

# [R통계] 분류분석

조성우

Department of E-business

2020-12-09

- **1. 개요**
    - 주제
    - 데이터
    - 연구법(분석기법)
  - **2. 본문**
    - **2.1 데이터 수집 및 전처리**
      - 2.1.1 패키지 준비
      - 2.1.2 데이터 불러오기
      - 2.1.3 데이터 전처리
    - **2.2 Logistic Regression Classification**
      - 2.2.1. 첫번째 Logistic Regression 모델 설계
      - 2.2.2. 단계적선택법 모델개선
- **Logistic Regression 개선모델에서 50%의 cutoff를 사용한 결과**
  - **Logistic Regression 개선모델에서 32%의 cutoff를 사용한 최종결과**
  - **Logistic Regression 개선모델에서 32%의 cutoff를 사용한 최종결과**

## 1. 개요

### 주제

이용서비스,계정정보,인구통계학정보를 독립변수로 이탈고객 train set을 학습시켜 Logistics Regression과 Naive Bayes 분류기를 만들어내고, 이 분류기로 validation set의 이탈고객 여부를 판별하는 분류모델 설계 및 분석 과정을 실시해보고자 한다.

### 데이터

#### 데이터셋

Kaggle에 공유된 "Telco Customer Churn" 데이터셋을 활용

#### 변수 및 특징

- **Churn** :  
이탈고객 여부
- 고객이 이용중인 서비스 :  
phone, multiple lines, internet, online security, online backup, device protection, tech support, and streaming TV and movies
- 고객계정 관련 정보 :  
how long they've been a customer(tenure), contract, payment method, paperless billing, monthly charges, and total charges

- 고객의 인구통계학 정보 :  
gender, age range, and if they have partners and dependents

연구법(분석기법)

- 1. EDA
- 2. Logistic Regression Classification
- 3. Naive Bayes Classification

2. 본문

2.1 데이터 수집 및 전처리

2.1.1 패키지 준비

```
library(tidyverse)
library(MASS)
library(car)
library(e1071) # 베이지
library(caret)
library(cowplot)
library(caTools)
library(pROC)
library(ggcorrplot)
library(ggplot2)
library(gmodels) # 베이지
```

2.1.2 데이터 불러오기

```
setwd("C:\\Users\\JSW\\Desktop\\아주대학교\\1_학기\\3학년 2학기_2020_2\\강의자료\\1_R통계\\R통계 과제\\고
telco=read.csv("Telco_CustomerChurn.csv")
```

```
knitr::kable(head(telco))
```

customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines
7590-VHVEG	Female	0	Yes	No	1	No	No phone service
5575-GNVDE	Male	0	No	No	34	Yes	No
3668-QPYBK	Male	0	No	No	2	Yes	No

customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines
7795-CFOCW	Male	0	No	No	45	No	No phone service
9237-HQITU	Female	0	No	No	2	Yes	No
9305-CDSKC	Female	0	No	No	8	Yes	Yes

**Telco Churn data** 원데이터의 모습

## 2.1.3 데이터 전처리

### 1. Categorical 특징 정리

```
telco = telco[complete.cases(telco),]

telco$SeniorCitizen = as.factor(ifelse(telco$SeniorCitizen==1,'YES','NO'))

telco <- data.frame(lapply(telco, function(x) {
  gsub("No internet service", "No", x)}))

telco <- data.frame(lapply(telco, function(x) {
  gsub("No phone service", "No", x)}))
```

### 2. 연속형 특징 표준화

```
num_columns <- c("tenure", "MonthlyCharges", "TotalCharges")
telco[num_columns] <- sapply(telco[num_columns], as.numeric)

telco_int <- telco[,c("tenure", "MonthlyCharges", "TotalCharges")]
telco_int <- data.frame(scale(telco_int))
```

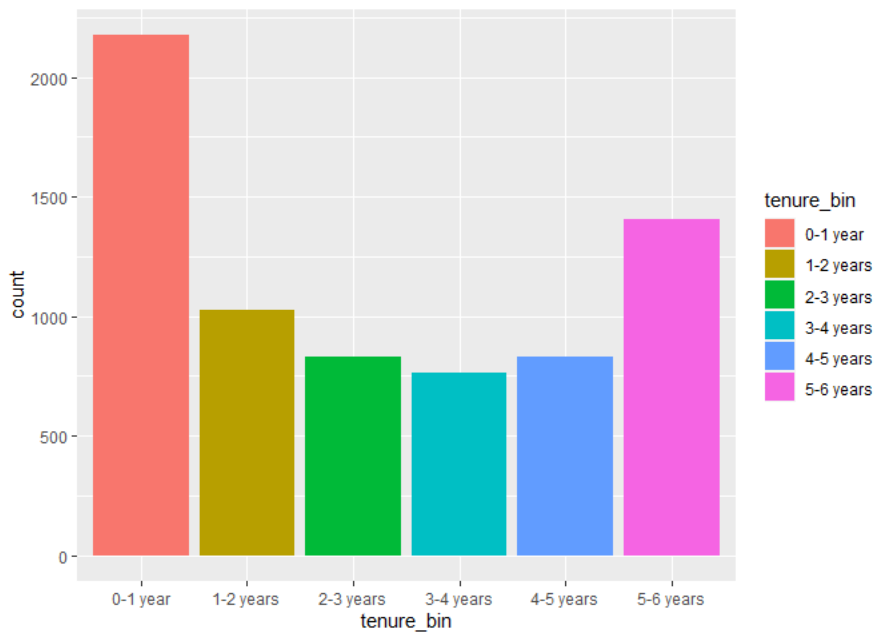
### 3. Creating derived features

```
telco <- mutate(telco, tenure_bin = tenure)

telco$tenure_bin[telco$tenure_bin >= 0 & telco$tenure_bin <= 12] <- '0-1 year'
telco$tenure_bin[telco$tenure_bin > 12 & telco$tenure_bin <= 24] <- '1-2 years'
telco$tenure_bin[telco$tenure_bin > 24 & telco$tenure_bin <= 36] <- '2-3 years'
telco$tenure_bin[telco$tenure_bin > 36 & telco$tenure_bin <= 48] <- '3-4 years'
telco$tenure_bin[telco$tenure_bin > 48 & telco$tenure_bin <= 60] <- '4-5 years'
telco$tenure_bin[telco$tenure_bin > 60 & telco$tenure_bin <= 72] <- '5-6 years'

telco$tenure_bin <- as.factor(telco$tenure_bin)

options(repr.plot.width = 6, repr.plot.height = 3)
ggplot(telco, aes(tenure_bin, fill = tenure_bin)) + geom_bar()
```



