mlr

2024-12-07

```
library(nnet)
train_data <- read.csv("train_data.csv")
test_data <- read.csv("test_data.csv")
#Multinomial logistic regression
mlog_reg_outcome <- multinom(result ~., data = train_data)
```

```
## # weights:  126 (82 variable)
## initial  value 112.058453
## iter  10 value 92.333609
## iter  20 value 69.003207
## iter  30 value 56.717895
## iter  40 value 43.522957
## iter  50 value 34.862586
## iter  60 value 28.895613
## iter  70 value 18.155438
## iter  80 value 0.435299
## iter  90 value 0.002427
## iter 100 value 0.000101
## final  value 0.000101
## stopped after 100 iterations
```

```
logregsum <- summary(mlog_reg_outcome)
```

```
## Warning in sqrt(diag(vc)): NaNs produced
```

```
predictions <- predict(mlog_reg_outcome, newdata = test_data)

conf_matrix <- table(test_data$result, predictions)
conf_matrix
```

```
##    predictions
##     0 1 2
##   0 2 2 2
##   1 4 3 0
##   2 1 5 7
```

```
write.csv(conf_matrix, "mlr_conf.csv")
#Accuracy
accuracy <- sum(diag(conf_matrix)) / sum(conf_matrix)
accuracy
```

```
## [1] 0.4615385
```

```
#Pvalues
p_values <- summary(mlog_reg_outcome)$coefficients / summary(mlog_reg_outcome)$standard.errors
```

```
## Warning in sqrt(diag(vc)): NaNs produced
## Warning in sqrt(diag(vc)): NaNs produced
```

```
p_values <- (1 - pnorm(abs(p_values), 0, 1)) * 2
p_values
```

```
##   (Intercept) X possession.team possession.in.contest total.attempts.team
## 1         NaN 0               0                     0                   0
## 2           0 0               0                     0                   0
##   on.target.attempts.team off.target.attempts.team
## 1                       0                       0
## 2                       0                       0
##   attempts.inside.the.penalty.area.team attempts.outside.the.penalty.area..team
## 1                                     0                                       0
## 2                                     0                                       0
##   left.channel.team left.inside.channel.team central.channel.team
## 1                 0                       0                    0
## 2                 0                       0                    0
##   right.inside.channel.team right.channel.team total.offers.to.receive.team
## 1                         0                  0                           0
## 2                         0                  0                           0
##   inbehind.offers.to.receive.team inbetween.offers.to.receive.team
## 1                               0                                0
## 2                               0                                0
##   infront.offers.to.receive.team
## 1                              0
## 2                              0
##   receptions.between.midfield.and.defensive.lines.team
## 1                                                    0
## 2                                                    0
##   attempted.line.breaks.team completed.line.breaks.team
## 1                          0                         0
## 2                          0                         0
##   attempted.defensive.line.breaks.team completed.defensive.line.breaks.team
## 1                                    0                                    0
## 2                                    0                                    0
##   yellow.cards.team red.cards.team fouls.committed fouls.drawn offsides.team
## 1               NaN              0               0           0             0
## 2                 0              0               0           0             0
##   passes.team passes.completed.team crosses.team crosses.completed.team
## 1           0                     0           0                      0
## 2           0                     0           0                      0
##   switches.of.play.completed.team corners.team free.kicks.team
## 1                               0            0               0
## 2                               0            0               0
##   goal.preventions.team forced.turnovers.team defensive.pressures.applied.team
## 1                     0                     0                                0
## 2                     0                     0                                0
##   total.goals total.attempts total.attempted.defensive.line.breaks
## 1           0              0                                     0
## 2           0              0                                     0
```

The accuracy of our model was 0.5. This means that we can predict the outcome of the match with an accuracy greater than random guessing (which would have an accuracy of 0.33).

However, the very high p-values indicate we may have a problem with multicollinearity. Let's fix this by using elastic net as a variable selection method.