNemar's Test P-Value : < 2.2e-16
##
##              Sensitivity : 0.4227
##              Specificity : 0.8463
##              Pos Pred Value : 0.6394
```

```
##          Neg Pred Value : 0.6945
##          Prevalence : 0.3920
##          Detection Rate : 0.1657
##          Detection Prevalence : 0.2591
##          Balanced Accuracy : 0.6345
##
##          'Positive' Class : 0
##
```

```
lg.cr.06 <- glm(GIO6 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
                INNAPOYO,
                family=binomial(link = "logit"),
                data = Training.cr)
summary(lg.cr.06)
```

```
##
## Call:
## glm(formula = GIO6 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.cr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8497  -1.0763  -0.7298   1.0667   1.7927
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.346e+00  9.005e-02 -14.952  < 2e-16 ***
## COOP1        2.791e-01  4.429e-02  6.301  2.96e-10 ***
## TAMANO       9.169e-06  9.431e-06  0.972   0.331
## MDOLOCAL1    -3.657e-02  6.982e-02 -0.524   0.600
## MDONAC1      2.352e-01  5.715e-02  4.115  3.86e-05 ***
## MDOUE1       2.105e-01  3.485e-02  6.039  1.55e-09 ***
## INNPROD1     -4.662e-02  6.713e-02 -0.695   0.487
## INNOBIEN1    3.474e-01  5.690e-02  6.105  1.03e-09 ***
## INNOSERV1    2.426e-01  4.375e-02  5.544  2.95e-08 ***
## INNPROC1     3.252e-01  5.737e-02  5.669  1.44e-08 ***
## INNFBABRI1   5.800e-01  4.386e-02 13.224  < 2e-16 ***
## INNLOGIS1    3.942e-01  4.482e-02  8.794  < 2e-16 ***
## INNAPOYO1    2.901e-01  4.090e-02  7.094  1.31e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 26712  on 19270  degrees of freedom
## Residual deviance: 24964  on 19258  degrees of freedom
## AIC: 24990
##
## Number of Fisher Scoring iterations: 4
```



```

probabs.cr.06 <- predict(lg.cr.06, Test.cr, type='response')
preds.cr.06 <- ifelse(probabs.cr.06 > 0.5, 1, 0)

confusionMatrix(factor(preds.cr.06), factor(Test.cr$GIO6))

```

```

## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 1634  913
##           1  823 1446
##
##           Accuracy : 0.6395
##           95% CI : (0.6258, 0.6531)
##       No Information Rate : 0.5102
##       P-Value [Acc > NIR] : < 2e-16
##
##           Kappa : 0.2782
##
##  Mcnemar's Test P-Value : 0.03267
##
##           Sensitivity : 0.6650
##           Specificity : 0.6130
##       Pos Pred Value : 0.6415
##       Neg Pred Value : 0.6373
##           Prevalence : 0.5102
##       Detection Rate : 0.3393
##   Detection Prevalence : 0.5289
##       Balanced Accuracy : 0.6390
##
##       'Positive' Class : 0
##

```

```

lg.cr.07 <- glm(GIO7 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
                INNAPPOYO,
                family=binomial(link = "logit"),
                data = Training.cr)
summary(lg.cr.07)

```

```

##
## Call:
## glm(formula = GIO7 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##     INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
##     INNAPPOYO, family = binomial(link = "logit"), data = Training.cr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9916  -1.1276   0.7088   1.0278   1.6792
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)

```

```
## (Intercept) -1.086e+00  8.963e-02 -12.119  < 2e-16 ***
## COOP1       3.494e-01  4.574e-02   7.638  2.20e-14 ***
## TAMANO      4.974e-06  9.604e-06   0.518  0.60450
## MDOLOCAL1   -4.372e-02  7.045e-02  -0.621  0.53482
## MDONAC1     1.734e-01  5.624e-02   3.082  0.00205 **
## MDOUE1      1.931e-01  3.495e-02   5.525  3.29e-08 ***
## INNPROD1    -7.385e-02  6.787e-02  -1.088  0.27656
## INNOBIEN1    3.778e-01  5.778e-02   6.539  6.18e-11 ***
## INNOSERV1    2.709e-01  4.507e-02   6.011  1.84e-09 ***
## INNPROC1     3.253e-01  5.778e-02   5.630  1.81e-08 ***
## INNFIABRI1   6.262e-01  4.484e-02  13.967  < 2e-16 ***
## INNLOGIS1    3.830e-01  4.654e-02   8.230  < 2e-16 ***
## INNAPYOY01   3.324e-01  4.222e-02   7.872  3.48e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 26538  on 19270  degrees of freedom
## Residual deviance: 24663  on 19258  degrees of freedom
## AIC: 24689
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.cr.07 <- predict(lg.cr.07, Test.cr, type='response')
preds.cr.07 <- ifelse(probabs.cr.07 > 0.5, 1, 0)

