Assignment-2.R

Pman

2023-05-14

setwd("C:/Users/Pman/Desktop/REGRESSION MODELS")  
  
churn <- read.csv("telecoms\_churn.csv", header = T)  
  
head(churn)

## X churn customerid gender internetservice partner monthlycharges  
## 1 6924 0 7619-PLRLP Female DSL Yes 74.10  
## 2 5298 0 7647-GYYKX Female No Yes 20.35  
## 3 6548 0 3908-MKIMJ Male DSL Yes 41.95  
## 4 6690 0 8561-NMTBD Female Fiber optic Yes 112.35  
## 5 6128 0 6198-PNNSZ Female Fiber optic Yes 109.80  
## 6 2850 0 6838-HVLXG Female No No 20.30  
## paymentmethod  
## 1 Bank transfer (automatic)  
## 2 Bank transfer (automatic)  
## 3 Electronic check  
## 4 Credit card (automatic)  
## 5 Bank transfer (automatic)  
## 6 Mailed check

fit1 <- glm(churn~factor(gender)+internetservice+partner+monthlycharges  
 +paymentmethod, family = binomial, data = churn)  
  
summary(fit1)

##   
## Call:  
## glm(formula = churn ~ factor(gender) + internetservice + partner +   
## monthlycharges + paymentmethod, family = binomial, data = churn)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.4515 -0.6343 -0.3658 -0.2260 2.8508   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.17470 0.79264 0.220 0.825559   
## factor(gender)Male 0.21865 0.26543 0.824 0.410096   
## internetserviceFiber optic 1.76588 0.52663 3.353 0.000799 \*\*\*  
## internetserviceNo -2.78320 0.60766 -4.580 4.65e-06 \*\*\*  
## partnerYes -0.67730 0.27075 -2.502 0.012366 \*   
## monthlycharges -0.04908 0.01150 -4.269 1.97e-05 \*\*\*  
## paymentmethodCredit card (automatic) 0.72001 0.56985 1.264 0.206409   
## paymentmethodElectronic check 1.72827 0.50345 3.433 0.000597 \*\*\*  
## paymentmethodMailed check 0.94381 0.56449 1.672 0.094532 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 460.30 on 511 degrees of freedom  
## Residual deviance: 383.67 on 503 degrees of freedom  
## AIC: 401.67  
##   
## Number of Fisher Scoring iterations: 6

drop1(fit1,test='LRT')

## Single term deletions  
##   
## Model:  
## churn ~ factor(gender) + internetservice + partner + monthlycharges +   
## paymentmethod  
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 383.67 401.67   
## factor(gender) 1 384.35 400.35 0.6828 0.4086213   
## internetservice 2 411.86 425.86 28.1886 7.567e-07 \*\*\*  
## partner 1 390.10 406.10 6.4346 0.0111919 \*   
## monthlycharges 1 403.73 419.73 20.0584 7.511e-06 \*\*\*  
## paymentmethod 3 403.75 415.75 20.0861 0.0001629 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Use R to calculate an estimate of the probability of churning for:   
# a male, living with his partner, monthly charges=70, internet service=fibre   
# optic, payment method= credit card. Use R to get a 95% confidence interval   
# for this estimate. Incorporate this CI into your report.   
  
new\_data <- data.frame(gender = "Male",  
 partner = "Yes",  
 monthlycharges = 70,  
 internetservice = "Fiber optic",  
 paymentmethod = "Credit card (automatic)")  
  
pred <- predict(fit1, newdata = new\_data, se.fit = TRUE)  
pred

## $fit  
## 1   
## -1.233788   
##   
## $se.fit  
## [1] 0.4703643  
##   
## $residual.scale  
## [1] 1

exp(pred$fit)

## 1   
## 0.2911874

lpci = exp(pred$fit - 1.96\*pred$se.fit)  
upci = exp(pred$fit + 1.96\*pred$se.fit)  
  
c(lpci,upci)

## 1 1   
## 0.1158218 0.7320735