#### 4.더미변수 변환

```
telco_cat = telco[, -c(1,6,19,20)]
```

```
#creating dummy variables
```

```
dummy<- data.frame(sapply(telco_cat,function(x) data.frame(model.matrix(~x-1,data =telco_cat))[, -1]))
```

```
head(dummy)
```

```
##   gender SeniorCitizen Partner Dependents PhoneService MultipleLines
## 1      0           0       1           0           0           0
## 2      1           0       0           0           1           0
## 3      1           0       0           0           1           0
## 4      1           0       0           0           0           0
## 5      0           0       0           0           1           0
## 6      0           0       0           0           1           1
##   InternetService.xFiber.optic InternetService.xNo OnlineSecurity OnlineBackup
## 1                        0           0           0           1
## 2                        0           0           1           0
## 3                        0           0           1           1
## 4                        0           0           1           0
## 5                        1           0           0           0
## 6                        1           0           0           0
##   DeviceProtection TechSupport StreamingTV StreamingMovies Contract.xOne.year
## 1                0           0           0           0           0
## 2                1           0           0           0           1
## 3                0           0           0           0           0
## 4                1           1           0           0           1
## 5                0           0           0           0           0
## 6                1           0           1           1           0
##   Contract.xTwo.year PaperlessBilling PaymentMethod.xCredit.card..automatic.
## 1                0           1           0
## 2                0           0           0
## 3                0           1           0
## 4                0           0           0
## 5                0           1           0
## 6                0           1           0
##   PaymentMethod.xElectronic.check PaymentMethod.xMailed.check Churn
```

```
## 1      1      0      0
## 2      0      1      0
## 3      0      1      1
## 4      0      0      0
## 5      1      0      1
## 6      1      0      1
## tenure_bin.x1.2.years tenure_bin.x2.3.years tenure_bin.x3.4.years
## 1      0      0      0
## 2      0      1      0
## 3      0      0      0
## 4      0      0      1
## 5      0      0      0
## 6      0      0      0
## tenure_bin.x4.5.years tenure_bin.x5.6.years
## 1      0      0
## 2      0      0
## 3      0      0
## 4      0      0
## 5      0      0
## 6      0      0
```

### #5. 최종 데이터셋 구성

```
telco_final <- cbind(telco_int, dummy)
head(telco_final)
```

```
## tenure MonthlyCharges TotalCharges gender SeniorCitizen Partner
## 1 -1.28015700 -1.1616113 -0.9941234 0 0 1
## 2 0.06429811 -0.2608594 -0.1737275 1 0 0
## 3 -1.23941594 -0.3638974 -0.9595809 1 0 0
## 4 0.51244982 -0.7477972 -0.1952338 1 0 0
## 5 -1.23941594 0.1961642 -0.9403906 0 0 0
## 6 -0.99496955 1.1584066 -0.6453233 0 0 0
## Dependents PhoneService MultipleLines InternetService.xFiber.optic
## 1 0 0 0 0
## 2 0 1 0 0
## 3 0 1 0 0
## 4 0 0 0 0
## 5 0 1 0 1
## 6 0 1 1 1
## InternetService.xNo OnlineSecurity OnlineBackup DeviceProtection TechSupport
## 1 0 0 1 0 0
## 2 0 1 0 1 0
## 3 0 1 1 0 0
## 4 0 1 0 1 1
## 5 0 0 0 0 0
## 6 0 0 0 1 0
## StreamingTV StreamingMovies Contract.xOne.year Contract.xTwo.year
## 1 0 0 0 0
## 2 0 0 1 0
## 3 0 0 0 0
## 4 0 0 1 0
## 5 0 0 0 0
## 6 1 1 0 0
## PaperlessBilling PaymentMethod.xCredit.card..automatic.
## 1 1 0
## 2 0 0
## 3 1 0
## 4 0 0
## 5 1 0
```

```
## 6          1          0
##  PaymentMethod.xElectronic.check PaymentMethod.xMailed.check Churn
## 1          1          0      0
## 2          0          1      0
## 3          0          1      1
## 4          0          0      0
## 5          1          0      1
## 6          1          0      1
##  tenure_bin.x1.2.years tenure_bin.x2.3.years tenure_bin.x3.4.years
## 1          0          0          0
## 2          0          1          0
## 3          0          0          0
## 4          0          0          1
## 5          0          0          0
## 6          0          0          0
##  tenure_bin.x4.5.years tenure_bin.x5.6.years
## 1          0          0
## 2          0          0
## 3          0          0
## 4          0          0
## 5          0          0
## 6          0          0
```

### #6.Train 데이터와 Validation(test) 데이터로 분할

```
set.seed(123)
indices = sample.split(telco_final$Churn, SplitRatio = 0.7)

train = telco_final[indices,] #훈련셋
validation = telco_final[!(indices),] #검증셋
```

## 2.2 Logistic Regression Classifiaction

로지스틱 회귀로 분류기를 설계하고, 단계적 선택법을 이용해 이를 반복적으로 개선한다.

### 2.2.1. 첫번째 Logistic Regression 모델 설계

이용서비스,계정정보,인구통계 정보를 독립변수로 Churn 종속변수를 분류하는 모델을 Logistic Regression로 설계

```
##
## Call:
## glm(formula = Churn ~ ., family = "binomial", data = train)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0959  -0.6765  -0.2674   0.6331   3.3919
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -4.711690    1.580952  -2.980 0.002880
##
## tenure        -2.419264    0.319733  -7.567 3.83e-14
## MonthlyCharges -1.523039    1.152472  -1.322 0.186321
## TotalCharges   0.236922    0.200225   1.183 0.236699
## gender         -0.004319    0.078799  -0.055 0.956292
```

```

## SeniorCitizen      0.342668    0.102039    3.358 0.000784
## Partner            0.049639    0.095330    0.521 0.602570
## Dependents        -0.148501    0.110427   -1.345 0.178693
## PhoneService      0.597310    0.785026    0.761 0.446729
## MultipleLines     0.529622    0.214832    2.465 0.013691
## InternetService.xFiber.optic  2.048017    0.960709    2.132 0.033025
## InternetService.xNo -2.254724    0.975826   -2.311 0.020856
## OnlineSecurity     -0.177982    0.215524   -0.826 0.408913
## OnlineBackup       0.126834    0.211103    0.601 0.547964
## DeviceProtection   0.255562    0.213872    1.195 0.232114
## TechSupport        -0.026070    0.215506   -0.121 0.903713
## StreamingTV        0.756980    0.393508    1.924 0.054396
## StreamingMovies     0.710913    0.394744    1.801 0.071711
## Contract.xOne.year -0.693044    0.130362   -5.316 1.06e-07
## Contract.xTwo.year -1.657976    0.233538   -7.099 1.25e-12
## PaperlessBilling   0.384841    0.090753    4.241 2.23e-05
## PaymentMethod.xCredit.card..automatic. -0.050640    0.137167   -0.369 0.711991
## PaymentMethod.xElectronic.check  0.249223    0.113791    2.190 0.028512
## PaymentMethod.xMailed.check -0.076660    0.138636   -0.553 0.580291
## tenure_bin.x1.2.years  0.147347    0.191048    0.771 0.440556
## tenure_bin.x2.3.years  0.964780    0.314795    3.065 0.002178
## tenure_bin.x3.4.years  1.955341    0.447108    4.373 1.22e-05
## tenure_bin.x4.5.years  2.801222    0.582303    4.811 1.50e-06
## tenure_bin.x5.6.years  3.598926    0.728585    4.940 7.83e-07
##
## (Intercept)          **
## tenure                ***
## MonthlyCharges
## TotalCharges
## gender
## SeniorCitizen        ***
## Partner
## Dependents
## PhoneService
## MultipleLines        *
## InternetService.xFiber.optic *
## InternetService.xNo  *
## OnlineSecurity
## OnlineBackup
## DeviceProtection
## TechSupport
## StreamingTV          .
## StreamingMovies      .
## Contract.xOne.year   ***
## Contract.xTwo.year   ***
## PaperlessBilling     ***
## PaymentMethod.xCredit.card..automatic. *
## PaymentMethod.xElectronic.check
## PaymentMethod.xMailed.check
## tenure_bin.x1.2.years
## tenure_bin.x2.3.years **
## tenure_bin.x3.4.years ***
## tenure_bin.x4.5.years ***
## tenure_bin.x5.6.years ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##
## Null deviance: 5699.5 on 4921 degrees of freedom
## Residual deviance: 3983.3 on 4893 degrees of freedom
## AIC: 4041.3
##