confusionMatrix(factor(preds.cr.07), factor(Test.cr$GI07))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 1232  708
##           1  931 1945
##
##              Accuracy : 0.6597
##              95% CI : (0.6461, 0.6731)
##      No Information Rate : 0.5509
##      P-Value [Acc > NIR] : < 2.2e-16
##
##              Kappa : 0.3056
##
##  Mcnemar's Test P-Value : 4.168e-08
##
##              Sensitivity : 0.5696
##              Specificity : 0.7331
##              Pos Pred Value : 0.6351
##              Neg Pred Value : 0.6763
##              Prevalence : 0.4491
##              Detection Rate : 0.2558
##      Detection Prevalence : 0.4028
##              Balanced Accuracy : 0.6514
##
```

```
##      'Positive' Class : 0
##
```

```
lg.cr.08 <- glm(GIO8 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFAFRI + INNLOGIS +
                INNAPOYO,
                family=binomial(link = "logit"),
                data = Training.cr)
summary(lg.cr.08)
```

```
##
## Call:
## glm(formula = GIO8 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFAFRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.cr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1280  -0.9975  -0.7417   1.1460   1.7640
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.234e+00  9.065e-02 -13.616  < 2e-16 ***
## COOP1        7.719e-01  4.469e-02  17.271  < 2e-16 ***
## TAMANO       4.302e-05  1.007e-05   4.273  1.93e-05 ***
## MDOLOCAL1    -4.725e-02  7.027e-02  -0.672   0.5013
## MDONAC1      6.853e-02  5.839e-02   1.174   0.2406
## MDOUE1       2.586e-01  3.534e-02   7.318  2.52e-13 ***
## INNPROD1     -9.433e-02  6.835e-02  -1.380   0.1675
## INNOBIEN1    6.063e-01  5.785e-02  10.480  < 2e-16 ***
## INNOSERV1    5.696e-02  4.322e-02   1.318   0.1875
## INNPROC1     -1.476e-01  5.778e-02  -2.554   0.0106 *
## INNFAFRI1    6.595e-01  4.448e-02  14.829  < 2e-16 ***
## INNLOGIS1    4.388e-01  4.461e-02   9.835  < 2e-16 ***
## INNAPOYO1    2.354e-01  4.092e-02   5.753  8.78e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 26491  on 19270  degrees of freedom
## Residual deviance: 24656  on 19258  degrees of freedom
## AIC: 24682
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.cr.08 <- predict(lg.cr.08, Test.cr, type='response')
preds.cr.08 <- ifelse(probabs.cr.08 > 0.5, 1, 0)

confusionMatrix(factor(preds.cr.08), factor(Test.cr$GIO8))
```

```
## Confusion Matrix and Statistics
##
```

```
##           Reference
## Prediction    0    1
##           0 1963 1037
##           1   700 1116
##
##           Accuracy : 0.6393
##           95% CI : (0.6256, 0.6529)
##           No Information Rate : 0.5529
##           P-Value [Acc > NIR] : < 2.2e-16
##
##           Kappa : 0.2594
##
## Mcnemar's Test P-Value : 7.51e-16
##
##           Sensitivity : 0.7371
##           Specificity : 0.5183
##           Pos Pred Value : 0.6543
##           Neg Pred Value : 0.6145
##           Prevalence : 0.5529
##           Detection Rate : 0.4076
##           Detection Prevalence : 0.6229
##           Balanced Accuracy : 0.6277
##
##           'Positive' Class : 0
##
```

```
lg.cr.09 <- glm(GIO9 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
  INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
  INNAPOYO,
  family=binomial(link = "logit"),
  data = Training.cr)
summary(lg.cr.09)
```

