**Telco Churn Project**

1. **Preprocessing the Data**

* We convert the dependent variable from string type to number.
* Creating three separate subsets of data for phone, Internet, and both to analyze each group separately.

1. **What predictors do you think contribute to the churn of (i) only telephone customers, (ii) only Internet service customers, and (iii) customers who subscribe to both phone and Internet services?**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Predictor** | **Rationale** | **Phone** | **Internet** | **Both** |
| Gender | Ideally, gender should not matter, but it is possible that men preferences may vary from women for phone and Internet | Yes (?) | Yes (?) | Yes (?) |
| Senior citizen | Senior citizens are less likely to be aware of alternative services and therefore less likely to churn. | Yes (-) | Yes (-) | Yes (-) |
| Partner and dependents | Family plans may have less propensity to churn, because of the difficulty to meet varied needs and preferences. | Yes (-) | Yes (-) | Yes (-) |
| Tenure | Customers with longer tenure may be satisfied with the service, and may be less likely to churn. | Yes (-) | Yes (-) | Yes (-) |
| Multiple lines | Having multiple lines will make it difficult to churn as replacing multiple lines may be expensive. Applies to phone | Yes (-) | No | Yes (-) |
| Contract | Longer term contracts will make it harder to churn. | Yes (-) | Yes (-) | Yes (-) |
| Paperless billing | It is unknown how this will impact churn, but it’s worth checking it out because this predictor is actionable | Yes (?) | Yes (?) | Yes (?) |
| Payment method | It is unknown how this will impact churn, but it’s worth checking it out because this predictor is actionable | Yes (?) | Yes (?) | Yes (?) |
| Monthly charges | More expensive charges may be a reason to churn quicker. | Yes (+) | Yes (+) | Yes (+) |
| Internet Services | Subscribers of DSL vs fiber optic services may have different propensity to churn; does not apply to phone service. | No | Yes (?) | Yes (?) |
| Online Security | Internet users may want network security. Does not apply to phone only users. | No | Yes (-) | Yes (-) |
| Online backup | Internet users may want online backup. Does not apply to phone only users. | No | Yes (-) | Yes (-) |
| Device protection | Internet users may want device protection. Does not apply to phone only users. | No | Yes (-) | Yes (-) |
| Techsupport | Internet users may need tech. Does not apply to phone only users. | No | Yes (-) | Yes (-) |
| Streaming TV | Internet users may want streaming TV. Does not apply to phone only users. | No | Yes (-) | Yes (-) |
| Streaming Movies | Internet users may want streaming movies. Does not apply to phone only users. | No | Yes (-) | Yes (-) |

**# Modeling the data**

#' Logistic models

phone\_model <- glm(Churn ~ gender + SeniorCitizen + Partner + Dependents +

tenure + MultipleLines + Contract + PaperlessBilling +

PaymentMethod + MonthlyCharges, data=train\_phone,

family=binomial (link="logit"))

internet\_model <- glm(Churn ~ gender + SeniorCitizen + Partner + Dependents +

tenure + OnlineSecurity + OnlineBackup + DeviceProtection +

TechSupport + StreamingTV + StreamingMovies + Contract +

PaperlessBilling + PaymentMethod + MonthlyCharges,

data=train\_internet, family=binomial (link="logit"))

both\_model <- glm(Churn ~ gender + SeniorCitizen + Partner + Dependents +

tenure + MultipleLines + OnlineSecurity + OnlineBackup +

DeviceProtection + TechSupport + StreamingTV + StreamingMovies +

Contract + PaperlessBilling + PaymentMethod + MonthlyCharges,

data=train\_both, family=binomial (link="logit"))

library(stargazer)

stargazer(phone\_model, internet\_model, both\_model, type="text", single.row=TRUE)

============================================================================================

Dependent variable: Churn

-----------------------------------------------------

Phone Only Internet Only Both phone & Internet

--------------------------------------------------------------------------------------------

genderMale -0.033 (0.249) 0.195 (0.251) -0.081 (0.083)

SeniorCitizen 0.975 (0.669) 0.196 (0.328) 0.195\* (0.104)

Partner -0.236 (0.375) 0.503\* (0.290) -0.076 (0.098)