```

```
## Number of Fisher Scoring iterations: 6
```

## 초기설계모델

### 2.2.2. 단계적선택법 모델개선

#### 단계적 선택법 첫번째 개선결과

```
## Start: AIC=4041.29
## Churn ~ tenure + MonthlyCharges + TotalCharges + gender + SeniorCitizen +
## Partner + Dependents + PhoneService + MultipleLines + InternetService.xFiber.optic +
## InternetService.xNo + OnlineSecurity + OnlineBackup + DeviceProtection +
## TechSupport + StreamingTV + StreamingMovies + Contract.xOne.year +
## Contract.xTwo.year + PaperlessBilling + PaymentMethod.xCredit.card..automatic. +
## PaymentMethod.xElectronic.check + PaymentMethod.xMailed.check +
## tenure_bin.x1.2.years + tenure_bin.x2.3.years + tenure_bin.x3.4.years +
## tenure_bin.x4.5.years + tenure_bin.x5.6.years
##
##
## Df Deviance AIC
## - gender 1 3983.3 4039.3
## - TechSupport 1 3983.3 4039.3
## - PaymentMethod.xCredit.card..automatic. 1 3983.4 4039.4
## - Partner 1 3983.6 4039.6
## - PaymentMethod.xMailed.check 1 3983.6 4039.6
## - OnlineBackup 1 3983.7 4039.7
## - PhoneService 1 3983.9 4039.9
## - tenure_bin.x1.2.years 1 3983.9 4039.9
## - OnlineSecurity 1 3984.0 4040.0
## - DeviceProtection 1 3984.7 4040.7
## - TotalCharges 1 3984.7 4040.7
## - MonthlyCharges 1 3985.0 4041.0
## - Dependents 1 3985.1 4041.1
## <none> 3983.3 4041.3
## - StreamingMovies 1 3986.5 4042.5
## - StreamingTV 1 3987.0 4043.0
## - InternetService.xFiber.optic 1 3987.8 4043.8
## - PaymentMethod.xElectronic.check 1 3988.1 4044.1
## - InternetService.xNo 1 3988.6 4044.6
## - MultipleLines 1 3989.4 4045.4
## - tenure_bin.x2.3.years 1 3992.7 4048.7
## - SeniorCitizen 1 3994.5 4050.5
## - PaperlessBilling 1 4001.4 4057.4
## - tenure_bin.x3.4.years 1 4002.5 4058.5
## - tenure_bin.x4.5.years 1 4006.6 4062.6
## - tenure_bin.x5.6.years 1 4007.9 4063.9
## - Contract.xOne.year 1 4012.9 4068.9
## - tenure 1 4042.6 4098.6
## - Contract.xTwo.year 1 4046.3 4102.3
##
##
## Step: AIC=4039.29
## Churn ~ tenure + MonthlyCharges + TotalCharges + SeniorCitizen +
## Partner + Dependents + PhoneService + MultipleLines + InternetService.xFiber.optic +
## InternetService.xNo + OnlineSecurity + OnlineBackup + DeviceProtection +
## TechSupport + StreamingTV + StreamingMovies + Contract.xOne.year +
## Contract.xTwo.year + PaperlessBilling + PaymentMethod.xCredit.card..automatic. +
## PaymentMethod.xElectronic.check + PaymentMethod.xMailed.check +
## tenure_bin.x1.2.years + tenure_bin.x2.3.years + tenure_bin.x3.4.years +
## tenure_bin.x4.5.years + tenure_bin.x5.6.years
##
##
## Df Deviance AIC
```



```

## - TechSupport 1 3983.3 4037.3
## - PaymentMethod.xCredit.card..automatic. 1 3983.4 4037.4
## - Partner 1 3983.6 4037.6
## - PaymentMethod.xMailed.check 1 3983.6 4037.6
## - OnlineBackup 1 3983.7 4037.7
## - PhoneService 1 3983.9 4037.9
## - tenure_bin.x1.2.years 1 3983.9 4037.9
## - OnlineSecurity 1 3984.0 4038.0
## - TotalCharges 1 3984.7 4038.7
## - DeviceProtection 1 3984.7 4038.7
## - MonthlyCharges 1 3985.0 4039.0
## - Dependents 1 3985.1 4039.1
## <none> 3983.3 4039.3
## - StreamingMovies 1 3986.6 4040.6
## - StreamingTV 1 3987.0 4041.0
## + gender 1 3983.3 4041.3
## - InternetService.xFiber.optic 1 3987.9 4041.9
## - PaymentMethod.xElectronic.check 1 3988.1 4042.1
## - InternetService.xNo 1 3988.6 4042.6
## - MultipleLines 1 3989.4 4043.4
## - tenure_bin.x2.3.years 1 3992.7 4046.7
## - SeniorCitizen 1 3994.6 4048.6
## - PaperlessBilling 1 4001.4 4055.4
## - tenure_bin.x3.4.years 1 4002.6 4056.6
## - tenure_bin.x4.5.years 1 4006.6 4060.6
## - tenure_bin.x5.6.years 1 4007.9 4061.9
## - Contract.xOne.year 1 4012.9 4066.9
## - tenure 1 4042.6 4096.6
## - Contract.xTwo.year 1 4046.3 4100.3
##
## Step: AIC=4037.31
## Churn ~ tenure + MonthlyCharges + TotalCharges + SeniorCitizen +
## Partner + Dependents + PhoneService + MultipleLines + InternetService.xFiber.optic +
## InternetService.xNo + OnlineSecurity + OnlineBackup + DeviceProtection +
## StreamingTV + StreamingMovies + Contract.xOne.year + Contract.xTwo.year +
## PaperlessBilling + PaymentMethod.xCredit.card..automatic. +
## PaymentMethod.xElectronic.check + PaymentMethod.xMailed.check +
## tenure_bin.x1.2.years + tenure_bin.x2.3.years + tenure_bin.x3.4.years +
## tenure_bin.x4.5.years + tenure_bin.x5.6.years
##
## Df Deviance AIC
## - PaymentMethod.xCredit.card..automatic. 1 3983.4 4035.4
## - Partner 1 3983.6 4035.6
## - PaymentMethod.xMailed.check 1 3983.6 4035.6
## - tenure_bin.x1.2.years 1 3983.9 4035.9
## - OnlineBackup 1 3984.6 4036.6
## - OnlineSecurity 1 3984.6 4036.6
## - TotalCharges 1 3984.7 4036.7
## - Dependents 1 3985.1 4037.1
## <none> 3983.3 4037.3
## - PhoneService 1 3986.2 4038.2
## + TechSupport 1 3983.3 4039.3
## + gender 1 3983.3 4039.3
## - DeviceProtection 1 3987.5 4039.5
## - PaymentMethod.xElectronic.check 1 3988.1 4040.1
## - MonthlyCharges 1 3992.2 4044.2
## - tenure_bin.x2.3.years 1 3992.8 4044.8
## - SeniorCitizen 1 3994.6 4046.6
## - StreamingMovies 1 3996.4 4048.4
## - StreamingTV 1 3997.9 4049.9
## - MultipleLines 1 4000.8 4052.8
##
## - PaperlessBilling 1 4001.4 4053.4
## - tenure_bin.x3.4.years 1 4002.6 4054.6