```
##
## Call:
## glm(formula = GIO9 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##       INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
##       INNAPOYO, family = binomial(link = "logit"), data = Training.cr)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9662  -1.0721  -0.7882   1.1224   1.6776
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.220e+00  8.977e-02 -13.586 < 2e-16 ***
## COOP1        6.339e-01  4.491e-02  14.114 < 2e-16 ***
## TAMANO       1.540e-05  9.425e-06   1.634  0.10227
## MDOLOCAL1    9.329e-02  7.000e-02   1.333  0.18261
## MDONAC1      1.230e-01  5.722e-02   2.150  0.03154 *
## MDOUE1       2.190e-01  3.487e-02   6.281 3.37e-10 ***
## INNPROD1     -5.764e-02  6.739e-02  -0.855  0.39239
## INNOBIEN1    6.197e-01  5.698e-02  10.875 < 2e-16 ***
## INNOSERV1    1.252e-01  4.349e-02   2.879  0.00399 **
```

```
## INNPORC1      -2.850e-01  5.721e-02  -4.981 6.32e-07 ***
## INNFBABR11    6.375e-01  4.435e-02  14.374 < 2e-16 ***
## INNLOGIS1     4.768e-01  4.522e-02  10.545 < 2e-16 ***
## INNAP0Y01     3.976e-01  4.119e-02   9.654 < 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: 26691  on 19270  degrees of freedom
## Residual deviance: 24945  on 19258  degrees of freedom
## AIC: 24971
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.cr.09 <- predict(lg.cr.09, Test.cr, type='response')
preds.cr.09 <- ifelse(probabs.cr.09 > 0.5, 1, 0)
```

```
confusionMatrix(factor(preds.cr.09), factor(Test.cr$GI09))
```

```
## Confusion Matrix and Statistics
##
##              Reference
## Prediction    0    1
##              0 1817 1051
##              1   724 1224
##
##              Accuracy : 0.6314
##              95% CI : (0.6176, 0.6451)
##      No Information Rate : 0.5276
##      P-Value [Acc > NIR] : < 2.2e-16
##
##              Kappa : 0.255
##
##      McNemar's Test P-Value : 1.011e-14
##
##              Sensitivity : 0.7151
##              Specificity : 0.5380
##              Pos Pred Value : 0.6335
##              Neg Pred Value : 0.6283
##              Prevalence : 0.5276
##              Detection Rate : 0.3773
##      Detection Prevalence : 0.5955
##              Balanced Accuracy : 0.6265
##
##              'Positive' Class : 0
##
```

```
lg.re.01 <- glm(GI01 ~ COOP + TAMANO + MDOLocal + MDONAC + MDOUE + INNPORC +
  INNOBIEN + INNOSERV + INNPORC + INNFBABR1 + INNLOGIS +
  INNAP0Y0,
  family=binomial(link = "logit"),
  data = Training.re)
summary(lg.re.01)
```

```
##
## Call:
## glm(formula = GIO1 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.re)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.4950  -1.0746   0.5610   0.7879   1.5356
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.519e-01  9.772e-02 -4.624 3.76e-06 ***
## COOP1        6.936e-01  5.260e-02 13.186 < 2e-16 ***
## TAMANO      -1.519e-05  1.019e-05 -1.491 0.136080
## MDOLOCAL1    3.117e-01  7.827e-02  3.983 6.82e-05 ***
## MDONAC1      2.056e-01  6.517e-02  3.154 0.001610 **
## MDOUE1       2.521e-01  4.116e-02  6.126 9.03e-10 ***
## INNPROD1     6.938e-01  9.068e-02  7.651 2.00e-14 ***
## INNOBIEN1    5.781e-01  8.124e-02  7.116 1.11e-12 ***
## INNOSERV1    2.347e-01  7.041e-02  3.333 0.000859 ***
## INNPROC1    -7.172e-01  6.825e-02 -10.509 < 2e-16 ***
## INNFBRI1     6.737e-01  5.742e-02 11.732 < 2e-16 ***
## INNLOGIS1    3.684e-01  6.035e-02  6.104 1.04e-09 ***
## INNAPOYO1    2.288e-01  5.562e-02  4.113 3.90e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 21846  on 18330  degrees of freedom
## Residual deviance: 19345  on 18318  degrees of freedom
## AIC: 19371
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.re.01 <- predict(lg.re.01, Test.re, type='response')
preds.re.01 <- ifelse(probabs.re.01 > 0.5, 1, 0)

confusionMatrix(factor(preds.re.01), factor(Test.re$GIO1))
```

```
## Confusion Matrix and Statistics
##
##              Reference
## Prediction    0    1
##              0 338 239
##              1 967 3036
##
##              Accuracy : 0.7367
##              95% CI : (0.7237, 0.7494)
##              No Information Rate : 0.7151
```

