Logistic Regression On Churn Based On Hour of Day, Days Since Prior Order, and Add to Cart Order

2024-04-19

# Load the necessary library

library(dplyr)

#Load Data Sets

order\_products\_train <- read.csv(“path\_to\_order\_products\_train.csv”) orders <- read.csv(“path\_to\_orders.csv”)

# Merge datasets on order\_id

combined\_data <- merge(order\_products\_train, orders, by = “order\_id”)

# Create churn variable

combined\_datareordered == 0)

# Select relevant columns for the logistic regression

final\_data <- combined\_data %>% select(churn, order\_hour\_of\_day, days\_since\_prior\_order, add\_to\_cart\_order) # Add other predictors as needed

# Logistic regression

logistic\_model <- glm(churn ~ ., data = final\_data, family = “binomial”)

# Summary of the logistic regression model

summary(logistic\_model)

A white screen with black text

Description automatically generated

* Orders made later in the day are slightly more likely to result in churn.
* The more days there are since a customer's prior order, the more likely they are to churn.
* Items added to the cart later in the order process are more likely to be associated with churn.