```

```

## - InternetService.xNo          1  4004.4 4056.4
## - InternetService.xFiber.optic 1  4005.7 4057.7
## - tenure_bin.x4.5.years        1  4006.7 4058.7
## - tenure_bin.x5.6.years        1  4008.0 4060.0
## - Contract.xOne.year           1  4013.1 4065.1
## - tenure                       1  4042.7 4094.7
## - Contract.xTwo.year           1  4046.8 4098.8
##
## Step: AIC=4035.44
## Churn ~ tenure + MonthlyCharges + TotalCharges + SeniorCitizen +
##   Partner + Dependents + PhoneService + MultipleLines + InternetService.xFiber.optic +
##   InternetService.xNo + OnlineSecurity + OnlineBackup + DeviceProtection +
##   StreamingTV + StreamingMovies + Contract.xOne.year + Contract.xTwo.year +
##   PaperlessBilling + PaymentMethod.xElectronic.check + PaymentMethod.xMailed.check +
##   tenure_bin.x1.2.years + tenure_bin.x2.3.years + tenure_bin.x3.4.years +
##   tenure_bin.x4.5.years + tenure_bin.x5.6.years
##
##                                     Df Deviance   AIC
## - PaymentMethod.xMailed.check      1   3983.6 4033.6
## - Partner                          1   3983.7 4033.7
## - tenure_bin.x1.2.years             1   3984.0 4034.0
## - OnlineBackup                     1   3984.7 4034.7
## - OnlineSecurity                   1   3984.7 4034.7
## - TotalCharges                     1   3984.9 4034.9
## - Dependents                      1   3985.3 4035.3
## <none>                             1   3983.4 4035.4
## - PhoneService                    1   3986.3 4036.3
## + PaymentMethod.xCredit.card..automatic. 1   3983.3 4037.3
## + TechSupport                     1   3983.4 4037.4
## + gender                          1   3983.4 4037.4
## - DeviceProtection                 1   3987.6 4037.6
## - PaymentMethod.xElectronic.check  1   3991.9 4041.9
## - MonthlyCharges                  1   3992.3 4042.3
## - tenure_bin.x2.3.years            1   3992.8 4042.8
## - SeniorCitizen                   1   3994.8 4044.8
## - StreamingMovies                  1   3996.5 4046.5
## - StreamingTV                     1   3998.1 4048.1
## - MultipleLines                    1   4000.9 4050.9
## - PaperlessBilling                 1   4001.6 4051.6
## - tenure_bin.x3.4.years            1   4002.7 4052.7
## - InternetService.xNo              1   4004.5 4054.5
## - InternetService.xFiber.optic     1   4005.8 4055.8
## - tenure_bin.x4.5.years            1   4006.8 4056.8
## - tenure_bin.x5.6.years            1   4008.0 4058.0
## - Contract.xOne.year               1   4013.3 4063.3
## - tenure                          1   4042.8 4092.8
## - Contract.xTwo.year               1   4047.0 4097.0
##
## Step: AIC=4033.63
## Churn ~ tenure + MonthlyCharges + TotalCharges + SeniorCitizen +
##   Partner + Dependents + PhoneService + MultipleLines + InternetService.xFiber.optic +
##   InternetService.xNo + OnlineSecurity + OnlineBackup + DeviceProtection +
##   StreamingTV + StreamingMovies + Contract.xOne.year + Contract.xTwo.year +
##   PaperlessBilling + PaymentMethod.xElectronic.check + tenure_bin.x1.2.years +
##   tenure_bin.x2.3.years + tenure_bin.x3.4.years + tenure_bin.x4.5.years +
##   tenure_bin.x5.6.years
##
##                                     Df Deviance   AIC
## - Partner                          1   3983.9 4031.9
## - tenure_bin.x1.2.years            1   3984.2 4032.2
## - OnlineSecurity                   1   3984.9 4032.9
##
## - OnlineBackup                     1   3984.9 4032.9
## - TotalCharges                     1   3985.0 4033.0

```

```

## - Dependents 1 3985.5 4033.5
## <none> 3983.6 4033.6
## - PhoneService 1 3986.5 4034.5
## + PaymentMethod.xMailed.check 1 3983.4 4035.4
## + TechSupport 1 3983.6 4035.6
## + PaymentMethod.xCredit.card..automatic. 1 3983.6 4035.6
## + gender 1 3983.6 4035.6
## - DeviceProtection 1 3987.8 4035.8
## - MonthlyCharges 1 3992.5 4040.5
## - tenure_bin.x2.3.years 1 3993.0 4041.0
## - SeniorCitizen 1 3995.0 4043.0
## - PaymentMethod.xElectronic.check 1 3995.4 4043.4
## - StreamingMovies 1 3996.8 4044.8
## - StreamingTV 1 3998.3 4046.3
## - MultipleLines 1 4001.3 4049.3
## - PaperlessBilling 1 4001.9 4049.9
## - tenure_bin.x3.4.years 1 4002.8 4050.8
## - InternetService.xNo 1 4005.0 4053.0
## - InternetService.xFiber.optic 1 4006.2 4054.2
## - tenure_bin.x4.5.years 1 4006.8 4054.8
## - tenure_bin.x5.6.years 1 4008.1 4056.1
## - Contract.xOne.year 1 4013.5 4061.5
## - tenure 1 4043.0 4091.0
## - Contract.xTwo.year 1 4047.1 4095.1
##
## Step: AIC=4031.93
## Churn ~ tenure + MonthlyCharges + TotalCharges + SeniorCitizen +
## Dependents + PhoneService + MultipleLines + InternetService.xFiber.optic +
## InternetService.xNo + OnlineSecurity + OnlineBackup + DeviceProtection +
## StreamingTV + StreamingMovies + Contract.xOne.year + Contract.xTwo.year +
## PaperlessBilling + PaymentMethod.xElectronic.check + tenure_bin.x1.2.years +
## tenure_bin.x2.3.years + tenure_bin.x3.4.years + tenure_bin.x4.5.years +
## tenure_bin.x5.6.years
##
##
## Df Deviance AIC
## - tenure_bin.x1.2.years 1 3984.5 4030.5
## - OnlineSecurity 1 3985.2 4031.2
## - OnlineBackup 1 3985.2 4031.2
## - TotalCharges 1 3985.3 4031.3
## - Dependents 1 3985.5 4031.5
## <none> 3983.9 4031.9
## - PhoneService 1 3986.8 4032.8
## + Partner 1 3983.6 4033.6
## + PaymentMethod.xMailed.check 1 3983.7 4033.7
## + PaymentMethod.xCredit.card..automatic. 1 3983.9 4033.9
## + TechSupport 1 3983.9 4033.9
## + gender 1 3983.9 4033.9
## - DeviceProtection 1 3988.2 4034.2
## - MonthlyCharges 1 3992.9 4038.9
## - tenure_bin.x2.3.years 1 3993.2 4039.2
## - PaymentMethod.xElectronic.check 1 3995.7 4041.7
## - SeniorCitizen 1 3995.7 4041.7
## - StreamingMovies 1 3997.2 4043.2
## - StreamingTV 1 3998.7 4044.7
## - MultipleLines 1 4001.7 4047.7
## - PaperlessBilling 1 4002.3 4048.3
## - tenure_bin.x3.4.years 1 4002.9 4048.9
## - InternetService.xNo 1 4005.4 4051.4
## - InternetService.xFiber.optic 1 4006.7 4052.7
## - tenure_bin.x4.5.years 1 4007.0 4053.0
## - tenure_bin.x5.6.years 1 4008.2 4054.2
##
## - Contract.xOne.year 1 4013.8 4059.8
## - tenure 1 4043.0 4089.0