```
##      P-Value [Acc > NIR] : 0.000582
##
##      Kappa : 0.2235
##
##      McNemar's Test P-Value : < 2.2e-16
##
##      Sensitivity : 0.2590
##      Specificity : 0.9270
##      Pos Pred Value : 0.5858
##      Neg Pred Value : 0.7584
##      Prevalence : 0.2849
##      Detection Rate : 0.0738
##      Detection Prevalence : 0.1260
##      Balanced Accuracy : 0.5930
##
##      'Positive' Class : 0
##
```

```
lg.re.02 <- glm(GIO2 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
  INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
  INNAPOYO,
  family=binomial(link = "logit"),
  data = Training.re)
summary(lg.re.02)
```

```
##
## Call:
## glm(formula = GIO2 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.re)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3925  -1.0805   0.5993   0.8300   1.7413
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -9.517e-01  9.598e-02  -9.915  < 2e-16 ***
## COOP1        5.689e-01  4.855e-02  11.718  < 2e-16 ***
## TAMANO       -2.230e-05  9.844e-06  -2.266  0.02348 *
## MDOLOCAL1    2.120e-01  7.657e-02  2.769  0.00561 **
## MDONAC1      5.752e-01  6.418e-02  8.963  < 2e-16 ***
## MDOUE1       3.654e-01  3.964e-02  9.217  < 2e-16 ***
## INNPROD1     5.071e-01  8.415e-02  6.026  1.68e-09 ***
## INNOBIEN1    6.148e-01  7.478e-02  8.222  < 2e-16 ***
## INNOSERV1    2.128e-01  6.456e-02  3.296  0.00098 ***
## INNPROC1     -6.462e-01  6.600e-02 -9.791  < 2e-16 ***
## INNFBABRI1   7.413e-01  5.545e-02  13.369  < 2e-16 ***
## INNLOGIS1    3.902e-01  5.812e-02  6.714  1.89e-11 ***
## INNAPOYO1    2.132e-01  5.357e-02  3.980  6.88e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
## Null deviance: 23072 on 18330 degrees of freedom
## Residual deviance: 20483 on 18318 degrees of freedom
## AIC: 20509
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.re.02 <- predict(lg.re.02, Test.re, type='response')
preds.re.02 <- ifelse(probabs.re.02 > 0.5, 1, 0)

confusionMatrix(factor(preds.re.02), factor(Test.re$GIO2))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  540  354
##           1  931 2755
##
##           Accuracy : 0.7194
##           95% CI : (0.7062, 0.7324)
##           No Information Rate : 0.6788
##           P-Value [Acc > NIR] : 1.433e-09
##
##           Kappa : 0.2824
##
## Mcnemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.3671
##           Specificity : 0.8861
##           Pos Pred Value : 0.6040
##           Neg Pred Value : 0.7474
##           Prevalence : 0.3212
##           Detection Rate : 0.1179
##           Detection Prevalence : 0.1952
##           Balanced Accuracy : 0.6266
##
##           'Positive' Class : 0
##
```

```
lg.re.03 <- glm(GIO3 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
  INNOBIEN + INNOSERV + INNPROC + INNFIABRI + INNLOGIS +
  INNAPOYO,
  family=binomial(link = "logit"),
  data = Training.re)
summary(lg.re.03)
```

```
##
## Call:
## glm(formula = GIO3 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
## INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFIABRI + INNLOGIS +
## INNAPOYO, family = binomial(link = "logit"), data = Training.re)
##
```



```
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.7422   0.2590   0.5628   0.7385   1.2140
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -7.701e-02  1.013e-01  -0.760   0.44700
## COOP1        8.050e-01  5.878e-02  13.694 < 2e-16 ***
## TAMANO       -5.125e-06  1.141e-05  -0.449   0.65347
## MDOLocal1    2.451e-01  8.141e-02   3.011   0.00260 **
## MDONAC1      1.282e-01  6.731e-02   1.905   0.05684 .
## MDOUE1       2.026e-01  4.304e-02   4.707 2.51e-06 ***
## INNPROD1     3.132e-01  1.016e-01   3.082   0.00206 **
## INNOBIEN1    5.629e-01  9.309e-02   6.047 1.48e-09 ***
## INNOSERV1    5.377e-01  7.967e-02   6.749 1.49e-11 ***
## INNPROC1    -4.944e-01  7.416e-02  -6.666 2.63e-11 ***
## INNFIABRI1   7.280e-01  6.378e-02  11.415 < 2e-16 ***
## INNLOGIS1    3.583e-01  6.724e-02   5.328 9.94e-08 ***
## INNAPYOY01   4.275e-01  6.194e-02   6.902 5.12e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 19802  on 18330  degrees of freedom
## Residual deviance: 18049  on 18318  degrees of freedom
## AIC: 18075
##
## Number of Fisher Scoring iterations: 5
```