Dependents -0.247 (0.365) -0.723\*\* (0.342) 0.017 (0.115)

tenure -0.047\*\*\* (0.014) -0.039\*\*\* (0.008) -0.034\*\*\* (0.003)

MultipleLines -0.525 (1.283) 0.128 (0.099)

OnlineSecurity -0.085 (0.812) -0.610\*\*\* (0.103)

OnlineBackup -0.015 (0.831) -0.287\*\*\* (0.096)

DeviceProtection 0.129 (0.830) -0.143 (0.098)

TechSupport -0.488 (0.855) -0.548\*\*\* (0.104)

StreamingTV 0.576 (1.602) -0.082 (0.109)

StreamingMovies 0.599 (1.565) -0.138 (0.109)

ContractOne year -1.290\*\*\* (0.443) -0.679\* (0.383) -0.503\*\*\* (0.138)

ContractTwo year -1.487\*\* (0.611) -1.811\*\* (0.795) -1.141\*\*\* (0.233)

PaperlessBilling 0.271 (0.263) 0.131 (0.276) 0.351\*\*\* (0.097)

PaymentMethodBank transfer (automatic) 0.405 (0.352) 0.341 (0.424) -0.195 (0.155)

PaymentMethodCredit card (automatic) -1.012\* (0.544) -0.041 (0.422) -0.048 (0.154)

PaymentMethodElectronic check 0.379 (0.377) 0.651\*\* (0.332) 0.272\*\* (0.128)

MonthlyCharges 0.031 (0.235) -0.029 (0.157) 0.035\*\*\* (0.004)

Constant -1.776 (4.699) 0.402 (3.914) -2.140\*\*\* (0.268)

--------------------------------------------------------------------------------------------

Observations 1,144 511 3,626

Log Likelihood -225.198 -207.096 -1,741.185

Akaike Inf. Crit. 478.396 452.191 3,522.369 ============================================================================================

1. **What are the top three predictors of churn of (i) only telephone customers, (ii) only Internet service customers, and (iii) customers who subscribe to both phone and Internet services. Explain using marginal effects how much each predictor contributes to churn probability.**

**Phone only:**

|  |  |  |  |
| --- | --- | --- | --- |
| Predictor | β | exp(β) | Interpretation |
| Two-year contract | -1.487 | 0.226 | Having a 2-year contract reduces the odds of churn by 77.4% over customers with month-to-month service. |
| One year contract | -1.290 | 0.275 | Having a 1-year contract reduces the odds of churn by 72.5% over customers with month-to-month service. |
| CC payment | -1.012 | 0.363 | Customers who pay by credit card (e.g., autopay) have 65.7% less odds of churn than people who pay using other means. |

**Internet only**

|  |  |  |  |
| --- | --- | --- | --- |
| Predictor | β | exp(β) | Interpretation |
| Two-year contract | -1.811 | 0.163 | Having a 2-year contract reduces odds of churn by 83.7% over customers with month-to-month service. |
| Dependents | -0.723 | 0.485 | Each additional dependent reduces odds of customer churn by 51.5%. However, this variable is not “actionable.” |
| One-year contract | -0.679 | 0.507 | Having a 1-year contract reduces odds of churn by 49.3% over customers with month-to-month service. |

**Both phone and Internet:**

|  |  |  |  |
| --- | --- | --- | --- |
| Predictor | β | exp(β) | Interpretation |
| Two year contract | -1.141 | 0.319 | Having a 2-year contract reduces odds of churn by 68.1% over customers with month-to-month service. |
| Online security | -0.610 | 0.543 | Customers with online security have 45.7% less odds of churn than customers without online security. |
| Tech support | -0.548 | 0.578 | Customers with tech support have 42.2% less odds of churn than customers without tech support. |

1. **Classification matrix against test data:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type of model | Recall | Precision | F1 score | AUC |
| Phone | 0.90 | 1 | 0.86 | 0.713 |
| Internet | 0.86 | 0.82 | 0.82 | 0.657 |
| Both | 0.86 | 0.82 | 0.82 | 0.683 |

The estimates below suggest that the phone model has greater predictive accuracy, and internet and both models have almost similar accuracy.