```

```
## - Contract.xTwo.year          1    4047.5 4093.5
##
## Step: AIC=4030.5
## Churn ~ tenure + MonthlyCharges + TotalCharges + SeniorCitizen +
##   Dependents + PhoneService + MultipleLines + InternetService.xFiber.optic +
##   InternetService.xNo + OnlineSecurity + OnlineBackup + DeviceProtection +
##   StreamingTV + StreamingMovies + Contract.xOne.year + Contract.xTwo.year +
##   PaperlessBilling + PaymentMethod.xElectronic.check + tenure_bin.x2.3.years +
##   tenure_bin.x3.4.years + tenure_bin.x4.5.years + tenure_bin.x5.6.years
##
##
##              Df Deviance    AIC
## - TotalCharges          1    3985.8 4029.8
## - OnlineBackup          1    3985.8 4029.8
## - OnlineSecurity        1    3985.8 4029.8
## - Dependents            1    3986.0 4030.0
## <none>                  3984.5 4030.5
## - PhoneService          1    3987.4 4031.4
## + tenure_bin.x1.2.years  1    3983.9 4031.9
## + Partner               1    3984.2 4032.2
## + PaymentMethod.xMailed.check 1    3984.3 4032.3
## + TechSupport           1    3984.5 4032.5
## + PaymentMethod.xCredit.card..automatic. 1    3984.5 4032.5
## + gender                1    3984.5 4032.5
## - DeviceProtection       1    3988.7 4032.7
## - MonthlyCharges        1    3993.4 4037.4
## - PaymentMethod.xElectronic.check 1    3996.2 4040.2
## - SeniorCitizen         1    3996.3 4040.3
## - StreamingMovies       1    3997.7 4041.7
## - StreamingTV           1    3999.2 4043.2
## - tenure_bin.x2.3.years  1    4000.3 4044.3
## - MultipleLines         1    4002.2 4046.2
## - PaperlessBilling       1    4002.9 4046.9
## - InternetService.xNo    1    4005.8 4049.8
## - InternetService.xFiber.optic 1    4007.2 4051.2
## - Contract.xOne.year     1    4014.3 4058.3
## - tenure_bin.x3.4.years  1    4022.8 4066.8
## - tenure_bin.x4.5.years  1    4032.5 4076.5
## - tenure_bin.x5.6.years  1    4033.0 4077.0
## - Contract.xTwo.year     1    4049.2 4093.2
## - tenure                 1    4099.3 4143.3
##
## Step: AIC=4029.75
## Churn ~ tenure + MonthlyCharges + SeniorCitizen + Dependents +
##   PhoneService + MultipleLines + InternetService.xFiber.optic +
##   InternetService.xNo + OnlineSecurity + OnlineBackup + DeviceProtection +
##   StreamingTV + StreamingMovies + Contract.xOne.year + Contract.xTwo.year +
##   PaperlessBilling + PaymentMethod.xElectronic.check + tenure_bin.x2.3.years +
##   tenure_bin.x3.4.years + tenure_bin.x4.5.years + tenure_bin.x5.6.years
##
##
##              Df Deviance    AIC
## - OnlineBackup          1    3987.0 4029.0
## - OnlineSecurity        1    3987.1 4029.1
## - Dependents            1    3987.3 4029.3
## <none>                  3985.8 4029.8
## - PhoneService          1    3988.5 4030.5
## + TotalCharges          1    3984.5 4030.5
## + tenure_bin.x1.2.years  1    3985.3 4031.3
## + Partner               1    3985.5 4031.5
## + PaymentMethod.xMailed.check 1    3985.6 4031.6
## + PaymentMethod.xCredit.card..automatic. 1    3985.7 4031.7
## + TechSupport           1    3985.7 4031.7
##
## + gender                1    3985.7 4031.7
## - DeviceProtection       1    3990.0 4032.0
```

```

## - MonthlyCharges          1  3993.8 4035.8
## - PaymentMethod.xElectronic.check 1  3997.1 4039.1
## - SeniorCitizen           1  3997.6 4039.6
## - StreamingMovies          1  3998.7 4040.7
## - StreamingTV              1  4000.2 4042.2
## - tenure_bin.x2.3.years    1  4002.5 4044.5
## - MultipleLines            1  4003.7 4045.7
## - PaperlessBilling         1  4003.9 4045.9
## - InternetService.xNo      1  4006.1 4048.1
## - InternetService.xFiber.optic 1  4007.9 4049.9
## - Contract.xOne.year       1  4015.9 4057.9
## - tenure_bin.x3.4.years    1  4028.0 4070.0
## - tenure_bin.x4.5.years    1  4041.1 4083.1
## - tenure_bin.x5.6.years    1  4046.5 4088.5
## - Contract.xTwo.year       1  4052.5 4094.5
## - tenure                   1  4131.0 4173.0
##
## Step: AIC=4029.04
## Churn ~ tenure + MonthlyCharges + SeniorCitizen + Dependents +
##   PhoneService + MultipleLines + InternetService.xFiber.optic +
##   InternetService.xNo + OnlineSecurity + DeviceProtection +
##   StreamingTV + StreamingMovies + Contract.xOne.year + Contract.xTwo.year +
##   PaperlessBilling + PaymentMethod.xElectronic.check + tenure_bin.x2.3.years +
##   tenure_bin.x3.4.years + tenure_bin.x4.5.years + tenure_bin.x5.6.years
##
##               Df Deviance   AIC
## - PhoneService      1  3988.5 4028.5
## - Dependents        1  3988.6 4028.6
## <none>              3987.0 4029.0
## + OnlineBackup      1  3985.8 4029.8
## + TotalCharges       1  3985.8 4029.8
## - DeviceProtection  1  3990.0 4030.0
## + TechSupport        1  3986.1 4030.1
## + tenure_bin.x1.2.years 1  3986.6 4030.6
## - OnlineSecurity     1  3990.7 4030.7
## + Partner            1  3986.7 4030.7
## + PaymentMethod.xMailed.check 1  3986.9 4030.9
## + PaymentMethod.xCredit.card..automatic. 1  3987.0 4031.0
## + gender             1  3987.0 4031.0
## - MonthlyCharges     1  3995.3 4035.3
## - PaymentMethod.xElectronic.check 1  3998.8 4038.8
## - SeniorCitizen      1  3999.4 4039.4
## - StreamingMovies     1  4001.0 4041.0
## - StreamingTV        1  4002.6 4042.6
## - tenure_bin.x2.3.years 1  4003.7 4043.7
## - MultipleLines      1  4004.6 4044.6
## - PaperlessBilling    1  4005.4 4045.4
## - InternetService.xNo 1  4011.2 4051.2
## - InternetService.xFiber.optic 1  4016.3 4056.3
## - Contract.xOne.year  1  4018.2 4058.2
## - tenure_bin.x3.4.years 1  4029.0 4069.0
## - tenure_bin.x4.5.years 1  4042.5 4082.5
## - tenure_bin.x5.6.years 1  4048.1 4088.1
## - Contract.xTwo.year  1  4056.9 4096.9
## - tenure              1  4131.3 4171.3
##
## Step: AIC=4028.46
## Churn ~ tenure + MonthlyCharges + SeniorCitizen + Dependents +
##   MultipleLines + InternetService.xFiber.optic + InternetService.xNo +
##   OnlineSecurity + DeviceProtection + StreamingTV + StreamingMovies +
##   Contract.xOne.year + Contract.xTwo.year + PaperlessBilling +
##
##   PaymentMethod.xElectronic.check + tenure_bin.x2.3.years +
##   tenure bin.x3.4.years + tenure bin.x4.5.years + tenure bin.x5.6.years