```
probabs.re.03 <- predict(lg.re.03, Test.re, type='response')
preds.re.03 <- ifelse(probabs.re.03 > 0.5, 1, 0)

confusionMatrix(factor(preds.re.03), factor(Test.re$GI03))
```

```
## Confusion Matrix and Statistics
##
##              Reference
## Prediction    0    1
##              0    8    3
##              1 1029 3540
##
##              Accuracy : 0.7747
##              95% CI : (0.7623, 0.7867)
##      No Information Rate : 0.7736
##      P-Value [Acc > NIR] : 0.4381
##
##              Kappa : 0.0106
##
##      McNemar's Test P-Value : <2e-16
##
##              Sensitivity : 0.007715
##              Specificity : 0.999153
##              Pos Pred Value : 0.727273
```

```
##          Neg Pred Value : 0.774787
##          Prevalence : 0.226419
##          Detection Rate : 0.001747
##          Detection Prevalence : 0.002402
##          Balanced Accuracy : 0.503434
##
##          'Positive' Class : 0
##
```

```
lg.re.04 <- glm(GIO4 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFAFRI + INNLOGIS +
                INNAPOYO,
                family=binomial(link = "logit"),
                data = Training.re)
summary(lg.re.04)
```

```
##
## Call:
## glm(formula = GIO4 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFAFRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.re)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3689  -1.0897   0.6537   0.9291   1.4604
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.881e-01  9.379e-02  -7.337 2.19e-13 ***
## COOP1        2.273e-01  4.324e-02   5.256 1.47e-07 ***
## TAMANO       4.105e-05  1.344e-05   3.055 0.00225 **
## MDOLOCAL1    1.794e-01  7.454e-02   2.407 0.01611 *
## MDONAC1      1.231e-01  6.348e-02   1.940 0.05243 .
## MDOUE1       4.314e-02  3.925e-02   1.099 0.27175
## INNPROD1     -1.517e-01  7.998e-02  -1.897 0.05779 .
## INNOBIEN1    2.830e-01  7.216e-02   3.921 8.81e-05 ***
## INNOSERV1    5.805e-01  5.771e-02  10.059 < 2e-16 ***
## INNPROC1     4.713e-01  6.529e-02   7.218 5.26e-13 ***
## INNFAFRI1    5.299e-01  5.610e-02   9.447 < 2e-16 ***
## INNLOGIS1    4.441e-01  5.836e-02   7.609 2.76e-14 ***
## INNAPOYO1    3.749e-01  5.253e-02   7.137 9.54e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 24260  on 18330  degrees of freedom
## Residual deviance: 22231  on 18318  degrees of freedom
## AIC: 22257
##
## Number of Fisher Scoring iterations: 4
```

```

probabs.re.04 <- predict(lg.re.04, Test.re, type='response')
preds.re.04 <- ifelse(probabs.re.04 > 0.5, 1, 0)

confusionMatrix(factor(preds.re.04), factor(Test.re$GI04))

```

```

## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  779  595
##           1  924 2282
##
##           Accuracy : 0.6683
##           95% CI : (0.6545, 0.682)
##       No Information Rate : 0.6282
##       P-Value [Acc > NIR] : 7.828e-09
##
##           Kappa : 0.2609
##
##  Mcnemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.4574
##           Specificity : 0.7932
##       Pos Pred Value : 0.5670
##       Neg Pred Value : 0.7118
##           Prevalence : 0.3718
##       Detection Rate : 0.1701
##   Detection Prevalence : 0.3000
##       Balanced Accuracy : 0.6253
##
##       'Positive' Class : 0
##

```

```

lg.re.05 <- glm(GI05 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFIABRI + INNLOGIS +
                INNAPPOYO,
                family=binomial(link = "logit"),
                data = Training.re)
summary(lg.re.05)

```

```

##
## Call:
## glm(formula = GI05 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##     INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFIABRI + INNLOGIS +
##     INNAPPOYO, family = binomial(link = "logit"), data = Training.re)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3344  -1.0817   0.6572   0.9168   1.3762
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)

```

```
## (Intercept) -4.264e-01  9.413e-02 -4.529 5.91e-06 ***
## COOP1       3.069e-01  4.356e-02  7.044 1.86e-12 ***
## TAMANO      3.727e-05  1.306e-05  2.854 0.004311 **
## MDOLocal1   5.473e-02  7.525e-02  0.727 0.467051
## MDONAC1     1.419e-01  6.353e-02  2.233 0.025542 *
## MDOUE1     -3.007e-02  3.935e-02 -0.764 0.444744
## INNPROD1    -2.143e-01  7.974e-02 -2.687 0.007207 **
## INNOBIEN1   2.397e-01  7.193e-02  3.332 0.000861 ***
## INNOSERV1   5.714e-01  5.736e-02  9.962 < 2e-16 ***
## INNPROC1    4.231e-01  6.493e-02  6.516 7.20e-11 ***
## INNFIABRI1  6.379e-01  5.564e-02 11.465 < 2e-16 ***
## INNLOGIS1   3.515e-01  5.748e-02  6.115 9.66e-10 ***
## INNAPYOY01  2.868e-01  5.255e-02  5.457 4.84e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 24095 on 18330 degrees of freedom
## Residual deviance: 22222 on 18318 degrees of freedom
## AIC: 22248
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.re.05 <- predict(lg.re.05, Test.re, type='response')
preds.re.05 <- ifelse(probabs.re.05 > 0.5, 1, 0)

confusionMatrix(factor(preds.re.05), factor(Test.re$GI05))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  709  536
##           1  959 2376
##
##               Accuracy : 0.6736
##               95% CI : (0.6598, 0.6872)
##               No Information Rate : 0.6358
##               P-Value [Acc > NIR] : 4.697e-08
##
##               Kappa : 0.2548
##
## Mcnemar's Test P-Value : < 2.2e-16
##
##               Sensitivity : 0.4251
##               Specificity : 0.8159
##               Pos Pred Value : 0.5695
##               Neg Pred Value : 0.7124
##               Prevalence : 0.3642
##               Detection Rate : 0.1548
##               Detection Prevalence : 0.2718
##               Balanced Accuracy : 0.6205
##
```

```
##      'Positive' Class : 0
##
```

```
lg.re.06 <- glm(GIO6 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFBARI + INNLOGIS +
                INNAPOYO,
                family=binomial(link = "logit"),
                data = Training.re)
summary(lg.re.06)
```

```
##
## Call:
## glm(formula = GIO6 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBARI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.re)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9432  -1.0478   0.6435   1.0462   1.6016
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.027e+00  9.222e-02 -11.139  < 2e-16 ***
## COOP1       2.159e-01  4.043e-02  5.339  9.32e-08 ***
## TAMANO      1.504e-05  9.893e-06  1.520  0.12840
## MDOLOCAL1   6.945e-02  7.271e-02  0.955  0.33949
## MDONAC1     1.859e-01  6.271e-02  2.965  0.00302 **
## MDOUE1      2.506e-01  3.777e-02  6.633  3.28e-11 ***
## INNPROD1    -1.763e-01  7.303e-02 -2.413  0.01581 *
## INNOBIEN1   3.824e-01  6.465e-02  5.915  3.31e-09 ***
## INNOSERV1   2.129e-01  5.146e-02  4.137  3.52e-05 ***
## INNPROC1    1.840e-01  6.018e-02  3.058  0.00223 **
## INNFBARI1   6.698e-01  5.037e-02 13.297  < 2e-16 ***
## INNLOGIS1   4.814e-01  5.126e-02  9.392  < 2e-16 ***
## INNAPOYO1   2.346e-01  4.703e-02  4.987  6.13e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 25367  on 18330  degrees of freedom
## Residual deviance: 23727  on 18318  degrees of freedom
## AIC: 23753
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.re.06 <- predict(lg.re.06, Test.re, type='response')
preds.re.06 <- ifelse(probabs.re.06 > 0.5, 1, 0)