```

```

##
##
## Df Deviance AIC
## - Dependents 1 3990.0 4028.0
## - DeviceProtection 1 3990.1 4028.1
## + TechSupport 1 3986.3 4028.3
## <none> 3988.5 4028.5
## + PhoneService 1 3987.0 4029.0
## + TotalCharges 1 3987.4 4029.4
## + tenure_bin.x1.2.years 1 3988.0 4030.0
## + Partner 1 3988.1 4030.1
## + PaymentMethod.xMailed.check 1 3988.3 4030.3
## + PaymentMethod.xCredit.card..automatic. 1 3988.4 4030.4
## + OnlineBackup 1 3988.5 4030.5
## + gender 1 3988.5 4030.5
## - OnlineSecurity 1 3996.2 4034.2
## - PaymentMethod.xElectronic.check 1 4000.2 4038.2
## - SeniorCitizen 1 4001.0 4039.0
## - MonthlyCharges 1 4001.6 4039.6
## - MultipleLines 1 4004.9 4042.9
## - StreamingMovies 1 4005.0 4043.0
## - tenure_bin.x2.3.years 1 4005.1 4043.1
## - PaperlessBilling 1 4006.5 4044.5
## - StreamingTV 1 4007.3 4045.3
## - Contract.xOne.year 1 4020.3 4058.3
## - tenure_bin.x3.4.years 1 4030.6 4068.6
## - InternetService.xFiber.optic 1 4033.8 4071.8
## - InternetService.xNo 1 4041.7 4079.7
## - tenure_bin.x4.5.years 1 4044.1 4082.1
## - tenure_bin.x5.6.years 1 4049.7 4087.7
## - Contract.xTwo.year 1 4060.4 4098.4
## - tenure 1 4137.1 4175.1
##
## Step: AIC=4028.03
## Churn ~ tenure + MonthlyCharges + SeniorCitizen + MultipleLines +
## InternetService.xFiber.optic + InternetService.xNo + OnlineSecurity +
## DeviceProtection + StreamingTV + StreamingMovies + Contract.xOne.year +
## Contract.xTwo.year + PaperlessBilling + PaymentMethod.xElectronic.check +
## tenure_bin.x2.3.years + tenure_bin.x3.4.years + tenure_bin.x4.5.years +
## tenure_bin.x5.6.years
##
## Df Deviance AIC
## - DeviceProtection 1 3991.7 4027.7
## + TechSupport 1 3987.8 4027.8
## <none> 3990.0 4028.0
## + Dependents 1 3988.5 4028.5
## + PhoneService 1 3988.6 4028.6
## + TotalCharges 1 3988.9 4028.9
## + tenure_bin.x1.2.years 1 3989.6 4029.6
## + PaymentMethod.xMailed.check 1 3989.9 4029.9
## + PaymentMethod.xCredit.card..automatic. 1 3990.0 4030.0
## + gender 1 3990.0 4030.0
## + OnlineBackup 1 3990.0 4030.0
## + Partner 1 3990.0 4030.0
## - OnlineSecurity 1 3997.8 4033.8
## - PaymentMethod.xElectronic.check 1 4002.1 4038.1
## - MonthlyCharges 1 4003.4 4039.4
## - SeniorCitizen 1 4004.4 4040.4
## - MultipleLines 1 4006.8 4042.8
## - tenure_bin.x2.3.years 1 4006.8 4042.8
## - StreamingMovies 1 4006.9 4042.9
## - PaperlessBilling 1 4008.4 4044.4
##
## - StreamingTV 1 4008.9 4044.9
## - Contract.xOne.year 1 4022.6 4058.6