confusionMatrix(factor(preds.re.06), factor(Test.re$GIO6))
```

```
## Confusion Matrix and Statistics
##
```

```
##           Reference
## Prediction    0    1
##           0 1400  928
##           1   792 1460
##
##           Accuracy : 0.6245
##           95% CI : (0.6102, 0.6385)
##           No Information Rate : 0.5214
##           P-Value [Acc > NIR] : < 2.2e-16
##
##           Kappa : 0.2494
##
## Mcnemar's Test P-Value : 0.001133
##
##           Sensitivity : 0.6387
##           Specificity : 0.6114
##           Pos Pred Value : 0.6014
##           Neg Pred Value : 0.6483
##           Prevalence : 0.4786
##           Detection Rate : 0.3057
##           Detection Prevalence : 0.5083
##           Balanced Accuracy : 0.6250
##
##           'Positive' Class : 0
##
```

```
lg.re.07 <- glm(GIO7 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFBARI + INNLOGIS +
                INNAPOYO,
                family=binomial(link = "logit"),
                data = Training.re)
summary(lg.re.07)
```

```
##
## Call:
## glm(formula = GIO7 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBARI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.re)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3417  -1.1482   0.6367   0.9298   1.4152
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.426e-01  9.477e-02  -6.781 1.20e-11 ***
## COOP1        3.876e-01  4.509e-02   8.597 < 2e-16 ***
## TAMANO       3.220e-05  1.274e-05   2.527 0.011501 *
## MDOLOCAL1    9.910e-02  7.593e-02   1.305 0.191849
## MDONAC1      2.457e-01  6.333e-02   3.879 0.000105 ***
## MDOUE1       1.057e-01  3.950e-02   2.675 0.007466 **
## INNPROD1     -2.641e-01  8.097e-02  -3.261 0.001108 **
## INNOBIEN1    4.170e-01  7.284e-02   5.725 1.04e-08 ***
## INNOSERV1    4.924e-01  5.958e-02   8.266 < 2e-16 ***
```

```
## INNPROC1      3.839e-01  6.679e-02   5.747 9.07e-09 ***
## INNFBRI1      7.350e-01  5.722e-02  12.844 < 2e-16 ***
## INNLOGIS1     3.982e-01  6.007e-02   6.629 3.38e-11 ***
## INNAPOYO1     3.004e-01  5.484e-02   5.478 4.31e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 23722  on 18330  degrees of freedom
## Residual deviance: 21658  on 18318  degrees of freedom
## AIC: 21684
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.re.07 <- predict(lg.re.07, Test.re, type='response')
preds.re.07 <- ifelse(probabs.re.07 > 0.5, 1, 0)
```

```
confusionMatrix(factor(preds.re.07), factor(Test.re$GIO7))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0  719  600
##           1  891 2370
##
##           Accuracy : 0.6745
##           95% CI : (0.6607, 0.688)
##      No Information Rate : 0.6485
##      P-Value [Acc > NIR] : 0.0001132
##
##           Kappa : 0.2551
##
##  Mcnemar's Test P-Value : 5.898e-14
##
##           Sensitivity : 0.4466
##           Specificity : 0.7980
##      Pos Pred Value : 0.5451
##      Neg Pred Value : 0.7268
##           Prevalence : 0.3515
##      Detection Rate : 0.1570
##      Detection Prevalence : 0.2880
##      Balanced Accuracy : 0.6223
##
##           'Positive' Class : 0
##
```

```
lg.re.08 <- glm(GIO8 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
                INNOBIEN + INNOSERV + INNPROC + INNFBRI + INNLOGIS +
                INNAPOYO,
                family=binomial(link = "logit"),
                data = Training.re)
summary(lg.re.08)
```

```
##
## Call:
## glm(formula = GIO8 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.re)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0966  -1.0708  -0.7673   1.1012   1.7853
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -8.821e-01  9.288e-02  -9.498  < 2e-16 ***
## COOP1        8.315e-01  4.110e-02  20.231  < 2e-16 ***
## TAMANO       4.124e-05  1.031e-05   3.999  6.37e-05 ***
## MDOLOCAL1    8.698e-02  7.333e-02   1.186  0.235608
## MDONAC1      1.582e-02  6.376e-02   0.248  0.804079
## MDOUE1       2.669e-01  3.816e-02   6.994  2.67e-12 ***
## INNPROD1    -2.803e-01  7.373e-02  -3.802  0.000143 ***
## INNOBIEN1    5.830e-01  6.520e-02   8.942  < 2e-16 ***
## INNOSERV1    3.203e-03  5.081e-02   0.063  0.949731
## INNPROC1    -4.246e-01  6.050e-02  -7.019  2.24e-12 ***
## INNFBRI1     9.014e-01  5.075e-02  17.762  < 2e-16 ***
## INNLOGIS1    4.766e-01  5.042e-02   9.454  < 2e-16 ***
## INNAPOYO1    2.010e-01  4.659e-02   4.315  1.60e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 25403  on 18330  degrees of freedom
## Residual deviance: 23559  on 18318  degrees of freedom
## AIC: 23585
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.re.08 <- predict(lg.re.08, Test.re, type='response')
preds.re.08 <- ifelse(probabs.re.08 > 0.5, 1, 0)