```

```

## - tenure_bin.x3.4.years          1  4032.5 4068.5
## - InternetService.xFiber.optic    1  4036.2 4072.2
## - InternetService.xNo             1  4043.8 4079.8
## - tenure_bin.x4.5.years          1  4046.4 4082.4
## - tenure_bin.x5.6.years          1  4052.0 4088.0
## - Contract.xTwo.year             1  4063.5 4099.5
## - tenure                          1  4141.0 4177.0
##
## Step: AIC=4027.71
## Churn ~ tenure + MonthlyCharges + SeniorCitizen + MultipleLines +
##   InternetService.xFiber.optic + InternetService.xNo + OnlineSecurity +
##   StreamingTV + StreamingMovies + Contract.xOne.year + Contract.xTwo.year +
##   PaperlessBilling + PaymentMethod.xElectronic.check + tenure_bin.x2.3.years +
##   tenure_bin.x3.4.years + tenure_bin.x4.5.years + tenure_bin.x5.6.years
##
##                                     Df Deviance   AIC
## + TechSupport                      1   3989.3 4027.3
## <none>                             1   3991.7 4027.7
## + DeviceProtection                 1   3990.0 4028.0
## + Dependents                       1   3990.1 4028.1
## + TotalCharges                     1   3990.5 4028.5
## + tenure_bin.x1.2.years            1   3991.3 4029.3
## + PaymentMethod.xMailed.check      1   3991.6 4029.6
## + PhoneService                     1   3991.6 4029.6
## + PaymentMethod.xCredit.card..automatic. 1   3991.7 4029.7
## + OnlineBackup                     1   3991.7 4029.7
## + gender                           1   3991.7 4029.7
## + Partner                           1   3991.7 4029.7
## - OnlineSecurity                   1   4000.6 4034.6
## - MonthlyCharges                   1   4003.4 4037.4
## - PaymentMethod.xElectronic.check  1   4003.8 4037.8
## - SeniorCitizen                    1   4006.2 4040.2
## - MultipleLines                    1   4007.2 4041.2
## - StreamingMovies                   1   4007.7 4041.7
## - tenure_bin.x2.3.years            1   4008.1 4042.1
## - StreamingTV                      1   4009.6 4043.6
## - PaperlessBilling                  1   4009.8 4043.8
## - Contract.xOne.year                1   4023.5 4057.5
## - tenure_bin.x3.4.years            1   4034.1 4068.1
## - InternetService.xFiber.optic     1   4036.8 4070.8
## - InternetService.xNo              1   4043.9 4077.9
## - tenure_bin.x4.5.years            1   4047.4 4081.4
## - tenure_bin.x5.6.years            1   4053.1 4087.1
## - Contract.xTwo.year                1   4064.2 4098.2
## - tenure                           1   4141.1 4175.1
##
## Step: AIC=4027.28
## Churn ~ tenure + MonthlyCharges + SeniorCitizen + MultipleLines +
##   InternetService.xFiber.optic + InternetService.xNo + OnlineSecurity +
##   StreamingTV + StreamingMovies + Contract.xOne.year + Contract.xTwo.year +
##   PaperlessBilling + PaymentMethod.xElectronic.check + tenure_bin.x2.3.years +
##   tenure_bin.x3.4.years + tenure_bin.x4.5.years + tenure_bin.x5.6.years +
##   TechSupport
##
##                                     Df Deviance   AIC
## <none>                             1   3989.3 4027.3
## - TechSupport                      1   3991.7 4027.7
## + Dependents                       1   3987.7 4027.7
## + DeviceProtection                 1   3987.8 4027.8
## + TotalCharges                     1   3987.9 4027.9
## + tenure_bin.x1.2.years            1   3988.9 4028.9
##
## + PhoneService                     1   3989.1 4029.1
## + PaymentMethod.xMailed.check      1   3989.2 4029.2

```

```
## + OnlineBackup 1 3989.2 4029.2
## + PaymentMethod.xCredit.card..automatic. 1 3989.3 4029.3
## + gender 1 3989.3 4029.3
## + Partner 1 3989.3 4029.3
## - MonthlyCharges 1 3996.2 4032.2
## - OnlineSecurity 1 3998.8 4034.8
## - PaymentMethod.xElectronic.check 1 4000.9 4036.9
## - MultipleLines 1 4002.2 4038.2
## - StreamingMovies 1 4002.8 4038.8
## - SeniorCitizen 1 4002.9 4038.9
## - StreamingTV 1 4004.6 4040.6
## - tenure_bin.x2.3.years 1 4005.8 4041.8
## - PaperlessBilling 1 4007.6 4043.6
## - Contract.xOne.year 1 4019.1 4055.1
## - InternetService.xFiber.optic 1 4021.1 4057.1
## - tenure_bin.x3.4.years 1 4031.5 4067.5
## - InternetService.xNo 1 4038.0 4074.0
## - tenure_bin.x4.5.years 1 4044.7 4080.7
## - tenure_bin.x5.6.years 1 4050.1 4086.1
## - Contract.xTwo.year 1 4055.8 4091.8
## - tenure 1 4137.1 4173.1
```

```
##
## Call:
## glm(formula = Churn ~ tenure + MonthlyCharges + SeniorCitizen +
##       MultipleLines + InternetService.xFiber.optic + InternetService.xNo +
##       OnlineSecurity + StreamingTV + StreamingMovies + Contract.xOne.year +
##       Contract.xTwo.year + PaperlessBilling + PaymentMethod.xElectronic.check +
##       tenure_bin.x2.3.years + tenure_bin.x3.4.years + tenure_bin.x4.5.years +
##       tenure_bin.x5.6.years + TechSupport, family = "binomial",
##       data = train)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1461  -0.6661  -0.2733   0.6249   3.3018
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -3.37123    0.27189 -12.399 < 2e-16 ***
## tenure         -2.07853    0.17751 -11.709 < 2e-16 ***
## MonthlyCharges -0.52335    0.19800  -2.643 0.008213 **
## SeniorCitizen   0.37111    0.10031   3.700 0.000216 ***
## MultipleLines   0.38516    0.10755   3.581 0.000342 ***
## InternetService.xFiber.optic 1.28144    0.22881   5.601 2.14e-08 ***
## InternetService.xNo -1.45816    0.20985  -6.949 3.69e-12 ***
## OnlineSecurity  -0.33178    0.10796  -3.073 0.002118 **
## StreamingTV      0.46276    0.11865   3.900 9.61e-05 ***
## StreamingMovies  0.42482    0.11585   3.667 0.000246 ***
## Contract.xOne.year -0.69055    0.12950  -5.333 9.68e-08 ***
##
## Contract.xTwo.year -1.68164    0.23166  -7.259 3.89e-13 ***
## PaperlessBilling   0.38535    0.09046   4.260 2.05e-05 ***
## PaymentMethod.xElectronic.check 0.28932    0.08445   3.426 0.000613 ***
## tenure_bin.x2.3.years 0.78164    0.19291   4.052 5.08e-05 ***
## tenure_bin.x3.4.years 1.72393    0.26800   6.432 1.26e-10 ***
## tenure_bin.x4.5.years 2.52442    0.34433   7.331 2.28e-13 ***
## tenure_bin.x5.6.years 3.30141    0.43049   7.669 1.73e-14 ***
## TechSupport       -0.17022    0.10928  -1.558 0.119308
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
```



```
## Null deviance: 5699.5 on 4921 degrees of freedom
## Residual deviance: 3989.3 on 4903 degrees of freedom
## AIC: 4027.3
##
## Number of Fisher Scoring iterations: 6
```

### 단계적 선택법 첫번째 개선모델

### VIF 적용

#####중복예측변수나 다중공선성이 높은 변수를 제거

### VIF적용1

```
## tenure MonthlyCharges
## 15.076710 20.162188
## SeniorCitizen MultipleLines
## 1.100707 1.838876
## InternetService.xFiber.optic InternetService.xNo
## 8.225529 2.489968
## OnlineSecurity StreamingTV
## 1.246308 2.206225
## StreamingMovies Contract.xOne.year
## 2.122071 1.386381
## Contract.xTwo.year PaperlessBilling
## 1.371229 1.125064
## PaymentMethod.xElectronic.check tenure_bin.x2.3.years
## 1.149683 2.656673
## tenure_bin.x3.4.years tenure_bin.x4.5.years
## 4.106103 6.527227
## tenure_bin.x5.6.years TechSupport
## 8.568998 1.332821
```

```
##
## Call:
## glm(formula = Churn ~ tenure + MonthlyCharges + SeniorCitizen +
## Partner + InternetService.xFiber.optic + InternetService.xNo +
## OnlineSecurity + OnlineBackup + TechSupport + StreamingTV +
## Contract.xOne.year + Contract.xTwo.year + PaperlessBilling +
## PaymentMethod.xElectronic.check + tenure_bin.x1.2.years +
## tenure_bin.x5.6.years, family = "binomial", data = train)
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
## -1.9863 -0.6589 -0.2775 0.6424 3.2758
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.48117 0.17354 -8.535 < 2e-16 ***
## tenure -0.88438 0.07316 -12.088 < 2e-16 ***
## MonthlyCharges 0.08404 0.15461 0.544 0.586731
## SeniorCitizen 0.37893 0.09935 3.814 0.000137 ***
## Partner -0.02734 0.08496 -0.322 0.747590
## InternetService.xFiber.optic 0.73441 0.20004 3.671 0.000241 ***
## InternetService.xNo -1.07362 0.19318 -5.558 2.73e-08 ***
## OnlineSecurity -0.46885 0.10462 -4.481 7.42e-06 ***
## OnlineBackup -0.12936 0.09465 -1.367 0.171691
## TechSupport -0.29675 0.10623 -2.793 0.005216 **
## StreamingTV 0.32029 0.11573 2.768 0.005647 **
```

```
## Contract.xOne.year      -0.69301    0.12783   -5.421 5.92e-08 ***
## Contract.xTwo.year      -1.72148    0.23296   -7.390 1.47e-13 ***
## PaperlessBilling         0.40142    0.08954    4.483 7.36e-06 ***
## PaymentMethod.xElectronic.check 0.32498    0.08352    3.891 9.99e-05 ***
## tenure_bin.x1.2.years   -0.52142    0.10707   -4.870 1.12e-06 ***
## tenure_bin.x5.6.years    0.54321    0.19325    2.811 0.004940 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 5699.5 on 4921 degrees of freedom
## Residual deviance: 4042.5 on 4905 degrees of freedom
## AIC: 4076.5
##
## Number of Fisher Scoring iterations: 6
```

```
##          tenure          MonthlyCharges
##          2.540970          12.310900
##          SeniorCitizen          Partner
##          1.092995          1.133093
##          InternetService.xFiber.optic          InternetService.xNo
##          6.398204          2.142763
##          OnlineSecurity          OnlineBackup
##          1.180975          1.288093
##          TechSupport          StreamingTV
##          1.271133          2.122655
##          Contract.xOne.year          Contract.xTwo.year
##          1.360444          1.392469
##          PaperlessBilling PaymentMethod.xElectronic.check
##          1.120396          1.141949
##          tenure_bin.x1.2.years          tenure_bin.x5.6.years
##          1.028052          1.757701
```

p-value가 높게 나오는 DeviceProtection제거

## VIF적용2

```
##
## Call:
## glm(formula = Churn ~ tenure + MonthlyCharges + SeniorCitizen +
## Partner + InternetService.xFiber.optic + InternetService.xNo +
## OnlineSecurity + OnlineBackup + TechSupport + Contract.xOne.year +
## Contract.xTwo.year + PaperlessBilling + PaymentMethod.xElectronic.check +
## tenure_bin.x1.2.years + tenure_bin.x5.6.years, family = "binomial",
## data = train)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9552  -0.6643  -0.2791   0.6644   3.2717
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.24627    0.15025  -8.295 < 2e-16 ***
## tenure         -0.87338    0.07296 -11.970 < 2e-16 ***
## MonthlyCharges  0.35114    0.12173   2.885 0.003918 **
## SeniorCitizen   0.37827    0.09923   3.812 0.000138 ***
## Partner        -0.02668    0.08487  -0.314 0.753275
## InternetService.xFiber.optic 0.45825    0.17270   2.653 0.007967 **
## InternetService.xNo -0.91779    0.18585  -4.938 7.88e-07 ***
```

```
## OnlineSecurity          -0.52247    0.10282   -5.081 3.75e-07 ***
## OnlineBackup            -0.16988    0.09346   -1.818 0.069109 .
## TechSupport             -0.33438    0.10525   -3.177 0.001488 **
## Contract.xOne.year      -0.67704    0.12751   -5.310 1.10e-07 ***
## Contract.xTwo.year      -1.70720    0.23269   -7.337 2.19e-13 ***
## PaperlessBilling         0.41657    0.08924    4.668 3.04e-06 ***
## PaymentMethod.xElectronic.check 0.34692    0.08304    4.178 2.95e-05 ***
## tenure_bin.x1.2.years   -0.52044    0.10684   -4.871 1.11e-06 ***
## tenure_bin.x5.6.years    0.52669    0.19290    2.730 0.006326 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 5699.5  on 4921  degrees of freedom
## Residual deviance: 4050.2  on 4906  degrees of freedom
## AIC: 4082.2
##
## Number of Fisher Scoring iterations: 6
```

```
##              tenure              MonthlyCharges
##              2.528728              7.598828
##      SeniorCitizen              Partner
##              1.092734              1.133297
##      InternetService.xFiber.optic      InternetService.xNo
##              4.776780              1.985836
##      OnlineSecurity              OnlineBackup
##              1.143833              1.258626
##      TechSupport              Contract.xOne.year
##              1.252283              1.358085
##      Contract.xTwo.year              PaperlessBilling
##              1.390401              1.114992
##      PaymentMethod.xElectronic.check      tenure_bin.x1.2.years
##              1.131205              1.027896
##      tenure_bin.x5.6.years
##              1.752778
```

p-value가 높게 나오는 StreamingTV제거

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.004518 0.041437 0.186763 0.268617 0.470467 0.868907
```

```
##      pred_churn
## actual_churn  No  Yes
##      No  1382  167
##      Yes   275  286
```

## Logistic Regression 개선모델에서 50%의 cutoff를 사용한 결과

- Accuracy = 0.79
- Sensitivity = 0.56
- Specificity = 0.89

```
## Accuracy  
## 0.7582938
```

```
## Sensitivity  
## 0.7308378
```

```
## Specificity  
## 0.7682376
```

### Logistic Regression 개선모델에서 32%의 cutoff를 사용한 최종결과

- Accuracy = 0.76
- Sensitivity = 0.76
- Specificity = 0.76

# 3. 결론 \*\*\* 해당 보고서에선 이용서비스,계정정보,인구통계학정보를 독립변수로 이탈고객 train set을 학습시켜 Logistics Regression 분류기를 만들어내고, 이 분류기로 validation set의 이탈고객 여부를 판별하는 분류모델 설계, 개선 및 분석 과정을 실시해보았다. 해당 최종 분류기의 성능은 아래와 같다.

### Logistic Regression 개선모델에서 32%의 cutoff를 사용한 최종결과

- Accuracy = 0.76
- Sensitivity = 0.76
- Specificity = 0.76