confusionMatrix(factor(preds.re.08), factor(Test.re$GIO8))
```

```
## Confusion Matrix and Statistics
##
##              Reference
## Prediction    0    1
##              0 1692 1053
##              1  618 1217
##
##              Accuracy : 0.6352
##              95% CI : (0.621, 0.6491)
##              No Information Rate : 0.5044
```



```
##      P-Value [Acc > NIR] : < 2.2e-16
##
##              Kappa : 0.269
##
##      McNemar's Test P-Value : < 2.2e-16
##
##              Sensitivity : 0.7325
##              Specificity : 0.5361
##              Pos Pred Value : 0.6164
##              Neg Pred Value : 0.6632
##              Prevalence : 0.5044
##              Detection Rate : 0.3694
##      Detection Prevalence : 0.5993
##              Balanced Accuracy : 0.6343
##
##      'Positive' Class : 0
##
```

```
lg.re.09 <- glm(GIO9 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE + INNPROD +
  INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
  INNAPOYO,
  family=binomial(link = "logit"),
  data = Training.re)
summary(lg.re.09)
```

```
##
## Call:
## glm(formula = GIO9 ~ COOP + TAMANO + MDOLOCAL + MDONAC + MDOUE +
##      INNPROD + INNOBIEN + INNOSERV + INNPROC + INNFBABRI + INNLOGIS +
##      INNAPOYO, family = binomial(link = "logit"), data = Training.re)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1268  -1.0817  -0.7369   1.1265   1.7378
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -8.787e-01  9.263e-02  -9.486  < 2e-16 ***
## COOP1        6.424e-01  4.077e-02  15.755  < 2e-16 ***
## TAMANO       1.369e-05  9.547e-06   1.434  0.15154
## MDOLOCAL1    9.491e-02  7.313e-02   1.298  0.19433
## MDONAC1      2.247e-02  6.356e-02   0.353  0.72376
## MDOUE1       2.395e-01  3.806e-02   6.294 3.09e-10 ***
## INNPROD1     -3.560e-01  7.366e-02  -4.833 1.35e-06 ***
## INNOBIEN1     7.698e-01  6.509e-02  11.826  < 2e-16 ***
## INNOSERV1     1.510e-01  5.112e-02   2.954  0.00313 **
## INNPROC1     -5.036e-01  6.055e-02  -8.317  < 2e-16 ***
## INNFBABRI1    9.199e-01  5.083e-02  18.099  < 2e-16 ***
## INNLOGIS1     4.705e-01  5.082e-02   9.258  < 2e-16 ***
## INNAPOYO1     3.043e-01  4.685e-02   6.495 8.31e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
##      Null deviance: 25411  on 18330  degrees of freedom
## Residual deviance: 23598  on 18318  degrees of freedom
## AIC: 23624
##
## Number of Fisher Scoring iterations: 4
```

```
probabs.re.09 <- predict(lg.re.09, Test.re, type='response')
preds.re.09 <- ifelse(probabs.re.09 > 0.5, 1, 0)

confusionMatrix(factor(preds.re.09), factor(Test.re$GI09))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 1662 1066
##           1   653 1199
##
##           Accuracy : 0.6247
##           95% CI   : (0.6105, 0.6387)
##      No Information Rate : 0.5055
##      P-Value [Acc > NIR] : < 2.2e-16
##
##           Kappa   : 0.2478
##
##  Mcnemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.7179
##           Specificity : 0.5294
##           Pos Pred Value : 0.6092
##           Neg Pred Value : 0.6474
##           Prevalence : 0.5055
##           Detection Rate : 0.3629
##      Detection Prevalence : 0.5956
##           Balanced Accuracy : 0.6236
##
##           'Positive' Class : 0
##
```